Scenarios Impact the Continued Influence Effect*

Alexander R. Hough[1], Benjamin Marrs[2], Alina Arakal[1], and Othalia Larue[3]

[1] Air Force Research Laboratory, Wright-Patterson AFB, OH 45433, USA
[2] ORISE at Air Force Research Laboratory, Wright-Patterson AFB, OH 45433 USA
[3] Parallax Advanced Research, Beavercreek, OH 45431, USA

Abstract. Human cognition is efficient, but vulnerable to misinformation and influence attempts. An example is the continued influence effect (CIE), where misinformation has a lasting effect even after presentation of corrections or discrediting facts. Experiments used a series of scenarios to explore manipulations that increase or decrease the CIE. However, there appears to be scenario effects that remain rather unexplained and may be a potential confound for experimental manipulations. We searched for and found six datasets from previous experiments in the literature and explored whether they had interactions between scenarios and experimental manipulations. All six datasets had either an interaction between scenario and experimental manipulation or a main effect for scenario. We present these analyses, discuss potential explanations, and discuss how to control for such effects in future research.

Keywords: Continued influence effect · misinformation · knowledge representation · memory · heuristics · biases · cognitive modeling

1 Introduction

Humans usually make good decisions using heuristics [14], but in some cases, these heuristics lead to systematic errors [19]. For instance, making decisions based on the ease an answer comes to mind [19] or relying on feelings [13]. Heuristics are often developed in stable and predictable environments [20] and with an understanding of them, an intelligent agent could exploit them to mislead or influence others [35]. One example is the continued influence effect (CIE), where misinformation has a lasting effect on decisions despite presentation of corrections or discrediting facts [18, 22]. Experimental research used narratives about different scenarios to determine if and how manipulations reduce the CIE. However, scenario effects are usually not considered. Here, we re-analyzed six preexisting datasets collected without our involvement [6, 7, 24, 26, 29, 33] and show scenario effects exist, which may be a potential confound for manipulations. We present analyses, potential explanations for findings, and discuss how to control for such effects in future research.

* This research was supported by the U. S. Air Force Research Laboratory's Cognition and Modeling Branch. Contents were reviewed and approved for public release: AFRL-2024-3912. The views expressed are those of the authors and do not reflect the official guidance or position of the United States Government, U.S. Department of Defense, U.S. Air Force, or any subsidiaries or employees.

1.1 The Continued Influence Effect

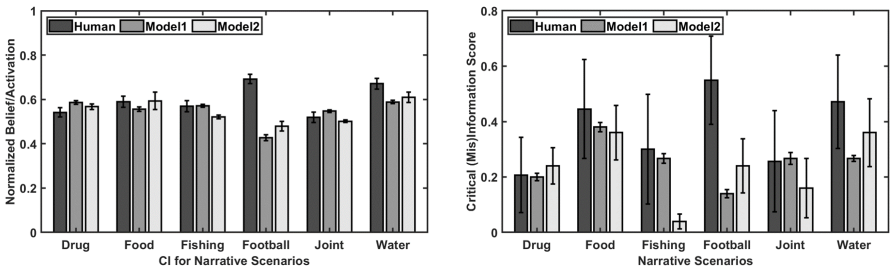
CIE experiments often present a narrative with a separate misinformation and correction article. Corrections can reduce but not eliminate the CIE [7, 18]. Episodic memory is used to frame the CIE. It suggests memory cannot be erased; only re-activated or associated with other information [43]. The CIE is generally thought to be due to a combination of competition between misinformation and corrections [1, 10], recency effects [8], and fluency [9, 39]. Research also includes emotional influence and found that correcting misinformation may create feelings of discomfort [36] and higher emotional responses increase belief in fake news [23]. Emotional experiences are also more accessible [5], remembered better [45], and often weighted more than experience [27]. Negative emotion appears to have more influence on memory, by enhancing recall [44].

Research covered a range of mitigation methods with some mixed findings. Retracting misinformation was found to create discomfort leading to greater reliance on misinformation [36], but providing an explanation with the correction helps to integrate the correction into a mental model [21]. There is evidence that discrediting an information source decreases reliance on misinformation [42]; however, there may be little or no effect without explicit, strong source discredit [23]. Corrections were found to be effective when misinformation aligned with one's worldview or political lean [7]; conversely, corrections were less effective when they challenge strong beliefs or attitudes [37], which may lead to some individuals doubling down on misinformation [30]. Research has suggested memory failure leads to decreased effectiveness of corrections; however, other research found enhancing memory of corrections does not impact its effectiveness [12]. There are reported benefits to presenting corrections after misinformation [2]. However, other research found little difference between fact-checks before or after misinformation, and when interleaved [40].

