The Role of Situation Models in the Continued Influence Effect

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Carolin Vanessa Hey
aus Düsseldorf

Düsseldorf, Mai 2026

aus dem Institut für Experimentelle Psychologie
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Berichterstattende:
1. Professorin Dr. Ute Johanna Bayen
2. Professor Dr. Raoul Bell

Tag der mündlichen Prüfung: 29.06.2026

Eigenständigkeitserklärung

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist. Die Dissertation wurde in der vorliegenden oder ähnlichen Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, den 18. Mai 2026 Unterschrift: _____

Table of Contents

Abstract 5
1 Introduction 7
1.1 Measuring the CIE 8
1.2 Theoretical Accounts of the CIE 8
1.3 The Current Work 10
2 A New Assessment of Situational Models 10
2.1 Experiment 1 11
3 Source Credibility Affects the CIE 12
3.1 Theoretical and Empirical Background 12
3.2 Overview of the Studies 13
3.2.1 Experiment 2.1 13
3.2.2 Experiment 2.2 14
3.3 Discussion 14
4 Determinants of Effects on the CIE 15
4.1 Overview of the Studies 15
4.1.1 Experiment 3.1 16
4.1.2 Experiment 3.2 19
4.2 Discussion 20
5 The Relationships Between the Situation Models and the CIE 21
5.1 Structural Equation Modelling of Experiment 3.2 21
5.1.1 Early Source Disclosure 22
5.1.2 Late Source Disclosure 23
5.2 Discussion 23
6 General Discussion 24
References 26
Appendix 32
Paper 1
Experiment 2.1
Experiment 2.2
Paper 2
Experiment 3.1
Experiment 3.2
Online Supplement 2 “Validation Experiments” of Paper 2
Experiment 1

Abstract

Information that is later proven to be false can continue to influence memory, beliefs, and inferences even after it has been retracted. This phenomenon is known as the *Continued Influence Effect* (CIE; e.g., Walter & Tukachinsky, 2020; Wilkes & Leatherbarrow, 1988). Despite extensive research demonstrating its robustness, there is ongoing debate about the underlying cognitive mechanisms. Two prominent accounts propose different cognitive processes to underly the effect (e.g., Ecker et al., 2022; Lewandowsky et al., 2012; Sanderson & Ecker, 2020): The *retrieval-failure* account attributes the CIE to failures during memory retrieval, whereas the *model-updating* account locates its causes in encoding processes, specifically in the construction and updating of situation models. Critically, direct empirical evidence for the model-updating account has been lacking.

The present dissertation addressed this gap by pursuing three objectives. First, I developed and validated a method for assessing situation models within a standard CIE paradigm. Second, I identified source credibility as a manipulation that affects situation-model construction and updating as well as the CIE, enabling experimental investigation of their relationship. Third, I used this manipulation to test predictions derived from the model-updating account against those of the retrieval-failure account.

In Experiment 1, I developed a question-based assessment of participants' situation models. The assessment was validated by experimentally manipulating agreement with the original information and demonstrating that it reliably differentiated between levels of agreement and predicted subsequent CIE responses.

Experiments 2.1 and 2.2 established source credibility as a suitable manipulation to test the model-updating account. Informant credibility and retractor credibility were manipulated orthogonally. Across both experiments, they exerted independent and antagonistic effects on the CIE: High informant credibility led to a stronger CIE, whereas high retractor credibility led to a weaker CIE. Importantly, these effects did not interact, indicating that credibility was applied selectively to the individual critical statements. This pattern is consistent with the assumption that source credibility can independently influence the integration of the original information into the situation model and its persistence therein following the retraction.

In Experiments 3.1 and 3.2, I used credibility-disclosure timing to disentangle encoding and retrieval processes and to test the assumption that the CIE arises during encoding. In Experiment 3.1, source credibility affected the CIE regardless of credibility disclosure timing, suggesting a possible influence at retrieval. However, this pattern may

reflect retrospective integration of credibility information or short-term activation at immediate retrieval. Experiment 3.2 addressed these issues by separating credibility disclosure and situation-model construction and introducing a delay before the CIE test. Under these conditions, source credibility affected the CIE only when it was available during encoding; when disclosed afterward, it had no effect on either the situation model or the CIE. Finally, I used a structural equation modeling to directly test the relationships between situation models and the CIE. The results showed that the situation model held after retraction consistently predicted the CIE and that source-credibility effects were mediated through the situation models when present.

Taken together, the findings provide direct empirical support for the model-updating account of the CIE. They demonstrate that the effect arises from the way information is integrated into and maintained within situation models, rather than from failures during retrieval. At the same time, some findings suggest that retrieval processes may still contribute under certain conditions. Overall, this dissertation advances the understanding of how misinformation continues to influence cognition despite retraction and clarifies the mechanisms underlying the CIE.

1 Introduction[1]

In an age of unprecedented information flow, misinformation (i.e., false or misleading information presented as fact) has emerged as a persistent cognitive and societal challenge. Cognitive psychology has identified a range of phenomena illustrating the memory's vulnerability to false information, for example, the misinformation effect where post-event misinformation alters memory for witnessed events (Loftus, 2005; Loftus et al., 1978). This dissertation focusses on the phenomenon that information continues to affect memory, belief, and inferences even after it has been retracted: the *Continued Influence Effect* (CIE; e.g., Ecker et al., 2010; H. M. Johnson & Seifert, 1994; Rich et al., 2023; for meta-analyses, see Chan et al., 2017; Walter & Tukachinsky, 2020). Typically, retractions lead to less influence of the original information compared to original information that was not retracted (e.g., Buczel et al., 2024; Ecker & Ang, 2019; Ecker et al., 2014, 2017, 2020; Guillory & Geraci, 2016; Rich & Zaragoza, 2016). Critically, however, in a myriad of studies with various materials, retractions did not eliminate the continued influence of the original information, that is, participants agreed with, believed in, and made inferences based on the original information despite the retraction (e.g., Buczel et al., 2022, 2024, 2025; Ecker & Antonio, 2021; Ecker et al., 2021, 2024; Ithisuphalap et al., 2020; Kan et al., 2021; McIlhiney et al., 2022; Miller et al., 2022; Rich et al., 2023; Sanderson et al., 2021, 2022; Susmann & Wegener, 2022, 2023; Tay et al., 2022; van Huijstee et al., 2022; for meta-analyses, see Chan et al., 2017; and Walter & Tukachinsky, 2020).

This lasting effect of misinformation after retraction is well established; however, there is an ongoing debate about the mechanisms underlying the CIE. In the present dissertation, I examined the cognitive mechanisms by testing predictions of the two prominent theoretical accounts of the CIE. The work pursued three main objectives: First, I developed an assessment of participants' agreement with the original information over the course of information accumulation. Second, I identified a manipulation that influences both information processing and the CIE, enabling experimental investigation of their relationship. Third, I used this manipulation to test predictions derived from the model-updating account against those of the retrieval-failure account.

[1] The introduction was partially adopted from Paper 1 and Paper 2 in the Appendix.

1.1 Measuring the CIE

The standard CIE paradigm consists of a study phase and a test phase. In the study phase, participants read an unfolding report about a fictitious event containing two critical pieces of information. Typically, the *original information* is an alleged cause of the described event. Following further non-critical information, the report later contains a retraction of the original information. For example, O'Rear and Radvansky (2020) used a news report about a food recall with the original information being that the recall was necessary because the food was contaminated due to unsanitary equipment in the processing plant and the retraction being that there was no evidence of unsanitary conditions. In the test phase, participants typically answer questions on the event that require them to go beyond the information explicitly mentioned in the report. Answers in agreement with the original information (e.g., that ensuring better sanitation in the processing plant could avoid such an event in the future) are scored as showing continued influence of the original information despite the retraction.

1.2 Theoretical Accounts of the CIE

Two cognitive accounts dominate the ongoing discussion about the mechanisms underlying the CIE: the retrieval-based *retrieval-failure account* and the encoding-based *model-updating account* (e.g., Ecker et al., 2022; Lewandowsky et al., 2012; Sanderson & Ecker, 2020). This dissertation focusses on the model-updating account. To distinguish this account from its retrieval-based alternative and to make transparent where the two accounts differ in their predictions for the current experiments, I will first briefly introduce the retrieval-failure account.

According to the retrieval-failure account, both original information and retraction are encoded in memory, associated with the presented context, and simultaneously retained in an activation-based semantic network (Ayers & Reder, 1998). When people make judgments and draw inferences, both memory representations compete for automatic activation. The automatic activation of the original information is likely to be misinterpreted as evidence that this information is true unless the retraction is also retrieved. However, retrieval of the retraction requires a strategic and cognitively demanding assessment of the validity of the original information (Ayers & Reder, 1998). The CIE results from failure to engage in or failure of these strategic retrieval processes, such as monitoring of source and veracity (cf. Begg et al., 1992), which would be necessary to identify the original information as false and to suppress its incorporation into judgments and inferences (Ecker, Lewandowsky, Swire, & Chang, 2011; Swire et al., 2017). It is also possible that strategic retrieval processes are

engaged, yet fail and thus, for example, lead to source-monitoring error (cf. M. K. Johnson et al., 1993). For instance, retrieving the original information, but misattributing it to the wrong source (i.e., the retraction) could lead to continued agreement with the original information (Lewandowsky et al., 2012). Thus, according to the retrieval-failure account, the causes of the CIE are located at retrieval of the critical statements during judgment formation. Empirical evidence for this account includes, for example, findings that greater familiarity of the original information and less recollection of the retraction led to a stronger CIE (Ecker, Lewandowsky, Swire, & Chang, 2011; Swire et al., 2017).

The model-updating account locates the causes of the CIE in the integration of information and inferences into the memory representation of the event at the processing and encoding of the critical statements. This account builds on the concept of *situation models*. A situation model is a structured and coherent memory representations of information about unfolding events (e.g., Johnson-Laird, 1980; van Dijk & Kintsch, 1983). The construction of a situation model involves the integration of new information from an event description and resulting elaborations including attitudes, opinions, and emotions, with existing knowledge about similar situations into a single coherent memory representation (cf. van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). Thus, a situation model represents an individual's interpretation of a situation. The model-updating account postulates that participants construct an initial situation model of the event including the original information and related elaborations. I refer to this initial version of the situation model as the *original situation model*. After the retraction, however, the original situation model is no longer accurate and requires updating. I refer to this post-retraction version of the situation model as the *updated situation model*.

According to the model-updating account, the CIE occurs because some agreement with the original information persists in the updated situation model despite the retraction (e.g., Wilkes & Leatherbarrow, 1988). Possible reasons for this persistence and participants' use of a situation model that continues to contain the original information are a) the difficulty and effortfulness of removing the retracted information or integrating a retraction into an otherwise coherent situation model, b) the avoidance of incomplete and implausible situation models resulting from removal of information without an alternative, and c) the feeling of discomfort from removing information (e.g., Gordon et al., 2017; Hamby et al., 2020; Susmann & Wegener, 2022). Empirical evidence for this account includes, for example, findings that pieces of information more relevant to the coherence of a situation model (i.e., original information central for the causal understanding of the event and retractions

including an alternative explanation for the event) affected participants' judgments more strongly than less relevant information (Ecker et al., 2010; Ecker, Lewandowsky, & Apai, 2011; Hamby et al., 2020; H. M. Johnson & Seifert, 1994; Kan et al., 2021). Critically, the existing empirical evidence only indirectly implies the relevance of situation models for the CIE. The relationships between the situation model and the CIE have never been empirically tested before which would significantly strengthen the evidence for the model-updating account.

1.3 The Current Work

Based on the existing literature and the gap identified therein, I derived the main goals for the present work. Specifically, I aimed to provide direct empirical evidence for the causal role of situation models in the formation of the CIE, as postulated by the model-updating account. To achieve this goal, I defined three objectives.

First, I developed an assessment of participants' situational models that could be embedded in a standard CIE experiment. This enabled me to examine the relationships between the original situation model, the updated situation model, and the CIE.

Second, to investigate these relationships experimentally, I identified a manipulation that affected the CIE and the construction of situation models. Ideally, such a manipulation should allow the original situation model and the updated situation model to be affected separately, enabling the investigation of both determinants of effects and mediations via the situation models on the CIE.

Third, I tested the model-updating account against the retrieval-failure account. From the model-updated account, three main predictions can be derived. First, the situation model should predict the CIE and, second, mediate the effects of variables on the CIE. Third, effects on the CIE should depend on whether or not a variable can influence the situation model. From the retrieval-failure account, in contrast, the opposing prediction can be derived: Variables should affect the CIE directly as long as they are manipulated prior to retrieval and judgment formation.

2 A New Assessment of Situational Models[2]

In this section, I will present the development of an assessment of situation models and Experiment 1, which I conducted to validate it. Following the principles of the *Question-*

[2] Section 2 was partially adopted from Paper 2 and its Online Supplement 2 "Validation Experiments" in the Appendix.

Answering method to assess situation models (Graesser & Clark, 1985), I designed a set of nine questions that probed participants' interpretations of the causal antecedents and consequences of a situation. The questions were tailored to the event of a waterworks shutdown described in a news report frequently used in the CIE paradigm (e.g., Brydges et al., 2018; Ecker & Antonio, 2021; Ecker et al., 2017; Ithisuphalap et al., 2020; McIlhiney et al., 2022, 2023; Miller et al., 2022; Rich et al., 2023; Sanderson et al., 2021). The original information was that a company's disposal of toxic chemicals in the water necessitated the shutdown; the retraction was that the company in fact complied with all standards and that there was no evidence incriminating it. The generated questions were designed such that participants had to go beyond the event description and infer from their interpretation of the situation to answer. Possible answers allowed but did not require references to the original information.

2.1 Experiment 1

Experiment 1 ($N = 140$) served to validate the new assessment of situation models by ensuring that all questions used in the assessment effectively differentiated between different levels of (dis-)agreement with the original information represented in the situation models. To this end, I manipulated participants' level of agreement with the original information by manipulating the information presented to them. Participants read the beginning of the news report about the waterworks shutdown followed by either the original information, or the retraction used in this CIE scenario (even though there was no original information to retract), or no information about the cause of the situation. The resulting situation models should, thus, include either agreement with, disagreement with, or indecision regarding the original information. Then, participants answered the nine questions to assess the situation model. The extent of (dis-)agreement with the original information in the open answers was rated by two independent raters. The inter-rater reliability indicated substantial agreement. Only one of 1,024 answers was rated as nonsensical. Thus, participants were generally willing to draw inferences from the little information provided in order to give meaningful answers.

The majority of questions differentiated between the three levels of agreement with the original information with participants showing the strongest agreement with the original information after reading it and the weakest after reading the retraction. Further, the answers to all questions in combination correctly classified the manipulated level of agreement with the original information for 100% of participants and predicted participants CIE measured at

the end of the experiment. Thus, the constructed set of questions validly assessed the level of agreement with the original information in participants' situation models and it can be used to investigate the relationships between the original situation model, the updated situation model, and the CIE.

3 Source Credibility Affects the CIE[3]

In this section, I will present the theoretical and empirical background of source credibility and its effects on the CIE to explain why source credibility can be a suitable manipulation to experimentally test determinants of effects on the CIE and mediations via the situation models on the CIE. Source credibility refers to the perceived willingness and capability of a source to provide reliable information within a specific context. It is a multidimensional construct influenced by factors such as perceived trustworthiness and expertise, which collectively shape perceptions of a source's credibility. Typically, high source credibility magnifies the persuasiveness and acceptance of a communication and low source credibility attenuates it (e.g., Greenberg & Miller, 1966; Guillory & Geraci, 2013; Lariscy & Tinkham, 1999; Nadarevic et al., 2020; Pratkanis et al., 1988; Skagerberg & Wright, 2009; for a narrative review, see Pornpitakpan, 2004, and for a meta-analysis, see Wilson & Sherrell, 1993).

Source credibility is thought to affect judgments and inferences by influencing the cognitive processes involved in forming the memory representation of encoded information, such as the depth of information processing and the nature of thoughts generated in response to the information (cf. Heesacker et al., 1983; Petty & Cacioppo, 1986; Priester et al., 1999; Sternthal, Dholakia, & Leavitt, 1978; Sternthal, Phillips, & Dholakia, 1978; Tobin & Raymundo, 2009). Low source credibility specifically may work as a forewarning to guard beliefs against persuasion (cf. Greenberg & Miller, 1966; Ward & McGinnies, 1974). Applied to the model-updating account of the CIE, this means that source credibility should affect the construction of situation models and, consequently, the CIE.

3.1 Theoretical and Empirical Background

There are two relevant sources in the CIE paradigm: the informant presenting the original information and the retractor presenting the retraction. Extant research suggests that informant and retractor credibility antagonistically affect how much participants agree with the original information in a CIE paradigm: High informant credibility led to a stronger CIE

[3] Section 3 was partially adopted from Paper 1 and Paper 2 in the Appendix.

than low informant credibility (Ecker et al., 2024; Westbrook et al., 2023). Conversely, high retractor credibility led to a weaker CIE that low retractor credibility (Ecker & Antonio, 2021; Guillory & Geraci, 2013; Pluviano et al., 2020; Susmann & Wegener, 2023). However, informant and retractor credibility have never been manipulated within the same experiment. Thus, for source credibility to be a suitable manipulation for the present cause, I needed to verify that the effects of informant and retractor credibility on the CIE did not interact and, preferably, were of similar strength. This would indicate that the credibility of one source does not influence the persuasiveness of the other. This pattern would suggest that source credibility is applied only to the corresponding critical statement, which would allow us to independently influence and observe the effects of the original information and the retraction separately.

3.2 Overview of the Studies

In two experiments with different materials, I manipulated informant credibility and retractor credibility orthogonally between participants in a standard CIE paradigm. In the unfolding news reports, original information was either presented by a high-credibility or a low-credibility informant. Later, this original information was either retracted by a high-credibility or a low-credibility retractor. The experiments also included a control group that was not exposed to the original information, the retraction, or the respective sources. Their results are not relevant for the present research question and are therefore not discussed here. I used a direct measure to assess participants' agreement with the original information and an indirect measure to assess their agreement with inferences consistent with the original information (following e.g., Rich & Zaragoza, 2016).

I predicted that high informant credibility should lead to more agreement with the original information than low informant credibility (e.g., Westbrook et al., 2023) and high retractor credibility should lead to less agreement with the original information than low retractor credibility (e.g., Guillory & Geraci, 2013). More importantly, for source credibility to be a suitable manipulation for the present cause, the effects of informant and retractor credibility should not interact and, preferably, be of similar size.

3.2.1 Experiment 2.1

For Experiment 2.1 ($N = 151$), I used a scenario about a food recall adapted from O'Rear and Radvansky (2020). The direct and indirect measure both consisted of one statement with which participants rated their agreement. For the direct measure, all predictions were confirmed. That is, informant and retractor credibility affected the CIE

antagonistically, independetly and to a similar extent, indicating that source credibility could be a suitable manipulation. Unexpectedly, however, the indirect measure failed to produce any effects of source credibility or a CIE altogether. Since indirect measures have reliably produced a CIE in past research (e.g., Ecker, Lewandowsky, Swire, & Chang, 2011; Wilkes & Leatherbarrow, 1988), I conducted Experiment 2.2 to verify whether the absence of effects was systematic or due to low reliability of the indirect measure.

3.2.2 Experiment 2.2

For Experiment 2.2 ($N = 145$), I used a scenario about the shutdown of a waterworks adapted from Ecker et al. (2017). The direct measure again consisted of one statement. To enhance the reliability of the indirect measure, I used four statements instead of one and combined the four agreement ratings into a single measure via Principal Component Analysis (e.g., DiStefano et al., 2009). With this modification, the predictions were confirmed in both the direct and the indirect measure. That is, informant and retractor credibility again affected the CIE antagonistically, independetly and to a similar extent.

3.3 Discussion

Across two experiments using different materials, informant credibility and retractor credibility had the predicted effects on the CIE. High informant credibility led to a stronger CIE, whereas high retractor credibility led to a weaker CIE (replicating Ecker & Antonio, 2021; Ecker et al., 2024; Guillory & Geraci, 2013; Pluviano et al., 2020; Susmann & Wegener, 2023; Westbrook et al., 2023). Importantly, the two credibility manipulations did not interact, suggesting that the credibility of one source did not influence the persuasiveness of the other. This pattern indicates that source credibility was applied selectively to the corresponding critical statement rather than to the overall narrative and is consistent with the assumption that each statement contributed separately to the construction and updating of the situation model.

These findings support the use of source credibility as an experimental manipulation to influence the strength of agreement with the original information in the original and updated situation model independently. Consequently, source credibility provides a suitable manipulation for experimentally examining how the construction and updating of situation models affect the magnitude of the CIE.

4 Determinants of Effects on the CIE[4]

As said before, the model-updating account assumes that the CIE arises from processes at encoding, whereas the retrieval-failure account attributes the CIE to processes at retrieval. To differentiate between these accounts using the source-credibility manipulation, it was necessary to ensure that source credibility could influence encoding (i.e., situation-model construction) but not retrieval. This can be achieved by manipulating the timing of credibility disclosure: When credibility is disclosed before the critical statements, it can affect situation-model construction during encoding; when disclosed only after processing, it cannot influence encoding. Therefore, when disclosed only after processing, it should not affect the CIE according to the model-updating account, but may still do so according to the retrieval-failure account. Persuasion research has shown that whereas credibility disclosure before a message produced source-credibility effects on judgments, source disclosure after a message but before judgment formation did not affect judgments (e.g., Allen et al., 2002; Greenberg & Miller, 1966; Mills & Harvey, 1972; O'keefe, 1987; Ward & McGinnies, 1974). Thus, once available, source credibility did not spontaneously affect the situation model or subsequent judgments.

4.1 Overview of the Studies

In two new experiments, I manipulated informant credibility and retractor credibility orthogonally between participants in a standard CIE paradigm. Critically, I additionally manipulated the credibility-disclosure timing. For half of the participants, high or low informant credibility was disclosed immediately before the original information, and high or low retractor credibility was disclosed immediately before the retraction, replicating the previously reported Experiments 2.1 and 2.2. Thus, source credibility was available to influence both the degree of integration and the updating of the situation model. For the other half of participants, informant and retractor credibility were disclosed only after the corresponding critical statements. In Experiments 3.1 and 3.2, I again used the scenario about the shutdown of a waterworks adapted from Ecker et al. (2017). To assess the CIE, I used six statements and computed a single CIE score as the mean of the standardized ratings of agreement with these statements of each participant.

I predicted that credibility-disclosure timing should moderate the source-credibility effects. Specifically, for the half of participants with early source disclosure, I predicted to

[4] Section 4 was partially adopted from Paper 2 in the Appendix.

replicate the source-credibility effects shown in the previously reported experiments, that is, high informant credibility should lead to more agreement than low informant credibility and high retractor credibility should lead to less agreement than low retractor credibility. For the other half of participants with late source disclosure, based on the model-updating account, I predicted the absence of source-credibility effects. This should hold even though source credibility was disclosed long before the CIE assessment and could thus affect the retrieval processes proposed to underlie the CIE according to the retrieval-failure account (e.g., Lewandowsky et al., 2012).

4.1.1 Experiment 3.1

Experiment 3.1 is not reported in the enclosed papers. Therefore, I provide a more detailed account of the relevant results. Data collection took place between December 14 and 17, 2021, and participants were recruited from the same pool as in Experiment 3.2. In the study phase of Experiment 3.1 ($N = 432$), the descriptions of informant and retractor credibility were presented either on the screen immediately preceding or immediately following the critical statements. After the study phase, participants completed a two-minutes filler task, as in Experiments 2.1 and 2.2. I assessed the CIE as the main dependent variable. Following the same procedure as in Experiment 3.2 (see Paper 2 for methodological details), participants further rated informant and retractor credibility for a manipulation check and I measured their memory for the details of the report.

Figure 1 shows the source-credibility ratings for all groups. Ratings of informant credibility were higher for the high-credibility informant than for the low-credibility informant, $F(1, 424) = 149.82$, $p < .001$, $\eta^2_p = .26$, $BF_{incl} = 1.94 \cdot 10^{26}$, and ratings of retractor credibility were higher for the high-credibility retractor than for the low-credibility retractor, $F(1, 424) = 223.68$, $p < .001$, $\eta^2_p = .35$, $BF_{incl} = 1.33 \cdot 10^{37}$. Thus, the manipulations of informant and retractor credibility had the intended selective effects on the credibility ratings.

Table 1 lists the mean memory for details of the report for all groups. There were no group differences in mean memory for details of the report, $F(7, 424) = 0.46$, $p = .865$, $\eta^2_p = .01$, $BF_{excl} = 298.12$. Thus, between-group CIE differences are not attributable to differences in memory for details of the report (cf. Sanderson et al., 2021).

Figure 2 shows the descriptive statistics of the CIE for all groups. Replicating Experiments 2.1 and 2.2, high informant credibility led to a stronger CIE than low informant credibility, $F(7, 424) = 42.78$, $p < .001$, $\eta^2_p = .09$, $BF_{incl} = 4.69 \cdot 10^7$, and high retractor credibility led to a weaker CIE that low retractor credibility, $F(7, 424) = 29.99$, $p < .001$, η^2_p

= .07, $BF_{incl} = 1.43 \cdot 10^5$. Unexpectedly, however, credibility-disclosure timing did not moderate these source-credibility effects, as indicated by the two non-significant two-way interactions of credibility-disclosure timing with informant credibility, $F(7, 424) = 0.36$, $p = .550$, $\eta^2_p < .01$, $BF_{excl} = 5.31$, and with retractor credibility, $F(7, 424) = 1.68$, $p = .196$, $\eta^2_p < .01$, $BF_{excl} = 3.08$.

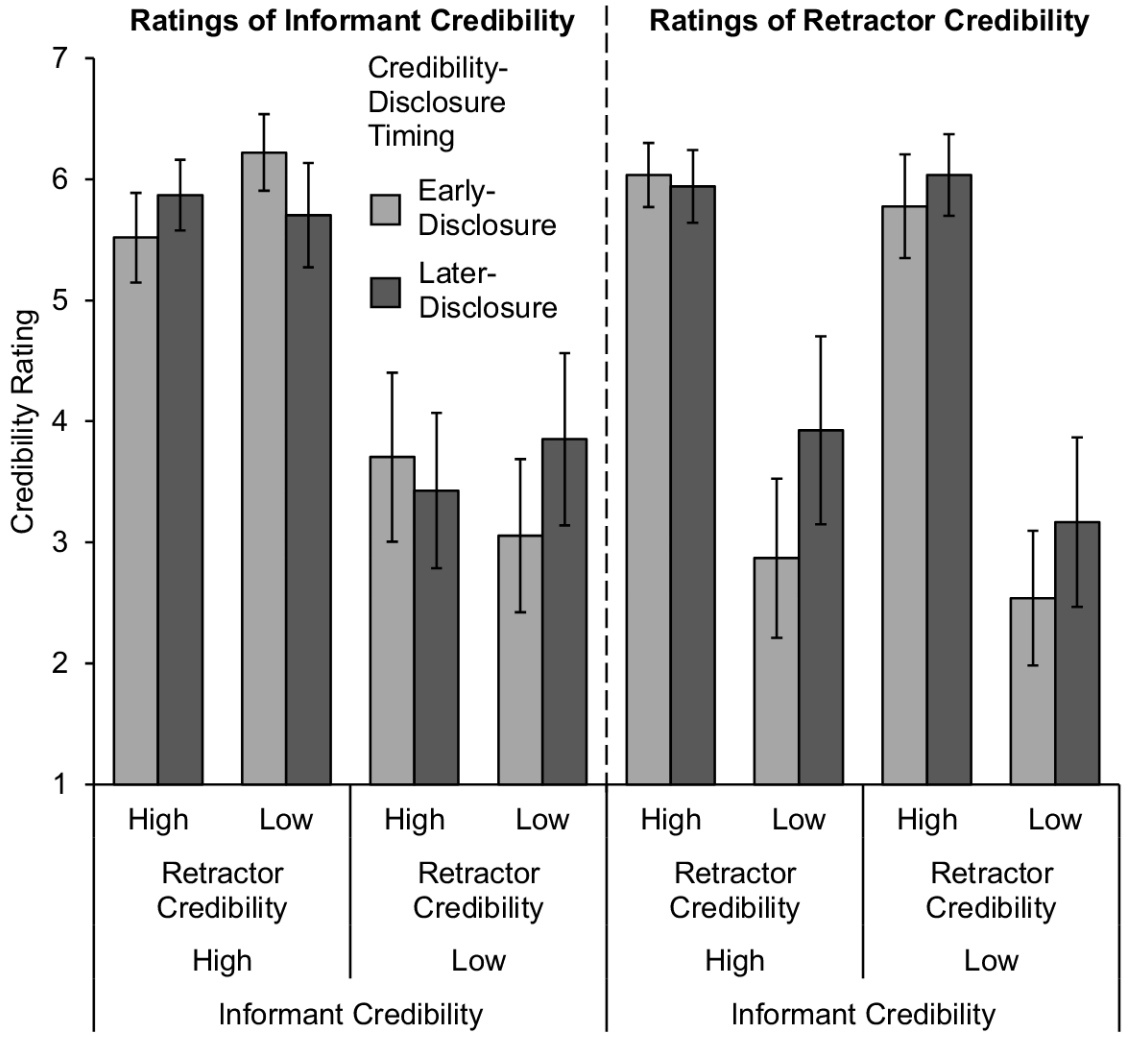


Figure 1. Credibility ratings of the informant and retractor on a Likert scale from 1 (*absolutely not credible*) to 7 (*absolutely credible*). Error bars indicate 95% confidence intervals.

Credibility-disclosure timing	Informant credibility	Retractor credibility	*n*	*M* [95% CI]
Early disclosure	High	High	54	5.88 [5.68, 6.07]
		Low	54	5.86 [5.66, 6.06]
	Low	High	54	5.89 [5.72, 6.07]
		Low	54	5.88 [5.72, 6.03]
Late disclosure	High	High	54	5.80 [5.55, 5.92]
		Low	54	5.81 [5.63, 5.98]
	Low	High	54	5.89 [5.71, 6.07]
		Low	54	5.95 [5.79, 6.12]

Table 1. Memory for details of the report was defined as mean rated agreement with four correct and four incorrect memory statements in the test questionnaire on a Likert scale from 1 (*strong disagreement*) to 7 (*strong agreement*). Ratings of incorrect statements were reverse-coded.

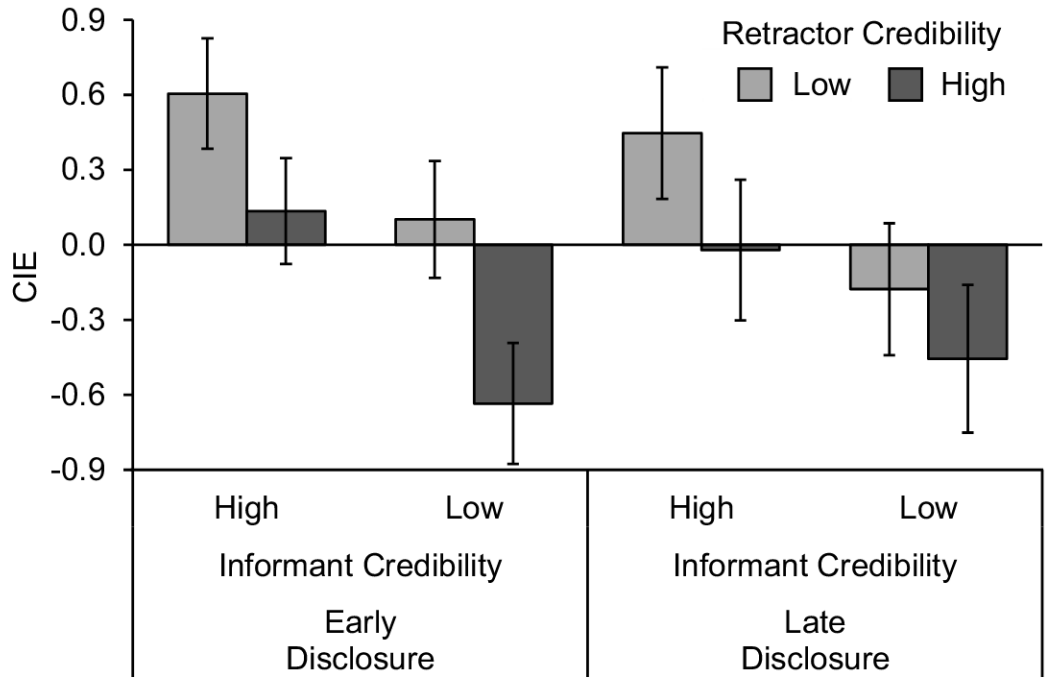


Figure 2. Mean agreement with the original information (in *z* scores) as the CIE. Higher scores indicate greater agreement. Error bars indicate 95% confidence intervals.

This absence of a moderating effect of credibility-disclosure timing is inconsistent with the encoding-based model-updating account. Instead, this pattern suggests that source credibility may also exert its effects at retrieval, in line with the retrieval-failure account. However, I considered two plausible alternative explanations for the absence of morderating

effects. First, the interval between the screens presenting the critical statements and the source credibility may have been too short to prevent retrospective integration of credibility into the situation model. Under such conditions, late disclosure may not have effectively isolated encoding of the critical statement from encoding of the source. Second, even with late source disclosure, the situation models may have been temporarily associated with informant and retractor credibility so that source credibility could have affected the agreement ratings in the immediate CIE test, even though the situation models were not affected by source credibility (cf. *sleeper effect*; Kumkale & Albarracín, 2004). Taken together, the present findings do not provide clear support for a strict dissociation between encoding- and retrieval-based accounts. I conducted Experiment 3.2 to test whether a modified manipulation of credibility-disclosure timing and a delayed CIE test were sufficient to fully disentangle encoding and retrieval processes.

4.1.2 Experiment 3.2

In Experiment 3.2 ($N = 435$), I assessed three dependent variables: participants' original situation model (after presentation of the original information), their updated situation model (after presentation of the retraction), and the CIE. For the situation-model assessments, I used the assessment generated and validated in Experiment 1. Thus, after reading the original information and retraction, respectively, the news report was interrupted by nine questions probing participants' interpretations of the situation. For the half of participants with early source disclosure, the descriptions of source credibility were already presented on the screen before the critical statements. For the half of participants with late source disclosure, the names of the sources were the only information disclosed with the critical statements before the situation-model assessments, as the descriptions of source credibility were presented only after the situation-model assessment on a separate screen. Thus, by the time source credibility was disclosed for this half of participants, the construction or updating of the situation model should already have been completed.

After the news report, I implemented a 45-minute delay with filler tasks before the test phase. Thereby, immediate source-credibility effects on the CIE should decay (cf. sleeper effect; Albarracín et al., 2017; Kumkale & Albarracín, 2004).

With these modifications, the predictions were fully confirmed. That is, credibility-disclosure timing moderated the source-credibility effects on the original and updated situation models and on the CIE. In the half of participants with early source disclosure, high informant credibility led to stronger integration of the original information into the original

situation model, stronger persistence in the updated situation model and a stronger CIE. Conversely, high retractor credibility led to weaker persistence of the orginial information in the updated situation model and a weaker CIE. Thus, when source credibility was disclosed before the integration of the original information into the original situation model and before the update of the situation model following the retraction, there were independent and antagonistic effects of informant and retractor credibility on the CIE, replicating Experiments 2.1, 2.2, and 3.1. In the half of participants with late source disclosure, neither dependent variable was affected by informant or retractor credibility.

These results suggest that the modifications were effective in isolating encoding from retrieval processes. By separating the late source disclosure from the critical statements through the situation-model assessments and delaying the CIE assessment, late-disclosed source credibility no longer influenced the CIE. This indicates that the effects observed in Experiment 3.1 were likely due to retrospective integration or residual activation of source credibility.

4.2 Discussion

Across Experiments 3.1 and 3.2, an inconsistent pattern emerged regarding the role of credibility-disclosure timing. In Experiment 3.1, source credibility affected the CIE regardless of credibility-disclosure timing, suggesting an influence beyond encoding. However, Experiment 3.2 showed that this pattern was likely due to insufficient separation of encoding and retrieval processes.

With optimized conditions in Experiment 3.2, as predicted by the model-updating account, the effects regarding source credibility depended on whether or not source credibility could affect the situation models. In line with my hypotheses, the timing of credibility disclosure moderated whether or not informant and retractor credibility affected the CIE. With early source disclosure, participants had the opportunity to consider source credibility when they processed the critical statements and, only in these groups, informant and retractor credibility affected the CIE. With late source disclosure, participants were unable to consider the source credibility of informant and retractor when they processed the respective critical statements so that source credibility neither affected the original situation model, nor the updated situation model, nor the CIE. After establishing the moderating effects of the timing of credibility disclosure on the effect of source credibility on the situation model and the CIE, the next section presents a *structural equation model* (SEM) to test the causal relationships between situation models and the CIE predicted by the model-

updating account.

5 The Relationships Between the Situation Models and the CIE[5]

To test the model-updating account and examine the role of situation models in the formation of the CIE beyond experimental effects on the CIE, I explicitly modeled the systematic relationships of situation-model construction and updating, and the CIE. Specifically, I derived two hypotheses and tested them with an SEM. First, participants' updated situation model should predict the CIE. Second, any effect of source credibility on the CIE should be mediated through the situation models.

5.1 Structural Equation Modelling of Experiment 3.2

I implemented these hypotheses in an SEM of serial mediation corresponding to the predicted relationships between the variables of Experiment 3.2. The SEM includes informant credibility and retractor credibility as predictors, the original situation model and the updated situation model as mediators, and the CIE as the outcome variable. I specified the effects included in the SEM corresponding to the predictions of the model-updating account.

The updated situation model should predict the CIE. Thus, I included a direct effect of the updated situation model on the CIE. Because the original situation model may predict the updated situation model, I included a direct effect of the original on the updated situation model. To test whether the original situation model continued to influence the CIE despite the update following the retraction, I additionally included a direct effect of the original situation model on the CIE. Any effect of source credibility on the CIE should be mediated through the situation models. Thus, I included direct effects of informant credibility on the original and the updated situation models and a direct effect of retractor credibility on the updated situation model. To test the prediction that source credibility would not affect the CIE beyond the mediated effects, I also included direct effects of informant and retractor credibility on the CIE. Figure 3 shows the SEM.

As demonstrated above, the effects of source credibility on the situation models and the CIE depended on credibility-disclosure timing. Therefore, I analyzed the data of participants with early and with late source disclosure in separate SEMs. Model fit was good in both groups.

---

[5] Section 5 was partially adopted from Paper 2 in the Appendix.

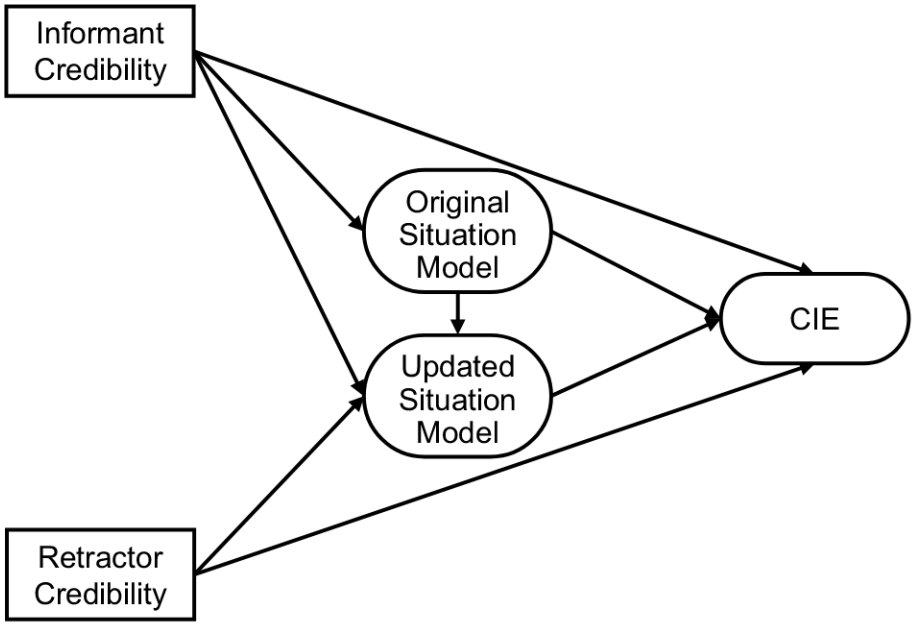


Figure 3. Structural equation model of effects of informant credibility and retractor credibility on the CIE, mediated through the original and updated situation models. The rectangles represent the manipulated predictor variables. The rounded rectangles represent the measured variables with the original and updated situation models as serial mediators and the CIE as the outcome variable. The arrows represent direct effects.