The CIE and misinformation-related effects are complex and need to be carefully studied. Despite the breadth of research, the impact of scenarios (i.e., topics) on the CIE and their interaction with manipulations are often overlooked. Here, we use several pre-existing datasets from published papers, which we were not authors, [3, 6, 7, 24, 26, 29, 33] to explore whether there are differences between scenarios and their interactions with experimental manipulations. We present results of these analyses, a hypothesis to explain the results, and discuss how to address the effects revealed through our analyses in future work.

2 Scenario Analyses

The scenario analyses were inspired by previous work [16, 17], where we developed a computational cognitive model of the CIE and were unable to capture differences across scenarios from a previous experiment [7]. Figure 1 shows model fits for a baseline memory-based model (Model1) and an extended model with emotion (Model2). We noted variation across scenarios for both belief ratings (a) and misinformation scores (b), particularly with football and water (see [16, 17] for more detail). This raised the question of whether scenario effects existed in



(a) Misinformation belief ratings (b) Misinformation score

Fig. 1: Comparison between human data and models for a) belief ratings and b) misinformation (CI) score.

the literature and whether they interacted with experimental manipulations. In the following section, we present six analyses that are grouped based on whether there were interactions between scenarios and experimental manipulations or just scenario main effects for misinformation reliance scores.

Ecker et al. (2024) [6] This experiment assessed the effectiveness of five misinformation discrediting conditions (i.e., control with generic information, misinformation/no correction, misinformation/correction, misinformation discredited/no correction, and combined misinformation discrediting/correction) to combat the CIE in five different scenarios (i.e., a band cancelling concert, suburb burglary, menopause treatments, mine construction, and a restaurant closing). Participants ($N = 292$) were recruited through the US based prolific platform (57% female and mean age 40). A two-way ANOVA (Figure 2) found a significant interaction between scenario and condition, $F(16, 1435) = 16.91$, $p < .001$ and a significant main effect of scenario, $F(4, 1435) = 93.98$, $p < .001$.

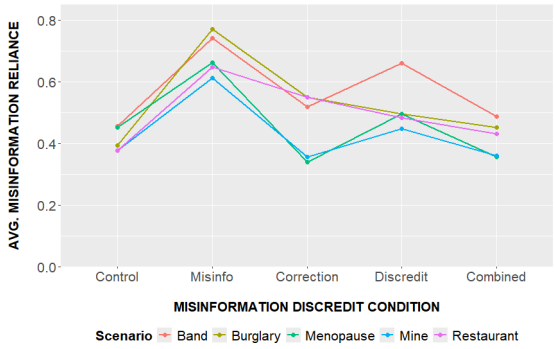


Fig. 2: Line plot with scenarios (lines) and misinformation discrediting conditions (x-axis) for misinformation reliance score from Ecker et al. (2024) [6]

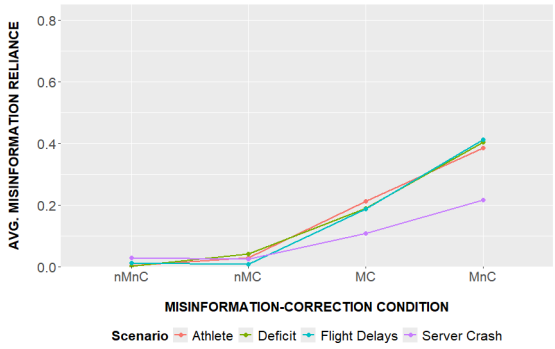


Fig. 3: Line plot with scenarios (lines) and misinformation/correction conditions (x-axis) for misinformation reliance score from Prike et al. (2023) [29]

Prike et al. (2023) [29] This experiment assessed backfire effects with four information conditions: no misinformation/no correction (nMnC), no misinformation/correction (nMC), misinformation/no correction (MC), and misinformation/correction (MnC). Conditions were assessed across four fictional news reports (i.e., flight delay, server crash, government budget deficit, and an athlete banned from playing). Participants residing in the US and proficient in English were recruited using Amazon Mechanical Turk (48% female and mean age of 43) A two-way ANOVA (Figure 3) revealed a significant interaction between information type and scenario, $F(9, 1088) = 30.8$, $p < .001$, and a significant main effect of scenario, $F(3, 1088) = 55.02$, $p < .001$. [29]

Sanderson et al. (2022) [33] This experiment focused on disruption of memory consolidation on CIE by physically relocating participants between presentation of misinformation and retractions. Participants ($N = 112$, 69% female,

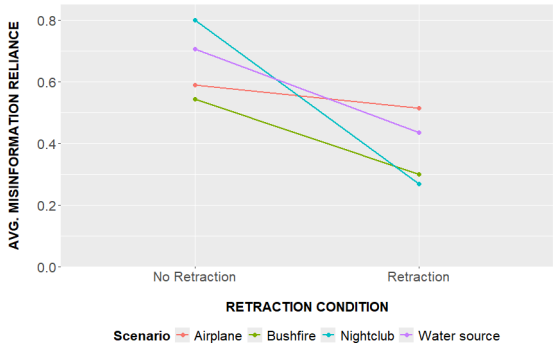


Fig. 4: Line plot with scenarios (lines) and retraction (x-axis) for misinformation reliance score from Sanderson et al. (2022) [33]

and mean age 21) recruited from the University of Western Australia were or were not physically relocated before reading the retraction, and a group of US based participants ($N = 56$, 46% female, and mean age 37) were recruited via Mechanical Turk and did not receive a retraction. All participants read the same fabricated news reports (i.e., emergency airplane landing, woman at nightclub, water source contamination, and a bushfire). For simplicity, data was condensed to a binary did/did not receive retraction for our analysis. A two-way ANOVA (Figure 4) found a significant interaction, $F(3,157) = 31.4$, $p < .001$, and a significant main effect of scenario, $F(3,157) = 23.69$, $p < .001$.

2.1 Main Effects for Scenarios

Ecker and Antonio (2021) [7] This experiment explored retraction source trustworthiness and expertise with six conditions ranging from least to most likely to reduce the CIE: no retraction (NoR), low expertise/trust (LELT), low expertise/high trust (LEHT), high expertise/low trust (HELT), high expertise/trust (HEHT), and highest expertise/trust (HEHT+). Fifty three participants (62% female and mean age 18.6) were recruited from the University of Western Australia and read stories about six scenarios (i.e., anti-viral drug, fishing restrictions, food additives, football scandal, joint condition treatments, and water contamination). We conducted a two-way ANOVA (Figure 5) with scenario and source conditions and only found a significant main effect for scenario, $F(5, 282) = 22.17$, $p < .001$.

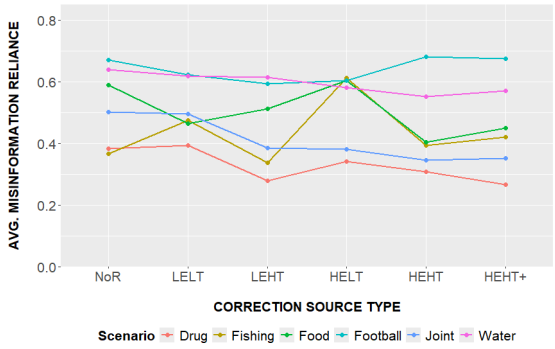


Fig. 5: Line plot with scenarios (lines) and correction source conditions (x-axis) for misinformation reliance scores from Ecker and Antonio (2021) [7].

2.2 Interactions between Scenarios and Manipulations

McIlhiney et al. (2022) [24] This experiment investigated whether individual's CIE susceptibility was stable over time. Two hundred and fifty US based

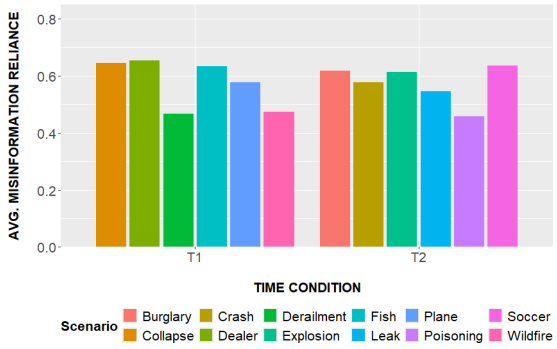


Fig. 6: Bar chart of mean misinformation reliance scores across twelve scenarios used in McIlhiney et al. (2022) [24] at Time 1 and Time 2 (x-axis).

participants (56% female and mean age 41) were recruited via Amazon Mechanical Turk and CloudResearch. Participants were given six scenarios (i.e., emergency airplane landing, wildfire, drug dealer death, woman collapsing at bar, train derailing, and fish deaths), then four weeks later they were given six different scenarios (i.e., explosion at warehouse, burglary, car crash, data leak, food poisoning, and soccer player suspension). The two sets of scenarios were believed to be equivalent. A one-way ANOVA (Figure 2b) revealed a significant difference between scenarios, $F(11, 3036) = 11.7$, $p < .001$.