5.1.1 Early Source Disclosure

For participants with early source disclosure, both predictions were confirmed in the SEM. First, the updated situation model predicted the CIE. Additionally, the original situation model predicted the updated situation model and the CIE. That is, the more strongly the original information was integrated into the original situation model, the more strongly it persisted in the updated situation model despite the retraction, and the stronger was the CIE. And the stronger the persistence of the original information in the updated situation model was, the stronger was the CIE.

Second, the effects of source credibility on the CIE were mediated through the situation models. That is, informant credibility affected the original situation model, and affected the updated situation model mediated through the original situation model. After inclusion of this indirect effect, the direct effect of informant credibility on the updated situation model was not significant. Thus, higher informant credibility led to stronger integration of the original information into the original situation model. This integration was maintained after the update (indicated by the significant indirect effect), but informant

credibility did not directly affect the degree to which the situation model was updated following the retraction. Critically, higher informant credibility led to a stronger CIE and this effect was mediated on the CIE through the original situation model and serially mediated through both situation models. After inclusion of these indirect effects, the direct effect of informant credibility on the CIE was not significant. Thus, the effect of informant credibility on the CIE was fully mediated through the situation models.

Retractor credibility affected the updated situation model and the CIE. Specifically, higher retractor credibility led to less persistence of the original information in the updated situation model and to a weaker CIE. The effect of retractor credibility on the CIE was partially mediated through the updated situation model, with a direct effect of retractor credibility on the CIE remaining. Thus, as predicted, retractor credibility affected the CIE through the degree to which the original information persisted in the situation model that was updated after the retraction. However, retractor credibility also affected the CIE directly, which was not predicted by the model-updating account.

5.1.2 Late Source Disclosure

For participants with late source disclosure, also both predictions were confirmed in the SEM. First, the updated situation model predicted the CIE. In contrast to participants with early source disclosure, the original situation model did not predict the updated situation model nor the CIE. Thus, for participants with late source disclosure, the CIE was exclusively predicted by the updated situation model.

Second, there were no source-credibility effects on the CIE. Accordingly, there was no mediation of such effects either. Specifically, informant credibility neither affected the original situation model, nor the updated situation model, nor the CIE. Likewise, retractor credibility did neither affect the updated situation model nor the CIE. Thus, source credibility did not affect the CIE when it could not affect the construction of the situation models.

5.2 Discussion

In line with my first hypothesis, the updated situation model predicted the CIE in the delayed test phase. This relationship existed regardless of whether or not source credibility could affect the situation model (i.e., was disclosed early or late). In line with my second set of hypotheses, the situation models mediated the source-credibility effects on the CIE if present (i.e., with early source disclosure only).

Beyond these primary findings, the SEM results provide further insights into the mechanisms underlying the CIE. For participants with early source disclosure, the original

situation model not only predicted the updated situation model but also exerted a direct effect on the CIE. This suggests that updating the situation model did not overwrite the original representation but rather resulted in the coexistence of both models, which jointly influenced participants' inferences. Thus, even after a retraction, the initially constructed representation continued to contribute to reasoning.

Furthermore, the full mediation of the informant-credibility effect and the partial mediation of the retractor-credibility effect refine the predictions of the model-updating account. While the effect of informant credibility on the CIE was entirely explained by its influence on the situation-model construction, the remaining direct effect of retractor credibility suggests that, under some conditions, retraction-related information may also influence judgments beyond its impact on the updated situation model. However, this direct effect was absent in the late-disclosure condition, which might indicate that it nevertheless depends on credibility being available during situation-model updating.

Taken together, these findings support the central assumption of the model-updating account that the CIE is driven by the structure and content of situation models. At the same time, they suggest that updating does not fully replace prior representations and that multiple representations can remain concurrently accessible and influence reasoning.

6 General Discussion

The present dissertation aimed to provide direct empirical evidence for the role of situation models in the CIE and to test encoding-based and retrieval-based accounts of the phenomenon against each other. Across five experiments, three main contributions emerge.

First, I introduced and validated a novel assessment of situation models within a standard CIE paradigm. This assessment allowed, for the first time, the direct measurement of participants' agreement with the original information during situation-model construction and updating. The validation results demonstrated that this assessment reliably differentiates between levels of agreement and predicts subsequent CIE responses, establishing it as a suitable tool for investigating the cognitive mechanisms underlying the CIE.

Second, I identified source credibility as a manipulation that affects both situation-model construction and the CIE. Across Experiments 2.1 and 2.2, informant and retractor credibility independently and antagonistically influenced the CIE, with no evidence of an interaction. This pattern is consistent with the idea that credibility is applied separately to different pieces of information, allowing it to selectively affect the integration and persistence of that information.

Third, I used this manipulation to test predictions derived from the model-updating and retrieval-failure accounts. Experiments 3.1 and 3.2 used credibility-disclosure timing to disentangle encoding and retrieval processes. Experiment 3.1 suggested that source credibility may also affect retrieval. However, Experiment 3.2 showed that, under conditions that effectively isolate encoding, source credibility only influenced the CIE when it was available during situation-model construction and updating. When this was not the case, no effects of source credibility emerged. This pattern is consistent with the model-updating account and constrains retrieval-based explanations. Structural equation modeling further showed that the updated situation model predicted the CIE and that source-credibility effects were mediated through the situation models when present.

Taken together, these findings suggest that the CIE arises from how information is integrated and maintained in situation models. Rather than arising primarily from failures at retrieval, the CIE appears to reflect the structure, persistence, and accessibility of the memory representations formed during encoding. At the same time, the results suggest that retrieval processes are not fully irrelevant. Under certain conditions, retrieval-related influences may still contribute to the CIE, for example when source information remains active.

Overall, the present dissertation advances the understanding of the CIE by providing direct evidence for the role of situation models. In doing so, it refines theoretical accounts of how misinformation continues to influence reasoning despite retraction. At a broader level, the findings speak to how people construct and maintain coherent mental representations while processing dynamic information over time.

References

Albarracín, D., Kumkale, G. T., & Poyner-Del Vento, P. (2017). How people can become persuaded by weak messages presented by credible communicators: Not all sleeper effects are created equal. *Journal of Experimental Social Psychology, 68*, 171–180. https://doi.org/10.1016/j.jesp.2016.06.009

Allen, M., Adamski, L., Bates, M., Bernhagen, M., Callendar, A., Casey, M., Czerwinski, A., Decker, L., Howard, G., Jordan, B., Kujawski, E., May, K., Olson, P., Parenteau, A., Reilly, S., Schmidt, J., Stebnitz, S., Thau, S., Tollefson, M., ... Zirbel, C. (2002). Effect of timing of communicator identification and level of source credibility on attitude. *Communication Research Reports, 19*(1), 46–55. https://doi.org/10.1080/08824090209384831

Ayers, M. S., & Reder, L. M. (1998). A theoretical review of the misinformation effect: Predictions from an activation-based memory model. *Psychonomic Bulletin & Review, 5*(1), 1–21. https://doi.org/10.3758/BF03209454

Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General, 121*(4), 446–458. https://doi.org/10.1037/0096-3445.121.4.446

Brydges, C. R., Gignac, G. E., & Ecker, U. K. (2018). Working memory capacity, short-term memory capacity, and the continued influence effect: A latent-variable analysis. *Intelligence, 69*, 117–122. https://doi.org/10.1016/j.intell.2018.03.009

Buczel, K. A., Siwiak, A., Kękuś, M., & Szpitalak, M. (2025). Do the protective effects last? The effectiveness of alternative + forewarning/inoculation techniques in reducing misinformation reliance and reliance regression in the continued influence effect procedure. *Quarterly Journal of Experimental Psychology, 78*(11), https://doi.org/10.1177/17470218251336232

Buczel, K. A., Siwiak, A., Szpitalak, M., & Polczyk, R. (2024). How do forewarnings and post-warnings affect misinformation reliance? The impact of warnings on the continued influence effect and belief regression. *Memory & Cognition, 52*(5), 1048–1064. https://doi.org/10.3758/s13421-024-01520-z

Buczel, K. A., Syszka, P. D., Siwiak, A., Szpitalak, M., & Polczyk, R. (2022). Vaccination against misinformation: The inoculation technique reduces the continued influence effect. *PLoS ONE, 17*(4), Article e0267463. https://doi.org/10.1371/journal.pone.0267463

Chan, M. S., Jones, C. R., Hall Jamieson, K., & Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological Science, 28*(11), 1531–1546. https://doi.org/10.1177%2F0956797617714579

DiStefano, C., Zhu, M., & Mîndrilã, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research, and Evaluation, 14*, Article 20. https://doi.org/10.7275/da8t-4g52

Ecker, U. K., & Antonio, L. M. (2021). Can you believe it? An investigation into the impact of retraction source credibility on the continued influence effect. *Memory & Cognition, 49*(4), 631–644. https://doi.org/10.3758/s13421-020-01129-y

Ecker, U. K., Hogan, J. L., & Lewandowsky, S. (2017). Reminders and repetition of misinformation: Helping or hindering its retraction? *Journal of Applied Research in Memory and Cognition, 6*(2), 185–192. https://doi.org/10.1016/j.jarmac.2017.01.014

Ecker, U. K., Lewandowsky, S., & Apai, J. (2011). Terrorists brought down the plane!—No, actually it was a technical fault: Processing corrections of emotive information. *The Quarterly Journal of Experimental Psychology, 64*(2), 283–310. https://doi.org/10.1080/17470218.2010.497927

Ecker, U. K., Lewandowsky, S., & Tang, D. T. (2010). Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Memory & Cognition, 38*(8), 1087–1100. https://doi.org/10.3758/MC.38.8.1087

Ecker, U. K., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., . . . Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology, 1*, 13–29. https://doi.org/10.1038/s44159-021-00006-y

Ecker, U. K., Lewandowsky, S., Swire, B., & Chang, D. (2011). Correcting false information in memory: Manipulating the strength of misinformation encoding and its retraction. *Psychonomic Bulletin & Review, 18*(3), 570–578. https://doi.org/10.3758/s13423-011-0065-1

Ecker, U. K., Prike, T., Paver, A. B., Scott, R. J., & Swire-Thompson, B. (2024). Don't believe them! Reducing misinformation influence through source discreditation. *Cognitive Research: Principles and Implications, 9*, Article 52. https://doi.org/10.1186/s41235-024-00581-7

Ecker, U. K., Sze, B. K., & Andreotta, M. (2021). Corrections of political misinformation: No evidence for an effect of partisan worldview in a US convenience sample. *Philosophical Transactions of the Royal Society B, 376*(1822), Article 20200145. https://doi.org/10.1098/rstb.2020.0145

Gordon, A., Brooks, J. C., Quadflieg, S., Ecker, U. K., & Lewandowsky, S. (2017). Exploring the neural substrates of misinformation processing. *Neuropsychologia, 106*, 216–224. https://doi.org/10.1016/j.neuropsychologia.2017.10.003

Graesser, A. C., & Clark, L. F. (1985). The generation of knowledge-based inferences during narrative comprehension. *Advances in Psychology, 29*, 53–94. https://doi.org/10.1016/S0166-4115(08)62732-6

Greenberg, B. S., & Miller, G. R. (1966). The effects of low-credible sources on message acceptance. *Speech Monographs, 33*(2), 127–136 https://doi.org/10.1080/03637756609375488.

Guillory, J. J., & Geraci, L. (2013). Correcting erroneous inferences in memory: The role of

source credibility. *Journal of Applied Research in Memory and Cognition, 2*(4), 201–209. https://doi.org/10.1016/j.jarmac.2013.10.001

Hamby, A., Ecker, U. K., & Brinberg, D. (2020). How stories in memory perpetuate the continued influence of false information. *Journal of Consumer Psychology, 30*(2), 240–259. https://doi.org/10.1002/jcpy.1135

Heesacker, M., Petty, R. E., & Cacioppo, J. T. (1983). Field dependence and attitude change: Source credibility can alter persuasion by affecting message-relevant thinking. *Journal of Personality, 51*(4), 653–666. https://doi.org/10.1111/j.1467-6494.1983.tb00872.x

Ithisuphalap, J., Rich, P. R., & Zaragoza, M. S. (2020). Does evaluating belief prior to its retraction influence the efficacy of later corrections? *Memory, 28*(5), 617–631. https://doi.org/10.1080/09658211.2020.1752731

Johnson, H. M., & Seifert, C. M. (1994). Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(6), 1420–1436. https://doi.org/10.1037/0278-7393.20.6.1420

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source Monitoring. *Psychological Bulletin, 114*(1), pp. 3–28. https://doi.org/10.1037/0033-2909.114.1.3.

Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science, 4*(1), 71–115. https://doi.org/10.1016/S0364-0213(81)80005-5

Kan, I. P., Pizzonia, K. L., Drummey, A. B., & Mikkelsen, E. J. (2021). Exploring factors that mitigate the continued influence of misinformation. *Cognitive Research: Principles and Implications, 6*(1), Article 76. https://doi.org/10.1186/s41235-021-00335-9

Kumkale, G. T., & Albarracín, D. (2004). The sleeper effect in persuasion: A meta-analytic review. *Psychological Bulletin, 130*(1), 143–172. https://doi.org/10.1037/0033-2909.130.1.143

Lariscy, R. A., & Tinkham, S. F. (1999). The sleeper effect and negative political advertising. *Journal of Advertising, 28*(4), 13–30. https://doi.org/10.1080/00913367.1999.10673593

Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest, 13*(3), 106–131. https://doi.org/10.1177%2F1529100612451018

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory, 12*(4), 361–366. https://doi.org/10.1101/lm.94705

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and*

Memory, 4(1), 19–31. https://doi.org/10.1037/0278-7393.4.1.19

McIlhiney, P., Gignac, G. E., Ecker, U. K., Kennedy, B. L., & Weinborn, M. (2023). Executive function and the continued influence of misinformation: A latent-variable analysis. *PLoS One, 18*(4), Article e0283951. https://doi.org/10.1371/journal.pone.0283951

McIlhiney, P., Gignac, G. E., Weinborn, M., & Ecker, U. K. (2022). Sensitivity to misinformation retractions in the continued influence paradigm: Evidence for stability. *Quarterly Journal of Experimental Psychology, 75*(7), 1259–1271. https://doi.org/10.1177/17470218211048986

Miller, A. L., Wissmann, K. T., & Peterson, D. J. (2022). The continued influence effect: Examining how age, retraction, and delay impact inferential reasoning. *Applied Cognitive Psychology, 36*(3), 708–723. https://doi.org/10.1002/acp.3939

Mills, J., & Harvey, J. (1972). Opinion change as a function of when information about the communicator is received and whether he is attractive or expert. *Journal of Personality and Social Psychology, 21*(1), 52–55. https://doi.org/10.1037/h0031939

Nadarevic, L., Reber, R., Helmecke, A. J., & Köse, D. (2020). Perceived truth of statements and simulated social media postings: an experimental investigation of source credibility, repeated exposure, and presentation format. *Cognitive Research, 5*(1), Article 56. https://doi.org/10.1186/s41235-020-00251-4

O'keefe, D. J. (1987). The persuasive effects of delaying identification of high-and low-credibility communicators: A meta-analytic review. *Central States Speech Journal, 38*(2), 63–72. https://doi.org/10.1080/10510978709368231

O'Rear, A. E., & Radvansky, G. A. (2020). Failure to accept retractions: A contribution to the continued influence effect. *Memory & Cognition, 48*(1), 127–144. https://doi.org/10.3758/s13421-019-00967-9

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology, 19*, 123–205. https://doi.org/10.1016/S0065-2601(08)60214-2

Pluviano, S., Sala, S. D., & Watt, C. (2020). The effects of source expertise and trustworthiness on recollection: the case of vaccine misinformation. *Cognitive Processing, 21*(3), 321–330. https://doi.org/10.1007/s10339-020-00974-8

Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology, 34*(2), 243–281. https://doi.org/10.1111/j.1559-1816.2004.tb02547.x

Pratkanis, A. R., Greenwald, A. G., Leippe, M. R., & Baumgardner, M. H. (1988). In search of reliable persuasion effects: III. The sleeper effect is dead. Long live the sleeper effect. *Journal of Personality and Social Psychology, 54*(2), 203–218. https://doi.org/10.1037/0022-3514.54.2.203

Priester, J., Wegener, D., Petty, R., & Fabrigar, L. (1999). Examining the psychological

process underlying the sleeper effect: The elaboration likelihood model explanation. *Media Psychology, 1*(1), 27–48. https://doi.org/10.1207/s1532785xmep0101_3

Rich, P. R., & Zaragoza, M. S. (2016). The continued influence of implied and explicitly stated misinformation in news reports. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(1), 62–74. https://doi.org/10.1037/xlm0000155

Rich, P. R., Donovan, A. M., & Rapp, D. N. (2023). Cause typicality and the continued influence effect. *Journal of Experimental Psychology: Applied, 29*(2), 221–238. https://doi.org/10.1037/xap0000454

Sanderson, J. A., & Ecker, U. K. (2020). The challenge of misinformation and ways to reduce its impact. In P. Van Meter, A. List, D. Lombardi, & P. Kendeou, *Handbook of learning from multiple representations and perspectives* (pp. 461–476). Routledge/Taylor & Francis Group. https://doi.org/10.4324/9780429443961-30

Sanderson, J. A., Farrell, S., & Ecker, U. K. (2022). Examining the role of information integration in the continued influence effect using an event segmentation approach. *PLoS ONE, 17*(7), Article e0271566. https://doi.org/10.1371/journal

Sanderson, J. A., Gignac, G. E., & Ecker, U. K. (2021). Working memory capacity, removal efficiency and event specific memory as predictors of misinformation reliance. *Journal of Cognitive Psychology, 33*(5), 518–532. https://doi.org/10.1080/20445911.2021.1931243

Skagerberg, E. M., & Wright, D. B. (2009). Susceptibility to postidentification feedback is affected by source credibility. *Applied Cognitive Psychology, 23*(4), 506–523. https://psycnet.apa.org/doi/10.1002/acp.1470

Sternthal, B., Dholakia, R., & Leavitt, C. (1978). The persuasive effect of source credibility: Tests of cognitive response. *Journal of Consumer Research, 4*(4), 252–260. https://doi.org/10.1086/208704

Sternthal, B., Phillips, L. W., & Dholakia, R. (1978). The persuasive effect of source credibility: A situational analysis. *Public Opinion Quarterly, 42*(3), 285–314. https://doi.org/10.1086/268454

Susmann, M. W., & Wegener, D. T. (2022). The role of discomfort in the continued influence effect of misinformation. *Memory & Cognition, 50*(2), 435–448. https://doi.org/10.3758/s13421-021-01232-8

Susmann, M. W., & Wegener, D. T. (2023). The independent effects of source expertise and trustworthiness on retraction believability: The moderating role of vested interest. *Memory & Cognition, 51*(4), 845–861. https://doi.org/10.3758/s13421-022-01374-3

Swire, B., Ecker, U. K., & Lewandowsky, S. (2017). The role of familiarity in correcting inaccurate information. *Journal of Experimental Psychology: Learning, Memory and Cognition, 43*(12), 1948–1961. https://doi.org/10.1037/xlm0000422

Tay, L. Q., Hurlstone, M. J., Kurz, T., & Ecker, U. K. (2022). A comparison of prebunking and debunking interventions for implied versus explicit misinformation. *British*

Journal of Psychology, 113(3), 591–607. https://doi.org/10.1111/bjop.12551

Tobin, S. J., & Raymundo, M. M. (2009). Persuasion by causal arguments: The motivating role of perceived causal expertise. *Social Cognition, 27*(1), 105–127. https://doi.org/10.1521/soco.2009.27.1.105

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of Discourse Comprehension.* Academic Press, Inc.

van Huijstee, D., Vermeulen, I., Kerkhof, P., & Droog, E. (2022). Continued influence of misinformation in times of COVID-19. *International Journal of Psychology, 57*(1), 136–145. https://doi.org/10.1002/ijop.12805

Walter, N., & Tukachinsky, R. (2020). A meta-analytic examination of the continued influence of misinformation in the face of correction: How powerful is it, why does it happen, and how to stop it? *Communication Research, 47*(2), 155–177. https://doi.org/10.1177%2F0093650219854600

Ward, C. D., & McGinnies, E. (1974). Persuasive effects of early and late mention of credible and noncredible sources. *The Journal of Psychology, 86*(1), 17–23. https://doi.org/10.1080/00223980.1974.9923879

Westbrook, V., Wegener, D. T., & Susmann, M. W. (2023). Mechanisms in continued influence: The impact of misinformation corrections on source perceptions. *Memory & Cognition, 51*(6), 1317–1330. https://doi.org/10.3758/s13421-023-01402-w

Wilkes, A. L., & Leatherbarrow, M. (1988). Editing episodic memory following the identification of error. *The Quarterly Journal of Experimental Psychology, 40*(2), 361–387. https://psycnet.apa.org/doi/10.1080/02724988843000168

Wilson, E. J., & Sherrell, D. L. (1993). Source effects in communication and persuasion research: A meta-analysis of effect size. *Journal of the Academy of Marketing Science, 21*(2), 101–112. https://doi.org/10.1007/BF02894421

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*(2), 162–185. https://doi.org/10.1037/0033-2909.123.2.162

Appendix

Papier 1 ist publiziert als:

Hey, C. V., Schaper, M. L., & Bayen, U. J. (2025). Relative source credibility affects the continued influence effect: Evidence of rationality in the CIE. *Cognition, 254*, Article 106000. https://doi.org/10.1016/j.cognition.2024.106000

Carolin V. Hey: Original Idea, Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Data curation, Writing – Original Draft, Writing – Review & Editing, Visualization, Project Administration
Marie Luisa Schaper: Methodology, Validation, Writing – Review & Editing, Supervision, Funding Acquisition
Ute Johanna Bayen: Resources, Writing – Review & Editing, Supervision, Funding Acquisition

Papier 2 ist zur Veröffentlichung angenommen:

Hey, C. V., Schaper, M. L., & Bayen, U. J. (2026). *The Continued Influence Effect and Situation Models: Evidence for the Model-Updating Account*. Manuscript accepted for publication pending minor revision in *Memory & Cognition*.

Carolin V. Hey: Original Idea, Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Data curation, Writing – Original Draft, Writing – Review & Editing, Visualization, Project Administration
Marie Luisa Schaper: Validation, Writing – Review & Editing, Supervision, Funding Acquisition
Ute Johanna Bayen: Methodology, Resources, Writing – Review & Editing, Supervision, Funding Acquisition

Die obigen Angaben bezüglich der Beiträge der Doktorandin zu den genannten Arbeiten sind zutreffend.

_____

Prof. Dr. Ute Johanna Bayen

Cognition 254 (2025) 106000



Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit



Relative source credibility affects the continued influence effect: Evidence of rationality in the CIE

Carolin V. Hey, Marie Luisa Schaper *, Ute J. Bayen

Institute for Experimental Psychology, Faculty of Mathematics and Natural Sciences, Heinrich Heine University Düsseldorf, Germany

ARTICLE INFO

Keywords:
Continued influence effect
Rationality
Source credibility
Misinformation
Memory

ABSTRACT

The *Continued Influence Effect* (CIE) is the phenomenon that retracted information often continues to influence judgments and inferences. The CIE is rational when the source that retracts the information (the *retractor*) is less credible than the source that originally presented the information (the *informant*; Connor Desai et al., 2020). Conversely, a CIE is not rational when the retractor is at least as credible as the informant. Thus, a rational account predicts that the CIE depends on the relative credibility of informant and retractor. In two experiments ($N = 151$, $N = 146$), informant credibility and retractor credibility were independently manipulated. Participants read a fictitious news report in which original information and a retraction were each presented by either a source with high credibility or a source with low credibility. In both experiments, when the informant was more credible than the retractor, participants showed a CIE compared to control participants who saw neither the information nor the retraction ($ds > 0.82$). When the informant was less credible than the retractor, participants showed no CIE, in line with a rational account. However, in Experiment 2, participants also showed a CIE when informant and retractor were equally credible ($ds > 0.51$). This cannot be explained by a rational account, but is consistent with error-based accounts of the CIE. Thus, a rational account alone cannot fully account for the complete pattern of results, but needs to be complemented with accounts that view the CIE as a memory-based error.

1. Introduction

The *Continued Influence Effect* (CIE) is the phenomenon that after information has been retracted, it continues to influence judgments and inferences toward agreement with the original information. For example, Seifert (2002) tells an anecdote about a restaurant suspected of causing a family's death by serving poisonous food. Even after learning that the actual cause of death was unrelated to the restaurant, customers continued to avoid the restaurant so that it went out of business. In empirical research, the CIE is usually studied with fictitious reports about an event that unfolds over time. The report contains a critical piece of original information that is subsequently retracted. Despite the retraction, participants continue to agree with the original information (e.g., Ecker et al., 2011; H. M. Johnson & Seifert, 1994; Rich & Zaragoza, 2016; for meta-analyses, see Chan et al., 2017; and Walter & Tukachinsky, 2020). In an era of fast and wide-spreading misinformation (Lewandowsky et al., 2017; Vosoughi et al., 2018), this limited effectiveness of retractions is worrisome. Agreement with incorrect information can lead to judgments and decisions (e.g., whether to support or oppose a policy) that can differ significantly from those that would be made based on facts alone (Kuklinksi et al., 2000). However, as troublesome as it may be in some situations, agreement with retracted information is not necessarily irrational.

The CIE has typically been explained as resulting from erroneous memory processes (Ecker et al., 2022; H. M. Johnson & Seifert, 1994; Lewandowsky et al., 2012; Seifert, 2002; Swire & Ecker, 2018). Connor Desai et al. (2020), however, proposed an explanation for the CIE that is based on rationality: According to their framework, a CIE is considered rational if the source that presented the original information (the *informant*) is deemed more credible than the source retracting the original information (the *retractor*). Conversely, a CIE would be considered irrational if the retractor is deemed more credible than the informant, as per our interpretation of Connor Desai's logic.

Generally, if the CIE is rational, judgments should be consistent with the information provided by the source that is relatively more credible, be it the informant or the retractor. The experiments reported here are

* Corresponding author at: Institute for Experimental Psychology, Building 23.02, Universitätsstr. 1, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany.
E-mail addresses: carolin.hey@hhu.de (C.V. Hey), marie.schaper@hhu.de (M.L. Schaper), sekretariat-bayen@hhu.de (U.J. Bayen).

https://doi.org/10.1016/j.cognition.2024.106000
Received 20 November 2023; Received in revised form 20 August 2024; Accepted 26 October 2024
Available online 18 November 2024
0010-0277/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

the first to put this rational account of the CIE to experimental test by investigating whether the CIE is affected by the *relative source credibility* of informant and retractor.

In the following sections, we first describe the standard CIE paradigm. Then, we review the theoretical accounts of the CIE. We present initial evidence for a rational account and derive hypotheses regarding effects of relative source credibility on the CIE.

1.1. The CIE paradigm

In their pioneering experiments on continued influence, Wilkes and Leatherbarrow (1988) and H. M. Johnson and Seifert (1994) introduced the paradigm that is now standard in CIE research. In a study phase, participants read a news report about a fictitious event. Within this report, participants are presented with a critical piece of original information, typically a possible cause of the described event. Later in the report, this original information is retracted. For example, O'Rear and Radvansky (2020) used a news report about a vegetable company that was forced to recall its product after it was found to be contaminated. Early in the report, participants received the critical piece of original information that the product was contaminated due to unsanitary equipment in the processing plant. Following further non-critical information on the incident, the report later contained the retraction that the product was not contaminated due to unsanitary equipment.

During a subsequent test phase, participants typically answer questions that require judgments and inferences about information that is not explicitly mentioned in the report but relates to the original information. Alternatively, the participants rate their agreement with statements that are consistent with the original information. For example, O'Rear and Radvansky (2020) asked participants how the event in question could be avoided in the future. Answers in agreement with the original information (e.g., ensuring better sanitation in the processing plant) were scored as showing continued influence of the original information despite the retraction. Typically, retractions led to less continued influence of the original information compared to original information that was not retracted (e.g., Ecker et al., 2020, 2017, 2014, 2011; Ecker & Ang, 2019; Guillory & Geraci, 2016; Rich & Zaragoza, 2016). Critically, however, in a myriad of studies with various materials, retractions did not eliminate the continued influence of the original information, that is, participants answered in agreement with the original information despite the retraction (e.g., Buczel et al., 2022; Ecker & Antonio, 2021; Ecker et al., 2021; Ithisuphalap et al., 2020; Kan et al., 2021; McIlhiney et al., 2022; Miller et al., 2022; Rich et al., 2023; Sanderson et al., 2022, 2021; Susmann & Wegener, 2022, 2023; Tay et al., 2022; van Huijstee et al., 2022; for meta-analyses, see Chan et al., 2017; and Walter & Tukachinsky, 2020).

1.2. Theoretical accounts of the CIE

Two cognitive accounts dominate the ongoing discussion about the mechanisms underlying the CIE (e.g., Ecker et al., 2022; Lewandowsky et al., 2012; Sanderson & Ecker, 2020). According to the *model-updating* account, unfolding events are represented as mental models in memory (Wilkes & Leatherbarrow, 1988). It takes effort to remove original information from a complete and coherent model and/or to integrate an additional retraction into such a model. Further, removing original information without plausible alternative information causes a gap in the model thereby rendering it implausible. According to this account, people fail to update information from complete models and are reluctant to use implausible models (Brydges et al., 2018; Hamby et al., 2020; McIlhiney et al., 2022). They, therefore, continue to use models that contain the original information to make judgments and inferences.

Another account is the *retrieval-failure account* (e.g., Sanderson & Ecker, 2020; also referred to as the *concurrent-storage hypothesis*, e.g., A. Gordon et al., 2019). According to this account, both original information and retraction are encoded in memory, associated with the presented context, and simultaneously retained in an activation-based semantic network (Ayers & Reder, 1998). When people make judgments and draw inferences, both memory representations compete for automatic activation. The automatic activation of the original information is likely to be misinterpreted as evidence that this information is true unless the retraction is also retrieved. However, retrieval of the retraction requires a strategic and cognitively demanding assessment of the validity of the original information (Ayers & Reder, 1998). The continued influence of the original information results from failure to engage in these strategic retrieval processes, such as monitoring of source and veracity (cf. Begg et al., 1992). These strategic processes would be necessary to identify the original information as false and to suppress its incorporation into judgments and inferences (Ecker et al., 2011; Swire, Ecker, & Lewandowsky, 2017). It is also possible that strategic retrieval processes are engaged, yet fail and thus, for example, lead to source-monitoring error (cf. M. K. Johnson et al., 1993). For instance, retrieving the original information, but misattributing it to the wrong source (i.e., the retractor) could lead to continued agreement with the original information (Lewandowsky et al., 2012).

As Connor Desai et al. (2020) pointed out, both the model-updating account and the retrieval-failure account describe the CIE as the result of errors in memory processes (i.e., mental-model integration or updating, or failures at retrieval). These accounts assume that retractions should normatively lead to a cancellation of agreement with retracted information. That is, in error-free judgments, the initial original information should always be disregarded in favor of the later retraction. In line with this notion, previous CIE researchers treated the original information as always and undoubtedly false (labeling it "misinformation") and treated the subsequent retraction as always and undoubtedly true. They thus established a normative interpretation of the CIE. Evaluated against this norm, the CIE is irrational.

However, such a norm is not necessarily appropriate for evaluating participants' judgments. Critically, rational people may not share the assumption that retractions are always truthful because, in real life, new information is not necessarily true just because it is more recent. For example, in March of 2022, Russian foreign minister Lavrov tried to incorrectly retract allegations of a military attack on Ukraine (Wintour, 2022). Thus, retractions may rectify false information, but they may also be false denials of accurate information. Accordingly, in CIE studies, not all participants believed the retractions (Guillory & Geraci, 2010; O'Rear & Radvansky, 2020; Rich & Zaragoza, 2016). Thus, the prevalent conceptualization of CIE scenarios (i.e., *incorrect* original information being *correctly* retracted) may conflict with participants' perspectives on the scenarios. Critically, O'Rear and Radvansky (2020) found that belief in the veracity of the retraction affected judgments and inferences: Participants who did not believe the retraction agreed as much with the original information as control participants who never encountered a retraction. By contrast, when participants believed the retraction, it reduced the agreement with the original information compared to control participants. O'Rear and Radvansky concluded that the continued influence of retracted original information is at least partially due to participants disbelieving the retraction.

Connor Desai et al. (2020) presented a conceptualization of CIE scenarios that appears closer to participants' perspectives and that is rooted in Bayesian rationality (for a general discussion on human Bayesian rationality, see Oaksford & Chater, 2007). Connor Desai et al. conceptualized the CIE paradigm as a scenario in which informant and retractor contradict each other regarding the veracity of the original

information. Participants may not immediately know whether the informant or the retractor provides incorrect information, resulting in a state of uncertainty. Critically, participants may evaluate the credibility of the informant and the retractor[1] and use the credibility to inform their judgments. In a variety of other paradigms (e.g., perceptual integration, Ernst & Banks, 2002; advice integration, Birnbaum et al., 1976), participants integrated pieces of information from different sources that varied in credibility by weighting each piece of information with the respective source credibility and then averaging the information to form a judgment. This integration procedure was deemed optimal by the researchers (cf. Ernst & Banks, 2002). Thus, in the CIE paradigm, participants could similarly use the available source credibility of the informant and retractor to weigh the original information and retraction against each other. According to Bayesian rationality, "unless one views both bodies of evidence [i.e., both sources' credibility] as entirely equal, in which case nothing changes, weighting the evidence will increase one's belief in the one at the expense of the other" (Hahn & Oaksford, 2006, p. 16). Thus, if the retractor is less credible than the informant, agreement with the original information and, consequently, the CIE may be entirely rational. In this case, people base their judgments and inferences on the original information as it originated from the more credible source. Conversely, if the retractor is more credible than the informant, a retraction should stop agreement with the original information. Thus, according to this account, a CIE is the result of a rational consideration of relative source credibility.

The aim of the current research was to conduct the first experimental test of this rational account of the CIE. To this end, we manipulated informant and retractor credibility and tested whether participants agreed with the information originating from the more credible source and thus rationally used relative source credibility as an indicator of statement veracity. In the following sections, we review the effects of source credibility on agreement with retracted original information. We then derive hypotheses regarding effects of (relative) source credibility on agreement with the original information and the CIE.

1.3. Effects of source credibility

The credibility of a source reflects the probability with which this source provides correct and reliable information. Source credibility is considered to be a multidimensional concept, with expertise and trustworthiness as the most prominent and widely accepted dimensions (Pornpitakpan, 2004). Expertise refers to the ability to provide correct information and can be achieved through the accumulation of experience, knowledge, skills, or training. Trustworthiness refers to the willingness to provide truthful information and can be achieved through moral integrity, demonstrated goodwill, and the conveyed impression that the source itself believes the provided information to be true (e.g., Pornpitakpan, 2004; Stadtler & Bromme, 2014). Examples of additional proposed dimensions of credibility are dynamism, competence, completeness, bias, transparency, objectivity, authoritativeness, and character (see Pornpitakpan, 2004; and Rieh & Danielson, 2007, for an overview of the debate). For the current research, we focused on an aggregate level of source credibility.

In general, source credibility affects reactions to information. The more credible a source, the greater its argument force (Hahn & Oaksford, 2007) and the stronger the persuasive influence of its message (for a narrative review, see Pornpitakpan, 2004; and for a meta-analysis, see Wilson & Sherrell, 1993). Effects of source credibility have, for example, been demonstrated in studies of explicit and implicit evaluations (Smith et al., 2013), reasoning performance (Copeland et al., 2011), and the subjective truth of statements (Nadarevic et al., 2020). Participants particularly attend to source information when confronted with contradicting information (Braasch et al., 2012), as for example, in the *misinformation* paradigm, where witnesses' testimonies and answers in memory tests can be distorted by post-event misinformation (Loftus, 2005). Only when post-event misinformation came from a low-credibility source were participants able to disregard some of the misleading information (Echterhoff et al., 2005).

Among the few studies of effects of source credibility conducted in the CIE paradigm, Guillory and Geraci (2013, Experiment 1) were the first to manipulate the source credibility of the retractor, henceforth *retractor credibility*. In their experiment, original information about a politician was retracted by either a high-credibility or a low-credibility retractor (e.g., a district attorney or a political blogger). Additionally, in some experiments, either only the trustworthiness dimension or only the expertise dimension of retractor credibility was manipulated. Guillory and Geraci (2013, Experiment 3) and Ecker and Antonio (2021) manipulated the trustworthiness of the retractor by using an independent and unbiased source (e.g., a water department biologist in a water contamination scenario) or a source with an ulterior motive (e.g., an engineer of the accused company; see also Pluviano et al., 2020), with both retractors being experts. Susmann and Wegener (2023) manipulated the expertise of the retractor by using a source with more than 20 years of experience and special training or a source with less than one year of experience and no training. All experiments showed similar effects: High retractor credibility, trustworthiness, and expertise led to less agreement with the original information than low retractor credibility, trustworthiness, and expertise. Further, participants who encountered the low-credibility/-trustworthiness retractor agreed as much with the original information as participants who did not encounter a retraction. By contrast, participants who encountered the high-credibility/-trustworthiness retractor agreed less with the original information than participants who did not encounter a retraction.

Recently, Westbrook et al. (2023) manipulated the source credibility of the informant, henceforth *informant credibility*, by introducing a source devaluation presented by the retractor. Following the retraction, the retractor either alleged that the informant had an ulterior motive for spreading the purportedly false information or did not comment on the informant's motive. Neither informant credibility nor retractor credibility were specified otherwise. Devalued informant credibility was rated lower than unvalued informant credibility, resulting in less agreement with the original information. Furthermore, Westbrook et al. showed that rated informant credibility was a strong positive predictor of agreement with the original information (see also Walter & Tukachinsky, 2020, for a meta-analysis on the impact of informant credibility, which showed that studies using high-credibility informants, e.g., a police report, showed more agreement than studies using low-credibility informants, e.g., a rumor). However, to the best of our knowledge, no previous study has directly manipulated informant credibility through different explicit source descriptions, as was done, for example, by Guillory and Geraci (2013) when examining retractor credibility.

To conclude, extant research suggests that informant credibility and retractor credibility antagonistically affect how much participants agree with the original information in a CIE paradigm. However, informant and retractor credibility have never been manipulated within the same experiment. Thus, experimental evidence of combined effects of informant and retractor credibility on agreement with the original information is missing. We will refer to the combined effects of informant and retractor credibility as the effect of their relative source credibility.

[1] Note that it is beyond the scope of the current research to assess whether an evaluation of source credibility itself is rational. Differences in evaluations of source credibility may emerge due to individual differences (e.g., culture; Morimoto & La Ferle, 2008), different attitudes toward source characteristics (e.g., Imhoff et al., 2018; Swire, Berinsky, et al., 2017; Traberg & van der Linden, 2022), use of different heuristics (e.g., Metzger et al., 2010), or biases (e.g., Patzer, 1983). Some factors may be valid indicators of source credibility and, by extension, of information veracity, whereas others may not (R. Gordon & Spears, 2012).

1.4. The current experiments

In two experiments, we conducted the first rigorous experimental tests of a rational account of the CIE by testing the effects of relative source credibility of informant and retractor on the CIE. We manipulated informant credibility and retractor credibility orthogonally between participants in a standard CIE paradigm and measured participants' agreement with the original information. In both experiments, we used informants and retractors who reported on the incident in question (e.g., a food recall, see Table 1). In the unfolding news reports, original information was either presented by a high-credibility or a low-credibility informant. Later, this original information was either retracted by a high-credibility or a low-credibility retractor. The design thus included four experimental groups with three levels of relative source credibility: In the *informant-higher* group, relative source credibility favored the informant. In the *retractor-higher* group, relative source credibility favored the retractor. In the *both-high* and *both-low* groups (the *equal* groups), relative source credibility favored neither source because both sources were about equal in credibility.

1.4.1. Measurement of the CIE

We used a *direct measure* to assess participants' agreement with the original information and an *indirect measure* to assess their agreement with inferences consistent with the original information (following Ithisuphalap et al., 2020; Rich & Zaragoza, 2016, 2020; Susmann & Wegener, 2022, 2023; Westbrook et al., 2023). Although these two aspects of agreement with the original information are inherently related and often highly correlated (e.g., Ithisuphalap et al., 2020; Rich & Zaragoza, 2016), Susmann and Wegener (2022) argued that they are not redundant. For example, the retraction may eliminate the belief in the original information, yet participants may agree with inferences consistent with the original information because it is the only information available. Further, the direct and indirect measures may depend to different extent on deliberate considerations and rationality and may, therefore, be influenced by source credibility to different extent. Nevertheless, we expected effects of informant credibility, retractor credibility, and their relative credibility on both measures.