Miller et al. (2022) [26] Miller et al. (2022) explored potential differences with age (i.e., young and old), retraction presence (i.e., retraction and no retraction), and retraction delay (i.e., no delay, 10 minutes, and 2 days) on the CIE. Participants were recruited via CloudResearch and were divided into younger (aged 18-35) and older adults (aged 60+). All participants read a series of six

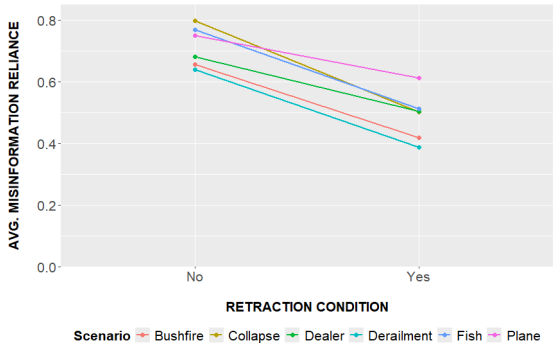


Fig. 7: Line plot for misinformation score with six scenarios (lines) and retraction presence (x-axis) from Miller et al. (2022) [26]

fictitious news reports (i.e., bushfire, woman collapsing at bar, drug dealer death, train derailing, fish deaths, and an emergency airplane landing), with half of each group also reading a retraction after a certain period of time. For simplicity, the data was condensed into a binary did/did not receive retraction, as there were no effects of retraction delay or age. We conducted a two-way ANOVA (Figure 7) with scenario and retraction presence and only found a significant main effect for scenario, $F(5, 978) = 11.32$, $p < .001$.

3 Discussion

We presented analyses on six CIE experiments in the existing literature to explore scenario main effects and interactions with their experimental manipulations. All six experiments examined showed a significant main effect of scenario type [7, 6, 24, 26, 29, 33]. Three of the six experiments also had an interaction between scenario and the experimental manipulation [6, 29, 33]. Manipulations ranged across variations on both the source of the misinformation as well as the method of its correction/retraction. Although it is unclear why misinformation about some scenarios create stronger beliefs or increased misinformation endorsement in open-ended and inference questions, we propose three potential explanations for these scenario effects. These explanations are supported by evidence from the existing literature, and we do not believe these to be independent: 1) context-driven reasoning, 2) degree of relevance or visceral response through imagination, and 3) previous experience or exposure to events/stories.

First, contextual cues often provide the basis upon which individuals reason and make judgments, particularly when there is a lack of direct experience with the situation [32, 11, 15]. One example of such a cue is the source of the retraction/correction [25]. In Ecker and Antonio (2021), the retraction from the football scenario refuted the initial misinformation that an athlete was taking performance-enhancing drugs; the potential correction sources from least to most credible were: fan club president, sports commentator, player's manager, team doctor, and director of an anti-doping authority. Some sources may have inadvertently reinforced misinformation beliefs (e.g., reasoning about the underlying motives for a cover-up). Second, judgment can also be influenced by the degree of relevance or visceral response to the scenario [41, 38]. In McIlhiney et al. (2022), the train derailment scenario occurred in France, while the participants were US-based; this is one such instance where the degree of direct relevance to the participant may have influenced responses to the scenario. Finally, previous experience and/or exposure to related events can impact judgments and reasoning [28]. For example, the correction statement for the server crash scenario used in Prike et al. (2023) clarified that the NYSE server crash was not caused by a cyber-attack as initially suggested. However, it is not uncommon for cyber-attack events to involve an initial denial by large corporations to avoid spreading panic. Prior exposure to cyber-attack related news reports could affect reasoning.