When assessing the CIE, choosing an appropriate control group is important. To test effects of retractions, some researchers used a *no-retraction* control group (e.g., Connor Desai & Reimers, 2019; Guillory & Geraci, 2013; O'Rear & Radvansky, 2020). Such control groups received the original information, but never encountered the retraction. They had, thus, no reason to doubt the veracity of the original information. Comparing agreement with the original information between such a control group and experimental groups that did receive a retraction, thus, indicates how much a specific retraction reduced agreement with the original information (i.e., the isolated *retraction effect*). In the current experiments, however, we wanted to know how strongly retracted original information would influence judgments and inferences (i.e., the combined effects of original information and retraction). Therefore, we

Table 1
Translated News Report in Experiment 1: Food Recall after Numerous Cases of Illness in Southern Germany.

Paragraph	Function	Group		Text	
1	Report	All		The German food company Zefter Vegetables Ltd. is recalling all packages of baby spinach from its Grüneri brand after the vegetable was found to be contaminated in a way that harms human health.	
2	Report	All		Zefter Vegetables Ltd. is one of the main employers in the Weißenhorn district and has enjoyed an excellent reputation so far. The company was founded in 1968 and has since been owned by the Zefter family, which now manages it in the third generation. This is the first case of contaminated goods of the company.	
3	Informant credibility manipulation	Informant higher, both high	Retractor higher, both low	*High-credibility informant* [Name A], investigative journalist, comments on the case. Recently, he had exposed the unsanitary conditions of a Dutch fast-food chain. As a result, he was awarded an important journalism prize.	*Low-credibility informant* [Name A], a self-proclaimed scandal blogger, comments on the case. Recently, on his website "My Scandals", he had made groundless accusations against a Dutch fast-food chain.
	Original information	All groups except control		[Name A] now publishes a post in which he writes: "The contamination of the spinach was caused by the unsanitary conditions in the production facility of Zefter Vegetables Ltd."	
4	Report	All		More than twenty people have fallen ill after eating the spinach. Cases have been reported in Bavaria and Baden-Wuerttemberg. All those affected were admitted to local hospitals with severe gastrointestinal complaints and acute dehydration. Four patients are in critical condition and their physicians cannot rule out the possibility of lasting damage.	
5	Report	All		The investigation by the responsible health department into the cause of contamination and possible culprits is still ongoing. An official statement by the authorities regarding the circumstances is not available at this time. Usually, it takes several weeks for investigations to be completed in such cases.	
6	Report	All		In a television interview, the CEO of Zefter Vegetables Ltd. expressed regret over the incident, stating, "We apologize to all those affected by the contaminated spinach. We will fully cooperate with the investigation of the authorities." Meanwhile, Zefter's 235 employees fear for their jobs.	
7	Retractor credibility manipulation	Retractor higher, both high	Informant higher, both low	*High-credibility retractor* A few days later, [Name B], journalist and fact-checker, gives a statement. He was recently awarded a prestigious journalism prize for numerous exposures of false reports and journalistic misconduct.	*Low-credibility retractor* A few days later, [Name B], food blogger from the website "The Gourmet World", gives a statement. He is known, among other things, for his dubious connections with the food industry lobby.
	Retraction	All groups except control		[Name B] publishes a post in which he firmly rejects the accusations made by [journalist/scandal blogger] [Name A]. [Name B] contradicts: "There is no evidence of unsanitary conditions at the Zefter production facility and the investigation by the health department will soon rule this out as the cause."	
8	Report	All		The responsible health department reports that the situation is under control and no more contaminated spinach is being distributed. All products packed in the same period as the contaminated spinach have been removed from grocery stores and destroyed.	
9	Report	All		Consumers are still urged to discard all baby green spinach products by the Grüneri brand packaged between June 12 and June 28. Products can also be returned to the grocery stores. For specific questions or concerns, customers can call the Zefter Info Hotline at (0800) 4890501.	

Note. The original German news report is available in the Open Science Framework at https://osf.io/hwxjy. It was adapted and translated from O'Rear and Radvansky (2020). The paragraphs were presented one at a time to the participants of the listed groups. Credibility manipulations are indicated in italics. In the presented reports "Uwe Schmidt" and "Ralf Müller" were assigned to "Name A" and "Name B" according to the counterbalancing scheme. All = informant higher, retractor higher, both high, both low, and control group. Informant higher = high-credibility informant and low-credibility retractor; both high = high-credibility informant and high-credibility retractor; both low = low-credibility informant and a low-credibility retractor; retractor higher = low-credibility informant and high-credibility retractor; control = participants encountered neither an informant nor a retractor.

compared agreement with the original information in the experimental groups (which all received both original information and retraction) with the spontaneous agreement with this information in a *no-information* control group. This control group encountered neither original information nor retraction (some researchers referred to such a control group as a *no-misinformation* control, e.g., Buczel et al., 2024; Ecker & Antonio, 2021; Ecker et al., 2011).

The no-information control group read the same news report as the experimental groups, but skipped the paragraphs containing the original information and the retraction, and answered the same questions as the experimental groups (e.g., Buczel et al., 2024; Ecker et al., 2011; Ecker & Antonio, 2021; Rich & Zaragoza, 2016; Thorson, 2016). Thereby, the control group provided a baseline measurement of spontaneous agreement (i.e., the extent of agreement with the original information induced by knowledge of just the general scenario, without knowing the specific allegations included in the original information). In the experimental groups, exposure to the original information should increase agreement with it beyond spontaneous agreement. Exposure to a perfectly effective retraction should revert this increased agreement with the original information at least to the level of spontaneous agreement. We only refer to an effect as a *CIE* when the agreement of an experimental group exceeded the spontaneous agreement of the control group. When an experimental group showed as little agreement with the original information as the control group (or less), this indicates that the effect of original information and retraction canceled each other out (or the retraction was even more effective than the original information). The difference in agreement between an experimental group and the no-information control group is thus an estimate of the CIE unconfounded by spontaneous agreement. Where experimental groups were compared to each other, we refer to differences in *agreement with the original information*.

1.4.2. Hypotheses

From a rational account of the CIE, we derived several hypotheses regarding the effects of (relative) source credibility on agreement with the original information and the CIE in our experiments. A rational account predicts antagonistic effects of informant and retractor credibility on agreement with the original information. High informant credibility should lead to more agreement than low informant credibility (e.g., Westbrook et al., 2023) and, as the retraction should reduce agreement with the original information, high retractor credibility should lead to less agreement than low retractor credibility (e.g., Guillory & Geraci, 2013).

Regarding the effects of relative source credibility on the CIE, the rational account predicts that participants should agree with the more credible of the two sources. When the informant is more credible than the retractor, participants should agree more with the original information than the control group. Thus, the informant-higher group should show a CIE. When the retractor is more credible than the informant, participants should agree with the retraction and, thus, disagree with the original information. The retractor-higher group should thus show the same level of agreement with the original information as the control group or less and show no CIE.

The rational account further predicts that in the two equal groups, where informant and retractor are equally credible so that relative source credibility favors neither source, the level of agreement with the original information should be the same, regardless of whether informant and retractor credibility is equally high or equally low. Further, a purely rational account predicts that when informant and retractor are equally credible, the effects of informant credibility and retractor credibility cancel each other out when weighting them against each other. Thus, participants' initial uncertainty would not change based on the available information (cf. Hahn & Oaksford, 2006). This uncertainty cannot justify agreement with the retracted information. Thus, the equal groups should show no CIE.

Effects of source credibility on the CIE are not necessarily inconsistent with the memory processes assumed in the model-updating and retrieval-failure accounts. Ecker and Antonio (2021) discussed that source credibility may affect the CIE because it may modulate the memory processes postulated to underlie the CIE. In general, the credibility of a source is thought to influence memory for the information it provides, because source credibility affects how deeply the information is elaborated and which thoughts are generated in response to the information (cf. Heesacker et al., 1983; Petty & Cacioppo, 1986; Priester et al., 1999; Sternthal, Dholakia, & Leavitt, 1978; Sternthal, Phillips, & Dholakia, 1978; Tobin & Raymundo, 2009).

A rational account of the CIE and accounts that attribute the CIE to irrational error diverge particularly in their predictions for the two equal groups and the retractor-higher group. As said, a rational account predicts no CIEs in these groups because the retractor is at least as credible as the informant. An irrational CIE, however, would occur in all groups, that is, even when the retractor is substantially more credible than the informant, there should be some irrational continued influence of the original information.

2. Experiment 1

2.1. Method

2.1.1. Design

The experiment had a 2 × 2 factorial design with informant credibility (high, low) and retractor credibility (high, low) as between-subjects factors resulting in four experimental groups: the informant-higher group, the retractor-higher group, the both-high group, and the both-low group. To obtain baseline measurements for all dependent variables, we included a control group that was not exposed to the original information, the retraction, or the respective sources. This resulted in five groups in total. Agreement with the original information was the main dependent variable. We used a direct and an indirect measure to assess participants' agreement (e.g., following Ithisuphalap et al., 2020). The *direct measure* was a rating of agreement with the verbatim original information. The *indirect measure* was a rating of agreement with a statement indirectly referring to the original information. Further, for manipulation checks, we had participants rate informant and retractor credibility and measured their memory for the contents of the report.

2.1.2. Participants

The local ethics committee approved the experiments. They were conducted online between September 22 and November 9, 2020. Participants were recruited via a participant pool maintained by the research group and social media. Participants who were psychology students at Heinrich Heine University Düsseldorf, Germany, received partial course credit as compensation. The others volunteered without compensation.

An experiment on retractor credibility by Guillory and Geraci (2013; Experiment 1) produced a medium-sized effect for the comparison of agreement with the original information between a high- and a low-credibility retractor. We, therefore, anticipated medium effect sizes for the main effects in a 2 × 2 ANOVA comparing agreement ratings for different levels of informant and retractor credibility. A power analysis (using G*Power, Faul et al., 2007) showed that 119 participants in the four experimental groups (30 in each group for even distribution) would be sufficient to detect medium effects ($f = 0.30$) with a power of $1 - \beta = .90$ and $\alpha = .05$. After adding the control group, the total target sample size was, therefore, 150 participants.

We collected data from 171 participants to account for exclusions. A priori exclusion criteria were failure in an attention check during the presentation of the report (8 excluded), incorrect summary of the report (1 excluded), and self-assessment as unserious or inattentive in a seriousness check (Aust et al., 2013; 7 excluded). We further excluded participants if they performed worse than three standard deviations

below the mean of memory for the content of the report (1 excluded), or if their study completion time exceeded the mean by more than three standard deviations (calculated separately for each group; 3 excluded). The pattern of results remained unchanged when all participants were included.

The final sample thus consisted of 151 participants (120 female, 31 male), with 119 participants in the experimental groups. Group sample sizes are listed in Table 2. All participants were native speakers of German with ages ranging between 17 and 32 years ($M = 21.75$, $SD = 3.65$) and had at least a high school diploma (German Abitur). Participants were randomly assigned to the five groups. All five groups were equivalent in age, gender, and educational degree obtained (all $ps > .308$).

2.1.3. Materials and procedure

2.1.3.1. Study phase. All materials of the two experiments were administered in German and are available in the Open Science Framework at https://osf.io/hwxjy. An English translation of the news report including the manipulated source descriptions is in Table 1 and a translation of the test items is in Appendix A. Participants completed the experiment at home on their personal computer or laptop. After giving informed consent, participants were informed that their memory and understanding of the following news report would be tested. They then read a fictitious news report. We adapted a report about a food recall from O'Rear and Radvansky (2020). In the report, an anonymous author described that a vegetable company had to recall its spinach after the product was found to be contaminated. The report consisted of nine paragraphs, each one to four sentences in length. Each paragraph appeared on a separate screen such that participants read the report one paragraph at a time, without backtracking. As in previous studies (e.g., Ecker et al., 2010; H. M. Johnson & Seifert, 1994; Kan et al., 2021), participants did not have the option to return to a previous paragraph so that they could not reread the original information once they had noticed the contradiction. To ensure that participants read the complete report without proceeding prematurely, each paragraph was presented for a minimum period of time based on its length and the average German reading speed of 0.34 s per word (Trauzettel-Klosinski et al., 2012). Only after the minimum period of time had elapsed could participants proceed to the next paragraph at their own pace. The original information was presented in paragraph 3 and retracted in paragraph 7.

In the experimental groups, the original information and the retraction were each presented by either a high-credibility or a low-credibility source. In the German version, each source description had a length of 20 words. To ensure that each source description could unfold its full effect, it immediately preceded (on the same screen) the original information or retraction (Ward & McGinnies, 1974). To highlight the high or low credibility of the sources, we manipulated both their capability and willingness to provide accurate information (i.e., expertise and trustworthiness, Pornpitakpan, 2004) by mentioning their profession (journalist or blogger) and an example of an occasion where the sources had shared accurate or false/misleading information in the past. Each participant of the four experimental groups encountered two sources with common male German names ("Ralf Müller" and "Uwe Schmidt"). Assignment of names to the original information or the retraction was approximately counterbalanced. We presented the original information and the retraction in direct speech to emphasize the individual sources as origins of the statements in contrast to the anonymous author of the report. The participants of the control group skipped the two manipulated paragraphs so that they did not encounter the original information, the retraction, or any source.

The final paragraph of the report contained an attention check that instructed participants to ignore the "Continue" button and instead click on a logo at the upper-right corner of the screen. Participants who clicked on "Continue" failed the check and were excluded from analyses. After the report, participants completed a two-minute filler task in which they classified simple arithmetic equations of the type A + B + C = D as correct or incorrect.

2.1.3.2. Test phase. In the ensuing test phase, participants rated their agreement with eight statements about the described incident on a 7-point Likert scale ranging from 1 (*strong disagreement*) to 7 (*strong agreement*). Each statement was presented on a separate screen, in fixed order (see Appendix A for English translations). Five of the statements served to probe memory for the contents of the report (Statements 1, 7, and 8 were correct; Statements 2 and 4 were incorrect). Statement 3 served as the direct measure of agreement with the original information and as such reproduced the original information verbatim. Statement 6 served as the indirect measure of agreement with the original information and as such was a possible inference from the original information. Participants should agree with these statements only if the original information continued to influence their judgments and inferences despite the retraction. Similar statements have been used previously to measure agreement with retracted original information (e.g., Ecker & Antonio, 2021). Statement 5 served to mask the subjective nature of the indirect measure by also requiring a subjective evaluation.

Next, participants rated the credibility of the sources on a 7-point Likert scale ranging from 1 (*absolutely not credible*) to 7 (*absolutely credible*). The experimental groups rated the credibility of their respective informant and retractor as shown during the experiment (e.g., the informant-higher group rated the high-credibility informant and the low-credibility retractor). The control group rated the credibility of all four sources. Note that the control group did not encounter any of these sources in their version of the report. Therefore, their ratings could not be influenced by informant and retractor contradicting each other and thus served as a manipulation check of informant credibility and retractor credibility. To introduce the sources to the control group and to reduce memory effects in the experimental groups, the complete source descriptions were presented along with the rating scale. However, the information which source issued the original information and which the retraction was not repeated. In the experimental groups, the order of the two credibility ratings (informant, retractor) was approximately counterbalanced. In the control group, the four credibility ratings were approximately counterbalanced in a Latin square design.

Then, participants gave a summary of the report and provided demographic information. Finally, they completed a seriousness check (Aust et al., 2013) in which they indicated whether they participated seriously and attentively and whether their data could be used in

Table 2
Discrimination index P_R and bias index B_R for the Memory Test in Experiments 1 and 2.

Experiment	Group	*n*	P_R		B_R	
			M	*SE*	*M*	*SE*
1	Control	30	.54	.02	.46	.02
	Informant higher	29	.54	.02	.46	.01
	Retractor higher	31	.53	.02	.47	.02
	Both high	29	.54	.02	.46	.02
	Both low	32	.53	.02	.48	.02
2	Control	31	.55	.02	.53	.02
	Informant higher	28	.51	.04	.54	.02
	Retractor higher	29	.51	.03	.53	.02
	Both high	28	.51	.04	.57	.03
	Both low	29	.49	.04	.55	.02

Note. Means and standard errors of the discrimination index P_R and the bias index B_R of the two-high threshold model of recognition memory (Corwin, 1994; Snodgrass & Corwin, 1988; see also Footnote 3) for each group. Control = participants encountered neither an informant nor a retractor; informant higher = high-credibility informant and low-credibility retractor; retractor higher = low-credibility informant and high-credibility retractor; both high = high-credibility informant and high-credibility retractor; both low = low-credibility informant and low-credibility retractor.

scientific analyses. They were assured that their compensation for the participation was unaffected by their answer on the seriousness check. Finally, participants were debriefed and, if eligible, received partial course credit. Mean completion time was about 13 min.

2.2. Results

The data of both experiments are available in the Open Science Framework at https://osf.io/hwxjy. For all analyses, we provide frequentist and Bayesian inference statistics. We set $\alpha = .05$ for all frequentist analyses unless noted otherwise. Effect-size interpretations follow Cohen (1988). We performed Bayesian analyses with JASP statistical software (JASP Team, 2024) using default priors. We report Bayes' Factors (BF) in such a way that they indicate support for the respective favored hypothesis. For *t* tests, one-way ANOVAs, correlations, and contrasts (Bayes' Factors for all contrasts are based on pairwise *t* tests) we report BF_{01} indicating support for the null hypothesis and BF_{10} indicating support for the alternative hypothesis. For the effects of two-way ANOVAs, we report BF_{incl} (BF_{excl}) indicating support for the inclusion (exclusion) of the effect based on matched models. For example, a Bayes' Factor of $BF_{10} = 5.00$ indicates that the observed data are five times more likely under the alternative hypothesis than under the null hypothesis. Based on the conventions by Raftery (1995), BFs larger than 3 indicate positive evidence, larger than 20 indicate strong evidence, and larger than 150 indicate very strong evidence.

In the following paragraphs, we first present results of the manipulation checks and compare memory for the content of the report between groups. Then, we test the effects of informant and retractor credibility and the effect of their relative credibility on agreement with the original information and the CIE.

2.2.1. Manipulation checks

2.2.1.1. Source credibility ratings in the control group. Fig. 1 shows descriptive statistics for the ratings of source credibility for all groups. For the manipulation check, we first performed three planned repeated-measures contrasts to compare the four source credibility ratings provided by the control group because these ratings were unaffected by possible reactive effects of original information and retraction presentation. As expected, the two high-credibility sources were rated as more credible than the two low-credibility sources, $F(1, 29) = 95.95$, $p < .001$, $\eta_p^2 = .77$, $BF_{10} = 4.42 \cdot 10^{14}$. As intended, there was no difference in credibility ratings of the two high-credibility sources, $F(1, 29) = 2.36$, $p = .136$, $\eta_p^2 = .08$, $BF_{01} = 1.80$, and the two low-credibility sources, $F(1, 29) = 1.39$, $p = .248$, $\eta_p^2 = .05$, $BF_{01} = 2.74$.

2.2.1.2. Source credibility ratings in the experimental groups. Next, we performed two 2 (informant credibility) × 2 (retractor credibility) ANOVAs to test for the effects of the source credibility manipulations on informant and retractor credibility ratings in the experimental groups. Informant credibility ratings showed the intended effect of the informant credibility manipulation, $F(1, 117) = 33.69$, $p < .001$, $\eta_p^2 = .22$, $BF_{incl} = 1.99 \cdot 10^6$. They were not affected by the retractor credibility manipulation, $F(1, 117) = 0.26$, $p = .612$, $\eta_p^2 < .01$, $BF_{excl} = 4.75$, and there was no interaction, $F(1, 117) = 9.04$, $p = .118$, $\eta_p^2 = .02$, $BF_{excl} = 1.39$. Likewise, retractor credibility ratings showed the intended effect of the retractor credibility manipulation, $F(1, 117) = 57.19$, $p < .001$, $\eta_p^2 = .33$, $BF_{incl} = 6.97 \cdot 10^8$. They were not affected by the informant credibility manipulation, $F(1, 117) = 3.26$, $p = .074$, $\eta_p^2 = .03$, $BF_{excl} = 1.22$, and there was no interaction, $F(1, 117) = 0.02$, $p = .889$, $\eta_p^2 < .01$, $BF_{excl} = 3.83$. To conclude, informant and retractor credibility manipulations were successful and selectively affected the intended source credibility.

2.2.1.3. No reactive effects of retraction presentation. Connor Desai et al. (2020) and Westbrook et al. (2023) suggested that presentation of a retraction may reactively reduce the credibility of the informant and/or the retractor. To test for a reactivity effect, we performed four planned contrasts in which we compared the credibility ratings of each of the four sources (high- and low-credibility informants, and high- and low-credibility retractors) from the two relevant experimental groups against the ratings in the control group. The experimental groups did not

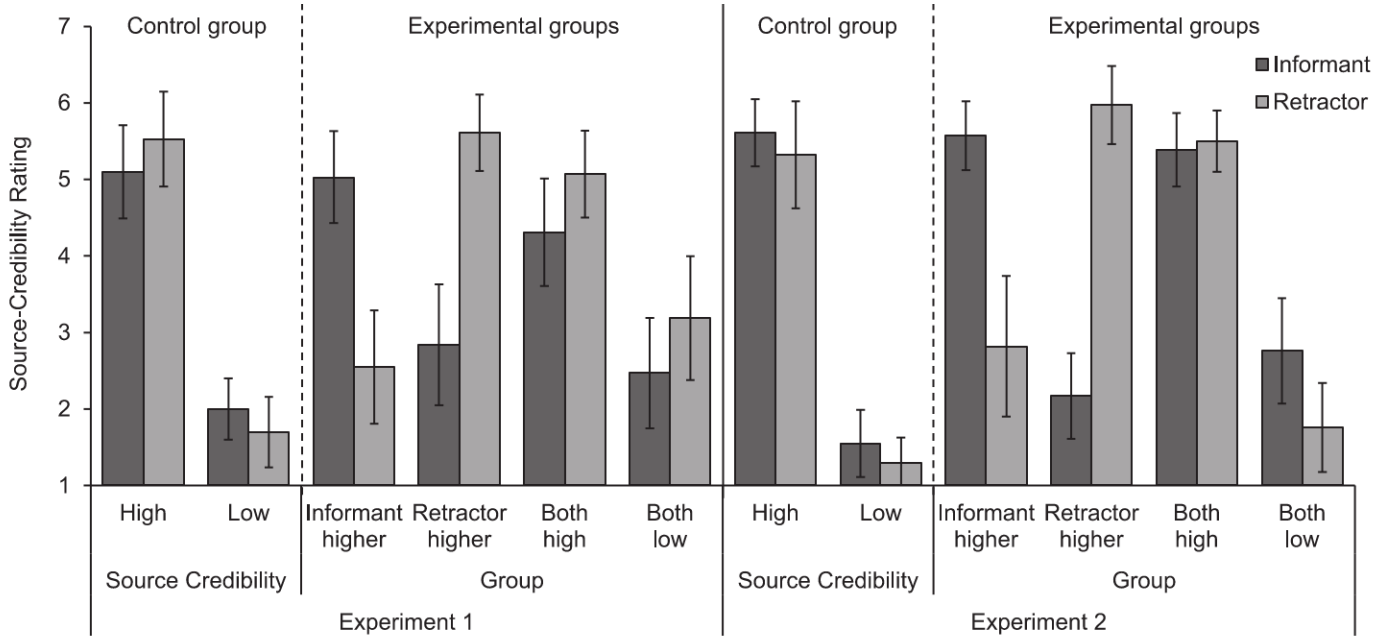


Fig. 1. Source-Credibility Ratings in Experiments 1 and 2.
Note. Mean source-credibility ratings on a scale from 1 (*absolutely not credible*) to 7 (*absolutely credible*). The control group rated all four sources. The experimental groups rated the respectively encountered sources: The informant-higher group rated the high-credibility informant and low-credibility retractor, the retractor-higher group rated the low-credibility informant and high-credibility retractor, the both-high group rated the high-credibility informant and high-credibility retractor, and the both-low group rated the low-credibility informant and low-credibility retractor. Error bars indicate 95 % credible intervals.

rate any source as less credible than the control group did. The low-credibility retractor was even rated as more credible in the experimental groups than in the control group, $F(1, 88) = 7.95$, $p = .006$, $\eta_p^2 = .08$, $BF_{10} = 7.23$ (all other $ps > .108$, $BF_{01} > 1.39$).[2] Thus, neither source was rated as less credible because a retraction was presented.

2.2.2. Memory for the contents of the report

Five statements probed participants' memory for the contents of the news report. Participants rated their agreement with three correct and two incorrect statements about the report on a 7-point scale. To compare memory for the contents of the report and response bias between groups, we calculated the discrimination index P_R and the bias index B_R of the two-high threshold model of recognition memory (e.g., Corwin, 1994; Snodgrass & Corwin, 1988) from these ratings following a procedure for responses on rating scales (Macmillan & Creelman, 2005).[3] Table 2 lists mean P_R and B_R for all groups. One-way ANOVAs across all groups showed no difference in the discrimination index P_R, $F(4, 146) = 0.05$, $p = .996$, $\eta_p^2 < .01$, $BF_{01} = 41.88$, and no difference in the bias index B_R, $F(4, 146) = 0.28$, $p = .888$, $\eta_p^2 = .01$, $BF_{01} = 29.77$. Thus, between-group differences in agreement with the critical items referring to the original information cannot be attributed to differences in memory for the report (Sanderson et al., 2021) nor to differences in response bias.

2.2.3. Effects of informant and retractor credibility

We now turn to the effects of source credibility on agreement with the original information and the appearance of a CIE. Fig. 2 shows descriptive statistics for ratings of agreement with the original information. We expected high informant credibility to lead to stronger agreement than low informant credibility and expected high retractor credibility to lead to less agreement than low retractor credibility. We performed two 2 (informant credibility) × 2 (retractor credibility) ANOVAs to test for the effects of informant and retractor credibility on the direct and indirect measures of agreement with the original information.

2.2.3.1. Direct measure. As predicted, both informant and retractor credibility yielded main effects on the direct agreement with the original information. High informant credibility led to stronger agreement than low informant credibility, $F(1, 117) = 15.77$, $p < .001$, $\eta_p^2 = .12$, $BF_{incl} = 182.27$. Conversely, high retractor credibility led to less agreement than low retractor credibility, $F(1, 117) = 18.80$, $p < .001$, $\eta_p^2 = .14$, $BF_{incl} = 539.46$. These effects replicate results of previous experiments (Ecker & Antonio, 2021; Guillory & Geraci, 2013; Pluviano et al., 2020; Susmann & Wegener, 2023; Westbrook et al., 2023). Informant and retractor credibility did not interact, $F(1, 117) = 1.00$, $p = .320$, $\eta_p^2 = .01$, $BF_{excl} = 2.53$.

Thus, informant and retractor credibility affected direct agreement with the original information independently and antagonistically. This confirms the first two predictions of the rational account for the direct measure and allows us to consider relative source credibility as a linear combination of informant and retractor credibility.

2.2.3.2. Indirect measure. As expected, high informant credibility led to stronger indirect agreement than low informant credibility, $F(1, 117) = 5.60$, $p = .020$, $\eta_p^2 = .05$, $BF_{incl} = 2.43$. Although this effect aligns with our prediction, due to the inconclusive BF, we refrain from a strong interpretation. In contrast to the direct measure, there was no main effect of retractor credibility on the indirect measure, $F(1, 117) = 0.12$, $p = .732$, $\eta_p^2 < .01$, $BF_{excl} = 5.01$. Thus, high retractor credibility did not lead to less indirect agreement with the original information than low retractor credibility. Informant and retractor credibility did not interact, $F(1, 117) = 0.40$, $p = .529$, $\eta_p^2 < .01$, $BF_{excl} = 2.98$.

Taken together, these results may indicate that informant and retractor credibility had little or no effect on indirect agreement with the original information. Possibly, however, our indirect measure did not adequately capture continued influence of the retracted original information on participants' agreement with inferences. Note that, contrary to previous studies (e.g., Ithisuphalap et al., 2020; Rich & Zaragoza, 2016), the direct and indirect measures were only weakly correlated, $r(149) = .24$, $p = .003$, $BF_{10} = 7.88$. Further, agreement with the indirect measure was significantly below the midpoint of the scale in all experimental groups except the informant-higher group, $t(29) = 1.97$, $p = .059$, $d = 0.37$, $BF_{10} = 1.06$ (all other $ps < .005$ and all $BF_{10} > 8.08$ in t tests against 4; see Fig. 2), indicating general disagreement. We discuss possible reasons for the suspected inadequacy of the indirect measure and possible remedies in the introduction to Experiment 2.

2.2.4. Effects of relative source credibility and comparisons to the control group

In our study, an experimental group showed a CIE when they agreed more strongly with the original information than the control group did spontaneously. Correcting for the baseline of spontaneous agreement was necessary because agreement in the control group exceeded the minimum possible agreement (scale floor of 1) in both the direct measure, $t(29) = 7.59$, $p < .001$, $d = 1.39$, $BF_{10} = 5.95 \cdot 10^5$, and the indirect measure, $t(29) = 8.50$, $p < .001$, $d = 1.55$, $BF_{10} = 5.18 \cdot 10^6$. The appearance of a CIE in comparison to the control group should depend on the relative source credibility of informant and retractor. We tested this via four planned contrasts (see Table 3).

2.2.4.1. Direct measure. First, we expected the informant-higher group to show a CIE because, in this group, participants should agree with the more credible informant. Contrast 1 demonstrated that, indeed, the informant-higher group agreed more strongly with the original information than the control group did spontaneously. That is, the informant-higher group showed a CIE.

Second, we expected the retractor-higher group to show no CIE because, in this group, the more credible retractor should stop participants from agreeing with the less credible informant. Contrast 2 demonstrated that, indeed, agreement with the original information in the retractor-higher group did not differ from the spontaneous agreement in the control group. That is, the retractor-higher group showed no CIE. This result speaks to the important influence of relative source credibility on the CIE. The combination of low informant credibility and high retractor credibility eliminated the continued influence of retracted original information, which was rarely achieved in prior research (Ecker et al., 2010; Walter & Tukachinsky, 2020). Together, Contrasts 1 and 2 demonstrate that the CIE depended on relative source credibility. Participants made a rational judgment by agreeing with the information provided by the more credible source.

Third, we expected equal levels of agreement with the original information in the equal groups because relative source credibility did not favor either source in both groups. Contrast 3 demonstrated that

[2] Correction for α inflation in the dependent contrasts using the Bonferroni-Holm adjustment (Holm, 1979) did not change the results.

[3] We determined the hit rates and false alarm rates at each of the six thresholds of the 7-point scale (1, 2, 3, 4, 5, and 6). The hit rate (H_i) was the proportion of correct statements on which a participant gave a rating above the threshold i, corrected using the log-linear rule (Corwin, 1994; Hautus, 1995). The false alarm rate (FA_i) was the corrected proportion of incorrect statements on which a participant gave a rating above the threshold i. For each threshold, we calculated $P_{Ri} = H_i - FA_i$. Overall discrimination was then calculated as the mean of P_{R1} to P_{R6}. Additionally, we calculated $B_{Ri} = FA_i / [1 - (H_i - FA_i)]$. Overall bias was calculated as the mean of B_{R1} to B_{R6}. A greater index value indicates a stronger bias to give higher rather than lower ratings. The null effects reported here and in Experiment 2 did not depend on the use of the indices based on the two-high threshold model of recognition memory. Alternative discrimination and bias indices calculated based on the signal detection model of recognition memory (d' and c; Green & Swets, 1966; Macmillan & Creelman, 2005) showed null effects as well.

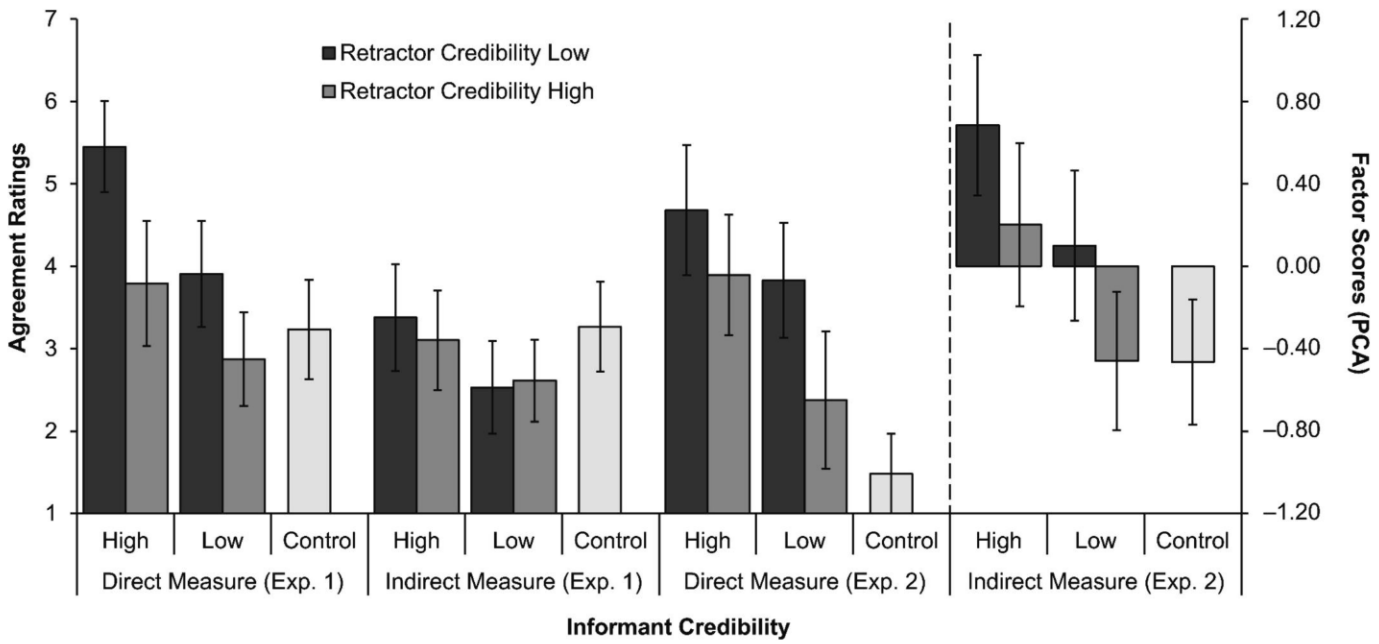


Fig. 2. Effects of Source Credibility on Agreement with the Original Information in Experiments 1 and 2.
Note. Agreement with the original information assessed with the direct and indirect measures by informant credibility and retractor credibility, and for the control group (which received neither original information nor retraction). Both measures of Experiment 1 and the direct measure of Experiment 2 represent the means of agreement ratings that were on a Likert scale from 1 to 7 (see *y*-axis to the left). The indirect measure of Experiment 2 represents the factor scores obtained from a Principal Component Analysis (PCA) on agreement ratings of four statements indirectly referring to the original information (see *y*-axis to the right). Error bars indicate 95 % credible intervals.

Table 3
Planned Contrasts on Agreement Ratings in Experiments 1 and 2.

Exp.	Contrast		Measure	t [a]	p (adj. α)	d	BF_{10}	BF_{01}
1	1	Informant higher vs. control	Direct	5.05	< .001 (.017) *	0.83	$1.60 \cdot 10^{4}$	$6.24 \cdot 10^{-5}$
			Indirect	0.28	.779 (.050)	0.06	0.27	3.67
	2	Retractor higher vs. control	Direct	0.84	.403 (.050)	0.14	0.37	2.74 [b]
			Indirect	1.66	.099 (.017)	0.28	1.02	0.98
	3	Both high vs. both low	Direct	0.26	.794 (.050)	< 0.01	0.27	3.75
			Indirect	1.45	.149 (.050)	0.24	0.60	1.66
	4	Equal vs. control	Direct	1.64	.104 (.025)	0.27	0.66	1.52
			Indirect	1.31	.192 (.025)	0.22	0.51	1.98
2	1	Informant higher vs. control	Direct	6.55	< .001 (.017) *	1.11	$5.98 \cdot 10^{6}$	$1.67 \cdot 10^{-7}$
			Indirect	4.83	< .001 (.017) *	0.82	4884.79	$2.05 \cdot 10^{-4}$
	2	Retractor higher vs. control	Direct	1.85	.066 (.050)	0.31	1.23	0.81
			Indirect	0.03	.981 (.050)	< 0.01	0.26	3.81
	3	Both high vs. both low	Direct	0.13	.895 (.050)	< 0.01	0.27	3.70
			Indirect	0.43	.672 (.050)	0.06	0.29	3.50
	4	Equal vs. control	Direct	5.69	< .001 (.025) *	0.96	$9.23 \cdot 10^{5}$	$1.08 \cdot 10^{-6}$
			Indirect	3.03	.003 (.025) *	0.51	9.53	0.11

Note. Planned contrasts on agreement ratings were performed to test for CIEs and group differences. We used the Bonferroni-Holm adjustment (Holm, 1979) to protect against α inflation in the dependent Contrasts 1, 2, and 4 (adjusted α-levels are in parentheses). Contrast 3 required no adjustment because it was independent of all other contrasts. Bayes' Factors (BFs) are based on pair-wise t tests. d = Cohen's d; Informant higher = high-credibility informant and low-credibility retractor; both high = high-credibility informant and high-credibility retractor; both low = low-credibility informant and low-credibility retractor; retractor higher = low-credibility informant and high-credibility retractor; control = participants encountered neither an informant nor a retractor; equal = joint examination of the both-high and both-low groups.

[a] df for Experiment 1 are 146. df for Experiment 2 are 140. [b] Note that the BF_{01} in favor of the null hypothesis is inconclusive only because the retractor-higher group showed even less agreement with the original information than the control group (in a one-tailed t test, the data are $BF_{0+} = 6.61$ times more likely under the null hypothesis that the retractor-higher group showed as little or less agreement than the control group).

* significant after Bonferroni-Holm adjustment (Holm, 1979).

agreement with the original information did, indeed, not differ between the both-high group and the both-low group.

Fourth, we tested whether the equal groups showed a CIE. Contrast 4 demonstrated that the equal groups agreed no more with the original information than the control group did spontaneously. Thus, these groups did not show a CIE. This result would be in line with a rational CIE based exclusively on relative source credibility and might mean that the effects of equally credible informants and retractors canceled each other out. However, with a weak BF_{01} of 1.52, there was no conclusive evidence for the null effect.

Taken together with the results of Contrasts 1 and 2, this indicates that participants only showed a CIE when relative source credibility favored the informant and that the CIE might be fully rational because it only appeared when relative source credibility provided definite

evidence to justify agreement with the retracted original information. This conclusion would stand in contrast to previous theoretical accounts that proposed erroneous memory processes as the cause of the CIE (Ecker et al., 2022; H. M. Johnson & Seifert, 1994; Lewandowsky et al., 2012; Seifert, 2002; Swire & Ecker, 2018). However, such a strong conclusion would be premature because the evidence for the null hypothesis of Contrast 4 was not conclusive. Experiment 2, therefore, included a replication of this condition.

2.2.4.2. Indirect measure. For the indirect measure, we calculated the corresponding four contrasts. All of them yielded non-significant results. Accordingly, a one-way ANOVA showed that all groups, including the control group, indirectly agreed to equal extent with the original information, $F(4, 146) = 1.92$, $p = .110$, $\eta_p^2 = .05$, $BF_{01} = 2.77$. This indicates that no experimental group showed a CIE compared to the control group. Thus, results from the indirect measure do not confirm the predictions of a rational account of the CIE. However, as said, the absence of relative source-credibility effects in this measure may mean that it did not adequately capture indirect agreement with the original information as a measure of the CIE. We will address this possibility in Experiment 2.

2.3. Discussion

The main objective of Experiment 1 was to examine relative source credibility of informant and retractor as a determinant of rational agreement with retracted original information and the CIE. Therefore, we tested predictions derived from relative source credibility and a rational account of the CIE. With a direct measure of agreement with the original information, these predictions were confirmed. We thus obtained the first experimental evidence that participants made rational judgments about retracted original information: Participants considered both informant and retractor credibility in their agreement with the original information. Interestingly, the main effects of informant and retractor credibility were of similar size (as indicated by their overlapping 95 % confidence intervals; informant credibility: $\eta_p^2 = .12$ [.03, .23], retractor credibility: $\eta_p^2 = .14$ [.04, .26]). This indicates that participants evenly considered both sources' credibility in their judgments and inferences. Further, the absence of an interaction indicates that informant and retractor credibility were considered independently of each other when weighting their respective statements (cf. Birnbaum et al., 1976). Importantly, a relatively more credible retractor stopped participants from showing a CIE altogether.

However, unexpectedly, our experiment did not show the corresponding effects in the indirect measure. Agreement ratings in this measure did not differ between any of the groups including the control group, even though indirect measures have reliably produced a CIE in past research (e.g., Ecker et al., 2011; H. M. Johnson & Seifert, 1994; Rich & Zaragoza, 2016; Wilkes & Leatherbarrow, 1988). In Experiment 2, we therefore aimed to increase the reliability of the indirect measure.

3. Experiment 2

Experiment 2 had two main objectives. First, we sought to replicate the result obtained with the direct measure of Experiment 1 that participants made rational judgments about retracted original information as predicted by relative source credibility and a rational account of the CIE. To this end, we used a standard CIE paradigm with a second set of materials adapted from the "Water Department Shutdown" scenario (e.g., Ecker et al., 2017; Rich et al., 2023). As in Experiment 1, we manipulated informant and retractor credibility in a 2 × 2 between-subjects design and added a no-information control group. Agreement with the original information was again assessed with a direct and an indirect measure, and participants again provided subjective source credibility ratings.

The second objective of Experiment 2 was to address the limitations of Experiment 1. The first limitation was the absence of CIEs and (relative) source-credibility effects on the indirect measure of Experiment 1. To address this, we made two methodological changes in Experiment 2. As reported, all experimental groups of Experiment 1 except for the informant-higher group rather disagreed than agreed with the statement indirectly referring to the original information, that is, that the accused company should be fined. Participants' hesitancy to impose a fine may arise from the principle *in dubio pro reo*, the benefit of the doubt. Even though participants may have agreed that the company was responsible for the incident, they may have rejected a punishment until the accused would be found legally guilty. Therefore, in Experiment 2, we used statements indirectly referring to the original information that did not reference legal actions against the accused party (see Appendix B). Further, to enhance the reliability of the indirect measure, we used four statements instead of one and combined the four agreement ratings into a single measure via Principal Component Analysis.

The second limitation of Experiment 1 was the relatively strong agreement of the control group with the original information (see Fig. 2), which may be the reason for the non-significant differences between the equal groups and the control group and, thus, the interpretation that the equal groups showed no CIE. The absence of a CIE in the equal groups is critical because it would indicate that, as predicted by a rational account, participants only showed a CIE when relative source credibility favored the informant. In Experiment 1, however, this particular result yielded inconclusive evidence in the Bayesian analyses, which makes it difficult to draw a clear conclusion. In Experiment 2, we therefore aimed for conclusive evidence regarding the presence or absence of a CIE in the equal groups. Ideally, a no-information control group should not be able to deduce the involvement of the accused party and show minimal judgments and inferences in agreement with the original information (e.g., Ecker et al., 2011; Guillory & Geraci, 2010). In Experiment 1, the control group may have suspected a causal involvement of the company because the company was mentioned throughout the report (for a similar result and discussion, see Rich & Zaragoza, 2016). In Experiment 2, we, therefore, used a report that mentioned the accused party exclusively in the paragraphs containing the original information and retraction. That is, the control group did not learn of the accused party at all.

With these modifications to our materials, we expected to replicate the results from the direct measure of Experiment 1 in both the direct and indirect measures of those groups of Experiment 2 in which relative source credibility favored one source. This would further corroborate the evidence for the influence of relative source credibility and a rational account of the CIE. However, replicating an absence of source credibility effects on the indirect measure would speak for differential effects of rationality on direct and indirect agreement with the original information. We further aspired to find conclusive evidence regarding the presence or absence of CIEs in the equal groups, which would help us discern rational and error-based accounts.

3.1. Method

3.1.1. Design and participants

The design, power considerations, recruitment, and exclusion criteria were the same as in Experiment 1. Between January 5 and January 24, 2021, we collected data from 160 new participants in this online experiment. We excluded participants because of failed attention checks ($n = 6$), an incorrect summary of the report ($n = 1$), failed seriousness checks ($n = 6$), and study completion times that exceeded the mean by more than three standard deviations ($n = 2$). The pattern of results remained unchanged when all participants were included. The final sample thus consisted of 145 participants (105 female, 38 male, 2 non-binary). The group sample sizes are listed in Table 2. All participants were native speakers of German between 17 and 32 years of age ($M = 23.11$, $SD = 3.77$). All five groups were equivalent in age, gender,

and educational degree obtained (all *ps* > .502, all BF_{01} > 12.62). Participants enrolled as psychology students at Heinrich Heine University Düsseldorf, Germany, received partial course credit as compensation. The other participants could enter a raffle of eight 75 € gift cards (with an expected value of about 5 € for each participant).

3.1.2. Materials and procedure

All materials were administered in German. An English translation of the fictitious news report can be found in Table 4, and a translation of the test items in Appendix B. For this experiment, we used a translated and adapted version of the "Water Department Shutdown" report (e.g., Brydges et al., 2018; Ecker et al., 2017; Ecker & Antonio, 2021; Ithisuphalap et al., 2020; McIlhiney et al., 2022; Miller et al., 2022; Rich et al., 2023; Sanderson et al., 2021). In eight separate text passages, each one to five sentences long, an anonymous author reported the shutdown of a waterworks due to several dead fish found in the supplying river. The critical original information, that the fish were killed by toxic chemical waste illegally disposed of by a pharmaceutical company, was presented in paragraph 3 and retracted in paragraph 6. We manipulated source credibility in the same way as in Experiment 1.

In the test phase, participants rated their agreement with 14 statements about the described incident on a 7-point Likert scale from 1 (*strong disagreement*) to 7 (*strong agreement*). Memory for the contents of the report was tested with eight statements (Statements 1, 6, 10, 11, and 13 were correct, Statements 2, 4, and 8 were incorrect). Agreement with the original information was again assessed with a direct measure and an indirect measure. For the direct measure, participants rated their agreement with the verbatim original information (Statement 9). For the indirect measure, participants rated their agreement with four possible inferences from the original information (Statements 3, 7, 12, and 14). Additionally, one statement presented an alternative cause for the fish kill as an exploratory reverse measure of agreement with the original information (Statement 5). However, agreement with this statement did not differ between groups, $F(4, 140) = 1.13$, $p = .346$, $\eta_p^2 = .03$, $BF_{01} = 8.33$, and was therefore excluded from further analyses.

3.2. Results

3.2.1. Manipulation checks

Fig. 1 shows the mean source-credibility ratings for all groups. For the manipulation checks, we conducted the same analyses as in Experiment 1.

Table 4
Translated News Report in Experiment 2: Temporary Shutdown of the Perlenbach Waterworks.

Paragraph	Function	Group		Text	
1	Report	All		The Perlenbach waterworks in Bavaria is forced to shut down the water supply from its main water source, the Danube. The reason is a worrying report of numerous deceased fish in the nearby section of the watercourse and the associated water protection area.	
2	Report	All		The Perlenbach plant normally supplies the greater Ingolstadt area with fresh water. Only recently, the facilities were upgraded with new technology to guarantee the highest quality standards for consumers. The plant supplies a total of 120,000 households with water.	
3	Informant credibility manipulation	Informant higher, both high	Retractor higher, both low	*High-credibility informant*	*Low-credibility informant*
				[Name A], investigative journalist, comments on the case. Recently, he had exposed the deliberate marine pollution caused by an oil rig. As a result, he was awarded a major journalism prize.	[Name A], a self-proclaimed scandal blogger, comments on the case. Recently, on his website "My Scandals", he had made groundless accusations against a Norwegian oil rig.
	Original information	All groups except control		[Name A] now publishes a post in which he writes: "The frightening fish die-off in this section of the Danube was caused by the illegal disposal of toxic chemical waste by the cosmetics and pharmaceutical company Obived Ltd." This company operates a plant a few kilometers upstream and is thus located within the water protection area.	
4	Report	All		Just a few months ago, another waterworks in North Rhine-Westphalia was briefly forced to shut down its water intake from the Ruhr River because numerous dead fish washed up on the bank near the intake point. As the investigations by the water protection authority revealed, there was, fortunately, no danger to consumers in this case, as the fish died of a viral disease that presented no danger to humans.	
5	Report	All		So far, no cases of nausea or symptoms of illness have been reported in the greater Ingolstadt area after consuming tap water. Nevertheless, the health department and the water protection authority continue to urge residents in the region: "Do not use your tap water for drinking or cooking." Bavarian authorities are in close contact with local hospitals.	
6	Retractor credibility manipulation	Retractor higher, both high	Informant higher, both low	*High-credibility retractor*	*Low-credibility retractor*
				A few days later, [Name B], journalist and fact-checker, gives a statement. He was recently awarded a prestigious journalism prize for numerous exposures of false reports and journalistic misconduct.	A few days later, [Name B], fitness blogger from the website "Melt Fat Fast", gives a statement. He is known, among other things, for numerous collaborations with dubious miracle pill manufacturers.
	Retraction	All groups except control		[Name B] publishes a post in which he firmly rejects the accusations made by [journalist/scandal blogger] [Name A]. He contradicts: "There is no evidence of illegal disposal of toxic chemical waste from Obived Cosmetics & Pharmaceuticals Ltd. The company complies with all legal and ecological standards."	
7	Report	All		The shutdown of the main water supply from the Danube poses problems for the Perlenbach waterworks and regional agriculture. The local water reserves were almost exhausted even before the shutdown due to a prolonged period of drought. Some local farmers are already complaining that their crops have dried up in the fields.	
8	Report	All		A spokesperson for the Perlenbach waterworks explains, "The water supply for private households is not at risk so far." The plant nevertheless publishes some tips on its website on how water could easily be saved in everyday life to counteract possible shortages.	

Note. The original German news report is available in the Open Science Framework at https://osf.io/hwxjy. It was translated and adapted from the supplement of "Reminders and repetition of misinformation: Helping or hindering its retraction?" by Ecker et al. (2017), *Journal of Applied Research in Memory and Cognition,* 6(2) with the permission from the authors. Copyright 2017 by the American Psychological Association. The paragraphs were presented one at a time to the participants of the listed groups. Credibility manipulations are indicated in italics. In the presented reports "Uwe Schmidt" and "Ralf Müller" were assigned to "Name A" and "Name B" according to the counterbalancing scheme. All = informant higher, retractor higher, both high, both low, and control group. Informant higher = high-credibility informant and low-credibility retractor; both high = high-credibility informant and high-credibility retractor; both low = low-credibility informant and low-credibility retractor; retractor higher = low-credibility informant and high-credibility retractor, control = participants encountered neither an informant nor a retractor.

3.2.1.1. Source credibility ratings in the control group. As expected, the high-credibility sources were rated as more credible than the low-credibility sources, $F(1,30) = 167.17$, $p < .001$, $\eta_p^2 = .85$, $BF_{10} = 1.47 \cdot 10^{21}$. Further, the two high-credibility sources did not differ in credibility, $F(1, 30) = 0.67$, $p = .418$, $\eta_p^2 = .02$, $BF_{01} = 3.83$, neither did the two low-credibility sources, $F(1, 30) = 1.11$, $p = .301$, $\eta_p^2 = .04$, $BF_{01} = 3.15$. In conclusion, the credibility manipulation worked as intended.

3.2.1.2. Source credibility ratings in the experimental groups. Informant credibility ratings showed the intended effect of the informant credibility manipulation, $F(1,110) = 123.47$, $p < .001$, $\eta_p^2 = .53$, $BF_{incl} = 3.72 \cdot 10^{16}$. They were not affected by the retractor credibility manipulation, $F(1, 110) = 1.98$, $p = .162$, $\eta_p^2 = .02$, $BF_{excl} = 2.13$, and there was no interaction, $F(1, 110) = 0.56$, $p = .454$, $\eta_p^2 = .01$, $BF_{excl} = 3.01$. Likewise, retractor credibility ratings showed the intended effect of the retractor credibility manipulation, $F(1, 110) = 124.91$, $p < .001$, $\eta_p^2 = .53$, $BF_{incl} = 1.98 \cdot 10^{16}$, and were not affected by the informant credibility manipulation, $F(1, 110) = 0.94$, $p = .334$, $\eta_p^2 = .01$, $BF_{excl} = 3.33$. There was an interaction, $F(1, 110) = 6.15$, $p = .015$, $\eta_p^2 = .05$, $BF_{incl} = 3.53$, indicating that the effect of the retractor credibility manipulation on retractor credibility ratings was smaller when participants encountered the high-credibility informant, $F(1, 54) = 29.62$, $p < .001$, $\eta_p^2 = .35$, $BF_{incl} = 1.03 \cdot 10^4$, than when they encountered the low-credibility informant, $F(1, 56) = 125.76$, $p < .001$, $\eta_p^2 = .69$, $BF_{incl} = 6.26 \cdot 10^{12}$. However, it was significant in both cases. To conclude, informant and retractor credibility manipulations were successful and affected the intended source credibility.

3.2.1.3. No reactive effects of retraction presentation. Again, the experimental groups did not rate any source as less credible than the control group did. Two sources were even rated as more credible in the relevant experimental groups than in the control group: the low-credibility informant, $F(1, 86) = 7.49$, $p = .009$, $\eta_p^2 = .08$, $BF_{10} = 5.37$, and the low-credibility retractor, $F(1, 85) = 6.94$, $p = .010$, $\eta_p^2 = .08$, $BF_{10} = 3.66$ (all other $ps > .218$, $BF_{01} > 2.19$; see also Footnote 2). Thus, again, neither source was rated as less credible because a retraction was presented.

3.2.2. Memory for the contents of the report

Participants rated their agreement with five correct and three incorrect statements about the contents of the report on a 7-point Likert scale and we again calculated the discrimination index P_R and the bias index B_R of the two-high threshold model of recognition memory (e.g., Corwin, 1994; Snodgrass & Corwin, 1988; see Table 2 and Footnote 3). One-way ANOVAs across all groups showed no differences in the discrimination index P_R, $F(4, 140) = 0.59$, $p = .673$, $\eta_p^2 = .02$, $BF_{01} = 18.21$, and no difference in the bias index B_R, $F(4, 140) = 0.73$, $p = .575$, $\eta_p^2 = .02$, $BF_{01} = 15.10$. Thus, as in Experiment 1, between-group differences in agreement with the critical items referring to the original information cannot be attributed to differences in memory for the report (Sanderson et al., 2021) nor to differences in response bias.

3.2.3. Effects of informant and retractor credibility

We expected main effects of informant and retractor credibility on agreement ratings in both the direct and the indirect measures. This would replicate the results from the direct measure in Experiment 1. In Experiment 2, both measures were highly correlated, $r(143) = .60$, $p < .001$, $BF_{10} = 3.83 \cdot 10^{12}$. Fig. 2 shows descriptive statistics of direct and indirect agreement in each group. We again performed two 2 (informant credibility) × 2 (retractor credibility) ANOVAs on the agreement ratings in the direct and indirect measures.

3.2.3.1. Direct measure. Replicating Experiment 1, there were main effects of informant credibility and retractor credibility on the direct measure. High informant credibility led to stronger agreement than low informant credibility, $F(1, 110) = 10.01$, $p = .002$, $\eta_p^2 = .08$, $BF_{incl} = 15.28$, and high retractor credibility led to less agreement than low retractor credibility, $F(1, 110) = 8.93$, $p = .003$, $\eta_p^2 = .08$, $BF_{incl} = 10.12$, replicating previous findings (Ecker & Antonio, 2021; Guillory & Geraci, 2013; Pluviano et al., 2020; Susmann & Wegener, 2023; Westbrook et al., 2023). There was no interaction, $F(1, 110) = 0.79$, $p = .377$, $\eta_p^2 = .01$, $BF_{excl} = 2.61$.

3.2.3.2. Indirect measure. For the indirect measure, participants rated their agreement with four possible inferences from the original information. To compute a single score of indirect agreement with the original information, we combined the four agreement ratings into one measure using a Principal Component Analysis. It identified a one-component solution with an eigenvalue >1 (Kaiser rule), which accounted for 50.33 % of the total variance in the four agreement ratings. All four ratings showed high factor loadings on this component (.80, .62, .85, and .52 for Statements 3, 7, 12, and 14, respectively). We calculated the factor score for each participant with the regression method, which weighs standardized agreement ratings for statements with their respective factor loadings (cf. DiStefano et al., 2009). These scores served as the indirect measure. Factor scores ranged from −1.96 to 2.31 ($M = 0.00$, $SD = 1.00$, $SE = 0.08$). Higher scores indicate stronger agreement with the original information.[4]

As predicted, there were main effects of informant credibility and retractor credibility on the indirect measure. High informant credibility led to stronger agreement than low informant credibility, $F(1, 110) = 12.63$, $p < .001$, $\eta_p^2 = .10$, $BF_{incl} = 49.62$. In Experiment 2, the BF indicated strong evidence for this effect (rather than inconclusive evidence as in Experiment 1). Conversely, high retractor credibility led to less agreement than low retractor credibility, $F(1, 110) = 8.81$, $p = .004$, $\eta_p^2 = .07$, $BF_{incl} = 10.01$, replicating Experiment 1. There was no interaction, $F(1, 110) = 0.05$, $p = .826$, $\eta_p^2 < .01$, $BF_{excl} = 3.71$. Thus, overall, Experiment 2 provided conclusive evidence in favor of the first two predictions derived from the rational account of the CIE.

3.2.4. Effects of relative source credibility and comparisons to the control group

Spontaneous agreement in the control group exceeded the minimum possible agreement in both the direct measure (tested against the scale floor of 1), $t(30) = 2.05$, $p = .049$, $d = 0.37$, $BF_{10} = 1.20$, and the indirect measure (tested against the factor score calculated from the lowest possible agreement ratings [1] for all four statements, −1.96), $t(30) = 10.04$, $p < .001$, $d = 1.81$, $BF_{10} = 2.37 \cdot 10^8$. Therefore, as in Experiment 1, we defined a CIE in an experimental group in comparison to the spontaneous agreement in the control group. A CIE should depend on the relative credibility of informant and retractor. We performed the same four planned contrasts on the direct and indirect measures as in Experiment 1 (see Table 3). Because these contrasts showed similar results for the direct and indirect measures, we report and discuss them jointly.

The informant-higher group showed CIEs in both measures when compared with the control group (Contrast 1). The retractor-higher group showed no CIE (Contrast 2; albeit with an inconclusive BF in the direct measure). These two results replicated the findings from the direct measure of Experiment 1 and demonstrated again that the CIE can be eliminated when relative source credibility favors the retractor. The CIE, thus, depends on the relative credibility of informant and retractor.

Concerning the equal groups, the levels of agreement with the original information in the both-high group and the both-low group did not differ in size (Contrast 3), replicating the finding from the direct measure of Experiment 1. However, in contrast to Experiment 1, the

[4] The factor scores were highly correlated with the mean of the four agreement ratings, $r(143) = .99$, $p < .001$, $BF_{10} = 4.45 \cdot 10^{129}$. Analysis with the mean as the indirect measure did not change the pattern of results.

levels of agreement in the equal groups exceeded the spontaneous agreement in the control group (Contrast 4). Thus, despite relative source credibility favoring neither source, the equal groups showed CIEs (with conclusive evidence for these effects from the Bayesian analyses).

3.3. Discussion

One objective of Experiment 2 was to use a second set of materials to replicate the findings obtained with the direct measure of Experiment 1 in support of a rational account of the CIE. We also sought to determine whether the null effects in the indirect measure of Experiment 1 were due to its specific adverse wording by using an improved measure of indirect agreement. The results of Experiment 2 confirmed most predictions regarding the effects of relative source credibility on the CIE in the direct measure and, critically, also the indirect measure. Thus, we deem it likely that the unexpected null effects in the indirect measure in Experiment 1 were artifacts of the specific materials used therein.

In Experiment 2, as in Experiment 1, when relative source credibility favored the retractor, participants showed no CIE. However, when relative source credibility favored neither source (equal groups), participants showed CIEs nonetheless. This pattern of results cannot be explained exclusively by a rational account of the CIE that is based on manipulated relative source credibility. Relative source credibility amplified or attenuated the size of the CIE, even eliminated it in the retractor-higher group. However, it cannot be the sole explanation of the CIE in Experiment 2 as a rational account would have predicted no CIE in the equal groups.

4. General discussion

Information that is later retracted often continues to influence judgments and inferences. A rational account of the CIE (Connor Desai et al., 2020) postulates that agreement with retracted original information should depend on the relative credibility of informant and retractor. To put this account to the first experimental test, we independently manipulated informant and retractor credibility in two experiments with a standard CIE paradigm. As predicted, informant credibility and retractor credibility both influenced the agreement with original information independently and antagonistically. Participants showed a CIE when the informant was more credible than the retractor. Critically, participants showed no CIE when the retractor was more credible than the informant. These results support the rational account of the CIE.

However, we also found a deviation from the predictions of a strictly rational account. When informant and retractor were equally credible, their effects on judgments and inferences should have traded off to yield no CIE. Whereas Experiment 1 was inconclusive in this regard, Experiment 2 clearly showed CIEs when informant and retractor were equally credible. Thus, equal source credibility resulted in levels of agreement that were inconsistent with rational judgments and inferences. Therefore, a rational account alone cannot fully account for our results.

There are several possible explanations for such a pattern of results. For one, an extended rational account that considers reactive effects of retraction presentation on source credibility ratings, as proposed by Connor Desai et al. (2020), could explain the CIEs in the equal groups. According to these authors, the presentation of the retraction introduced a contradiction between original information and retraction in their study, resulting in a reduction in the perceived credibility of both the informant and the retractor (similar to the reactive effect on informant credibility reported by Westbrook et al., 2023). Connor Desai et al. argued that such reactive effects may alter the relative credibility of informant and retractor. In our study, however, we found no such reactive effects of retraction presentation on source credibility ratings when we compared the experimental groups to the control group. In neither of our experiments did any source receive lower credibility ratings when involved in the contradiction compared to when no contradiction occurred. Nonetheless, our results do not necessarily conflict with the findings of reactive effects reported by Connor Desai et al. and Westbrook et al. (2023). In these studies, source credibility was unspecified, and the presence of a contradiction in the form of a retraction was the sole cue available to participants for making source credibility estimates. In our experiments, the strong manipulations of source credibility likely overrode any potential reactive effects stemming from retraction presentation. We conclude that this extended rational account cannot explain the results obtained in our experiments.

Instead, the CIEs in the equal groups of Experiment 2 are consistent with explanations based on erroneous memory processes, that is, the model-updating account and the retrieval-failure account (H. M. Johnson & Seifert, 1994; Lewandowsky et al., 2012). According to the model-updating account, participants retain the original information in their mental model because they fail to update the model after the retraction and/or are reluctant to use an incomplete and implausible model instead (Brydges et al., 2018; Hamby et al., 2020). The CIE results from judgments and inferences based on these mental models, which still contain the original information. According to the retrieval-failure account, a CIE results from failure to engage in strategic retrieval processes that would be necessary to detect the original information as false and suppress its incorporation into judgments and inferences (Ecker et al., 2011; Swire, Ecker, & Lewandowsky, 2017). The cognitive errors proposed in these two accounts may have led participants to show a CIE when informant and retractor were equally credible.

Furthermore, the memory processes described in the model-updating account and the retrieval-failure account (H. M. Johnson & Seifert, 1994; Lewandowsky et al., 2012) may mediate the effects of source credibility on the CIE. Ecker and Antonio (2021) suggested that source credibility may impact the CIE by modulating the memory and integration processes postulated in the two error-based accounts of the CIE (H. M. Johnson & Seifert, 1994; Lewandowsky et al., 2012). Thus, Ecker and Antonio's considerations offer a potential reconciliation between the rational and the error-based theoretical accounts. For example, they argue that information from more credible sources may be encoded more deeply. As high source credibility increases processing effort (e.g., Heesacker et al., 1983; Hoogeveen et al., 2022), it may lead to more elaborate encoding. According to the model-updating account, this may facilitate the integration of original information into a mental model or facilitate a model update in line with a retraction. According to the retrieval-failure account, high source credibility may facilitate the automatic activation of either information during judgment formation. Further, monitoring of source credibility may affect judgments at the time of retrieval and guide the evaluation of retrieved information. For example, in the illusory truth paradigm, participants evaluated statements as more likely to be true when they attributed them to a high-credibility source than to a low-credibility source, even when this attribution was incorrect (Begg et al., 1992). These mechanisms may explain how source credibility leads to the observed effects on the CIE. Investigating the processes underlying source-credibility effects on the CIE may, thus, also help to advance the debate concerning the model-updating account and the retrieval-failure account.

The CIEs in the equal groups of Experiment 2 are also consistent with explanations based on motivational factors. Susmann and Wegener (2022) showed that participants continued to agree with the original information to reduce feelings of psychological discomfort produced by the processing of retractions. The relative source credibility of informant and retractor may modulate the extent of this discomfort. In the current experiments, this could mean that when the retractor was more credible than the informant, discomfort may have been reduced or absent because relative source credibility may have justified the retraction. In this case, no discomfort would have motivated continued agreement with the original information in the retractor-higher group.

Therefore, we deem it likely that the CIE is not exclusively explained by relative source credibility. Overall, our results are best explained by a combination of rational and irrational processes: Erroneous memory

processes or motivational factors likely led to an irrational CIE, and relative source credibility changed the size of this CIE when it favored one of the sources. This change resulted in a strong CIE (informant-higher group) or the rational elimination of a CIE (retractor-higher group). The rational account thus complements error-based and motivational theoretical accounts of the CIE (e.g., H. M. Johnson & Seifert, 1994; Lewandowsky et al., 2012; Susmann & Wegener, 2022). Future research should be aimed at clarifying the mechanisms through which source credibility influences judgments and inferences in the CIE paradigm.

Another open question is how other indicators of statement veracity besides source credibility may lead to a rational CIE. One interesting candidate is message credibility or plausibility. Consistent with a rational account and mirroring the pattern of antagonistic source-credibility effects in the current experiments, Kim et al. (2020) showed that higher credibility ratings of the original information or the retraction were associated with more agreement with the original information or less agreement with it, respectively. Moreover, when Rich et al. (2023) manipulated the plausibility of original information and retraction (e.g., a fire caused by an explosion of a gas cylinder vs. by fallen satellite debris), the effects of message plausibility were limited to the original information. Consistent with a rational account, more plausible original information led to more agreement with this original information, but retraction plausibility did not affect agreement with the original information. Thus, retraction plausibility and retractor credibility seem to have different effects. Future research should compare the effects of different indicators of statement veracity on the CIE and examine whether and why conceptually equivalent properties of the original information and retraction can have different effects on the CIE.

In addition to its theoretical implications, our research has important implications for real-world contexts as the findings imply that enhancing retraction effectiveness necessitates highly credible retractors. This is particularly critical in real-world scenarios, where retracting misinformation is markedly more challenging than in controlled laboratory settings (Walter & Murphy, 2018). Outside the laboratory, however, it may be less likely to find a retractor (or any source) that is undisputedly of high credibility. The availability of greater amounts of information about real-life sources can make it easier to assess their credibility, but it also renders real-life sources more controversial and complex. This creates more variation in the evaluation of sources due to individual biases and prior beliefs about sources (e.g., partisanship; Traberg & van der Linden, 2022). Further, real-world scenarios are likely more strongly influenced by domain knowledge, prior beliefs about the conveyed message, and emotions, and they are likely more personally relevant than a fictitious report. This may lead to less reliance on source credibility over other information (e.g., stereotypes; Ecker et al., 2014). Therefore, future research should investigate how source credibility shapes people's judgments in real-world scenarios and compare its influence to other factors.

To conclude, the experiments presented here show that source credibility as an indicator of information veracity can rationally inform judgments about the information. This is the first rigorous experimental test of the notion that the CIE is rational, particularly when veracity indicators favor the original information. Thus, our results contribute to a comprehensive theoretical understanding of the continued influence of retracted information.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to revise individual sentences in the manuscript for clarity and word choice. After using this tool, the authors reviewed and edited the text as needed and take full responsibility for the content of the publication.

CRediT authorship contribution statement

Carolin V. Hey: Writing – original draft, Writing – review & editing, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Marie Luisa Schaper:** Writing – review & editing, Validation, Supervision, Methodology, Funding acquisition. **Ute J. Bayen:** Writing – review & editing, Supervision, Resources, Funding acquisition.

Acknowledgment

The experiments were not preregistered. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The research reported in this article was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project No. 436585458 granted to Marie Luisa Schaper and Ute Johanna Bayen. Parts of this research were presented at the 64th Tagung Experimentell arbeitender Psychologen und Psychologinnen (Conference of Experimental Psychologists), March 2022, University of Cologne, virtual meeting. We thank Siegmund Switala for programming the experiments.

Appendix A. Test Statements in Experiment 1 (translated from German)

Participants rated their agreement with each statement in the presented order on a Likert scale from 1 (*strong disagreement*) to 7 (*strong agreement*). The original German statements are available in the Open Science Framework.

1. The cases of the disease occurred in the German states of Bavaria and Baden-Wuerttemberg.
2. The recalled spinach was of the Greenerous brand. *(incorrect)*.
3. *(Direct measure)* The contamination of the spinach was caused by the unsanitary conditions in the production facility.
4. The diseased consumers showed only mild symptoms. *(incorrect)*.
5. The managing director of Zefter Vegetables Ltd. gave a candid television interview. (*subjective judgment, excluded from analyses*).
6. *(Indirect measure)* The health department should impose a heavy fine on Zefter Vegetables Ltd.
7. The risk of buying contaminated spinach now is low.
8. The contaminated spinach was packed in the month of June.

Appendix B. Test Statements in Experiment 2 (translated from German)

Participants rated their agreement with each statement in the presented order on a Likert scale from 1 (*strong disagreement*) to 7 (*strong agreement*). The original German statements are available in the Open Science Framework.

1. The name of the waterworks concerned is Perlenbach.
2. The technology of the waterworks dates from 25 years ago. *(incorrect)*.
3. *(Indirect measure 1)* This incident could have been prevented by stricter oversight of the companies within the water protection area.
4. So far, there are some known cases of complaints and illnesses after consumption of the water. *(incorrect)*.
5. *(Exploratory reverse measure)* The fish in the Danube died of natural causes.
6. The waterworks normally supplies 120,000 households with fresh water.

7. *(Indirect measure 2)* The Bavarian water protection authority should focus its investigations on human-made causes for the fish deaths.
8. The water reserves in the Ingolstadt region are still in very good supply. *(incorrect)*.
9. *(Direct measure)* The reason for the fish die-off in the Danube was illegally disposed toxic chemical waste from a local cosmetics and pharmaceutical company.
10. The other case of a sudden fish die-off in North Rhine-Westphalia was triggered by a natural cause.
11. The shutdown of the water supply from the Danube has severely impacted local agriculture.
12. *(Indirect measure 3)* The local cosmetic and pharmaceutical company should take responsibility for the incident.
13. The water supply to private households has not been in jeopardy so far.
14. *(Indirect measure 4)* The news report named the real cause of the fish die-off in the Danube.

Data availability

The data and materials from the two experiments are available in the Open Science Framework at https://osf.io/hwxjy.

References

Aust, F., Diedenhofen, B., Ullrich, S., & Musch, J. (2013). Seriousness checks are useful to improve data validity in online research. *Behavior Research Methods, 45*(2), 527–535. https://doi.org/10.3758/s13428-012-0265-2

Ayers, M. S., & Reder, L. M. (1998). A theoretical review of the misinformation effect: Predictions from an activation-based memory model. *Psychonomic Bulletin & Review, 5*(1), 1–21. https://doi.org/10.3758/BF03209454

Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General, 121*(4), 446–458. https://doi.org/10.1037/0096-3445.121.4.446

Birnbaum, M. H., Wong, R., & Wong, L. K. (1976). Combining information from sources that vary in credibility. *Memory & Cognition, 4*(3), 330–336. https://doi.org/10.3758/BF03213185

Braasch, J. L., Rouet, J., Vilbert, N., & Britt, M. A. (2012). Readers' use of source information in text comprehension. *Memory & Cognition, 40*, 450–465. https://doi.org/10.3758/s13421-011-0160-6

Brydges, C. R., Gignac, G. E., & Ecker, U. K. H. (2018). Working memory capacity, short-term memory capacity, and the continued influence effect: A latent-variable analysis. *Intelligence, 69*, 117–122. https://doi.org/10.1016/j.intell.2018.03.009

Buczel, K. A., Siwiak, A., Szpitalak, M., & Polczyk, R. (2024). How do forewarnings and post-warnings affect misinformation reliance? The impact of warnings on the continued influence effect and belief regression. *Memory & Cognition, 52*(5), 1048–1064. https://doi.org/10.3758/s13421-024-01520-z

Buczel, K. A., Syszka, P. D., Siwiak, A., Szpitalak, M., & Polczyk, R. (2022). Vaccination against misinformation: The inoculation technique reduces the continued influence effect. *PLoS One, 17*(4), Article e0267463. https://doi.org/10.1371/journal.pone.0267463

Chan, M. S., Jones, C. R., Hall Jamieson, K., & Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological Science, 28*(11), 1531–1546. https://doi.org/10.1177/0956797617714579

Cohen, J. (1988). *Statistical power analysis for the behavioral science* (2. ed.). Routledge. https://doi.org/10.4324/9780203771587

Connor Desai, S. A., Pilditch, T. D., & Madsen, J. K. (2020). The rational continued influence of misinformation. *Cognition, 205*. https://doi.org/10.1016/j.cognition.2020.104453. Article 104453.

Connor Desai, S. A., & Reimers, S. (2019). Comparing the use of open and closed questions for web-based measures of the continued-influence effect. *Behavior Research Methods, 51*(3), 1426–1440. https://doi.org/10.3758/s13428-018-1066-z

Copeland, D. E., Gunawan, K., & Bies-Hernandez, N. J. (2011). Source credibility and syllogistic reasoning. *Memory & Cognition, 39*(1), 117–127. https://doi.org/10.3758/s13421-010-0029-0

Corwin, J. (1994). On measuring discrimination and response bias: Unequal numbers of targets and distractors and two classes of distractors. *Neuropsychology, 8*(1), 110–117. https://doi.org/10.1037/0894-4105.8.1.110

DiStefano, C., Zhu, M., & Mîndrilă, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research and Evaluation, 14*. https://doi.org/10.7275/da8t-4g52. Article 20.

Echterhoff, G., Hirst, W., & Hussy, W. (2005). How eyewitnesses resist misinformation: Social postwarnings and the monitoring of memory characteristics. *Memory & Cognition, 33*(5), 770–782. https://doi.org/10.3758/BF03193073

Ecker, U. K. H., & Ang, L. C. (2019). Political attitudes and the processing of misinformation corrections. *Political Psychology, 40*(2), 241–260. https://doi.org/10.1111/pops.12494

Ecker, U. K. H., & Antonio, L. M. (2021). Can you believe it? An investigation into the impact of retraction source credibility on the continued influence effect. *Memory & Cognition, 49*(4), 631–644. https://doi.org/10.3758/s13421-020-01129-y

Ecker, U. K. H., Butler, L. H., & Hamby, A. (2020). You don't have to tell a story! A registered report testing the efectiveness of narrative versus non-narrative misinformation corrections. *Cognitive Research: Principles and Implications, 5*. https://doi.org/10.1186/s41235-020-00266-x. Article 64.

Ecker, U. K. H., Hogan, J. L., & Lewandowsky, S. (2017). Reminders and repetition of misinformation: Helping or hindering its retraction? *Journal of Applied Research in Memory and Cognition, 6*(2), 185–192. https://doi.org/10.1016/j.jarmac.2017.01.014

Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., … Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology, 1*, 13–29. https://doi.org/10.1038/s44159-021-00006-y

Ecker, U. K. H., Lewandowsky, S., Fenton, O., & Martin, K. (2014). Do people keep believing because they want to? Preexisting attitudes and the continued influence of misinformation. *Memory & Cognition, 42*(2), 292–304. https://doi.org/10.3758/s13421-013-0358-x

Ecker, U. K. H., Lewandowsky, S., Swire, B., & Chang, D. (2011). Correcting false information in memory: Manipulating the strength of misinformation encoding and its retraction. *Psychonomic Bulletin & Review, 18*(3), 570–578. https://doi.org/10.3758/s13423-011-0065-1

Ecker, U. K. H., Lewandowsky, S., & Tang, D. T. (2010). Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Memory & Cognition, 38*(8), 1087–1100. https://doi.org/10.3758/MC.38.8.1087

Ecker, U. K. H., Sze, B. K., & Andreotta, M. (2021). Corrections of political misinformation: No evidence for an effect of partisan worldview in a US convenience sample. *Philosophical Transactions of the Royal Society B, 376*(1822), Article 20200145. https://doi.org/10.1098/rstb.2020.0145

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415*(6870), 429–433. https://doi.org/10.1038/415429a

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146

Gordon, A., Quadflieg, S., Brooks, J. C., Ecker, U. K. H., & Lewandowsky, S. (2019). Keeping track of 'alternative facts': The neural correlates of processing misinformation corrections. *NeuroImage, 193*, 46–56. https://doi.org/10.1016/j.neuroimage.2019.03.014

Gordon, R., & Spears, K. (2012). You don't act like you trust me: Dissociations between behavioural and explicit measures of source credibility judgement. *Quarterly Journal of Experimental Psychology, 65*(1), 121–134. https://doi.org/10.1080/17470218.2011.591534

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Wiley & Sons, Inc. https://psycnet.apa.org/record/1967-02286-000.

Guillory, J. J., & Geraci, L. (2010). The persistence of inferences in memory for younger and older adults: Remembering facts and believing inferences. *Psychonomic Bulletin & Review, 17*(1), 73–81. https://doi.org/10.3758/PBR.17.1.73

Guillory, J. J., & Geraci, L. (2013). Correcting erroneous inferences in memory: The role of source credibility. *Journal of Applied Research in Memory and Cognition, 2*(4), 201–209. https://doi.org/10.1016/j.jarmac.2013.10.001

Guillory, J. J., & Geraci, L. (2016). The persistence of erroneous information in memory: The effect of valence on the acceptance of corrected information. *Applied Cognitive Psychology, 30*(2), 282–288. https://doi.org/10.1002/acp.3183

Hahn, U., & Oaksford, M. (2006). A normative theory of argument strength. *Informal Logic, 26*(1), 1–24. https://doi.org/10.22329/il.v26i1.428

Hahn, U., & Oaksford, M. (2007). The rationality of informal argumentation: A Bayesian approach to reasoning fallacies. *Psychological Review, 114*(3), 707–732. https://doi.org/10.1037/0033-295x.114.3.704

Hamby, A., Ecker, U. K. H., & Brinberg, D. (2020). How stories in memory perpetuate the continued influence of false information. *Journal of Consumer Psychology, 30*(2), 40–259. https://doi.org/10.1002/jcpy.1135

Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d′. *Behavior Research Methods, Instruments, & Computers, 27*(1), 46–51. https://doi.org/10.3758/BF03203619

Heesacker, M., Petty, R. E., & Cacioppo, J. T. (1983). Field dependence and attitude change: Source credibility can alter persuasion by affecting message-relevant thinking. *Journal of Personality, 51*(4), 653–666. https://doi.org/10.1111/j.1467-6494.1983.tb00872.x

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65–70. http://www.jstor.org/stable/4615733.

Hoogeveen, S., Haaf, J. M., Bulbulia, J. A., Ross, R. M., McKay, R., Altay, S., … van Elk, M. (2022). The Einstein effect provides global evidence for scientific source credibility effects and the influence of religiosity. *Nature Human Behavior, 6*(4), 523–535. https://doi.org/10.1038/s41562-021-01273-8

van Huijstee, D., Vermeulen, I., Kerkhof, P., & Droog, E. (2022). Continued influence of misinformation in times of COVID-19. *International Journal of Psychology, 57*(1), 136–145. https://doi.org/10.1002/ijop.12805

Imhoff, R., Lamberty, P., & Klein, O. (2018). Using power as a negative cue: How conspiracy mentality affects epistemic trust in sources of historical knowledge. *Personality and Social Psychology Bulletin, 44*(9), 1364–1379. https://doi.org/10.1177/0146167218768779

Ithisuphalap, J., Rich, P. R., & Zaragoza, M. S. (2020). Does evaluating belief prior to its retraction influence the efficacy of later corrections? *Memory, 28*(5), 617–631. https://doi.org/10.1080/09658211.2020.1752731

JASP Team. (2024). *JASP (Version 0.18.3) [Computer software]*. https://jasp-stats.org/.

Johnson, H. M., & Seifert, C. M. (1994). Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(6), 1420–1436. https://doi.org/10.1037/0278-7393.20.6.1420

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*(1), 3–28. https://doi.org/10.1037/0033-2909.114.1.3

Kan, I. P., Pizzonia, K. L., Drummey, A. B., & Mikkelsen, E. J. (2021). Exploring factors that mitigate the continued influence of misinformation. *Cognitive Research: Principles and Implications, 6*(1). https://doi.org/10.1186/s41235-021-00335-9. Article 76.

Kim, S. C., Vraga, E. K., & Cook, J. (2020). An eye tracking approach to understanding misinformation and correction strategies on social media: The mediating role of attention and credibility to reduce HPV vaccine misperceptions. *Health Communication, 36*(13), 1687–1696. https://doi.org/10.1080/10410236.2020.1787933

Kuklinksi, J. H., Quirk, P. J., Jerit, J., Schwieder, D., & Rich, R. F. (2000). Misinformation and the currency of democratic citizenship. *The Journal of Politics, 62*(3), 790–816. https://doi.org/10.1111/0022-3816.00033

Lewandowsky, S., Ecker, U. K. H., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the "post-truth" era. *Journal of Applied Research in Memory and Cognition, 6*(4), 353–369. https://doi.org/10.1016/j.jarmac.2017.07.008

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

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory, 12*(4), 361–366. https://doi.org/10.1101/lm.94705

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2 ed.). Lawrence Erlbaum Associates https://psycnet.apa.org/record/2004-19022-000.

McIlhiney, P., Gignac, G. E., Weinborn, M., & Ecker, U. K. H. (2022). Sensitivity to misinformation retractions in the continued influence paradigm: Evidence for stability. *Quarterly Journal of Experimental Psychology, 75*(7), 1259–1271. https://doi.org/10.1177/17470218211048986

Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication, 60*(3), 413–439. https://doi.org/10.1111/j.1460-2466.2010.01488.x

Miller, A. L., Wissmann, K. T., & Peterson, D. J. (2022). The continued influence effect: Examining how age, retraction, and delay impact inferential reasoning. *Applied Cognitive Psychology, 36*(3), 708–723. https://doi.org/10.1002/acp.3939

Morimoto, M., & La Ferle, C. (2008). Examining the influence of culture on perceived source credibility of Asian Americans & the mediating role of similarity. *Journal of Current Issues and Research in Advertising, 30*(1), 49–60. https://doi.org/10.1080/10641734.2008.10505237

Nadarevic, L., Reber, R., Helmecke, A. J., & Köse, D. (2020). Perceived truth of statements and simulated social media postings: An experimental investigation of source credibility, repeated exposure, and presentation format. *Cognitive Research, 5*(1). https://doi.org/10.1186/s41235-020-00251-4. Article 56.

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198524496.001.0001

O'Rear, A. E., & Radvansky, G. A. (2020). Failure to accept retractions: A contribution to the continued influence effect. *Memory & Cognition, 48*(1), 127–144. https://doi.org/10.3758/s13421-019-00967-9

Patzer, G. L. (1983). Source credibility as a function of communicator physical attractiveness. *Journal of Business Research, 11*(2), 229–241. https://doi.org/10.1016/0148-2963(83)90030-9

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology, 19*, 123–205. https://doi.org/10.1016/S0065-2601(08)60214-2

Pluviano, S., Sala, S. D., & Watt, C. (2020). The effects of source expertise and trustworthiness on recollection: The case of vaccine misinformation. *Cognitive Processing, 21*(3), 321–330. https://doi.org/10.1007/s10339-020-00974-8

Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology, 34*(2), 243–281. https://doi.org/10.1111/j.1559-1816.2004.tb02547.x

Priester, J., Wegener, D., Petty, R., & Fabrigar, L. (1999). Examining the psychological process underlying the sleeper effect: The elaboration likelihood model explanation. *Media Psychology, 1*(1), 27–48. https://doi.org/10.1207/s1532785xmep0101_3

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology, 25*, 111–163. https://doi.org/10.2307/271063

Rich, P. R., Donovan, A. M., & Rapp, D. N. (2023). Cause typicality and the continued influence effect. *Journal of Experimental Psychology: Applied, 29*(2), 221–238. https://doi.org/10.1037/xap0000454

Rich, P. R., & Zaragoza, M. S. (2016). The continued influence of implied and explicitly stated misinformation in news reports. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(1), 62–74. https://doi.org/10.1037/xlm0000155

Rich, P. R., & Zaragoza, M. S. (2020). Correcting misinformation in news stories: An investigation of correction timing and correction durability. *Journal of Applied Research in Memory and Cognition, 9*(3), 310–322. https://doi.org/10.1016/j.jarmac.2020.04.001

Rieh, S. Y., & Danielson, D. R. (2007). Credibility: A multidisciplinary framework. *Annual Review of Information Science and Technology, 41*(1), 307–364. https://doi.org/10.1002/aris.2007.1440410114

Sanderson, J. A., & Ecker, U. K. H. (2020). The challenge of misinformation and ways to reduce its impact. In P. Van Meter, A. List, D. Lombardi, & P. Kendeou (Eds.), *Handbook of learning from multiple representations and perspectives* (pp. 461–476). Routledge/Taylor & Francis Group. https://doi.org/10.4324/9780429443961-30.

Sanderson, J. A., Farrell, S., & Ecker, U. K. H. (2022). Examining the role of information integration in the continued influence effect using an event segmentation approach. *PLoS One, 17*(7), Article e0271566. https://doi.org/10.1371/journal

Sanderson, J. A., Gignac, G. E., & Ecker, U. K. H. (2021). Working memory capacity, removal efficiency and event specific memory as predictors of misinformation reliance. *Journal of Cognitive Psychology, 33*(5), 518–532. https://doi.org/10.1080/20445911.2021.1931243

Seifert, C. M. (2002). The continued influence of misinformation in memory: What makes a correction effective? In B. H. Ross (Ed.), *Vol. 41. The psychology of learning, and motivation: Advances in research and theory* (pp. 265–292). Academic Press. https://psycnet.apa.org/record/2003-02136-007.

Smith, C. T., De Houwer, J., & Nosek, B. A. (2013). Consider the source: Persuasion of implicit evaluations is moderated by source credibility. *Personality and Social Psychology Bulletin, 39*(2), 193–205. https://doi.org/10.1177/0146167212472374

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*(1), 34–50. https://doi.org/10.1037/0096-3445.117.1.34

Stadtler, M., & Bromme, R. (2014). The content-source integration model: A taxonomic description of how readers comprehend conflicting scientific information. In D. N. Rapp, & J. L. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (pp. 379–402). MIT Press. https://doi.org/10.7551/mitpress/9737.003.0023.

Sternthal, B., Dholakia, R., & Leavitt, C. (1978). The persuasive effect of source credibility: Tests of cognitive response. *Journal of Consumer Research, 4*(4), 252–260. https://doi.org/10.1086/208704

Sternthal, B., Phillips, L. W., & Dholakia, R. (1978). The persuasive effect of source credibility: A situational analysis. *Public Opinion Quarterly, 42*(3), 285–314. https://doi.org/10.1086/268454

Susmann, M. W., & Wegener, D. T. (2022). The role of discomfort in the continued influence effect of misinformation. *Memory & Cognition, 50*(2), 435–448. https://doi.org/10.3758/s13421-021-01232-8

Susmann, M. W., & Wegener, D. T. (2023). The independent effects of source expertise and trustworthiness on retraction believability: The moderating role of vested interest. *Memory & Cognition, 51*(4), 845–861. https://doi.org/10.3758/s13421-022-01374-3

Swire, B., Berinsky, A. J., Lewandowsky, S., & Ecker, U. K. H. (2017). Processing political misinformation: Comprehending the Trump phenomenon. *Royal Society Open Science, 4*(3), Article 160802. https://doi.org/10.1098/rsos.160802

Swire, B., & Ecker, U. K. H. (2018). Misinformation and its correction: Cognitive mechanisms and recommendations for mass communication. In B. G. Southwell, E. A. Thorson, & L. Sheble (Eds.), *Misinformation and mass audiences* (pp. 195–211). University of Texas Press. https://doi.org/10.7560/314555-013.

Swire, B., Ecker, U. K. H., & Lewandowsky, S. (2017). The role of familiarity in correcting inaccurate information. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 43*(12), 1948–1961. https://doi.org/10.1037/xlm0000422

Tay, L. Q., Hurlstone, M. J., Kurz, T., & Ecker, U. K. H. (2022). A comparison of prebunking and debunking interventions for implied versus explicit misinformation. *British Journal of Psychology, 113*(3), 591–607. https://doi.org/10.1111/bjop.12551

Thorson, E. (2016). Belief echoes: The persistent effects of corrected misinformation. *Political Communication, 33*(3), 460–480. https://doi.org/10.1080/10584609.2015.1102187

Tobin, S. J., & Raymundo, M. M. (2009). Persuasion by causal arguments: The motivating role of perceived causal expertise. *Social Cognition, 27*(1), 105–127. https://doi.org/10.1521/soco.2009.27.1.105

Traberg, C. S., & van der Linden, S. (2022). Birds of a feather are persuaded together: Perceived source credibility mediates the effect of political bias on misinformation susceptibility. *Personality and Individual Differences, 185*. https://doi.org/10.1016/j.paid.2021.111269. Article 111269.

Trauzettel-Klosinski, S., Dietz, K., & IReST-Study-Group. (2012). Standardized assessment of reading performance: The new international reading speed texts IReST. *Investigative Ophthalmology & Visual Science, 53*(9), 5452–5461. https://doi.org/10.1167/iovs.11-8284

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science, 359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Walter, N., & Murphy, S. T. (2018). How to unring the bell: A meta-analytic approach to correction of misinformation. *Communication Monographs, 85*(3), 423–441. https://doi.org/10.1080/03637751.2018.1467564

Walter, N., & Tukachinsky, R. (2020). A meta-analytic examination of the continued influence of misinformation in the face of correction: How powerful is it, why does it happen, and how to stop it? *Communication Research, 47*(2), 155–177. https://doi.org/10.1177/0093650219854600

Ward, C. D., & McGinnies, E. (1974). Persuasive effects of early and late mention of credible and noncredible sources. *The Journal of Psychology, 86*(1), 17–23. https://doi.org/10.1080/00223980.1974.9923879

Westbrook, V., Wegener, D. T., & Susmann, M. W. (2023). Mechanisms in continued influence: The impact of misinformation corrections on source perceptions. *Memory & Cognition, 51*(6), 1317–1330. https://doi.org/10.3758/s13421-023-01402-w

Wilkes, A. L., & Leatherbarrow, M. (1988). Editing episodic memory following the identification of error. *The Quarterly Journal of Experimental Psychology, 40*(2), 361–387. https://doi.org/10.1080/02724988843000168

Wilson, E. J., & Sherrell, D. L. (1993). Source effects in communication and persuasion research: A meta-analysis of effect size. *Journal of the Academy of Marketing Science, 21*(2), 101–112. https://doi.org/10.1007/BF02894421

Wintour, P. (2022). *Sergei Lavrov prefers propaganda over reality in Ukraine talks*. The Guardian. March 15th. Retrieved from https://www.theguardian.com/world/2022/mar/10/sergei-lavrov-russia-foreign-minister-propaganda-in-ukraine-talks.

The Continued Influence Effect and Situation Models:

Evidence for the Model-Updating Account

Carolin V. Hey, Marie Luisa Schaper, & Ute J. Bayen

Heinrich Heine University Düsseldorf, Faculty of Mathematics and Natural Sciences,

Institute for Experimental Psychology

Author Note

Carolin V. Hey, https://orcid.org/0000-0001-6431-5093

Marie Luisa Schaper, https://orcid.org/0000-0001-8912-8569

Ute Johanna Bayen, https://orcid.org/0000-0003-0062-3155

Correspondence concerning this article should be addressed to Marie Luisa Schaper, Institute for Experimental Psychology, Building 23.02, Universitätsstr. 1, Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany. Email: marie.schaper@hhu.de

Abstract

Information often continues to influence judgments even after retraction (*Continued Influence Effect*, CIE). According to the *model-updating account*, the CIE occurs because humans base their judgments on situation models that are not completely updated and, thus, continue to contain the original information despite retraction. These situation models should determine the CIE. However, direct evidence for this account is lacking. The aim of the current research was to test the model-updating account by assessing participants' situation models and testing whether these models mediated effects of the credibility of the original informant and of the retractor on the CIE. Participants ($N = 435$) read a fictitious news report in which original information and retraction were presented by sources of either high or low credibility. Source credibility was disclosed or not disclosed before processing of the original information and retraction. Thus, source credibility could or could not affect situation-model construction, respectively. Supporting the model-updating account, structural equation modeling showed that, when source credibility was disclosed before information processing, the situation model updated after retraction fully mediated the effect of informant credibility and partially mediated the effect of retractor credibility on the CIE. Higher source credibility led to stronger effects of original information and retraction on judgments than lower source credibility. When source credibility was not disclosed before information processing, it did not affect the CIE. These results strongly support the model-updating account and are inconsistent with purely retrieval-based CIE accounts. Some unexpected effects, however, could not be explained by the model-updating account alone.

Keywords: Continued Influence Effect, misinformation, situation model, source credibility, memory (5/5 keywords)

The Continued Influence Effect and Situation Models: Evidence for the Model-Updating Account

The *Continued Influence Effect* (CIE) describes the phenomenon that retractions of information often do not reverse the influence of this original information on judgments and inferences. Despite the retraction, people typically show continued agreement with the original information (e.g., Ecker et al., 2010; Johnson & Seifert, 1994; Rich et al., 2023; for meta-analyses, see Chan et al., 2017; Walter & Tukachinsky, 2020). In typical CIE experiments, participants read a fictitious news report about an unfolding event, for example, about the shutdown of a waterworks due to numerous dead fish in the water (e.g., Ecker et al., 2017). The report contains a critical piece of original information, which is typically an alleged cause for the reported event, for example, that the shutdown was necessary because a company disposed of toxic chemicals in the water. Subsequently, the report contains a retraction of the original information, for example, that there was no evidence incriminating this company and that it complied with all standards. Even though the original information was retracted, participants continued to agree with it and with related inferences, for example, that the company disposed of the chemicals and that the company should take responsibility for the incident (e.g., Hey et al., 2025).

Since the CIE has first been described, it has attracted increasing attention (see review by Ecker et al., 2022). Numerous studies have addressed applied questions to develop strategies that mitigate the impact of the CIE (e.g., Buczel et al., 2024; Ecker et al., 2024; Kan et al., 2021; Swire-Thompson et al., 2021; Tay et al., 2022). However, support for theoretical accounts of the CIE has so far been rather indirect. A thorough theoretical understanding, however, is essential in discerning the roles of cognitive processes such as (biased) information processing, memory (re-)construction, retrieval mechanisms, judgment formation, and decision making in the CIE. Therefore, in the current research, we combined

approaches from research on text comprehension, judgment formation, and persuasion to provide evidence regarding one prominent theoretical account of the CIE.

Two not mutually exclusive cognitive accounts dominate the ongoing discussion about the causes of the CIE: the *model-updating account* and the *retrieval-failure account* (e.g., Ecker et al., 2022; Lewandowsky et al., 2012; Sanderson & Ecker, 2020). In the current research, we focused on the model-updating account, but we first briefly introduce the retrieval-failure account to compare and contrast how the predictions of both accounts differ for the current experiment. According to the retrieval-failure account, the original information and retraction are simultaneously retained in memory. The CIE occurs when the original information is automatically activated and retrieved, but the strategic and cognitively demanding retrieval of the retraction fails (Ecker et al., 2010; Swire et al., 2017). Thus, the retrieval-failure account locates the causes of the CIE at retrieval of the critical statements during judgment formation. This account is supported, for example, by findings of stronger CIEs due to greater familiarity of the original information (Ecker, Lewandowsky, Swire, & Chang, 2011) and due to less recollection of the retraction (Swire et al., 2017).

The Model-Updating Account

The model-updating account was proposed with the first experiments on the CIE by Wilkes and Leatherbarrow (1988). According to this account, the key process of the CIE is the integration of information and inferences about an unfolding event into the memory representation of the event, the *situation model* (a concept of memory representations that was first introduced in the text-comprehension literature, e.g., Johnson-Laird, 1980; van Dijk & Kintsch, 1983; also referred to as a *mental model*, e.g., Rich et al., 2023). A situation model integrates all information relevant to the personal interpretation of an event: new information from the event description and about known similar situations, as well as elaborations beyond the description, such as inferences, beliefs, attitudes, and emotions (e.g.,

the inference that consumption of the water would harm people's health or the emotion of anger about a company endangering the health of citizens; cf. van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998).

The model-updating account postulates that the CIE occurs because, after integration of the original information, participants do not completely update and revise their situation model following the retraction. Consequently, some influence of the original information persists in the updated situation model (e.g., Wilkes & Leatherbarrow, 1988). Participants may find it difficult and effortful to remove retracted information from a complete situation model or fully integrate a retraction into a coherent situation model. Further, they may avoid to remove the original information because the situation model would be incomplete and implausible without an alternative, and because removal would produce discomfort (Susmann & Wegener, 2022). Participants, therefore, make judgments and inferences based on a situation model that continues to contain the original information (e.g., Gordon et al., 2017; Hamby et al., 2020; Susmann & Wegener, 2022). According to the model-updating account, the causes of the CIE are, thus, located at the processing and encoding of the critical statements.

To our knowledge, no study has directly assessed situation models in the context of the CIE so far, thus the support for the model-updating account consists primarily of results that align with predictions of the account and indirectly imply situation models as the basis of the CIE. For example, when the presumed integration of either original information or retraction could have contributed to the situation model's coherence and completeness, this piece of information influenced participants' judgments and inferences more strongly. The CIE was stronger when the original information was central for the causal understanding of the event (presumably filling a gap in the situation model; Hamby et al., 2020; Johnson & Seifert, 1994; Kan et al., 2021) and when it was coherent with the situation (Rich et al.,

2023). Conversely, the CIE was weaker when the retraction included an alternative explanation that supplanted the original information (presumably preventing a gap in the situation model), but even then, the CIE was often not fully eliminated (Ecker et al., 2010; Ecker, Lewandowsky, & Apai, 2011; Johnson & Seifert, 1994; Kan et al., 2021).

To the best of our knowledge, the contents of situation models and their relationship with the CIE has never been examined. The aim of the current research was to test the model-updating account by providing experimental evidence regarding the causal role of situation models in the formation of the CIE. To this end, we introduced a situation-model assessment into the CIE paradigm. This assessment allowed us to test experimentally whether the strength of the CIE was determined by two factors: the degree to which the original information was integrated into the situation model and the persistence of this integration despite retraction. To influence this integration and persistence, we needed a manipulation known to affect the strength of the CIE and theorized to affect the formation of situation models. Furthermore, we sought a manipulation that allowed us to independently influence integration and persistence. We chose to manipulate source credibility because it is, to our knowledge, the only manipulation in the CIE literature shown to influence both the effects of original information and retraction with similar strength and without interaction (Hey et al., 2025).

Source Credibility

Source credibility refers to the perceived willingness and capability of a source to provide reliable information within a specific context. It is a multidimensional construct influenced by factors such as perceived trustworthiness, expertise, authority, and attractiveness, which collectively shape perceptions of a source's credibility. Typically, high source credibility magnifies the persuasiveness and acceptance of a communication and low source credibility attenuates it (e.g., Greenberg & Miller, 1966; Guillory & Geraci, 2013;

Lariscy & Tinkham, 1999; Nadarevic et al., 2020; Pratkanis et al., 1988; Skagerberg & Wright, 2009; for a narrative review, see Pornpitakpan, 2004, and for a meta-analysis, see Wilson & Sherrell, 1993).

Source credibility is thought to affect judgments and inferences by influencing the cognitive processes involved in forming the memory representation of encoded information, such as the depth of information processing and the nature of thoughts generated in response to the information (cf. Heesacker et al., 1983; Petty & Cacioppo, 1986; Priester et al., 1999; Sternthal, Dholakia, & Leavitt, 1978; Sternthal, Phillips, & Dholakia, 1978; Tobin & Raymundo, 2009). Low source credibility specifically may work as a forewarning to guard beliefs against persuasion (cf. Greenberg & Miller, 1966; Ward & McGinnies, 1974) and may disrupt knowledge aquisition and revision (Van Boekel et al., 2017). Applied to the model-updating account of the CIE, this means that source credibility should affect the construction of situation models and, consequently, the CIE.

There are two relevant sources in the CIE paradigm: the *informant* presenting the original information and the *retractor* presenting the retraction. Informant credibility and retractor credibility can be manipulated independently of each other within the same experiment and have been shown to affect the CIE independently, antagonistically, and to a similar extent: In two experiments by Hey et al. (2025), the CIE was stronger when informant credibility was higher rather than lower, whereas the CIE was weaker when retractor credibility was higher rather than lower (for similar experiments manipulating only either informant credibility or retractor credibility, see Ecker & Antonio, 2021; Guillory & Geraci, 2013; Pluviano et al., 2020; Susmann & Wegener, 2023; Westbrook et al., 2023). These simultaneous effects of informant and retractor credibility on the CIE were of similar strength and did not interact, indicating that the credibility of one source did not influence the persuasiveness of the other. This pattern suggests that source credibility was applied only to

the corresponding critical statement, which allows us to independently influence and observe the effects of the original information and the retraction separately.

The model-updating account can explain these effects of informant credibility and retractor credibility on the CIE. Specifically, informant credibility should affect the degree to which agreement with the original information is integrated into a situation model, such that higher rather than lower informant credibility should lead to stronger integration, leading to a stronger CIE. Conversely, low informant credibility might even lead to the integration of disagreement with the original information. However, in our previous research, low informant credibility alone was not sufficient to eliminate the CIE (Hey et al., 2025), suggesting that even information presented by a low-credibility source was considered to some degree. Analogously, retractor credibility should affect the degree to which agreement with the original information persists in the situation model despite the retraction, such that higher rather than lower retractor credibility should lead to stronger updating of the situation model, in turn reducing the persistence and leading to a weaker CIE.

However, the retrieval-failure account can also explain the effects of source credibility on the CIE. Instead of source credibility altering the memory representation of the described event, higher informant credibility could facilitate the activation and retrieval of the original information during judgment formation, and likewise higher retractor credibility of the retraction. Additionally, knowledge of differences in source credibility between informant and retractor could increase the likelihood of strategic source monitoring during judgment formation (for a similar argument regarding the mechanism by which warnings could affect the CIE, see Ecker et al., 2010).

Therefore, a demonstration of source-credibility effects on the CIE does not suffice as support for the model-updating account. Rather, it is necessary to demonstrate three additional effects: First, source credibility should affect the degree of integration and the

degree of persistence of the original information in the situation model, approximated by how strongly participants relied on it when interpreting the situation, both after its initial presentation and after its retraction. Second, these effects of source credibility on the situation model should be mediated on the CIE. Third, source credibility should not affect the CIE if it is not known at the time when the critical information is processed and can thus not affect the situation model at encoding. This should hold even when source credibility is known at the time of retrieval and may, thus, affect retrieval and source monitoring during judgment formation, in which case the retrieval-failure account, by contrast, predicts that source credibility should affect the CIE (see Buczel et al., 2024, for a similar argument about the ineffectiveness of warnings presented after the original information but before judgment formation).

To effectively prevent source-credibility effects on the construction of situation models, one may disclose source credibility only after the construction of the situation model is completed. Persuasion research has shown that whereas source disclosure before a message produced source-credibility effects on judgments, source disclosure after a message but before judgment formation did not affect judgments (e.g., Allen et al., 2002; Greenberg & Miller, 1966; Mills & Harvey, 1972; O'keefe, 1987; Ward & McGinnies, 1974). Thus, once available, source credibility did not spontaneously affect the situation model. Consequently, disclosing source credibility after situation-model construction but before judgment formation allows researchers to differentiate whether source credibility affects encoding, retrieval at judgment formation, or both.

The Current Experiment

The aim of the current experiment was to rigorously test the model-updating account of the CIE by assessing participants' situation models and examining situation-model integration and updating as fundamental processes underlying the CIE (e.g., Lewandowsky et

al., 2012; Wilkes & Leatherbarrow, 1988). We used a standard CIE paradigm with between-subjects manipulations of informant and retractor credibility. We further manipulated the timing of the disclosure of source credibility (hereafter: credibility-disclosure timing). Informant credibility should affect the integration of the original information into the situation model; retractor credibility should affect the persistence of the original information in this situation model. This should, however, hold only if the credibility of the respective source is known to participants at the time of information delivery and consequently during the construction of the situation model.

Participants read an unfolding news report adapted from the "Water Department Shutdown" scenario (e.g., Ecker et al., 2017), which included original information and a retraction. Replicating Hey et al. (2025), informants and retractors were high-credibility journalists or low-credibility bloggers. Critically, in the *early-disclosure* groups, informant credibility was disclosed immediately before the original information, and retractor credibility was disclosed immediately before the retraction. Thus, source credibility was available to influence both the degree of integration and the updating of the situation model. In the *late-disclosure* groups, informant and retractor credibility were disclosed only after the corresponding critical statements. Thus, integration of the original information into the situation model and subsequent updating following the retraction should occur without source credibility. In line with previous research, source credibility should not retrospectively influence the degree of integration or the persistence of the original information in the situation model (cf. Allen et al., 2002; Greenberg & Miller, 1966; Mills & Harvey, 1972; O'keefe, 1987; Ward & McGinnies, 1974).

We assessed participants' situation model twice during the news report. The first assessment, immediately after the original information, measured the degree to which the original information was integrated into the *original situation model*. The second assessment,

immediately after the retraction, measured the degree to which the original information persisted in the *updated situation model*. For these situation-model assessments, we developed a set of questions following the the *Question-Answering method* (Graesser & Clark, 1985). Participants were asked about their thoughts and opinions on the news report, requiring inferences from their understanding of the situation. The questions could be answered using the original information, but also allowed unrelated answers. Beyond assessment of the situation model, these questions encouraged elaboration and integration of the information presented up to this point. Finally, we assessed the CIE as agreement with the original information in the test phase, using statements consistent with the original information that were embedded in a memory test for details of the report.

According to the model-updating account, the CIE occurs because participants base their judgments on a situation model in which the original information persists after the retraction, reflecting incomplete updating. From this account, three predictions can be derived for the current experiment. First, the updated situation model should predict the CIE (regression hypothesis). Specifically, stronger persistence of the original information in the updated situation model should lead to a stronger CIE. We also tested whether the degree of integration of the original information into the original situation model predicted its persistence in the updated situation model and the CIE.

Second, the effect of a variable such as source credibility on the CIE should depend on whether or not this variable can affect the situation model. In our experiment, the credibility-disclosure timing should thus moderate whether informant and retractor credibility affected the situation model and, consequently, the CIE (moderation hypotheses).

Third, any effect of source credibility on the CIE should be mediated through the situation models (mediation hypotheses), but only if source credibility is disclosed early enough to affect them. Thus, in the early-disclosure groups, higher informant credibility

should lead to a stronger integration of the original information into the original situation model. Higher retractor credibility should enhance situation-model updating and, thus, lead to weaker persistence of the original information in the updated situation model. Consequently, higher informant credibility should lead to a stronger CIE, and higher retractor credibility should lead to a weaker CIE. Importantly, these effects should be fully mediated through the original and updated situation models, rather than affecting the CIE directly. By contrast, effects of source credibility on the CIE not mediated through the situation models would indicate that the situation models were not the sole determinants of the CIE.

Conversely, in the late-disclosure groups, source credibility cannot affect the situation models and should thus not affect the CIE. According to the model-updating account, this should hold even though source credibility was disclosed long before the CIE assessment and could thus affect the retrieval processes proposed to underlie the CIE according to the retrieval-failure account (e.g., Lewandowsky et al., 2012). Therefore, the absence of source-credibility effects on the CIE in these groups would contradict the retrieval-failure account.

Method

Design

The experiment had a 2 × 2 × 2 factorial design with informant credibility (high, low), retractor credibility (high, low), and credibility-disclosure timing (early disclosure, late disclosure) as between-subjects factors resulting in eight experimental groups. We assessed participants' original situation model (after presentation of the original information) and their updated situation model (after presentation of the retraction), as shown in Figure 1. We also assessed the CIE. We further had participants rate informant and retractor credibility for a manipulation check and measured their memory for the details of the report.

Participants

The local ethics committee approved the experiment. Participants were recruited in

the United Kingdom via Prolific (www.prolific.com) and took part in the online experiment between June 12 and July 22, 2023. They had at least a high-school diploma and each received £12 as compensation.

Our hypotheses pertain to the moderation of source-credibility effects by credibility-disclosure timing (i.e., the interaction). To determine the sample size for our experiment, we broke this interaction down by credibility-disclosure timing. In the early-disclosure groups, we anticipated medium-sized effects of informant and retractor credibility on the CIE, based on previous CIE research (e.g., Guillory & Geraci, 2013; Hey et al., 2025). Conversely, in the late-disclosure groups, we anticipated null effects of informant and retractor credibility on the CIE. Therefore, we aimed for a relatively high power of $1 - \beta = .95$. A power analysis (using G*Power; Faul et al., 2007) showed that, in the 2 × 2 design of the early-disclosure groups (i.e., in four groups), 210 participants would be sufficient to detect medium effects ($f = 0.25$) with a power of .95 and $\alpha = .05$. For even distribution among four groups and two counterbalancing conditions (see below), the number of participants needed to be divisible by eight, resulting in 216 required participants. For the total sample including all eight groups, we doubled the target sample size to 432 participants. This sample size allowed us to detect small effects ($f_{min} = 0.17$) for the predicted two-way interaction of source credibility and credibility-disclosure timing with a power of .95 and $\alpha = .05$.

In total, we collected data from 478 participants who we randomly assigned to the eight groups. Participants were excluded if they met at least one of the following exclusion criteria: they failed one of the attention checks during the study phase ($n = 19$) and test phase ($n = 0$), or their processing speed indicated inattentiveness in one of three ways. First, their combined completion time for the study phase (excluding the situation-model assessments) and the test phase exceeded the sample mean by more than three standard deviations ($n = 7$). Second, participants spent more than 10 min on the screen within the situation-model

assessment where their intraindividual typing rate per character was slowest (n = 3). Third, they spent more than 2 min on the screen with the slowest intraindividual typing rate and this rate deviated by more than five interquartile ranges from their intraindividual median typing rate (n = 14).[1]

Of the final sample of 435 native English speakers, 197 self-reported as female, 230 as male, and eight as non-binary, with ages ranging between 18 and 32 years (M = 26.25, SD = 3.76). Group sample sizes are in Appendix A.

Materials and Procedure

All materials were administered in English. During recruitment, participants were markedly and repeatedly urged to complete the experiment at home on their personal computer or laptop. The procedure of the experiment is shown in Figure 1.

Study Phase

We first describe the general procedure of the study phase, and then the details of the source-credibility manipulations and the assessment of the situation models. After giving informed consent, participants began with the study phase. They were informed that their memory and understanding of the following text would be tested. Then they read a fictitious news report adapted from the "Water Department Shutdown" scenario (the full news report is in Online Supplement 1 in the OSF repository at https://osf.io/hpm5y/?view_only=25267025ef0e4cff968094d88f76e68c; Brydges et al., 2018; Ecker & Antonio, 2021; Ecker et al., 2017; Hey et al., 2025; Ithisuphalap et al., 2020; McIlhiney et al., 2022, 2023; Miller et al., 2022; Rich et al., 2023; Sanderson et al., 2021). The report consisted of eight paragraphs, each two to four sentences long. Each paragraph

[1] Further a priori exclusion criteria were an incorrect summary of the report, a self-assessment as unserious or inattentive in a seriousness check (Aust et al., 2013), and performance in the memory test more than three standard deviations below the mean, but no further exclusions were necessary based on these criteria. All participants were reimbursed for their participation regardless of whether they were excluded or not.

was presented on a separate screen, without the option to go back. To ensure that participants read each paragraph in full before proceeding, a minimum presentation time was set for each paragraph based on its word count and the average English reading speed (0.26 seconds per word, Trauzettel-Klosinski et al., 2012). After the minimum presentation time, participants could proceed to the next paragraph at their own pace.

In the report, an anonymous author described that a waterworks had to stop their water intake from a river due to numerous dead fish in the water, as well as the implications of this shutdown. To clarify that the original information and the retraction were not provided by the anonymous author of the news report but by other sources, we presented both critical statements in direct speech, attributing them to common male English names ("Paul Wilson" or "Mark Brown") to indicate that these individuals were making the statements. Assignment of names to the critical statements was approximately counterbalanced.

The original information was presented in paragraph 3 and stated that the fish deaths were caused by the willful illegal dumping of toxic chemical waste by a pharmaceutical company operating a factory near the river and water department. Immediately after the presentation of the original information, the report was interrupted to assess the original situation model, which reflects the degree to which the original information was integrated into the situation model. The assessment method is described in detail below. Critically, in the early-disclosure groups, the description of informant credibility (see below) was presented on a separate screen before the original information, ensuring that informant credibility was already disclosed when they encountered the original information and elaborated on it in the situation-model assessment. In the late-disclosure groups, the name of the informant was the only information about the source disclosed with the original information before the situation-model assessment, as the description of informant credibility was presented after the situation-model assessment on a separate screen. Thus, we assume

that by the time informant credibility was disclosed in the late-disclosure groups, the integration of the original information into the original situation model had already been completed. Afterward, the report continued in all groups.

The retraction was presented in paragraph 6 and stated that there was no evidence that the pharmaceutical company illegally dumped chemical waste into the river, that their factory complied with all standards, and that it was not to be blamed for the fish deaths. Immediately after this retraction, the report was again interrupted to assess the current situation model, which reflected the degree to which the original information persisted in the updated situation model despite the retraction. Retractor credibility was disclosed in the same manner as informant credibility (i.e., on a separate screen, before the retraction in the early-disclosure groups and after the assessment of the updated situation model in the late-disclosure groups). The final paragraph of the report contained the first attention check, which instructed participants to ignore the "Continue" button and instead click on the progress bar at the bottom of the screen. Participants who clicked on "Continue" failed the check and were excluded.

Source-Credibility Manipulations. We manipulated informant and retractor credibility in a similar way as in our prior experiments, that is, by providing source descriptions translated from the German materials used by Hey et al. (2025). The verbatim English descriptions are in Online Supplement 1 in the OSF repository at https://osf.io/hpm5y/?view_only=25267025ef0e4cff968094d88f76e68c. All sources were reporters. The descriptions of high- or low-credibility sources differed in the professionalism of the respective sources, that is, they were journalists or bloggers, and each description included an example of an occasion where the source had shared truthful or incorrect/misleading information in the past. This highlighted both their (in-)capability and (un-)willingness to provide correct and reliable information (i.e., expertise and

trustworthiness; Pornpitakpan, 2004). All four source descriptions had a length of 30 words.

We wanted to ensure that the English source descriptions would produce the anticipated differences in credibility ratings. Therefore, we conducted Validation Experiment 1 with 20 participants recruited on Prolific. Details and data of the validation experiment are in Online Supplement 2 in the OSF repository at https://osf.io/hpm5y/?view_only=25267025ef0e4cff968094d88f76e68c. Participants rated the credibility of all four described sources without encountering them in the context of a news report. As intended, high-credibility sources received higher ratings than low-credibility sources, and there was no difference between ratings for informants and retractors, and no interaction. The credibility of the sources was thus validated.

Assessment of the Situation Models. We used the principles of the Question-Answering method for assessing situation models (Graesser & Clark, 1985) to generate questions that probed participants' interpretations of the causal antecedents and consequences of the situation. The nine generated Why-, How-, and What-Happened-Next questions probed participants' thoughts and opinions about six aspects of the water department shutdown situation (e.g., the health risk posed by drinking the water or potential legal consequences): two rating questions on a scale from 0% to 100% with higher ratings indicating greater agreement with the original information and seven open-ended questions. They required participants to go beyond the event description and to make general inferences based on their understanding of the situation, but they could be answered without explicit evaluation of the belief in the original information or the likelihood that it caused the event. This minimized potential confounding effects of the situation-model assessment on the effectiveness of the retraction and on the CIE (cf. Ithisuphalap et al., 2020). The questions used to assess the situation models are listed in Appendix B.

To validate this assessment of situation models, we wanted to ensure that all questions

used in the assessments effectively differentiated between situation models that reflected either agreement with, disagreement with, or indecision regarding the original information. Furthermore, we wanted to validate the CIE statements used in the CIE test (as described below). To these ends, we conducted Validation Experiment 2 with 127 participants recruited on Prolific. Details and data are in Online Supplement 2 in the OSF repository at https://osf.io/hpm5y/?view_only=25267025ef0e4cff968094d88f76e68c. In this validation experiment, we presented the first two paragraphs of the "Water Department Shutdown" scenario (see Online Supplement 1) and manipulated which information followed: either the original information from a highly credible source, or the retraction from a highly credible source, or no information about the cause of the situation. The resulting situation models should, thus, include either agreement with, disagreement with, or indecision regarding the original information. Then, participants answered the nine questions to assess the situation model. For the situation-model assessment, the extent of (dis-)agreement with the original information reflected in each open answer was rated by two independent and blind raters on a Likert scale from –2 (strong disagreement) to +2 (strong agreement) with substantial inter-rater reliability ($\kappa_{weighted} = .62$; Landis & Koch, 1977). ANOVAs and post-hoc *t* tests showed that the majority of questions validly differentiated between agreement with, disagreement with, and indecision regarding the original information. Importantly, a multinomial logistic regression with all nine questions as predictors classified the group membership (agreement, disagreement, or indecision regarding the original information) correctly for 100% of the participants. That is, the combination of the nine questions allowed for a perfect differentiation of the three types of situation models.

In the main experiment, we therefore used all nine questions for the situation-model assessments. We presented them grouped according to the six aspects of the water department shutdown situation. The six aspects were probed in the same order at both time points of the

situation-model assessment (i.e., after the original information and after the retraction), with the order of presentation randomized for each participant. The answers to the open questions were again rated by independent and blind raters on a Likert scale from –2 (strong disagreement) to +2 (strong agreement). The total of five raters were familiar with the experimental design, but unfamiliar with the hypotheses of the experiment. Before the rating, they underwent extensive training based on the ratings from the validation experiment. For each question, five lists each containing a subset of answers were constructed in a way that made the association of a single answer to an experimental group, level of manipulation, time of administration, or the answer of the same participant to another question impossible. Each rater rated the answers of one list, which were presented simultaneously one below the other on a computer screen. All answers were rated by two raters.

Less than 0.1% of all answers were rated as nonsensical and excluded. Raters agreed on the exact rating in 69% of all answers (e.g., both ratings were +2) and on a positive or negative rating in an additional 13% (e.g., one rating was +1 and the other was +2). The inter-rater reliability was $\kappa_{weighted} = .72$, which indicates substantial agreement (Landis & Koch, 1977). Conflicts were resolved by using the median rating. Descriptive statistics, differences between times of administration, and correlations for all nine questions and the CIE are in Online Supplement 3 in the OSF repository at https://osf.io/hpm5y/?view_only=25267025ef0e4cff968094d88f76e68c. To compute a single score for each situation model, we first standardized the data using a *z*-score transformation and then computed each participant's mean of the *z*-scores for the original situation model and the updated situation model (Cronbach's $\alpha = .70$ and .77 respectively).

Delay With Filler Tasks

In the early-disclosure groups, the situation models should be affected by source credibility. In the late-disclosure groups, the situation models, informant, and retractor

credibility should be separate memory representations. However, these representations may be temporarily associated so that source credibility would affect the late-disclosure groups' agreement ratings in an immediate test, even though their situation models were not affected by source credibility. Critically, this association decays over time so that source credibility should not affect the late-disclosure groups' agreement ratings in a delayed test (cf. the *sleeper effect*, Kumkale & Albarracín, 2004).

Therefore, we implemented a 45-minute delay before the test phase which should be sufficiently long for immediate source-credibility effects to disappear (Albarracín et al., 2017). As filler tasks, we used a classification task where participants evaluated simple arithmetic equations as correct or incorrect, three measures of analytical thinking (including an attention check which was passed by 99% of participants), and reading paragraphs of The Great Gatsby (Fitzgerald, 1925) until 45 minutes had elapsed. With this delay, the CIE could be assessed based on the situation models alone, unaffected by immediate source-credibility effects on judgment.

Test Phase

After the filler tasks, participants proceeded with the test phase. The CIE and memory for details of the report were assessed with a questionnaire that consisted of 14 statements about the report and utilized a Likert scale from 1 (*strong disagreement*) to 7 (*strong agreement*). The 14 statements were each presented on a separate screen, in a fixed order (see Online Supplement 1 in the OSF repository at https://osf.io/hpm5y/?view_only=25267025ef0e4cff968094d88f76e68c). Four correct and four incorrect statements probed memory for details of the news report (e.g., "The water department normally supplies 120,000 households with fresh water."). This memory test served to ensure that possible differences in the CIE could not be attributed to differences in memory for details of the report (cf. Sanderson et al., 2021). The remaining six statements

measured the CIE. One statement presented the original information verbatim. Five statements were possible inferences consistent with the original information that one could draw from the news report (e.g., "The riverside pharmaceutical company should take responsibility for the incident.").

As mentioned above, these CIE statements had been validated in Validation Experiment 2. In this experiment, we manipulated whether the original information, the retraction, or no further information was presented regarding the cause of the water department shutdown so that the resulting situation model should include either agreement with, disagreement with, or indecision regarding the original information. As in the main experiment, participants rated their agreement with the CIE statements on a Likert scale from 1 (*strong disagreement*) to 7 (*strong agreement*). A multinomial logistic regression with all six CIE ratings as predictors classified correctly, for 100% of the participants of the validation experiment, whether they (dis-)agreed with the original information. That is, the validation experiment showed that a combination of the CIE ratings allowed for a perfect differentiation between agreement and disagreement with the original information (for details and data, see Online Supplement 2 in the OSF repository at https://osf.io/hpm5y/?view_only=25267025ef0e4cff968094d88f76e68c).

In the main experiment, to compute a single score for the CIE, we standardized the ratings of agreement with the six CIE statements using a *z*-score transformation and computed the mean of the *z*-scores for each participant (Cronbach's $\alpha = .87$).[2] Next, for a

[2] We also computed scores for the dependent variables as factor scores from Principal Component Analyses (PCA) using the regression method which weighs the standardized data with their respective factor loadings (cf. DiStefano et al., 2009). For the CIE, we also computed a simple mean of the six raw agreement ratings. Due to the near-perfect correlations of the reported averaged *z* scores with the factor scores for the original situation model ($r = .99$, $p < .001$, $BF_{10} > 10^{500}$), the updated situation model ($r = .98$, $p < .001$, $BF_{10} > 10^{500}$), the CIE ($r > .99$, $p < .001$, $BF_{10} > 10^{500}$), and with the mean raw agreement ratings ($r > .99$, $p < .001$, $BF_{10} > 10^{500}$), we present the results of the simpler computation method, which also allowed us to report all dependent variables on the same scale, in the main text. The

manipulation check, participants rated the credibility of their respective informant and retractor, in approximately counterbalanced order. Credibility was rated on a Likert scale ranging from 1 (*absolutely not credible*) to 7 (*absolutely credible*). To avoid memory effects, the complete source descriptions were presented alongside the rating scale, without repeating which source provided the original information and which provided the retraction. These ratings were followed by the second attention check. In the same layout as the preceding rating items, participants were instructed to ignore the rating scale and click on the progress bar to continue. Next, participants summarized the report in one sentence, provided their demographic information, and indicated whether they had participated attentively and seriously, and whether their data could be used in scientific analyses (seriousness check; Aust et al., 2013). They were assured that their answer on the seriousness check would not affect their compensation for the participation. Finally, participants were debriefed and received their compensation via Prolific. Mean completion time of the experiment was 68 minutes.

Results

The raw data and the analysis code are available in the OSF repository at https://osf.io/hpm5y/?view_only=25267025ef0e4cff968094d88f76e68c. We used JASP (JASP Team, 2024) to obtain frequentist and Bayesian inference statistics. Bayesian statistics allow us to quantify evidence for the alternative and *for* the null hypotheses. We report Bayes' Factors (BF) in such a way that they indicate support for the respectively favored hypothesis. For *t* tests, correlations, and one-way ANOVAs, BF_{01} indicates support for the null hypothesis and BF_{10} indicates support for the alternative hypothesis. For the effects of multi-factor ANOVAs, we report BF_{incl} (BF_{excl}) indicating support for the inclusion (exclusion) of the effect based on matched models. For example, $BF_{01} = 10.00$ indicates that the observed data are ten times more likely under the null than under the alternative

---

results and conclusions are the same for all computation methods.

hypothesis. BFs larger than 3, 20, and 150 indicate positive, strong, and very strong evidence, respectively (Raftery, 1995). For frequentist analyses, we set $\alpha = .05$. Effect-size interpretations follow Cohen (Cohen, 1988). We first present a check of the source-credibility manipulations, then show that there were no unintended group differences in general memory for the report.

Manipulation Check

Figure 2 shows the source-credibility ratings for all groups. To check if the source-credibility manipulations were successful, we performed two 2 × 2 × 2 ANOVAs with the between-subjects factors informant credibility (high, low), retractor credibility (high, low), and credibility-disclosure timing (early disclosure, late disclosure) on ratings of informant credibility and retractor credibility. Ratings of informant credibility were higher for the high-credibility informant than for the low-credibility informant, $F(1, 427) = 165.57$, $p < .001$, $\eta^2_p = .28$, $BF_{incl} = 9.45 \cdot 10^{28}$. There were no other effects on informant-credibility ratings (all other ps $> .059$, $\eta^2_p < .01$, $BF_{excl} > 1.71$). Ratings of retractor credibility were higher for the high-credibility retractor than for the low-credibility retractor, $F(1, 427) = 235.18$, $p < .001$, $\eta^2_p = .36$, $BF_{incl} = 2.66 \cdot 10^{39}$. There were no other effects on retractor-credibility ratings (all other ps $> .135$, $\eta^2_p < .01$, $BF_{excl} > 3.08$). Thus, the manipulations of informant and retractor credibility had the intended selective effects on the credibility ratings with no other group differences.

Memory for Details of the Report

To assess memory for details of the report, we had participants rate their agreement with the four correct and four incorrect memory statements. Ratings of incorrect statements were reverse-coded so that higher ratings indicate greater (dis-)agreement with (in-)correct facts. The groups' mean ratings are in Appendix A. A one-way ANOVA across all groups showed no differences in mean ratings, $F(7, 427) = 1.29$, $p = .254$, $\eta^2_p = .02$, $BF_{excl} = 35.11$.

Thus, between-group CIE differences are not attributable to differences in memory for details of the report (cf. Sanderson et al., 2021).

The Timing of Credibility Disclosure Moderated the Effects of Source Credibility

As mentioned in the Methods section, the scores for the dependent variables, original situation model, updated situation model, and CIE, were computed as the means of standardized agreement ratings. Descriptive statistics for these scores are in Figure 3; bivariate correlations in Table 1. As the standardization eliminated mean differences between the original and the updated situation model, we used paired *t* tests on the unstandardized ratings for each question to check whether the retraction reduced agreement with the original information. The retraction reduced the rated agreement with the original information in all nine questions (all ts > 5.93, all ps < .001, all $d_z > 0.28$, all $BF_{10} > 10^6$). Thus, the agreement with the original information was overall reduced following the retraction, which indicates that the retraction induced an update of the original situation model.

The model-updating account predicts a causal relationship between the updated situation model and the CIE. Therefore, the effect of source credibility on the CIE should depend on whether or not source credibility can affect the situation model (moderation hypotheses). We used tests of interactions in ANOVAs to determine if credibility-disclosure timing moderated the effects of informant credibility and retractor credibility on the original and updated situation models, such that effects of source credibility should only occur in the early-disclosure groups. Importantly, we then tested whether credibility-disclosure timing likewise moderated the effects of informant credibility and retractor credibility on the CIE. We conducted three 2 × 2 × 2 ANOVAs with the between-subjects factors informant credibility (high, low), retractor credibility (high, low), and credibility-disclosure timing (early disclosure, late disclosure) on the original situation model, the updated situation model, and the CIE. To follow up on significant moderating effects (i.e., interactions) of credibility-

disclosure timing, we analyzed the early-disclosure and late-disclosure groups separately with two 2 × 2 ANOVAs with the between-subjects factors informant credibility (high, low) and retractor credibility (high, low).

The Original Situation Model

Test statistics for the original situation model are in Table 2 (top third). As expected, credibility-disclosure timing moderated the effect of informant credibility on the original situation model as indicated by a two-way interaction effect of informant credibility and credibility-disclosure timing. To follow up on the interaction, we analyzed the early-disclosure and late-disclosure groups separately. In the early-disclosure groups, as expected, high informant credibility led to greater agreement with the original information than low informant credibility, that is, with high informant credibility, the original information was integrated more strongly into the original situation model. In the late-disclosure groups, informant credibility was not yet disclosed when the original situation model was assessed, and thus, trivially, did not affect agreement with the original information. Interestingly, an exploratory contrast comparing all late-disclosure groups against the early-disclosure groups with high informant credibility showed that the informant with undisclosed credibility led to similar agreement with the original information as the high-credibility informant, $F(1, 427) = 0.24$, $p = .623$, $\eta^2_p < .01$, $BF_{01} = 6.71$.[3] This finding is consistent with previous research showing that information from unknown sources and from known and trusted sources is treated similarly (Sterrett et al., 2019; Ward & McGinnies, 1974).

The Updated Situation Model and the CIE

According to the model-updating account, the timing of credibility disclosure should moderate the effects of both informant and retractor credibility on the updated situation model. These effects should be mediated through the updated situation model on the CIE.

[3] The BF for the contrast analysis is based on a pair-wise *t* test.

Therefore, we predicted similar patterns of results for the updated situation model and the CIE and report them jointly. Test statistics are in Table 2 (mid third for the updated situation model, bottom third for the CIE).

As predicted, the credibility-disclosure timing moderated the effects of informant credibility and retractor credibility on the updated situation model and on the CIE, as indicated by two-way interaction effects of informant credibility and credibility-disclosure timing and of retractor credibility and credibility-disclosure timing on both dependent variables. To follow up on these interactions, we again analyzed the early-disclosure and late-disclosure groups separately.

In the early-disclosure groups, as expected, high informant credibility led to greater agreement with the original information than low informant credibility. That is, with high informant credibility, the original information persisted more strongly in the updated situation model. High retractor credibility led to less agreement with the original information than low retractor credibility, that is, with high retractor credibility, the original information persisted less strongly in the updated situation model. High informant credibility also led to a stronger CIE than low informant credibility and high retractor credibility also led to a weaker CIE than low retractor credibility. Thus, when source credibility was disclosed before the integration of the original information into the original situation model and before the update of the situation model following the retraction, there were independent and antagonistic effects of informant and retractor credibility on the CIE (replicating Hey et al., 2025).

In the late-disclosure groups, retractor credibility was not yet disclosed when the updated situation model was assessed, and thus, trivially, did not affect agreement with the original information after retraction. Informant credibility did not affect agreement with the original information after retraction either. The CIE was also not affected by informant or retractor credibility. Thus, when source credibility was disclosed only after the construction

of the situation models, there were no effects of source credibility on the updated situation model or the CIE.

To summarize, these results are consistent with the predictions of the model-updating account that source credibility affected the CIE by affecting the integration of the original information into the situation model and by affecting the update of the situation model following the retraction. After establishing the moderating effects of the timing of credibility disclosure on the effect of source credibility on the situation model and the CIE, we used structural equation modeling to test the causal relationships of the situation models and the CIE predicted by the model-updating account.

Testing the Model-Updating Account with Structural Equation Modeling

The model-updating account of the CIE leads to a set of hypotheses that we tested with a structural equation model (SEM). First, participants' updated situation model should predict the CIE (regression hypothesis). Second, any effect of source credibility on the CIE should be mediated through the situation models (mediation hypotheses). Third, the effect of a variable on the CIE should depend on whether or not it can affect the situation model (moderation hypotheses). We implemented these hypotheses in an SEM corresponding to the predicted relationships between variables.

Specification and Implementation of the Structural Equation Model

An SEM allows for the analysis of complex relationships between multiple variables simultaneously. Based on multiple regressions, they enable the estimation of direct effects between several predictors and variables as well as the estimation of indirect effects through one or more mediators. SEMs are used to test theoretical models for which the relationships between the variables are specified a priori, as the three predictions of the model-updating account do for the current experiment. The SEM specifies the paths of direct effects between predictors, moderators, mediators, and the outcome variable. The path of an indirect effect is

the sequence of several direct paths through one or more mediators. Mathematically, indirect effects are the product of these direct effects. The total effect of a predictor, moderator, or mediator is the sum of all their direct and indirect effects on a variable (labeled Σ).

The predictions for the current experiment correspond to an SEM of moderated serial mediation. In an SEM, the moderation of an effect is indicated by a significant interaction effect of a predictor and a moderator (labeled M with respective indices). The mediation of an effect is indicated by a significant indirect effect through a mediator (cf. Baron & Kenny, 1986). The mediation is full if, after the inclusion of the indirect effect, the direct effect of the predictor is non-significant. The mediation is only partial if the indirect and direct effects of a predictor are both significant.

The SEM we used includes informant credibility (I) and retractor credibility (R) as predictors (high = 1, low = –1), credibility-disclosure timing (T) as a moderator (early disclosure = 1, late disclosure = –1), the original situation model (O) and the updated situation model (U) as mediators, and the CIE (C) as the outcome variable. We specified the effects included in the SEM corresponding to the predictions of the model-updating account. Figure 4 shows the conceptual SEM including the paths relevant for the predictions of the model-updating account; the figure in Online Supplement 4 shows the complete SEM including all paths.

The updated situation model should predict the CIE (regression hypothesis). Thus, we included a direct effect of the updated situation model on the CIE (path $U \rightarrow C$). Because the original situation model may predict the updated situation model, we included a direct effect of the original on the updated situation model (path $O \rightarrow U$). To test whether the original situation model continued to influence the CIE despite the update following the retraction, we additionally included a direct effect of the original situation model on the CIE (path $O \rightarrow C$).

Any effect of source credibility on the CIE should be mediated through the situation

models (mediation hypotheses). Thus, we included direct effects of informant credibility on the original and the updated situation models (paths $I \rightarrow O$ and $I \rightarrow U$) and a direct effect of retractor credibility on the updated situation model (path $R \rightarrow U$). To test the prediction that source credibility would not affect the CIE beyond the mediated effects, we also included direct effects of informant and retractor credibility on the CIE (paths $I \rightarrow C$ and $R \rightarrow C$, both predicted to be zero). We did not allow retractor credibility to affect the original situation model because retractor credibility was manipulated only after the assessment of the original situation model. We further did not include the interaction term of informant credibility and retractor credibility in the SEM because Hey et al. (2025) and the presented ANOVAs (see Table 2) consistently showed the absence of such an interaction. Inclusion of these effects in an overparameterized SEM led to slightly worse model fit (see below), but to identical result patterns.

The effect of source credibility on the CIE should depend on whether or not it can affect the situation models; that is, the effects of source credibility should depend on the timing of credibility disclosure (moderation hypotheses, as shown by the ANOVAs reported above). Thus, we included moderating effects in the SEM. The credibility-disclosure timing could moderate the effects of informant credibility on the original situation model (moderation path $M_{I \rightarrow O}$), on the updated situation model (path $M_{I \rightarrow U}$), and on the CIE (path $M_{I \rightarrow C}$). Likewise, it could moderate the effects of retractor credibility on the updated situation model (path $M_{R \rightarrow U}$) and on the CIE (path $M_{R \rightarrow C}$). As mathematically, a moderating effect in the SEM is defined as the interaction of two predictors (e.g., informant credibility × credibility-disclosure timing), credibility-disclosure timing also needed to be included in the SEM as a predictor. It could have the same direct effects on the situation models and the CIE as informant credibility.

There were no predictions as to whether the credibility-disclosure timing should

moderate the effects of the original and updated situation models. To evaluate the trade-off between model parsimony and comprehensiveness, we tested both a parsimonious SEM with the moderation of source-credibility effects only and a comprehensive SEM in which we allowed the moderation of all effects that were included in the SEM. Both SEMs led to identical patterns of results for the predicted effects, but the parsimonious SEM showed slightly worse model fit (see below). Because further the additionally included moderating effects were significant, we report the results of the comprehensive SEM to provide a more complete picture of the relationships in the data.

We implemented the SEM in R using the *lavaan* package (Rosseel, 2012) and used the maximum-likelihood procedure to estimate the effects (the analysis code is available in the OSF repository at https://osf.io/hpm5y/?view_only=25267025ef0e4cff968094d88f76e68c). We evaluated the fit of the SEM to the sample data against the following criteria: a non-significant χ^2 test of overall fit, a Comparative Fit Index (CFI) larger than .96 (Hu & Bentler, 1999), a Tucker Lewis Index (TLI) larger than .95 (Schumacker & Komax, 2004), a Root Mean Square Error of Approximation (RMSEA), and a Standardized Root Mean Square Residual (RSMS) smaller than .08 (Byrne, 1994). Additionally, we report the Akaike and Bayesian Information Criteria (AIC and BIC) to allow comparisons of model fit (with smaller values indicating better fit). The model fit of the SEM was good, $\chi^2(5) = 7.04$, $p = .218$, CFI = 1.00, TLI = 0.98, RMSEA = 0.03, SRMR = 0.02, AIC = 2105.42, BIC = 2195.07.[4]

To test the significance of effects, we estimated 95% confidence intervals with a

[4] Worse model fit after inclusion of a retractor-credibility effect on the original situation model and the interaction term of informant credibility and retractor credibility (overparameterized vs. comprehensive SEM): ΔAIC = 3.50, ΔBIC = 19.81. Worse model fit after exclusion of credibility-disclosure timing moderating the effects of the original and updated situation model (parsimonious vs. comprehensive SEM): ΔAIC = 14.09, ΔBIC = 1.87.

bootstrapping procedure using 10,000 samples, as recommended by Hayes (2015). Confidence intervals for the unstandardized regression weights are bias-corrected. An effect is significant if the 95% confidence interval reported in brackets does not include zero.

The Timing of Credibility Disclosure Moderated the Effects of Source Credibility

Table 3 lists the unstandardized and standardized regression weights of all direct, indirect, and total effects. The model-updating account predicts that the effect of source credibility on the CIE should depend on whether or not source credibility can affect the situation model (moderation hypotheses). As expected, and mirroring the ANOVA results, credibility-disclosure timing moderated the source-credibility effects: The total effects of informant credibility were stronger in the early-disclosure groups than in the late-disclosure groups, both on the original situation model ($\Sigma M_{I \rightarrow O}$) and on the CIE ($\Sigma M_{I \rightarrow C}$). Further, the total effects of retractor credibility were stronger in the early-disclosure groups than in the late-disclosure groups, both on the updated situation model ($\Sigma M_{R \rightarrow U}$) and on the CIE ($\Sigma M_{R \rightarrow C}$). Thus, consistent with the model-updating account, the effects of source credibility on the CIE depended on whether or not the source credibility was disclosed such that it could affect the situation models.

Moreover, the credibility-disclosure timing unexpectedly moderated the effects of the situation models: The effects of the original situation model were stronger in the early-disclosure than in the late-disclosure groups, both on the updated situation model (path $M_{O \rightarrow U}$) and on the CIE (path $M_{O \rightarrow C}$). The effect of the updated situation model on the CIE was weaker in the early-disclosure than in the late-disclosure groups (path $M_{U \rightarrow C}$). That is, early disclosure of source credibility led to stronger relationships of the original situation model with the updated situation model and the CIE, but led to a weaker relationship of the updated situation model with the CIE. Interestingly, the stronger effect of the original situation model on the CIE and the weaker effect of the updated situation model on the CIE in the early-

disclosure groups did not lead to a stronger CIE compared to the late-disclosure groups (indicated by the absence of an effect of credibility-disclosure timing, $\Sigma T \rightarrow C$, see also Table 2 for the corresponding ANOVA result).

The Updated Situation Model Predicted the CIE

Due to the multiple moderating effects of the credibility-disclosure timing, we refrain from interpreting the remaining regression weights of the full SEM. Instead, to test for the causal effect of the updated situation model on the CIE and the mediation of source-credibility effects through the situation models, we analyzed the data of the early-disclosure and late-disclosure groups in separate SEMs that included informant and retractor credibility as predictors, the situation models as mediators, and the CIE as the outcome variable. Model fit was good in the early-disclosure groups, $\chi^2(1) = 0.10$, $p = .756$, CFI = 1.00, TLI = 1.03, RMSEA = 0.00, SRMR = 0.01, AIC = 1065.37, BIC = 1102.60, and the late-disclosure groups, $\chi^2(1) < 0.01$, $p = .956$, CFI = 1.00, TLI = 1.10, RMSEA = 0.00, SRMR < 0.01, AIC = 1039.23, BIC = 1076.41. Figure 5 shows the two serial-mediation models; Table 4 lists the unstandardized and standardized regression weights of all direct, indirect, and total effects.

The moderating effects in the joint model showed that there were the expected group differences in the effects of source credibility and unexpected group differences in the effects of the situation models. Here, we examined the unexpected differences in the effects of the situation models. The updated situation model was a better predictor of the CIE in the late-disclosure groups than in the early-disclosure groups (indicated by the moderating effect $M_U \rightarrow C$ reported above). Critically, as predicted by the model-updating account, the updated situation model predicted the CIE in both the early-disclosure and late-disclosure groups (paths $U \rightarrow C$; regression hypothesis). That is, the stronger the persistence of the original information in the updated situation model was, the stronger was the CIE, in both groups.

Further, the original situation model predicted the updated situation model and the

CIE in the early-disclosure groups only, but not in the late-disclosure groups. Specifically, in the early-disclosure groups, the more strongly the original information was integrated into the original situation model, the more strongly it persisted in the updated situation model despite the retraction (path $O \rightarrow U$), and the stronger was the CIE (path $O \rightarrow C$). There was no difference between the direct effects of the original situation model and of the updated situation model on the CIE ($[O \rightarrow C] - [U \rightarrow C] = -0.03$ [–0.31, 0.26], $\Delta\beta = -0.06$ [–0.26, 0.15]). In the late-disclosure groups, by contrast, the original situation model did not predict the updated situation model nor the CIE. Thus, in the late-disclosure groups, the CIE was exclusively predicted by the updated situation model.

The Situation Models Mediated the Source-Credibility Effects on the CIE

The model-updating account predicts that the effect of source credibility on the CIE should be fully mediated through the situation models (mediation hypotheses). However, this prediction only holds if source credibility is disclosed such that it can affect the situation models, that is, in the early-disclosure groups. Indeed, the moderating effects of credibility-disclosure timing reported above confirmed that the effects of source credibility differed between the early-disclosure and late-disclosure groups.

Early-Disclosure Groups. In the early-disclosure groups, higher informant credibility should lead to stronger integration of the original information into the original situation model and to its stronger persistence in the updated situation model despite retraction. As expected, informant credibility affected the original situation model (path $I \rightarrow O$), and affected the updated situation model mediated through the original situation model (path $I \rightarrow O \rightarrow U$). After inclusion of this indirect effect, the direct effect of informant credibility on the updated situation model was not significant (path $I \rightarrow U$). Thus, higher informant credibility led to stronger integration of the original information into the original situation model. This integration was maintained after the update (indicated by the significant

indirect effect), but informant credibility did not directly affect the degree to which the situation model was updated following the retraction.

These effects of informant credibility on the situation models should be fully mediated on the CIE so that higher informant credibility should lead to a stronger CIE. Informant credibility affected the CIE ($\Sigma I \rightarrow C$). Consistent with the ANOVA results (see Table 2 and Figure 3), higher informant credibility led to a stronger CIE, replicating Hey et al. (2025). As expected, this effect was mediated on the CIE through the original situation model (path $I \rightarrow O \rightarrow C$) and serially mediated through both situation models (path $I \rightarrow O \rightarrow U \rightarrow C$). After inclusion of these indirect effects, the direct effect of informant credibility on the CIE was not significant (path $I \rightarrow C$). Thus, as predicted by the model-updating account, the effect of informant credibility on the CIE was fully mediated through the situation models.

Higher retractor credibility should lead to a stronger update of the original situation model and, thus, to weaker persistence of the original information in the updated situation model. Confirming this prediction, retractor credibility affected the updated situation model (path $R \rightarrow U$). The updated situation model should fully mediate this effect of retractor credibility on the CIE so that higher retractor credibility should lead to a weaker CIE. Retractor credibility affected the CIE ($\Sigma R \rightarrow C$). Consistent with the ANOVA results and replicating Hey et al. (2025), higher retractor credibility led to a weaker CIE. As in Hey et al. (2025), the effects of informant and retractor credibility on the CIE were of similar size (indicated by the non-significant difference of effects, $|\Sigma I \rightarrow C| - |\Sigma R \rightarrow C| = 0.02$ [–0.12, 0.16], $\Delta\beta = 0.03$ [–0.14, 0.20]). The total effect of retractor credibility on the CIE was partially mediated through the updated situation model (path $R \rightarrow U \rightarrow C$), with a direct effect of retractor credibility on the CIE remaining (path $R \rightarrow C$). Thus, as predicted by the model-updating account, retractor credibility affected the CIE through the degree to which the original information persisted in the situation model that was updated after the retraction.

However, retractor credibility also affected the CIE directly, which was not predicted by the model-updating account.

Late-Disclosure Groups. In the late-disclosure groups, informant credibility could not affect the original situation model because it was not disclosed before the processing of the original information, but it might affect the updated situation model. Consequently, informant credibility should only affect the CIE if its effect was mediated by the updated situation model. Retractor credibility could affect neither the original nor the updated situation model because it was not yet disclosed to participants during the processing of either critical information. Consequently, retractor credibility should not affect the CIE. Confirming these predictions and consistent with the ANOVA results, informant credibility neither affected the original situation model (path $I \rightarrow O$), nor the updated situation model (path $I \rightarrow U$). Accordingly, informant credibility did not affect the CIE either ($\Sigma I \rightarrow C$). Likewise, retractor credibility did not affect the updated situation model (path $R \rightarrow U$) and, accordingly, did not affect the CIE either ($\Sigma R \rightarrow C$). To summarize, informant and retractor credibility did not have any direct or indirect effects in the late-disclosure groups (for all indirect effects, $|B| < 0.02$, $|\beta| < 0.03$). Thus, consistent with the model-updating account, source credibility did not affect the CIE when it could not affect the construction of the situation models.

Discussion

The model-updating account proposes that the CIE occurs because participants make judgments and inferences based on an incompletely updated situation model that continues to contain the original information despite retraction (e.g., Wilkes & Leatherbarrow, 1988). The main aim of the experiment presented here was to put this account to empirical test. To this end, we developed a method for the assessment of situation models in the CIE paradigm. In combination with an SEM, we used this method to directly test the mediating role of situation models in the CIE. In support of the model-updating account, we found that reliance on

situation models in judgments and inferences is a cause of the CIE. Here, we discuss the evidence we found in support of the model-updating account of the CIE and why alternative accounts cannot explain our results. Then, we discuss unexpected findings and practical implications.

Support for the Model-Updating Account: Situation Models Mediated Source-Credibility Effects on the CIE

Overall, our research provides direct evidence for the model-updating account of the CIE by empirically demonstrating the relationships between situation models during processing of a news report and the CIE. In line with our regression hypothesis, the updated situation model predicted the CIE in the delayed test phase. The stronger the original information persisted in the updated situation model, the stronger was the CIE. This relationship existed regardless of whether or not source credibility could affect the situation model (i.e., was disclosed early or late).

As predicted by the model-updating account, the effects regarding source credibility depended on whether or not source credibility could affect the situation model. In line with our moderation hypotheses, the timing of credibility disclosure moderated whether or not informant and retractor credibility affected the CIE. In the early-disclosure groups, participants had the opportunity to consider source credibility when they processed the critical statements and, only in these groups, informant and retractor credibility affected the CIE. In the late-disclosure groups, participants were unable to consider the source credibility of informant and retractor when they processed the respective critical statements so that source credibility neither affected the original situation model, nor the updated situation model, nor the CIE.[5]

[5] The absence of an informant-credibility effect on the CIE in the late-disclosure groups may seem inconsistent with studies in which informant discreditation at retraction (e.g., the retractor described that the informant had an ulterior motive to lie) reduced the CIE beyond

In line with our mediation hypotheses, the situation models mediated the source-credibility effects on the CIE in the early-disclosure groups. Using an SEM, we mapped the paths of the source-credibility effects and showed that, consistent with the model-updating account, the situation models played a causal role in the CIE. Participants constructed an original situation model that was influenced by informant credibility. Low informant credibility attenuated the integration of the original information and related inferences compared to the groups with high informant credibility and the late-disclosure groups. The original situation model predicted the updated situation model assessed after the retraction. This update was stronger for participants who received the retraction from a highly credible retractor than for participants who received the retraction from a retractor with low credibility. Consequently, the degree to which the original information persisted in the updated situation model after the retraction was influenced by two opposing factors, namely, the degree of initial integration of the original information into the original situation model and the strength of the update. Finally, as described above, the updated situation model predicted the CIE in the delayed test phase. Taken together, the results support our three key hypotheses derived from the model-updating account concerning the relationships between the situation models and the CIE and thus provide strong support for this account. Future research could further strengthen the support for the causal role of the situation model for the CIE by using the assessment of situation models combined with an SEM to investigate whether these results generalize to other CIE scenarios and whether the situation model mediates the effects of other variables on the CIE as well.

---

the effect of the retraction alone (Ecker et al., 2024; Westbrook et al., 2023). In these studies, participants did consider informant credibility that was disclosed after the original information when it was explicitly linked to the retraction. In our experiment, there was no direct connection between informant credibility and retraction, which may explain the different results.

Boundaries of Source-Credibility Effects on the CIE: A Retrieval-Based Account Alone Is Not Sufficient

The other prominent account of the CIE, the retrieval failure account, presumes that the CIE is the result of failure to retrieve the retraction or the source of the original information during judgment formation (Ecker et al., 2010; Swire et al., 2017). This account thus locates the causes of the CIE at retrieval during judgment formation. Therefore, one may derive the prediction that as long as informant and retractor credibility are available for retrieval during judgment formation, they should affect the CIE. In all groups of the current experiment, source credibility was disclosed long before the CIE test and thus could have affected, for example, the evaluation of the retrieved critical pieces of information or the likelihood of strategic source monitoring. A retrieval-based account thus predicts effects of source credibility on the CIE in both the early-disclosure and late-disclosure groups. One may argue that, in the early-disclosure groups, the assessment of the situation model provided an opportunity to rehearse both the event and source credibility, whereas in the late-disclosure groups, such rehearsal could not include source credibility. This may result in weaker source-credibility effects in the late-disclosure groups. Nevertheless, a retrieval-based account predicts at least some effects of source credibility on the CIE in the late-disclosure groups. Critically, however, source credibility did not affect the CIE in the late-disclosure groups, which is inconsistent with purely retrieval-based accounts of the CIE and renders explanations of source-credibility effects on the CIE (e.g., Hey et al., 2025) that are exclusively based on retrieval processes improbable.

It should be noted, however, that although we consider it highly unlikely, the possibility remains that source credibility did not affect the CIE in the late-disclosure groups because participants did not retrieve source credibility. Kumkale and Albarracín (2004) describe some anecdotal evidence from experiments on the sleeper effect that later rather than

earlier source disclosure may impede delayed recall and recognition of sources. Nevertheless, our data rather support the causal relevance of the situation models as an explanation for the absence of source-credibility effects in the late-disclosure groups. Future research could investigate the influence of actually presented and correctly remembered versus falsely remembered source credibility on the CIE in order to experimentally test the retrieval-failure account.

Effects not Predicted by the Model-Updating Account

In addition to the effects predicted by the model-updating account, we found two unexpected effects on the CIE in the early-disclosure groups and an unexpected disparity between the early-disclosure and late-disclosure groups. These effects cannot be explained by the model-updating account alone, but are also not inconsistent with a causal role of situation models in the CIE. Critically, the retrieval-failure account cannot explain these effects either. In the following, we describe these effects and speculate about the underlying causes where possible. However, future research that extends beyond the scope of the current experiment would need to replicate the effects and, if replicated, explain their causes and implications.

The Updated Situation Model Did Not Fully Explain the CIE

The first unexpected effect in the early-disclosure groups was the remaining direct effect of retractor credibility on the CIE after statistical control of its indirect effect mediated through the updated situation model. This may indicate that, during the delayed test of the CIE, participants retrieved an independent memory of retractor credibility in addition to the situation model, which affected the agreement with the original information directly. That is, in the early-disclosure groups, the CIE was not exclusively based on the memory representation of the situation but on additional information as well. To explain this, the model-updating account would need to be expanded to include other sources of influence on the CIE in addition to the situation models. It remains unclear why only retractor credibility

had this direct effect on the CIE, whereas the effect of informant credibility was fully mediated through both situation models, as predicted. Although informant and retractor credibility affected the CIE through different paths, their total effects on the CIE were of similar size. To our knowledge, neither the literature on the CIE nor on source credibility offers a theoretical explanation for this disparity between the effects of informant and retractor credibility. Further, it remains unclear why the corresponding effect did not also occur in the late-disclosure groups.

The second unexpected effect in the early-disclosure groups was the remaining direct effect of the original situation model on the CIE after statistical control of its indirect effect mediated through the updated situation model. The more the original information was initially integrated into the original situation model, the stronger was the CIE. The original situation model directly predicted the CIE to similar extent as the updated situation model. This indicates that the update following the retraction did not overwrite the original situation model but created a new version of it, while the original situation model remained accessible. During the delayed test of the CIE, presumably, both situation models were retrieved, considered concurrently, and equally influenced the judgments in the test of the CIE.

Interestingly, this remaining direct effect of the original situation model on the CIE aligns well with the Knowledge Revision Components (KReC) framework (Kendeou & O'Brien, 2014), which broadly describes how new information modifies existing knowledge representations.[6] According to the Encoding and Passive Activation principles, encoded information cannot be deleted from long-term memory and can later be reactivated, consistent with our finding that the original situation model remained accessible and influenced the CIE alongside the updated model in the early-disclosure groups. Furthermore, the Competing Activation Principle posits that old and newly encoded information compete

[6] We thank a reviewer for suggesting the inclusion of the KReC framework in this discussion.

for activation depending on their degree of elaboration. In this light, the joint effect of both situation models on the CIE in the early-disclosure groups suggests similar degrees of elaboration, whereas the exclusive effect of the updated situation model in the late-disclosure groups suggests a greater degree of elaboration of the updated situation model. This could reflect the different amount of information available at situation-model assessment: In the early-disclosure groups, both situation models were based on one critical statement and source-credibility information, whereas in the late-disclosure groups, only the updated situation model could include informant-credibility information. However, this explanation is weakened by that fact that this informant-credibility information was empirically not considered in the updated model. In sum, although our experiment primarily tested the model-updating account, our findings are also broadly consistent with the KreC framework (Kendeou & O'Brien, 2014).

Because the original situation model could not be affected by the retraction, one could presume that its additional effect on the CIE in the early-disclosure groups would lead to a stronger CIE in these groups compared to the late-disclosure groups, in which only the updated situation model affected the CIE. However, the strength of the CIE did not differ between the early-disclosure and late-disclosure groups. Thus, the timing of source-credibility disclosure affected the relative contribution of the original and updated situation models to the CIE, but it did not affect the strength of the CIE itself.

The Timing of Credibility Disclosure Affected the Relationship Between the Situation Models

The third unexpected finding is the disparity in the relationship of the original and the updated situation model that depended on the timing of source-credibility disclosure. In the early-disclosure groups, the original situation model predicted the updated situation model. Thus, the persistence of the original information in the situation model depended on the

strength of its initial integration. By contrast, in the late-disclosure groups, the original situation model did not predict the updated situation model. Thus, the degree to which the original information persisted in the situation model after the retraction did not depend on the strength of its initial integration. Taken together, this disparity in the relationships indicates that early disclosure of source credibility may have helped to structure the updating of the situation model. With late disclosure of source credibility, on the other hand, there may have been more room for individual interpretations regarding (the cause of) the contradiction which made the update less systematic, more variable, and possibly driven more strongly by individual differences.

It should be noted, however, that the generalizability of the three unexpected findings in this experiment and their implications are open to question, as the effects of retractor credibility and of the original situation model only occurred in the early-disclosure groups. Due to the (deliberate) absence of source-credibility effects on any dependent variable in the late-disclosure groups, these groups were arguably more similar to the groups in standard CIE experiments without the manipulation of source credibility (e.g., Ecker et al., 2010; McIlhiney et al., 2022; Rich & Zaragoza, 2016; Susmann & Wegener, 2022; Wilkes & Leatherbarrow, 1988). This may imply that in standard CIE experiments, the effect of the updated situation model on the CIE is primarily relevant and that relationships with the original situation model and other variables may vary depending on the manipulation.

Practical Implications

There are practical implications of the finding that source credibility, a reliable indicator of statement veracity, did not affect participants' interpretations of a situation if credibility was disclosed only after participants elaborated on the situation. It is noteworthy that participants in the late-disclosure groups did not consider source credibility once it became available. In particular, informant credibility was disclosed before participants

processed the retraction and before the assessment of the updated situation model. Thus, regarding the processing of the retraction, participants had the opportunity to interpret the retraction in light of the credibility of the original informant. However, informant credibility did not affect the updated situation model. Likewise, both informant and retractor credibility were disclosed long before the CIE test. However, neither informant nor retractor credibility affected the CIE in the late-disclosure groups. For information consumers, these findings emphasize the importance of researching the source of information before elaborating on the information itself. For information providers, it emphasizes the importance of disclosing a source as early within a report as possible.

An additional finding with important practical implications is that a statement from a source with unknown credibility in the late-disclosure groups had the same effect on participants' interpretations as a statement from a highly credible source in the early-disclosure groups (cf. Sterrett et al., 2019; Ward & McGinnies, 1974). Thus, as long as no contrary evidence is provided, information consumers appear to regard sources as highly credible and to treat their statements accordingly. This emphasizes that information providers should not only identify sources early on, but also provide enough information about the source so that information consumers can assess whether it is a source they want to believe.

Conclusion

The current research significantly advances the theoretical understanding of the CIE by providing experimental evidence that situation models predicted the CIE and mediated the effects of source credibility on the CIE. The model-updating account provides comprehensive explanations for the majority of our results, whereas purely retrieval-based accounts of the CIE are inconsistent with the results. To explain all our results, however, more advanced theories are needed.

Declarations

Funding: The research reported in this article was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project No. 436585458 granted to Marie Luisa Schaper and Ute Johanna Bayen.

Conflicts of interest/Competing interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethics approval: This study has been approved by the Ethics Committee of the Faculty of Mathematics and Natural Sciences of Heinrich Heine University Düsseldorf.

Consent to participate/for publication: All participants provided informed consent to participate in the study and to the publication of their anonymized data.

Availability of data and materials/Code availability: The materials, data and analysis code are provided in the appendices or the Online Supplements in the Open Science Framework at https://osf.io/hpm5y/?view_only=5773b05f290c40d0bb4613094526b92d.

Authors' contributions (CRediT author statement):

Carolin V. Hey: Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Data curation, Writing – Original Draft, Writing – Review & Editing, Visualization, Project Administration

Marie Luisa Schaper: Validation, Writing – Review & Editing, Supervision, Funding Acquisition

Ute J. Bayen: Methodology, Resources, Writing – Review & Editing, Supervision, Funding Acquisition

Acknowledgments

Parts of this research were presented at the 66th Tagung Experimentell arbeitender Psychologen und Psychologinnen (Conference of Experimental Psychologists), March 2024,

Regensburg, Germany. We thank Siegmund Switala for programming the experiment and thank Hannah Tryba, Marie Christine Hetgens, Verena Peick, and Vincent Pablo Votava Socada for rating the situation models.

Open Practices Statement

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study, and the study follows JARS (Appelbaum et al., 2018). All research materials are available in the appendices. All data and analysis code are available at https://osf.io/hpm5y/?view_only=25267025ef0e4cff968094d88f76e68c. Data were analyzed using JASP, version 0.19.2 (JASP Team, 2024), R, version 4.3.1 (R Core Team, 2023), and the package *lavaan*, version 0.6-17 (Rosseel, 2012). The experiment was not preregistered.

References

Albarracín, D., Kumkale, G. T., & Poyner-Del Vento, P. (2017). How people can become persuaded by weak messages presented by credible communicators: Not all sleeper effects are created equal. *Journal of Experimental Social Psychology, 68*, 171–180. https://doi.org/10.1016/j.jesp.2016.06.009

Allen, M., Adamski, L., Bates, M., Bernhagen, M., Callendar, A., Casey, M., Czerwinski, A., Decker, L., Howard, G., Jordan, B., Kujawski, E., May, K., Olson, P., Parenteau, A., Reilly, S., Schmidt, J., Stebnitz, S., Thau, S., Tollefson, M., ... Zirbel, C. (2002). Effect of timing of communicator identification and level of source credibility on attitude. *Communication Research Reports, 19*(1), 46–55. https://doi.org/10.1080/08824090209384831

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *The American Psychologist, 73*(1), 3–25. https://doi.org/10.1037/amp0000191

Aust, F., Diedenhofen, B., Ullrich, S., & Musch, J. (2013). Seriousness checks are useful to improve data validity in online research. *Behavior Research Methods, 45*(2), 527–535. https://doi.org/10.3758/s13428-012-0265-2

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Brydges, C. R., Gignac, G. E., & Ecker, U. K. (2018). Working memory capacity, short-term memory capacity, and the continued influence effect: A latent-variable analysis. *Intelligence, 69*, 117–122. https://doi.org/10.1016/j.intell.2018.03.009

Buczel, K. A., Siwiak, A., Szpitalak, M., & Polczyk, R. (2024). How do forewarnings and post-warnings affect misinformation reliance? The impact of warnings on the continued influence effect and belief regression. *Memory & Cognition, 52*(5), 1048–1064. https://doi.org/10.3758/s13421-024-01520-z

Byrne, B. M. (1994). *Structural equation modeling with EQS and EQS/Windows.* Sage Publications.

Chan, M. S., Jones, C. R., Hall Jamieson, K., & Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological Science, 28*(11), 1531–1546. https://doi.org/10.1177%2F0956797617714579

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

DiStefano, C., Zhu, M., & Mîndrilã, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research, and Evaluation, 14*, Article 20. https://doi.org/10.7275/da8t-4g52

Ecker, U. K., & Antonio, L. M. (2021). Can you believe it? An investigation into the impact of retraction source credibility on the continued influence effect. *Memory & Cognition, 49*(4), 631–644. https://doi.org/10.3758/s13421-020-01129-y

Ecker, U. K., Hogan, J. L., & Lewandowsky, S. (2017). Reminders and repetition of misinformation: Helping or hindering its retraction? *Journal of Applied Research in Memory and Cognition, 6*(2), 185–192. https://doi.org/10.1016/j.jarmac.2017.01.014

Ecker, U. K., Lewandowsky, S., & Apai, J. (2011). Terrorists brought down the plane!—No, actually it was a technical fault: Processing corrections of emotive information. *The Quarterly Journal of Experimental Psychology, 64*(2), 283–310. https://doi.org/10.1080/17470218.2010.497927

Ecker, U. K., Lewandowsky, S., Cook, S., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P., Vraga, E. K., Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology, 1*, 13–29. https://doi.org/10.1038/s44159-021-00006-y

Ecker, U. K., Lewandowsky, S., Swire, B., & Chang, D. (2011). Correcting false information in memory: Manipulating the strength of misinformation encoding and its retraction. *Psychonomic Bulletin & Review, 18*(3), 570–578. https://doi.org/10.3758/s13423-011-0065-1

Ecker, U. K., Lewandowsky, S., & Tang, D. T. (2010). Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Memory & Cognition, 38*(8), 1087–1100. https://doi.org/10.3758/MC.38.8.1087

Ecker, U. K., Prike, T., Paver, A. B., Scott, R. J., & Swire-Thompson, B. (2024). Don't believe them! Reducing misinformation influence through source discreditation. *Cognitive Research: Principles and Implications, 9*, Article 52. https://doi.org/10.1186/s41235-024-00581-7

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fitzgerald, F. S. (1925). *The Great Gatsby.* Charles Scribner's Sons.

Gordon, A., Brooks, J. C., Quadflieg, S., Ecker, U. K., & Lewandowsky, S. (2017). Exploring the neural substrates of misinformation processing. *Neuropsychologia, 106*, 216–224. https://doi.org/10.1016/j.neuropsychologia.2017.10.003

Graesser, A. C., & Clark, L. F. (1985). The generation of knowledge-based inferences during narrative comprehension. *Advances in Psychology, 29*, 53–94. https://doi.org/10.1016/S0166-4115(08)62732-6

Greenberg, B. S., & Miller, G. R. (1966). The effects of low-credible sources on message acceptance. *Speech Monographs, 33*(2), 127–136. https://doi.org/10.1080/03637756609375488

Guillory, J. J., & Geraci, L. (2013). Correcting erroneous inferences in memory: The role of source credibility. *Journal of Applied Research in Memory and Cognition, 2*(4), 201–209. https://doi.org/10.1016/j.jarmac.2013.10.001

Hamby, A., Ecker, U. K., & Brinberg, D. (2020). How stories in memory perpetuate the continued influence of false information. *Journal of Consumer Psychology, 30*(2), 240–259. https://doi.org/10.1002/jcpy.1135

Hayes, A. F. (2015). An index and test of linear moderated mediation. *Multivariate Behavioral Research, 50*(1), 1–22. https://doi.org/10.1080/00273171.2014.962683

Heesacker, M., Petty, R. E., & Cacioppo, J. T. (1983). Field dependence and attitude change: Source credibility can alter persuasion by affecting message-relevant thinking. *Journal of Personality, 51*(4), 653–666. https://doi.org/10.1111/j.1467-6494.1983.tb00872.x

Hey, C. V., Schaper, M. L., & Bayen, U. J. (2025). Relative source credibility affects the continued influence effect: Evidence of rationality in the CIE. *Cognition, 254*, Article 106000. https://doi.org/10.1016/j.cognition.2024.106000

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Ithisuphalap, J., Rich, P. R., & Zaragoza, M. S. (2020). Does evaluating belief prior to its retraction influence the efficacy of later corrections? *Memory, 28*(5), 617–631. https://doi.org/10.1080/09658211.2020.1752731

JASP Team. (2024). *JASP (Version 0.19.2) [Computer Software]*. https://jasp-stats.org/

Johnson, M. J., & Seifert, C. M. (1994). Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(6), 1420–1436. https://doi.org/10.1037/0278-7393.20.6.1420

Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science, 4*(1), 71–115. https://doi.org/10.1016/S0364-0213(81)80005-5

Kan, I. P., Pizzonia, K. L., Drummey, A. B., & Mikkelsen, E. J. (2021). Exploring factors that mitigate the continued influence of misinformation. *Cognitive Research: Principles and Implications, 6*(1), Article 76. https://doi.org/10.1186/s41235-021-00335-9

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). The MIT Press. https://doi.org/10.7551/mitpress/9737.001.0001

Kumkale, G. T., & Albarracín, D. (2004). The sleeper effect in persuasion: A meta-analytic review. *Psychological Bulletin, 130*(1), 143–172. https://doi.org/10.1037/0033-2909.130.1.143

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. https://doi.org/10.2307/2529310

Lariscy, R. A., & Tinkham, S. F. (1999). The sleeper effect and negative political advertising. *Journal of Advertising, 28*(4), 13–30. https://doi.org/10.1080/00913367.1999.10673593

Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012).

Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest, 13*(3), 106–131. https://doi.org/10.1177%2F1529100612451018

McIlhiney, P., Gignac, G. E., Ecker, U. K., Kennedy, B. L., & Weinborn, M. (2023). Executive function and the continued influence of misinformation: A latent-variable analysis. *PLoS One, 18*(4), e0283951. https://doi.org/10.1371/journal.pone.0283951

McIlhiney, P., Gignac, G. E., Weinborn, M., & Ecker, U. K. (2022). Sensitivity to misinformation retractions in the continued influence paradigm: Evidence for stability. *Quarterly Journal of Experimental Psychology, 75*(7), 1259–1271. https://doi.org/10.1177/17470218211048986

Miller, A. L., Wissmann, K. T., & Peterson, D. J. (2022). The continued influence effect: Examining how age, retraction, and delay impact inferential reasoning. *Applied Cognitive Psychology, 36*(3), 708–723. https://doi.org/10.1002/acp.3939

Mills, J., & Harvey, J. (1972). Opinion change as a function of when information about the communicator is received and whether he is attractive or expert. *Journal of Personality and Social Psychology, 21*(1), 52–55. https://doi.org/10.1037/h0031939

Nadarevic, L., Reber, R., Helmecke, A. J., & Köse, D. (2020). Perceived truth of statements and simulated social media postings: an experimental investigation of source credibility, repeated exposure, and presentation format. *Cognitive Research, 5*(1), Article 56. https://doi.org/10.1186/s41235-020-00251-4

O'keefe, D. J. (1987). The persuasive effects of delaying identification of high-and low-credibility communicators: A meta-analytic review. *Central States Speech Journal, 38*(2), 63–72. http://doi.org/10.1080/10510978709368231

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology, 19*, 123–205.

https://doi.org/10.1016/S0065-2601(08)60214-2

Pluviano, S., Sala, S. D., & Watt, C. (2020). The effects of source expertise and trustworthiness on recollection: the case of vaccine misinformation. *Cognitive Processing, 21*(3), 321–330. https://doi.org/10.1007/s10339-020-00974-8

Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology, 34*(2), 243–281. https://doi.org/10.1111/j.1559-1816.2004.tb02547.x

Pratkanis, A. R., Greenwald, A. G., Leippe, M. R., & Baumgardner, M. H. (1988). In search of reliable persuation effects: III. The sleeper effect is dead. Long live the sleeper effect. *Journal of Personality and Social Psychology, 54*(2), 203–218. https://doi.org/10.1037/0022-3514.54.2.203

Priester, J., Wegener, D., Petty, R., & Fabrigar, L. (1999). Examining the psychological process underlying the sleeper effect: The elaboration likelihood model explanation. *Media Psychology, 1*(1), 27–48. https://doi.org/10.1207/s1532785xmep0101_3

R Core Team. (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology, 25*, 111–163. https://doi.org/10.2307/271063

Rich, P. R., Donovan, A. M., & Rapp, D. N. (2023). Cause typicality and the continued influence effect. *Journal of Experimental Psychology: Applied, 29*(2), 221–238. https://doi.org/10.1037/xap0000454

Rich, P. R., & Zaragoza, M. S. (2016). The continued influence of implied and explicitly stated misinformation in news reports. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(1), 62–74. https://doi.org/10.1037/xlm0000155

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of*

Statistical Software, 48(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Sanderson, J. A., & Ecker, U. K. (2020). The challenge of misinformation and ways to reduce its impact. In P. Van Meter, A. List, D. Lombardi, & P. Kendeou (Eds.), *Handbook of Learning from Multiple Representations and Perspectives* (pp. 461–476). Routledge. https://doi.org/10.4324/9780429443961

Sanderson, J. A., Gignac, G. E., & Ecker, U. K. (2021). Working memory capacity, removal efficiency and event specific memory as predictors of misinformation reliance. *Journal of Cognitive Psychology, 33*(5), 518–532. https://doi.org/10.1080/20445911.2021.1931243

Schumacker, R. E., & Komax, R. G. (2004). *A beginner's guide to structural equation modeling* (2nd ed.). Lawrence Erlbaum Associates.

Skagerberg, E. M., & Wright, D. B. (2009). Susceptibility to postidentification feedback is affected by source credibility. *Applied Cognitive Psychology, 23*(4), 506–523. https://doi.org/10.1002/acp.1470

Sternthal, B., Dholakia, R., & Leavitt, C. (1978). The persuasive effect of source credibility: Tests of cognitive response. *Journal of Consumer Research, 4*, 252–260. https://doi.org/10.1086/208704

Sternthal, B., Phillips, L. W., & Dholakia, R. (1978). The persuasive effect of source credibility: A situational analysis. *Public Opinion Quarterly, 42*(3), 285–314. https://doi.org/10.1086/268454

Sterrett, D., Malato, D., Benz, J., Kantor, L., Tompson, T., Rosenstiel, T., Sonderman, J., & Loker, K. (2019). Who shared it?: Deciding what news to trust on social media. *Digital Journalism, 7*(6), 783–801. https://doi.org/10.1080/21670811.2019.1623702

Susmann, M. W., & Wegener, D. T. (2022). The role of discomfort in the continued influence effect of misinformation. *Memory & Cognition, 50*(2), 435–448.

https://doi.org/10.3758/s13421-021-01232-8

Susmann, M. W., & Wegener, D. T. (2023). The independent effects of source expertise and trustworthiness on retraction believability: The moderating role of vested interest. *Memory & Cognition, 51*(4), 845–861. https://doi.org/10.3758/s13421-022-01374-3

Swire, B., Ecker, U. K., & Lewandowsky, S. (2017). The role of familiarity in correcting inaccurate information. *Journal of Experimental Psychology: Learning, Memory and Cognition, 43*(12), 1948–1961. https://doi.org/10.1037/xlm0000422

Swire-Thompson, B., Cook, J., Butler, L. H., Sanderson, J. A., Lewandowsky, S., & Ecker, U. K. (2021). Correction format has a limited role when debunking misinformation. *Cognitive Research: Principles and Implications, 6*, Article 83. https://doi.org/10.1186/s41235-021-00346-6

Tay, L. Q., Hurlstone, M. J., Kurz, T., & Ecker, U. K. (2022). A comparison of prebunking and debunking interventions for implied versus explicit misinformation. *British Journal of Psychology, 113*(3), 591–607. https://doi.org/10.1111/bjop.12551

Tobin, S. J., & Raymundo, M. M. (2009). Persuasion by causal arguments: The motivating role of perceived causal expertise. *Social Cognition, 27*(1), 105–127. https://doi.org/10.1521/soco.2009.27.1.105

Trauzettel-Klosinski, S., Dietz, K., & IReST-Study-Group. (2012). Standardized assessment of reading performance: The new international reading speed texts IReST. *Investigative Ophthalmology & Visual Science, 53*(9), 5452–5461. https://doi.org/10.1167/iovs.11-8284

Van Boekel, M., Lassonde, K. A., O'Brien, E. J., & Kendeou, P. (2017). Source credibility and the processing of refutation texts. *Memory & Cognition, 45*, 168–181. https://doi.org/10.3758/s13421-016-0649-0.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of Discourse Comprehension.* Academic

Press, Inc.

Walter, N., & Tukachinsky, R. (2020). A meta-analytic examination of the continued influence of misinformation in the face of correction: How powerful is it, why does it happen, and how to stop it? *Communication Research, 47*(2), 155–177. https://doi.org/10.1177%2F0093650219854600

Ward, C. D., & McGinnies, E. (1974). Persuasive effects of early and late mention of credible and noncredible sources. *The Journal of Psychology, 86*(1), 17–23. https://doi.org/10.1080/00223980.1974.9923879

Westbrook, V., Wegener, D. T., & Susmann, M. W. (2023). Mechanisms in continued infuence: The impact of misinformation corrections on source perceptions. *Memory & Cognition*, *51*(6), 1317–1330. https://doi.org/10.3758/s13421-023-01402-w

Wilkes, A. L., & Leatherbarrow, M. (1988). Editing episodic memory following the identification of error. *The Quarterly Journal of Experimental Psychology, 40A*(2), 361–387. https://doi.org/10.1080/02724988843000168

Wilson, E. J., & Sherrell, D. L. (1993). Source effects in communication and persuasion research: A meta-analysis of effect size. *Journal of the Academy of Marketing Science, 21*(2), 101–112. https://doi.org/10.1007/BF02894421

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*(2), 162–185. https://doi.org/10.1037/0033-2909.123.2.162

Table 1

Bivariate Correlations of the Original and the Updated Situation Models, and the CIE

Group	Variables	*r*	BF_{10}
All [a]	Original Situation Model ~ Updated Situation Model	**.28**	**> 10^6**
	Original Situation Model ~ CIE	**.33**	**> 10^9**
	Updated Situation Model ~ CIE	**.57**	**> 10^{36}**
Early-Disclosure [b]	Original Situation Model ~ Updated Situation Model	**.43**	**> 10^8**
	Original Situation Model ~ CIE	**.50**	**> 10^{12}**
	Updated Situation Model ~ CIE	**.56**	**> 10^{16}**
Late-Disclosure [c]	Original Situation Model ~ Updated Situation Model	.11	0.29 [d]
	Original Situation Model ~ CIE	.12	0.35 [e]
	Updated Situation Model ~ CIE	**.60**	**> 10^{19}**

Note. [a] $df = 433$, [b] $df = 216$, [c] $df = 215$, [d] corresponds to $BF_{01} = 3.39$, [e] corresponds to $BF_{01} = 2.83$. Significant effects are printed in bold ($p < .001$).

Table 2

Moderation Analyses of the Original and the Updated Situation Models, and the CIE

Dependent Variable	Group	Effect	F	p	η^2_p	BF_{incl}	BF_{excl}
Original Situation Model	All	Informant credibility	**44.61**	**< .001**	**.09**	**> 10^6**	**< 10^{-6}**
		Retractor credibility	0.07	.798	< .01	0.11	9.08
		Credibility-disclosure timing	**47.26**	**< .001**	**.10**	**> 10^7**	**< 10^{-7}**
		Informant credibility × retractor credibility	0.85	.356	< .01	0.23	4.44
		Informant credibility × credibility-disclosure timing	**68.62**	**< .001**	**.14**	**> 10^{12}**	**< 10^{-12}**
		Retractor credibility × credibility-disclosure timing	0.03	.853	< .01	0.15	6.53
		Informant credibility × retractor credibility × credibility-disclosure timing	0.43	.514	< .01	0.21	4.80
	Early-disclosure	Informant credibility	**113.34**	**< .001**	**.35**	**> 10^{18}**	**< 10^{-18}**
		Retractor credibility	0.10	.754	< .01	0.16	6.28
		Informant credibility × retractor credibility	0.04	.848	< .01	0.19	5.26
	Late-disclosure	Informant credibility	1.27	.260	.01	0.27	3.75
		Retractor credibility	< 0.01	.960	< .01	0.15	6.70
		Informant credibility × retractor credibility	1.23	.269	.01	0.32	3.08
Updated Situation Model	All	Informant credibility	**14.67**	**< .001**	**.03**	**112.62**	**0.01**
		Retractor credibility	**9.76**	**.002**	**.02**	**11.64**	**0.09**
		Credibility-disclosure timing	0.42	.517	< .01	0.13	7.90
		Informant credibility × retractor credibility	0.97	.325	< .01	0.21	4.68
		Informant credibility × credibility-disclosure timing	**8.70**	**.003**	**.02**	**9.04**	**0.11**
		Retractor credibility × credibility-disclosure timing	**4.53**	**.034**	**.01**	**1.15**	**0.87**
		Informant credibility × retractor credibility × credibility-disclosure timing	0.83	.363	< .01	0.26	3.78
	Early-disclosure	Informant credibility	**21.87**	**< .001**	**.09**	**2987.45**	**< 10^{-3}**
		Retractor credibility	**13.13**	**< .001**	**.06**	**55.19**	**0.02**
		Informant credibility × retractor credibility	1.71	.192	.01	0.43	2.33
	Late-disclosure	Informant credibility	0.41	.524	< .01	0.18	5.55
		Retractor credibility	0.52	.471	< .01	0.19	5.26
		Informant credibility × retractor credibility	< 0.01	.957	< .01	0.19	5.35

CIE	All	Informant credibility	**19.62**	**< .001**	**.04**	**1076.28**	**$< 10^{-3}$**
		Retractor credibility	**19.15**	**< .001**	**.04**	**1044.21**	**$< 10^{-3}$**
		Credibility-disclosure timing	0.12	.726	< .01	0.12	8.69
		Informant credibility × retractor credibility	0.83	.362	< .01	0.20	4.97
		Informant credibility × credibility-disclosure timing	**14.22**	**< .001**	**.03**	**113.48**	**0.01**
		Retractor credibility × credibility-disclosure timing	**10.95**	**.001**	**.02**	**25.69**	**0.04**
		Informant credibility × retractor credibility × credibility-disclosure timing	0.18	.668	< .01	0.20	5.04
	Early-disclosure	Informant credibility	**33.69**	**< .001**	**.14**	**$> 10^{5}$**	**$< 10^{-5}$**
		Retractor credibility	**29.58**	**< .001**	**.12**	**$> 10^{4}$**	**$< 10^{-4}$**
		Informant credibility × retractor credibility	0.90	.343	< .01	0.31	3.28
	Late-disclosure	Informant credibility	0.22	.643	< .01	0.16	6.09
		Retractor credibility	0.57	.452	< .01	0.19	5.15
		Informant credibility × retractor credibility	0.12	.733	< .01	0.20	5.08

Note. Test statistics of the 2 × 2 × 2 ANOVAs with all groups and the 2 × 2 ANOVAs in the late-disclosure and early-disclosure groups for the original situation model, updated situation model, and the CIE. $df_{effect} = 1$ for all analyses. $df_{error} = 427$ for the 2 × 2 × 2 ANOVAs with all groups and $df_{error} = 213$ and 214 for the 2 × 2 ANOVAs in the late-disclosure and early-disclosure groups, respectively. Significant effects are printed in bold.

Table 3

Regression Weights of the Structural Equation Model of Moderated Serial Mediation with All Groups

Direct Effect	*B* [95% CI]			β [95% CI]		
$I \rightarrow O$	**0.15**	**[0.11,**	**0.19]**	**0.28**	**[0.20,**	**0.35]**
$I \rightarrow U$	0.04	[–0.02,	0.10]	0.06	[–0.04,	0.16]
$I \rightarrow C$	0.03	[–0.03,	0.10]	0.04	[–0.04,	0.12]
$R \rightarrow U$	**–0.08**	**[–0.14,**	**–0.03]**	**–0.14**	**[–0.23,**	**–0.05]**
$R \rightarrow C$	**–0.11**	**[–0.17,**	**–0.05]**	**–0.14**	**[–0.21,**	**–0.06]**
$T \rightarrow O$	**–0.15**	**[–0.20,**	**–0.11]**	**–0.28**	**[–0.36,**	**–0.21]**
$T \rightarrow U$	**0.06**	**[0.01,**	**0.12]**	**0.10**	**[0.01,**	**0.19]**
$T \rightarrow C$	0.01	[–0.04,	0.07]	0.02	[–0.06,	0.09]
$O \rightarrow U$	**0.28**	**[0.17,**	**0.39]**	**0.26**	**[0.16,**	**0.36]**
$O \rightarrow C$	**0.25**	**[0.13,**	**0.38]**	**0.18**	**[0.09,**	**0.26]**
$U \rightarrow C$	**0.62**	**[0.50,**	**0.72]**	**0.47**	**[0.39,**	**0.55]**
$M_{I \rightarrow O}$	**0.18**	**[0.14,**	**0.23]**	**0.34**	**[0.26,**	**0.42]**
$M_{I \rightarrow U}$	0.01	[–0.06,	0.07]	0.01	[–0.09,	0.12]
$M_{I \rightarrow C}$	0.03	[–0.04,	0.09]	0.03	[–0.05,	0.12]
$M_{R \rightarrow U}$	**–0.06**	**[–0.11,**	**–0.004]**	**–0.09**	**[–0.18,**	**–0.01]**
$M_{R \rightarrow C}$	**–0.09**	**[–0.15,**	**–0.03]**	**–0.12**	**[–0.19,**	**–0.04]**
$M_{O \rightarrow U}$	**0.15**	**[0.03,**	**0.27]**	**0.13**	**[0.03,**	**0.23]**
$M_{O \rightarrow C}$	**0.17**	**[0.05,**	**0.30]**	**0.12**	**[0.03,**	**0.20]**
$M_{U \rightarrow C}$	**–0.16**	**[–0.27,**	**–0.05]**	**–0.12**	**[–0.20,**	**–0.04]**
Total Effect						
$\Sigma M_{I \rightarrow O}$	**0.18**	**[0.14,**	**0.23]**	**0.34**	**[0.26**	**0.42]**
$\Sigma M_{I \rightarrow C}$	**0.11**	**[0.04,**	**0.18]**	**0.14**	**[0.05,**	**0.23]**
$\Sigma M_{R \rightarrow U}$	**–0.06**	**[–0.11,**	**–0.004]**	**–0.09**	**[–0.18,**	**–0.01]**
$\Sigma M_{R \rightarrow C}$	**–0.13**	**[–0.19,**	**–0.06]**	**–0.16**	**[–0.25,**	**–0.08]**
$\Sigma T \rightarrow C$	0.03	[–0.04,	0.09]	0.03	[–0.05,	0.12]

Note. *I* = informant credibility; *O* = original situation model; *U* = updated situation model; *C* = CIE; *R* = retractor credibility; *T* = credibility-disclosure timing; *M* = moderating effects of credibility-disclosure timing. Unstandardized and standardized regression weights were estimated with the maximum-likelihood procedure implemented in the R package *lavaan* (Rosseel, 2012); 95% confidence intervals were estimated with a bootstrapping procedure using 10,000 samples, as recommended by Hayes (2015). Confidence intervals for the unstandardized regression weights were bias-corrected. A regression weight is significant if the 95% confidence interval reported in brackets does not include zero. Significant regression weights are printed in bold. Total effects of a predictor or moderator are the sum of all their direct and indirect effects on a variable (labeled Σ).

Table 4

Standardized Regression Weights of the Structural Equation Models in the Early-Disclosure and Late-Disclosure Groups

Direct Effect	Early-Disclosure Groups				Late-Disclosure Groups			
	B [95% CI]		β [95% CI]		*B* [95% CI]		β [95% CI]	
$I \rightarrow O$	**0.33**	**[0.27, 0.39]**	**0.59**	**[0.50, 0.68]**	–0.04	[–0.10, 0.03]	–0.08	[–0.21, 0.06]
$I \rightarrow U$	0.04	[–0.06, 0.14]	0.07	[–0.09, 0.22]	0.03	[–0.04, 0.10]	0.05	[–0.08, 0.11]
$I \rightarrow C$	0.06	[–0.04, 0.16]	0.07	[–0.05, 0.19]	0.01	[–0.07, 0.09]	0.01	[–0.10, 0.12]
$R \rightarrow U$	**–0.14**	**[–0.21, –0.06]**	**–0.22**	**[–0.34, –0.10]**	–0.03	[–0.10, 0.05]	–0.05	[–0.19, 0.09]
$R \rightarrow C$	**–0.20**	**[–0.29, –0.11]**	**–0.24**	**[–0.35, –0.14]**	–0.02	[–0.10, 0.06]	–0.02	[–0.13, 0.09]
$O \rightarrow U$	**0.43**	**[0.26, 0.62]**	**0.39**	**[0.25, 0.53]**	0.13	[–0.02, 0.28]	0.11	[–0.01, 0.24]
$O \rightarrow C$	**0.43**	**[0.25, 0.61]**	**0.30**	**[0.17, 0.42]**	0.08	[–0.09, 0.25]	0.05	[–0.06, 0.16]
$U \rightarrow C$	**0.46**	**[0.30, 0.61]**	**0.35**	**[0.23, 0.47]**	**0.77**	**[0.63, 0.92]**	**0.59**	**[0.50, 0.68]**
Indirect Effect								
$I \rightarrow O \rightarrow U$	**0.14**	**[0.08, 0.22]**	**0.23**	**[0.13, 0.33]**	0.00	[–0.02, 0.002]	0.00	[–0.01, 0.01]
$I \rightarrow O \rightarrow C$	**0.14**	**[0.08, 0.22]**	**0.17**	**[0.09, 0.25]**	0.00	[–0.02, 0.002]	0.00	[–0.02, 0.01]
$I \rightarrow O \rightarrow U \rightarrow C$	**0.07**	**[0.04, 0.11]**	**0.08**	**[0.04, 0.12]**	0.00	[–0.02, 0.001]	–0.01	[–0.03, 0.01]
$R \rightarrow U \rightarrow C$	**–0.06**	**[–0.11, –0.03]**	**–0.08**	**[–0.12, –0.03]**	–0.02	[–0.08, 0.03]	–0.03	[–0.11, 0.05]
Total Effect								
$\Sigma I \rightarrow C$	**0.29**	**[0.19, 0.38]**	**0.35**	**[0.24, 0.45]**	0.02	[–0.07, 0.12]	0.03	[–0.10, 0.16]
$\Sigma R \rightarrow C$	**–0.26**	**[–0.35, –0.18]**	**–0.32**	**[–0.43, –0.21]**	–0.04	[–0.13, 0.06]	–0.05	[–0.18, 0.08]

Note. *I* = informant credibility; *O* = original situation model; *U* = updated situation model; *C* = CIE; *R* = retractor credibility. Standardized regression weights were estimated with the maximum-likelihood procedure implemented in the R package *lavaan* (Rosseel, 2012); 95% confidence intervals were estimated with a bootstrapping procedure using 10,000 samples, as recommended by Hayes (2015). A regression weight is significant if the 95% confidence interval reported in brackets does not include zero. Significant regression weights are printed in bold. Indirect effects of a predictor or mediator are the sequence of several direct effects through one or more mediators and are calculated as the product of these direct effects. Total effects of a predictor are the sum of all their direct and indirect effects on a variable (labeled Σ).

Figure 1

Procedure

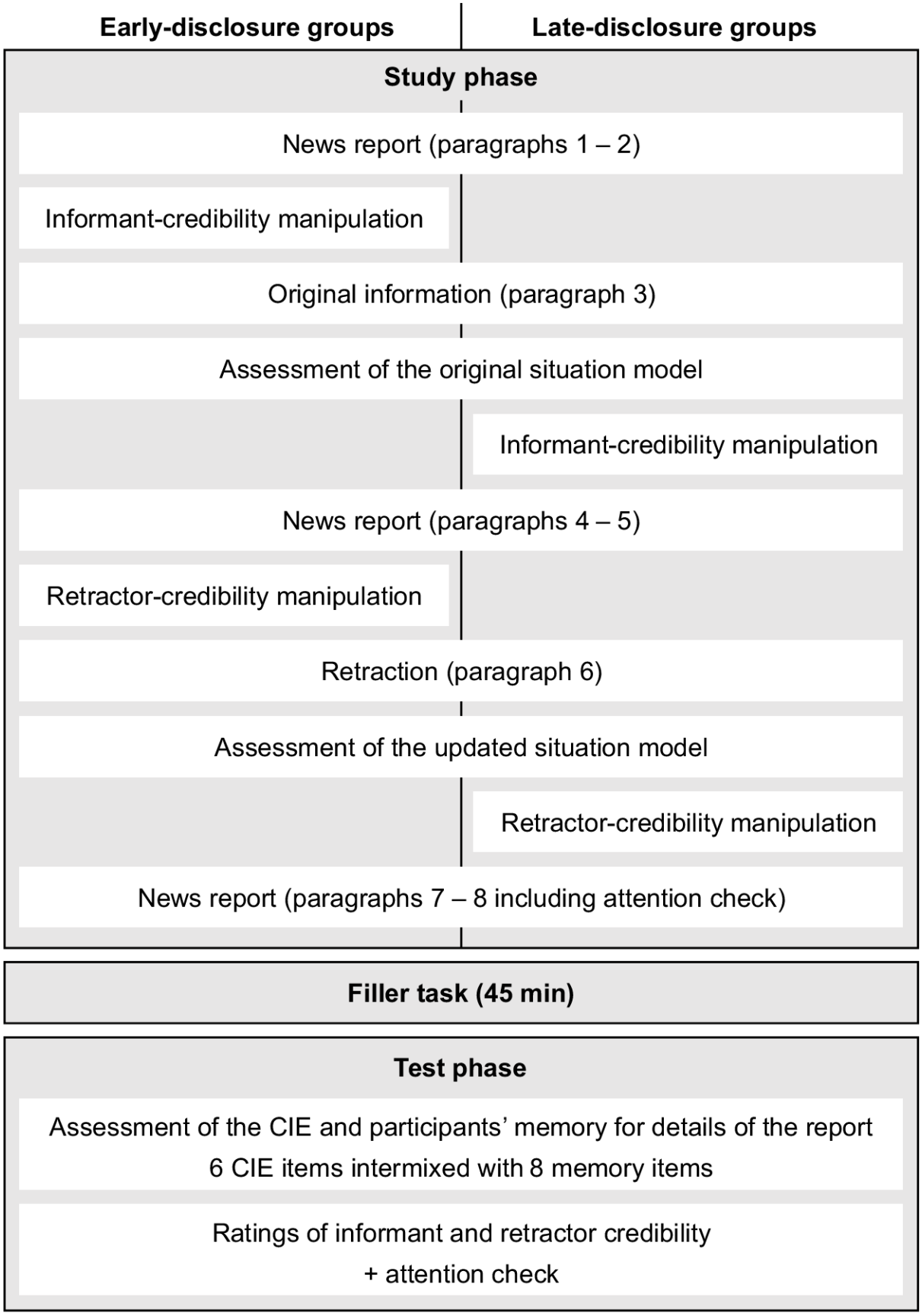
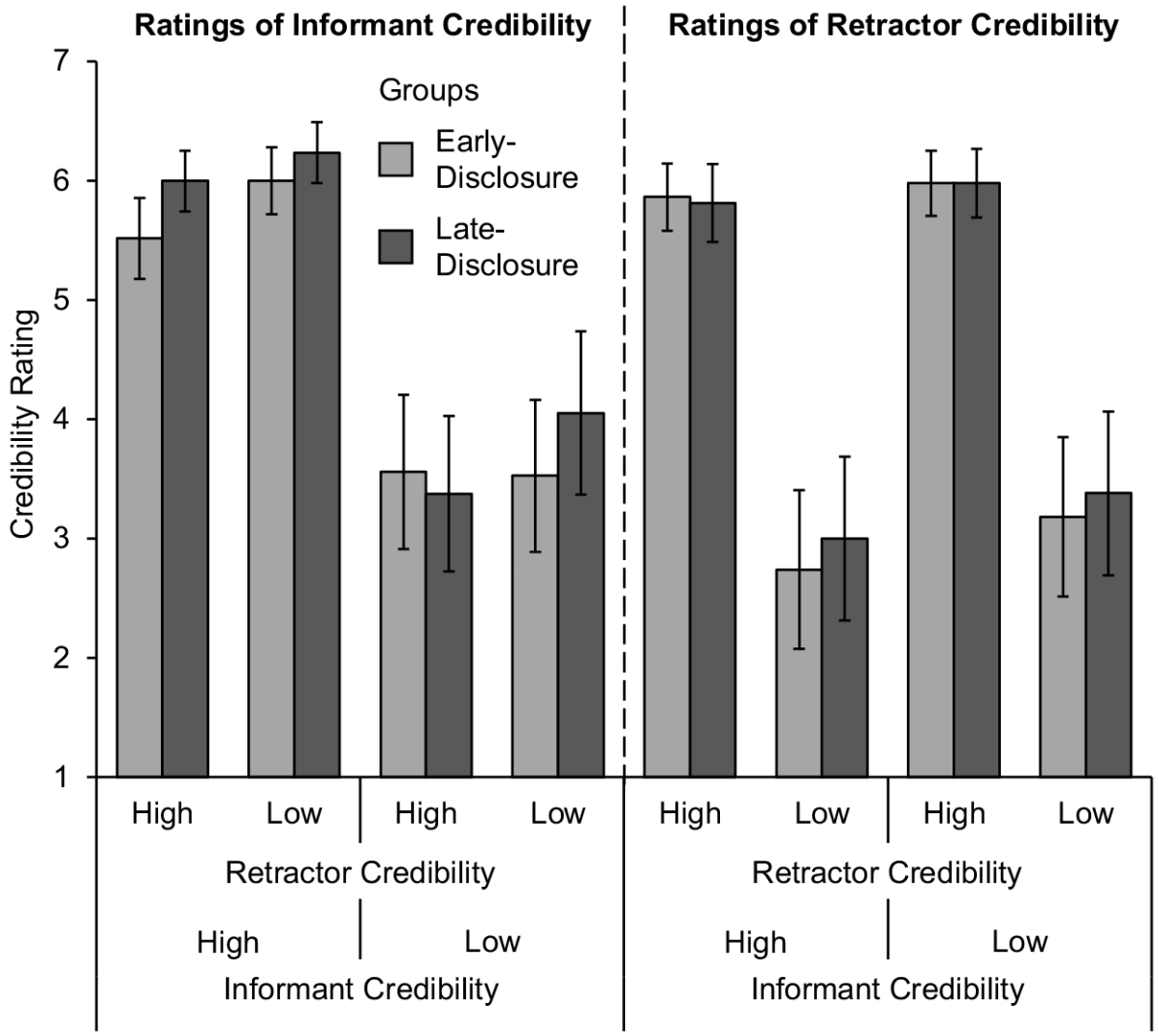


Figure 2

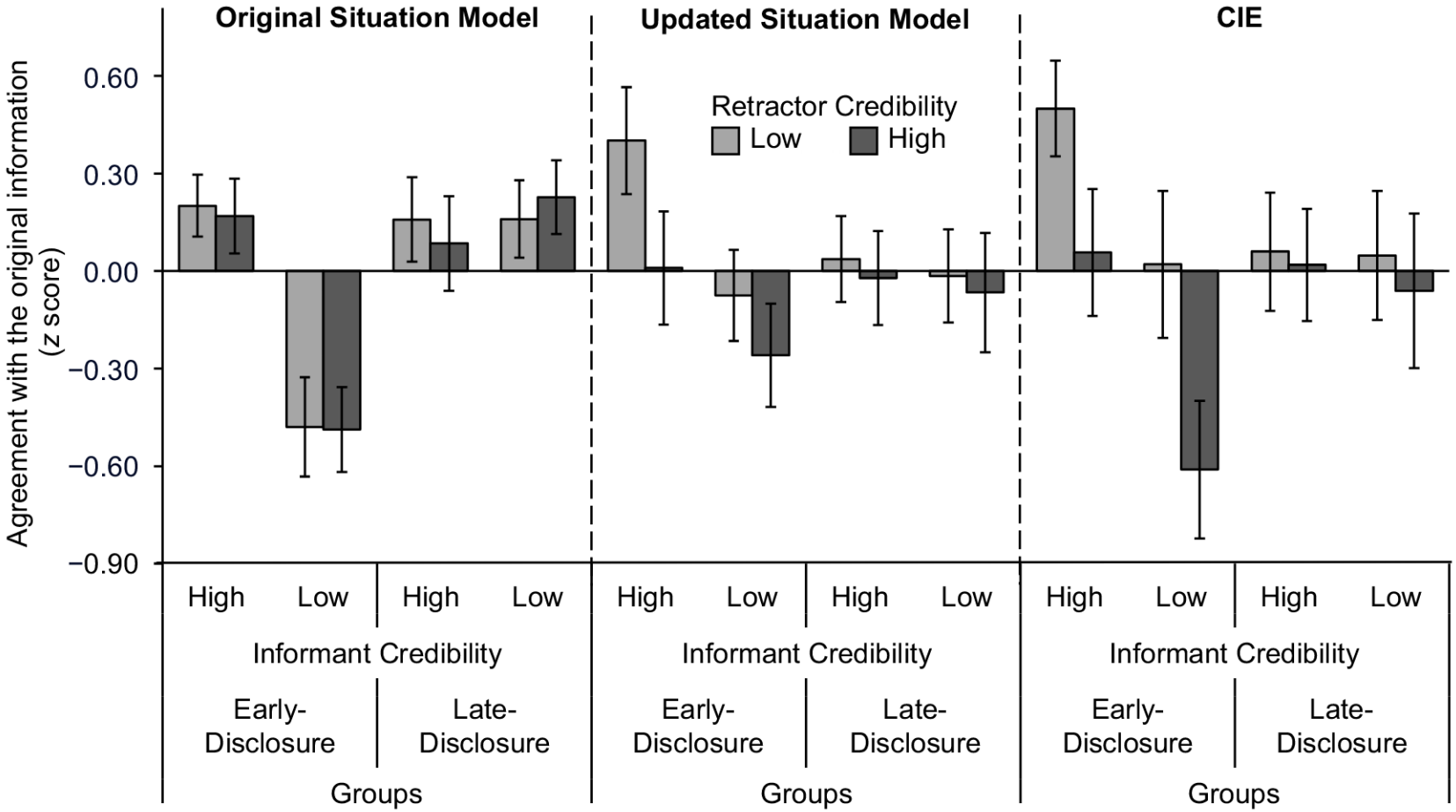
Source-Credibility Ratings



Note. Credibility ratings of the informant and retractor. Error bars indicate 95% confidence intervals.

Figure 3

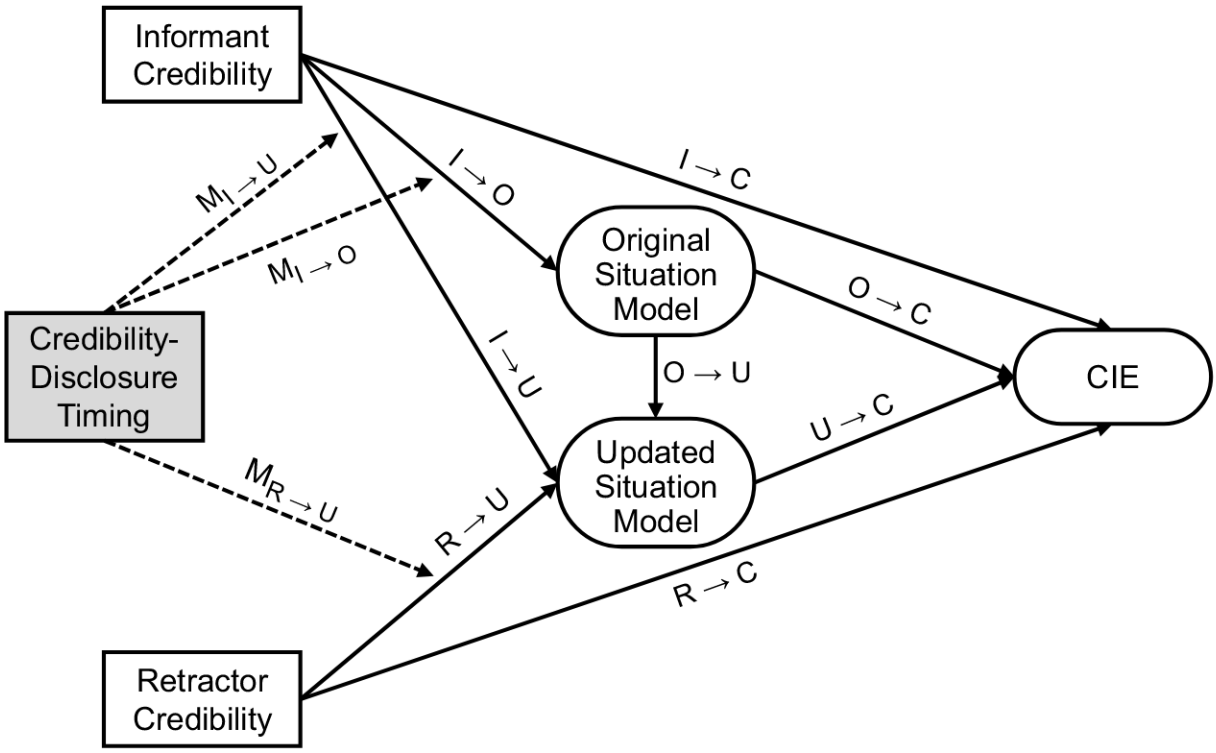
Effects of Informant and Retractor Credibility, and Credibility-Disclosure Timing on the Original and Updated Situation Models, and the CIE



Note. Mean agreement with the original information (in *z* scores) for the original situation model, the updated situation model, and the CIE. Higher scores indicate greater agreement. Error bars indicate 95% confidence intervals.

Figure 4

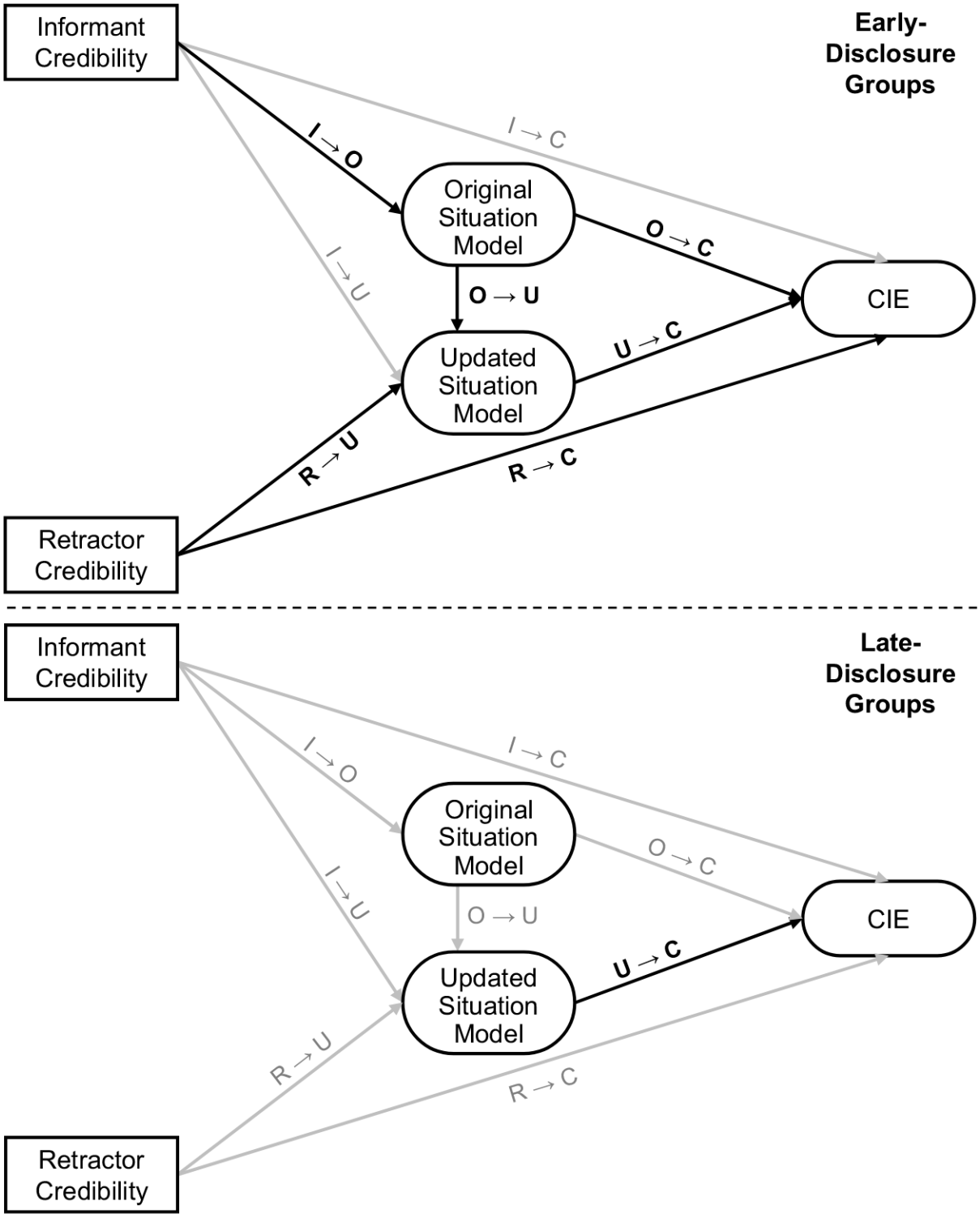
Conceptual Structural Equation Model of Moderated Serial Mediation



Note. Conceptual structural equation model corresponding to the predictions of the model-updating account for the current experiment: Informant credibility and retractor credibility may affect the CIE mediated through the original and the updated situation models. Credibility-disclosure timing moderates the effects of informant and retractor credibility on the situation models. The white rectangles represent the manipulated predictor variables. The grey rectangle represents the manipulated moderator variable. The rounded rectangles represent the measured variables with the original and updated situation models as serial mediators and the CIE as the outcome variable. Solid arrows represent direct effects, dashed arrows represent moderating effects.

Figure 5

Structural Equation Models of Serial Mediation in the Early-Disclosure and Late-Disclosure Groups



Note. Structural equation models of effects of informant credibility and retractor credibility on the CIE, mediated through the original and updated situation models in the early-disclosure groups (n = 218) and late-disclosure groups (n = 217). The rectangles represent the manipulated predictor variables. The rounded rectangles represent the measured variables with the original and updated situation models as serial mediators and the CIE as the outcome variable. The arrows represent direct effects. Regression weights were estimated with the maximum-likelihood procedure implemented in the R package *lavaan* (Rosseel, 2012), 95% bias-corrected confidence intervals were estimated with a bootstrapping procedure using 10,000 samples, as recommended by Hayes (2015). A regression weight is significant if the 95% confidence interval reported in brackets does not include zero, the corresponding path is shown in black. Paths corresponding to non-significant regression weights are shown in gray.

Appendix A

Mean Memory for Details of the Report

Credibility-disclosure timing	Informant credibility	Retractor credibility	*n*	*M* [95% CI]
Early disclosure	High	High	52	5.89 [5.74, 6.05]
		Low	54	5.76 [5.62, 5.90]
	Low	High	57	5.86 [5.72, 5.99]
		Low	55	5.62 [5.45, 5.79]
Late disclosure	High	High	54	5.74 [5.55, 5.92]
		Low	55	5.81 [5.67, 5.96]
	Low	High	53	5.77 [5.59, 5.94]
		Low	55	5.67 [5.50, 5.84]

Note. Memory for details of the report was defined as mean rated agreement with the four correct and four incorrect memory statements in the test questionnaire on a Likert scale from 1 (*strong disagreement*) to 7 (*strong agreement*). Ratings of incorrect statements were reverse-coded.

Appendix B

Questions of the Situation-Model Assessments

Probed aspect		Question
General reason	1	Why did the Freemont water department shut down the water supply from the Denroy River?
Specific reason	2	Why did so many fish die in the Denroy River?
Residents' reaction	3	How did the Shelby residents react to the shutdown of the water department?
	4	Why did the residents react that way?
Potential health risk	5	How likely will people's health be harmed if they drink water from the Denroy River? [a]
	6	Why is the health risk the way you just rated it?
Potential legal consequences	7	How likely is it that a responsible party will be legally charged? [a]
	8	Why is it as likely as you just indicated that a responsible party will be legally charged?
Sequel	9	Considering all you know so far, what will happen next?

Note. Questions asked during the situation-model assessment. The following instructions was displayed above each question: "Now, we are interested in your thoughts and opinions about the described situation in the Shelby region. Please take a moment to reflect on the information you have read. Please answer the following questions according to your subjective understanding of the situation. What do you think, ..." Questions relating to the same aspect were always presented together. The order of aspects was randomized for each participant.

[a] These questions were answered on a rating scale from 0% to 100%.

Online Supplement 2

to "The Continued Influence Effect and Situation Models: Evidence for the Model-Updating Account", Authors: Carolin V. Hey, Marie Luisa Schaper, and Ute Bayen

Validation Experiments

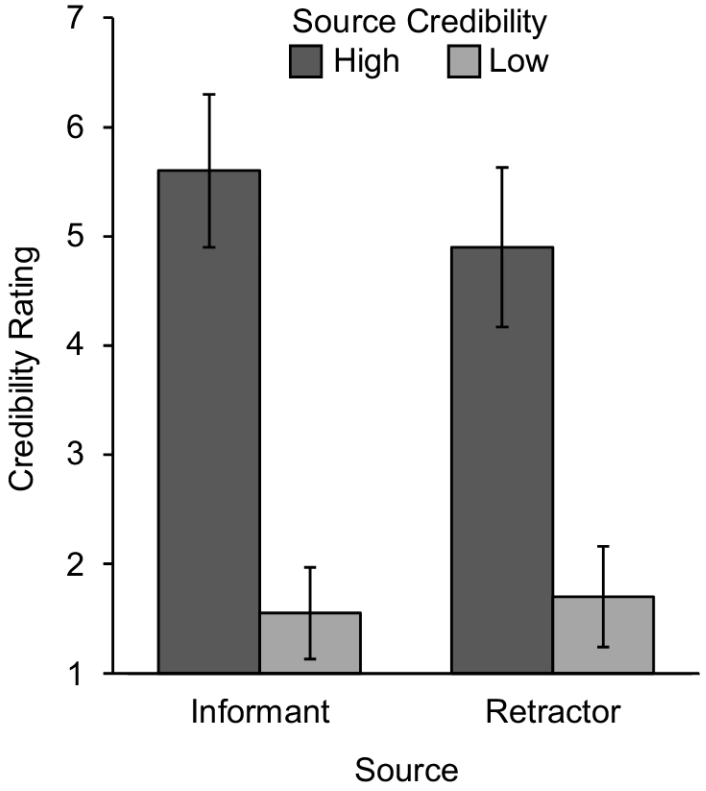
We aimed to validate the source descriptions that we translated from German to English based on Hey et al. (2025) and the newly developed items for the assessment of the situation model and the CIE. To these ends, we conducted two validation experiments with participants recruited on Prolific (www.prolific.com). Inclusion criteria were the same as those for the main experiment. Participants of the validation experiments were excluded from participation in the main experiment.

Validation Experiment 1: The Source-Credibility Manipulations

Validation Experiment 1 served to ensure that the English source descriptions would produce the anticipated differences in credibility ratings. On December 13, 2021, we recruited 20 participants with ages ranging between 20 and 32 years (M = 26.80, SD = 3.83, 18 self-reported as female, 2 as male) who received £1 as compensation. They rated the credibility of all four source descriptions without encountering them in the context of a news report on a Likert scale ranging from 1 (*absolutely not credible*) to 7 (*absolutely credible*). The order of the four credibility ratings was counterbalanced in a Latin square design. Verbatim source descriptions and the source-credibility ratings are in Figure S2.1. As intended, high-credibility sources received higher ratings than low-credibility sources, $F(1, 19) = 98.03$, $p < .001$, $\eta^2_p = .84$, and there was no difference between ratings for informants and retractors, $F(1, 19) = 1.24$, $p = .280$, $\eta^2_p = .06$, and no interaction, $F(1, 19) = 2.63$, $p = .122$, $\eta^2_p = .12$. The credibility of the sources was thus validated.

Figure S1.1

Source Descriptions and Source Credibility Ratings in Validation Experiment 1



Source	Source Credibility	
	High	Low
Informant	An investigative journalist. He has recently exposed the wanton marine pollution caused by a Norwegian oil rig. As a result, he was awarded a major reporter's prize.	A self-proclaimed scandal blogger. He has recently made groundless allegations against a Norwegian oil rig on his website. As a result, he was flagged for spreading misinformation.
Retractor	A journalist and fact checker. He was recently awarded a prestigious journalism prize. He is known for exposing numerous cases of fabricated reporting, journalistic misconduct, and misinformation.	A fitness blogger from the website "Melt fat fast", which targets young women. He is known for collaborations with shady miracle pill manufacturers and devious marketing tactics.

Note. Error bars indicate 95% confidence intervals.

Validation Experiment 2: The Situation-Model Assessment and CIE Test

Validation Experiment 2 served to validate the assessment of participants' situation models. We generated nine Why, How, and What-Happened-Next questions according to the principles of the Question-Answering method (Graesser & Clark, 1985; hereafter: situation-model assessment [SMA] questions). The SMA questions were tailored to the fictitious news report from the "Water Department Shutdown" scenario (e.g., Brydges et al., 2018; Ecker et al., 2017; Hey et al., 2025; McIlhiney et al., 2023) and required participants to draw inferences based on their understanding of the situation.

We conducted Validation Experiment 2 for three reasons: First, we wanted to check whether participants were generally willing to draw inferences from the little information provided in order to give meaningful answers to the SMA questions, even though they did not have enough information to give definitive answers. Second, we wanted to identify those SMA questions that allowed us to differentiate between situation models that indicate agreement with, disagreement with, and indecision regarding the original information. Third, we wanted to validate the CIE statements used in the CIE test. To this end, we manipulated participants' situation models by presenting information that would unequivocally induce agreement with the original information (*original-information* group) or disagreement with the original information (*retraction* group), or by presenting no information related to the original information (*uninformed* group). Useful SMA questions and CIE statements should be able to differentiate between the groups and to predict participants' group membership based on their answers.

Method

Design. The validation experiment had the between-subjects factor *manipulated situation model* resulting in three groups: the original-information group, the retraction group, and the uninformed group. The main dependent variables were the ratings of agreement with

the original information indicated by the answers to each SMA question and each CIE statement.

Participants. Between December 9 and 12, 2022, we recruited 140 new participants. Participants were excluded if they met at least one of the following a priori exclusion criteria: they failed the attention check (n = 2), they did not report the responsibility for the situation according to the information provided to their respective group (n = 5), they self-assessed as unserious or inattentive in a seriousness check (n = 4; Aust et al., 2013), their individual completion time exceeded the mean completion time by more than three standard deviations (n = 2). Of the final sample of 127 native English speakers, 61 self-reported as female, 62 as male, and four as non-binary, with ages ranging between 19 and 32 years (M = 26.59, SD = 3.81). They received £3 as compensation for participation. Participants were randomly assigned to the three groups: original-information group (n = 44), retraction group (n = 40), and uninformed group (n = 43).

Materials and Procedure. Participants completed the experiment at home on their personal computer or laptop. After giving informed consent, participants began with the study phase. They read the first two paragraphs of the "Water Department Shutdown" scenario as presented in the main experiment. Then, the manipulation was administered. For the uninformed group, the report ended after the first two paragraphs. Thus, they did not receive any information on the cause of the water department shutdown situation, the original information, or retraction. For the original-information group, the original information was presented in the third paragraph. For the retraction group, the retraction information was presented in the third paragraph. The statements were preceded by the sentence "*The authorities have now published the first results of their investigation.*" Thus, both original information and retraction were presented by a highly credible source so that participants should integrate the presented information into their situation models. Otherwise, the

statements were presented in the same way as in the main experiment.

Next, all participants were instructed to take a moment to reflect on the information they had read and to answer the following questions according to their subjective understanding of the situation. Appendix B of the main article lists all SMA questions. For convenience, we repeat them in the note of Figure S2.2 in this supplement. The SMA questions were presented in the same way as in the main experiment, that is, grouped by the six aspects of the water department shutdown situation. However, they were not presented in a random order for each participant but in the fixed order presented in Figure S2.2. This was done so that possible carry-over effects between questions were constant across groups. After the nine SMA questions, the question "*Who is responsible for the fish death in the Denroy river*?" was presented as a manipulation check. Then, participants indicated their agreement with the original information via the six CIE statements that were presented in the same way as in the main experiment. Lastly, participants provided their demographics, completed the seriousness check (Aust et al., 2013), and received their compensation via Prolific. Mean completion time was 8 minutes.

Rating Procedure. For the two SMA questions that were answered with a rating from 0% to 100%, higher ratings indicated greater agreement with the original information. The extent of (dis-)agreement with the original information in the open answers was rated by two independent raters on a Likert scale from –2 (*strong disagreement*) to +2 (*strong agreement*). The raters were familiar with the experimental design and the hypotheses of the experiment but blind regarding which answer belonged to a specific participant or group. Before the rating, they underwent extensive training based on simulated answers.

Only one of 1,024 answers was rated as nonsensical. Thus, participants were generally willing to draw inferences from the little information provided in order to give meaningful answers to the SMA questions. Raters agreed on the exact rating for 48% of all answers (e.g.,

both ratings were +2) and on a positive or negative rating for an additional 36% (e.g., one rating was +1 and one was +2). The inter-rater reliability was $\kappa_{weighted} = .62$, which indicates substantial agreement (Landis & Koch, 1977). Conflicts where the raters agreed on a positive or negative rating were resolved by using the median rating; conflicts where the raters disagreed more strongly (e.g., one rating was positive or neutral, whereas the other one was negative) were resolved by discussion between the raters.

Results and Discussion

Validation of the SMA Questions. We expected the original-information group to have the strongest rated agreement with the original information in each SMA question, followed by the uninformed group and the retraction group. Descriptive statistics of the nine SMA questions are in Figure S1.2. They conform to the expected ranking of the groups. We tested the agreement ratings for group differences with one-way ANOVAs and post-hoc *t* tests. Post-hoc tests were corrected with Bonferroni-Holm correction for α error inflation (Holm, 1979). The ANOVAs showed group differences in the rated agreement of all SMA questions (all *p*s < .004). Table S2.1 lists the test statistics of the post-hoc tests. For all SMA questions, the rated agreement in the original-information group was stronger than the rated agreement in the retraction group and, for all but SMA question 5, stronger than the rated agreement in the uninformed group. Additionally, for the SMA questions 8 and 9, the rated agreement in the uninformed group was stronger than the rated agreement in the retraction group.

As said, we sought to identify those SMA questions that differentiate between situation models that indicate agreement with, disagreement with, or indecision regarding the original information. Therefore, we entered each question as a sole predictor in a multinomial logistic regression to predict participants' situation model according to their group membership, that is, whether participants agreed with the original information, disagreed with

it, or had no opinion about it. Each question improved the model fit of its regression model compared to a null model, all $\chi^2(2) > 11.95$, ps $< .003$. The rates of correct classification were between 44% (question 7) and 71% (question 8). This means that each question individually contributed to a correct prediction of participants' group membership.

In a next step, we tested whether a combination of all nine SMA questions accurately predicted the group membership of the participants. We entered all SMA questions in one multinomial logistic regression to predict participants' group membership. Together, the SMA questions correctly classified 100% of the participants into their groups. Thus, we conclude that the SMA questions we constructed validly assess the (dis-)agreement with the original information in participants' situation models.

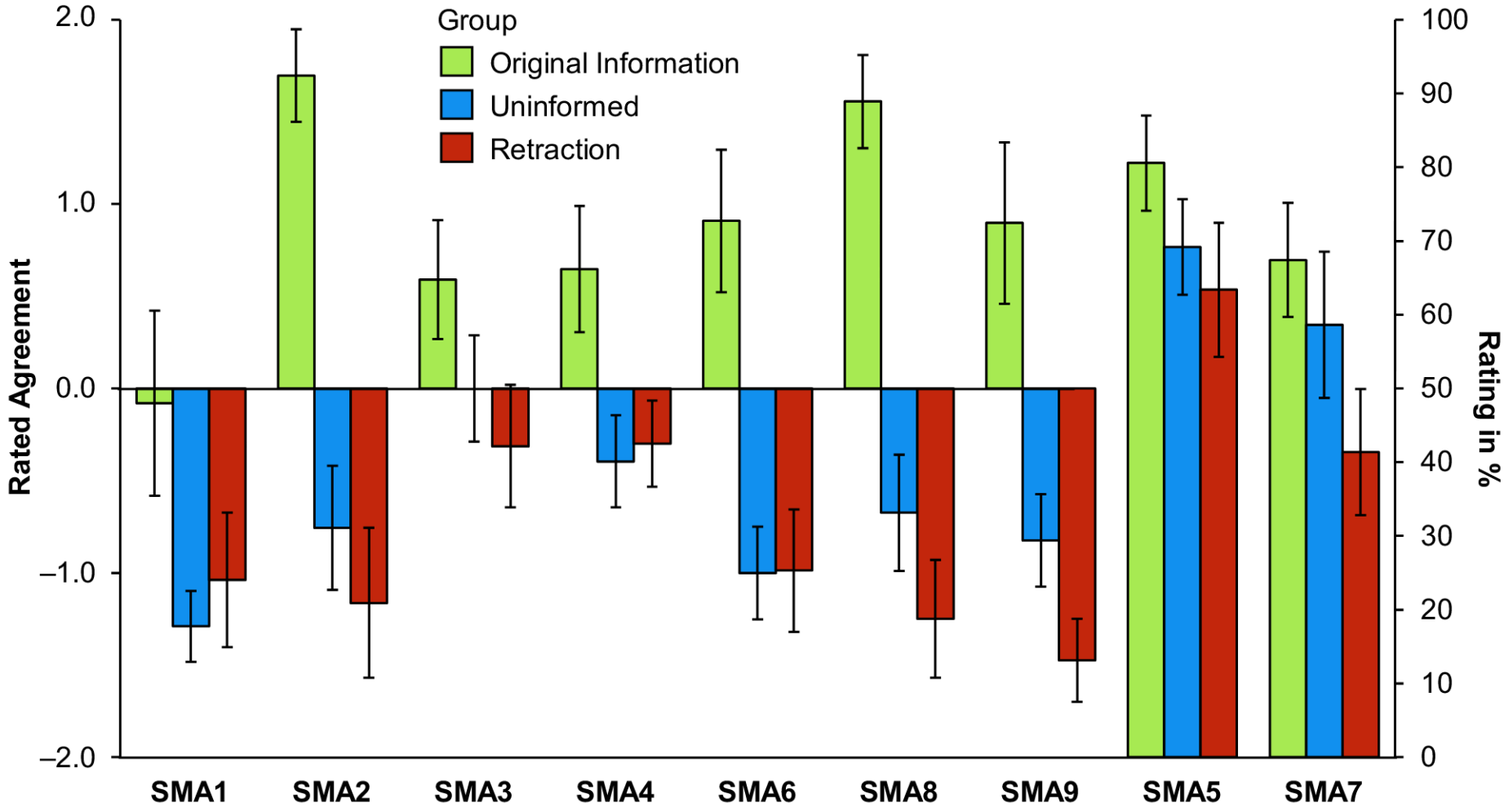
Finally, we wanted to ensure that the ratings of the nine SMA questions could be combined to a single score of agreement with the original information in the situation model. As the answers to the SMA questions were measured on different scales (i.e., –2 to +2 or 0% to 100%), we first standardized them using the z-score transformation.[1] The standardized ratings had good internal consistency (Cronbach's $\alpha = .81$). We computed the mean of the z-scores for each participant (original-information group: $M = 0.71$, 95% CI [0.61, 0.81], uninformed group: $M = –0.29$, 95% CI [–0.39, –0.19], retraction group: $M = –0.47$, 95% CI [–0.59, –0.34]). These agreement scores differed between the groups, $F(2, 124) = 142.25$, $p < .001$, $\eta_p^2 = .70$, $BF_{incl} = 2.78 \cdot 10^{29}$. The original-information group had higher agreement scores than the retraction group, $t(124) = 15.43$, $p < .001$, $d = 3.37$, $BF_{10} = 4.50 \cdot 10^{21}$, and the uninformed group, $t(124) = 13.44$, $p < .001$, $d = 2.88$, $BF_{10} = 1.16 \cdot 10^{21}$. Further, the

[1] We also computed a score for the situation model as factor scores from Principal Component Analyses (PCA) using the regression method which weighs the standardized data with their respective factor loadings (cf. DiStefano et al., 2009). Due to the near-perfect correlation of the reported averaged z score with the factor score ($r > .99$, $p < .001$, $BF_{10} > 10^{125}$), we present the results of the simpler computation method. The results and conclusions are the same for both computation methods.

uninformed group had higher agreement scores that the retraction group, $t(124) = 2.23$, $p = .027$, $d = 0.49$, $BF_{10} = 1.76$. Thus, we conclude that a scale consisting of all nine SMA questions validly assesses the (dis-)agreement with the original information in participants' situation models. We, therefore, used these nine SMA questions to assess the situation models in the main experiment.

Figure S1.2

Rated Agreement with the Original Information in the Responses to the SMA Questions in Validation Experiment 2



Note. Rated agreement with the original information in the original-information, uninformed, and retraction groups. The x-axes represent the rating scale from –2 (*strong disagreement*) to +2 (*strong agreement*) for the SMA questions 1 to 4, 6, 8, and 9, and the rating scale from 0% to 100% for the SMA questions 5 and 7. Error bars indicate 95% confidence intervals. General reason: SMA 1 = *Why did the Freemont water department shut down the water supply from the Denroy River?*; Specific reason: SMA 2 = *Why did so many fish die in the Denroy River?*; Residents' reaction: SMA 3 = *How did the Shelby residents react to the shutdown of the water department?* and SMA 4 = *Why did the residents react that way?*; Potential health risk: SMA 5 = *How likely will people's health be harmed if they drink water from the Denroy River?* and SMA 6 = *Why is the health risk the way you just rated it?*; Potential legal consequences: SMA 7 = *How likely is it that a responsible party will be legally charged?* and SMA 8 = *Why is it as likely as you just indicated that a responsible party will be legally charged?*; Sequel: SMA 9 = *Considering all you know so far, what will happen next?*

Table S1.1

Post-hoc t *Test of Group Differences in Responses to the SMA Questions*

Question	Comparison	*t*(124)	*p*	*d*
SMA 1	Original Information vs. Retraction	3.60	.001*	0.79
	Original Information vs. Uninformed	4.63	< .001*	0.99
	Retraction vs. Uninformed	0.95	.613	0.21
SMA 2	Original Information vs. Retraction	12.16 [a]	< .001*	2.67
	Original Information vs. Uninformed	10.63 [a]	< .001*	2.29
	Retraction vs. Uninformed	1.73 [a]	.198	0.38
SMA 3	Original Information vs. Retraction	4.07	< .001*	0.89
	Original Information vs. Uninformed	2.72	.020*	0.58
	Retraction vs. Uninformed	1.41	.341	0.31
SMA 4	Original Information vs. Retraction	4.76	< .001*	1.04
	Original Information vs. Uninformed	5.34	< .001*	1.14
	Retraction vs. Uninformed	0.48	.883	0.10
SMA 5	Original Information vs. Retraction	3.33	.003*	0.73
	Original Information vs. Uninformed	2.25	.067	0.48
	Retraction vs. Uninformed	1.12	.505	0.25
SMA 6	Original Information vs. Retraction	8.19	< .001*	1.79
	Original Information vs. Uninformed	8.40	< .001*	1.80
	Retraction vs. Uninformed	0.05	.998	0.01
SMA 7	Original Information vs. Retraction	4.21	< .001*	0.92
	Original Information vs. Uninformed	3.08	.007*	0.66
	Retraction vs. Uninformed	1.18	.469	0.26
SMA 8	Original Information vs. Retraction	13.46	< .001*	2.94
	Original Information vs. Uninformed	10.90	< .001*	2.34
	Retraction vs. Uninformed	2.74	.019*	0.60
SMA 9	Original Information vs. Retraction	10.36	< .001*	2.26
	Original Information vs. Uninformed	7.67	< .001*	1.64
	Retraction vs. Uninformed	2.82	.015*	0.62

Note. [a] *df* = 123 because one answer was rated as nonsensical. The wording of the SMA questions are provided in the note to Figure S1.2.

* significant after Bonferroni-Holm correction (Holm, 1979).

Validation of the CIE Items. We expected the original-information group to show the strongest agreement with each item, followed by the uninformed group and the retraction group. Descriptive statistics of the rated agreement with the six CIE items are in Figure S1.3. They conform to the expected ranking of the groups. We tested the rated agreement for group differences with one-way ANOVAs and post-hoc *t* tests with Bonferroni-Holm correction for α error inflation (Holm, 1979). The ANOVAs showed group differences in mean agreement for all CIE items (all *p*s < .024). Table S2.2 lists the test statistics of the post-hoc *t* tests. For all CIE statements but statement 4, the agreement in the original-information group was stronger than the agreement in the retraction group. For all CIE statements but statements 4 and 5, the agreement in the original-information group was stronger than the agreement in the uninformed group, which was stronger than the agreement in the retraction group.

We tested whether the CIE items predicted the group membership of the participants, that is, whether they (dis-)agreed with the original information or had no opinion about it. To this end, we entered the six agreement ratings in a multinomial logistic regression to predict participants' group membership. Each statement significantly improved the model fit over and above the other items with the exception of statement 1, $\chi^2(2) = 0.91$, $p = .635$, all other $\chi^2(2) > 6.83$, *p*s < .033. Table S1.3 maps the predicted group memberships on the actual group memberships. Overall, the CIE items correctly classified 87% of the participants into their respective group. Importantly, all incorrect classifications involved the uninformed group. Thus, no participant of the original-information group was misclassified as belonging to the retraction group or vice versa. Therefore, we conclude that the CIE items used in the main experiment validly assess (dis-)agreement with the original information.

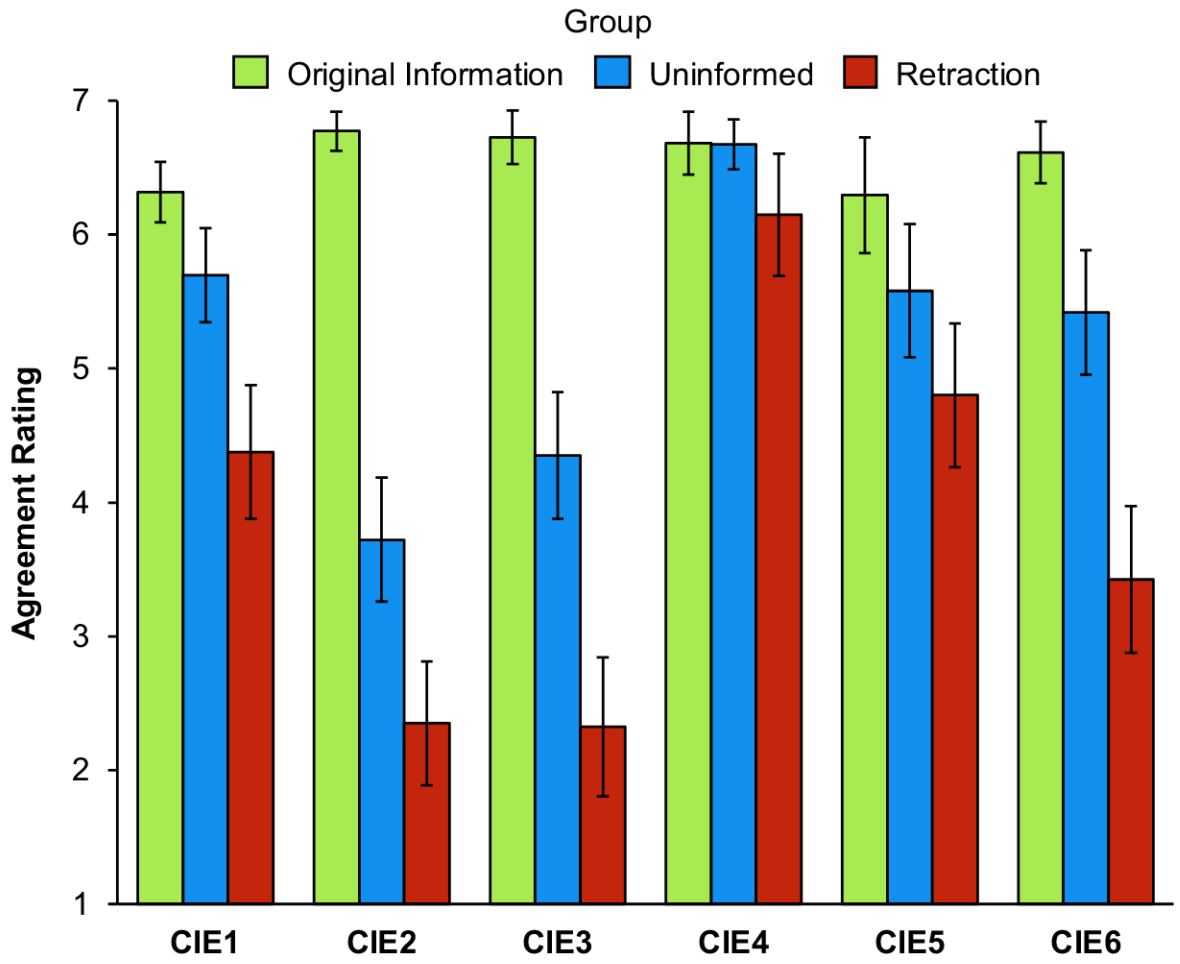
Further, we wanted to ensure that the ratings of the six CIE statements could be combined to a single score of agreement with the original information. We wanted to report all dependent variables on the same scale, so we first standardized the agreement ratings of

the CIE statements using the *z*-score transformation.[2] The standardized ratings of the six CIE statements had good internal consistency (Cronbach's $\alpha = .89$). We computed the mean of the *z*-score for each participant (original-information group: $M = 0.66$, 95% CI [0.56, 0.77], uninformed group: $M = 0.01$, 95% CI [–0.16, 0.18], retraction group: $M = –0.74$, 95% CI [–0.98, –0.50]). The agreement scores differed between the groups, $F(2, 124) = 63.82$, $p < .001$, $\eta_p^2 = .51$, $BF_{incl} = 4.71 \cdot 10^{16}$. The original-information group had higher agreement scores than the retraction group, $t(124) = 11.30$, $p < .001$, $d = 2.47$, $BF_{10} = 9.21 \cdot 10^{14}$, and the uninformed group, $t(124) = 5.38$, $p < .001$, $d = 1.15$, $BF_{10} = 2.32 \cdot 10^{6}$. Further, the uninformed group had higher agreement scores that the retraction group, $t(124) = 5.99$, $p < .001$, $d = 1.31$, $BF_{10} = 9.99 \cdot 10^{4}$. Thus, we conclude that a scale consisting of all six CIE statements validly assesses the (dis-)agreement with the original information. We, therefore, used these six CIE statements to assess the situation models in the main experiment. Finally, we used a linear regression to ensure that the agreement score of the situation model predicted the agreement score of the CIE, $B = 0.84$, 95% CI [0.66, 1.01], $\beta = 0.65$, $t(125) = 9.64$, $p < .001$ (even after entering group as a second predictor, $B = 0.47$, 95% CI [0.19, 0.75], $\beta = 0.37$, $t(124) = 3.35$, $p = .001$).

[2] We also computed scores for the CIE as factor scores from Principal Component Analyses (PCA) using the regression method which weighs the standardized data with their respective factor loadings (cf. DiStefano et al., 2009) and a simple mean of the six raw agreement ratings. Due to the near-perfect correlations of the reported averaged *z* scores with the factor score ($r > .99$, $p < .001$, $BF_{10} > 10^{142}$) and the mean raw agreement ratings ($r = .99$, $p < .001$, $BF_{10} > 10^{116}$), we present the results of the simpler computation method, which also allowed us to report all dependent variables on the same scale, in the main text. The results and conclusions are the same for all computation methods.

Figure S1.3

Ratings of Agreement with the CIE Statements in Validation Experiment 2



Note. Mean agreement ratings of the original-information, uninformed, and retraction groups on the scale from 1 (*strong disagreement*) to 7 (*strong agreement*) for the six statements used to assess the CIE in the main experiment. Error bars indicate 95% confidence intervals. CIE1 = *This incident could have been prevented by stricter control of the companies within proximity to the river.* CIE2 = *The reason for the death of fish in the Denroy River was the illegal dumping of toxic chemical waste by a cosmetics and pharmaceutical company.* CIE3 = *The riverside pharmaceutical company should take responsibility for the incident.* CIE4 = *Fines for industrial water pollution should be increased to drastically deter companies from polluting the water.* CIE5 = *The riverside pharmaceutical company should start an internal investigation and review its procedures.* CIE6 = *The riverside pharmaceutical company*

should prepare for civil lawsuits from harmed consumers and farmers who have been financially damaged.

Table S1.2

Post-hoc t *Test for the CIE items*

Statement	Comparison	*t*(124)	*p*	*d*
CIE1	Original Information vs. Retraction	7.51	< .001*	1.64
	Original Information vs. Uninformed	2.44	.042*	0.52
	Retraction vs. Uninformed	5.08	< .001*	1.12
CIE2	Original Information vs. Retraction	16.55	< .001*	3.62
	Original Information vs. Uninformed	11.64	< .001*	2.50
	Retraction vs. Uninformed	5.10	< .001*	1.12
CIE3	Original Information vs. Retraction	15.08	< .001*	3.30
	Original Information vs. Uninformed	8.30	< .001*	1.78
	Retraction vs. Uninformed	6.90	< .001*	1.52
CIE4	Original Information vs. Retraction	2.47	.039	0.54
	Original Information vs. Uninformed	0.04	> .999	0.01
	Retraction vs. Uninformed	2.43	.044	0.53
CIE5	Original Information vs. Retraction	4.34	< .001*	0.95
	Original Information vs. Uninformed	2.11	.091	0.45
	Retraction vs. Uninformed	2.26	.066	0.50
CIE6	Original Information vs. Retraction	10.61	< .001*	2.32
	Original Information vs. Uninformed	4.05	< .001*	0.87
	Retraction vs. Uninformed	6.60	<.001*	1.45

Note. For the wording of the items, see the note to Figure S1.3.

* significant after Bonferroni-Holm correction (Holm, 1979).

Table S1.3

Group Classifications Based on the Agreement Ratings with the CIE Statements

Actual group membership	Predicted group membership		
	Original Information	Uninformed	Retraction
Original Information	41	3	0
Uninformed	1	34	8
Retraction	0	5	35

References in Online Supplement 1

Aust, F., Diedenhofen, B., Ullrich, S., & Musch, J. (2013). Seriousness checks are useful to improve data validity in online research. *Behavior Research Methods, 45*(2), 527–535. https://doi.org/10.3758/s13428-012-0265-2

Brydges, C. R., Gignac, G. E., & Ecker, U. K. (2018). Working memory capacity, short-term memory capacity, and the continued influence effect: A latent-variable analysis. *Intelligence, 69*, 117–122. https://doi.org/10.1016/j.intell.2018.03.009

DiStefano, C., Zhu, M., & Mîndrilã, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research, and Evaluation, 14*, Article 20. https://doi.org/10.7275/da8t-4g52

Ecker, U. K., Hogan, J. L., & Lewandowsky, S. (2017). Reminders and repetition of misinformation: Helping or hindering its retraction? *Journal of Applied Research in Memory and Cognition, 6*(2), 185–192. https://doi.org/10.1016/j.jarmac.2017.01.014

Graesser, A. C., & Clark, L. F. (1985). The generation of knowledge-based inferences during narrative comprehension. *Advances in Psychology, 29*, 53–94. https://doi.org/10.1016/S0166-4115(08)62732-6

Hey, C. V., Schaper, M. L., & Bayen, U. J. (2025). Relative source credibility affects the continued influence effect: Evidence of rationality in the CIE. *Cognition, 254*, Article 106000. https://doi.org/10.1016/j.cognition.2024.106000

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65–70. http://www.jstor.org/stable/4615733

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. https://doi.org/10.2307/2529310

McIlhiney, P., Gignac, G. E., Ecker, U. K., Kennedy, B. L., & Weinborn, M. (2023). Executive function and the continued influence of misinformation: A latent-variable analysis. *PLoS One, 18*(4), e0283951. https://doi.org/10.1371/journal.pone.0283951