Each of these potential explanations involve individual differences, which has only been partially addressed in the CIE. For instance, individuals with lower

working memory capacity [4] or verbal ability [31] may be more susceptible to the CIE, suggesting they have more difficulty updating mental models. However, others were unable to replicate some of these findings and suggested the encoding of episodic memory is more important than working memory capcity [34]. It is important to understand why these scenario effects exist by more thoroughly addressing individual differences through the lens of the potential explanations presented here.

Future research could combat potential confounds for scenario effects by: 1) using scenario as a covariate in analyses, 2) asking or gathering information about personal experience, attitudes, and if they were exposed to similar stories, or 3) conducting thorough pilot testing to ensure that scenario effects do not exist or at least average out. In addition to empirical research, computational models could be a tool to test hypotheses about different explanations through model comparisons. Recent cognitive models [16, 17] could be used as a baseline and compared to unique models with added features, such as contextual cues, emotional responses to certain events and sources, and previous knowledge or additional related narratives.

References

1. Ayers, M.S., Reder, L.M.: A theoretical review of the misinformation effect: Predictions from an activation-based memory model. Psychonomic Bulletin and Review **5**, 1–21 (1998)
2. Brashier, N.M., Pennycook, G., Berinsky, A.J., Rand, D.G.: Timing matters when correcting fake news. Proceedings of the National Academy of Sciences **118**(5), e2020043118 (2021)
3. Bruns, H., Lewandowsky, S., Pennycook, G., Pantazi, M., Schmid, P., Krawczyk, M.W., Dessart, F.J., Smillie, L.: The role of (trust in) the source of prebunks and debunks of misinformation. evidence from online experiments in four eu countries. https://doi.org/10.31219/osf.io/vd5qt (2023)
4. Brydges, C.R., Gignac, G.E., Ecker, U.K.: Working memory capacity, short-term memory capacity, and the continued influence effect: A latent-variable analysis. Intelligence **69**, 117–122 (2018)
5. Buchanan, T.W.: Retrieval of emotional memories. Psychological bulletin **133**(5), 761 (2007)
6. Ecker, U., Prike, T., Paver, A., Scott, R., Swire-Thompson, B.: Don't believe them! reducing misinformation influence through source discreditation. Tech. rep., Working Paper. https://doi. org/10.31234/osf. io/5yj2v (2024)
7. Ecker, U.K., Antonio, L.M.: Can you believe it? An investigation into the impact of retraction source credibility on the continued influence effect. Memory & Cognition **49**, 631–644 (2021)
8. Ecker, U.K., Lewandowsky, S., Cheung, C.S., Maybery, M.T.: He did it! she did it! no, she did not! multiple causal explanations and the continued influence of misinformation. Journal of memory and language **85**, 101–115 (2015)
9. Ecker, U.K., Lewandowsky, S., Swire, B., Chang, D.: Correcting false information in memory: Manipulating the strength of misinformation encoding and its retraction. Psychonomic Bulletin & Review **18**, 570–578 (2011)

10. Ecker, U.K., Lewandowsky, S., Tang, D.T.: Explicit warnings reduce but do not eliminate the continued influence of misinformation. Memory & Cognition **38**, 1087–1100 (2010)
11. Ecker Ullrich K. H., Lewandowsky Stephan, F.O.M.K.: Do people keep believing because the want to? preexisting attitudes and the continued influence of misinformation. Memory Cognition **12**, 292–304 (2013)
12. Fazio, L.K., Hong, M.K., Pillai, R.M.: Combatting rumors around the french election: the memorability and effectiveness of fact-checking articles. Cognitive Research: Principles and Implications **8**(1), 44 (2023)
13. Finucane, M.L., Alhakami, A., Slovic, P., Johnson, S.M.: The affect heuristic in judgments of risks and benefits. Journal of behavioral decision making **13**(1), 1–17 (2000)
14. Gigerenzer, G., Gaissmaier, W.: Heuristic decision making. Annual review of psychology **62**, 451–482 (2011)
15. Hornsey, M.J., Fielding, K.S.: Attitude roots and jiu jitsu persuasion: Understanding and overcoming the motivated rejection of science. American Psychologist **72**, 459–473 (2017)
16. Hough, A.R., Larue, O.: A model of memory and emotion mechanisms underlying the continued influence effect. In: 17th SBP-BRiMS Conference (2024)
17. Hough, A.R., Larue, O.: Exploring memory mechanisms underlying the continued influence effect. In Proceedings of the 22nd International Conference on Cognitive Modeling. Via mathpsych.org/presentation/1605 (2024)
18. Johnson, H.M., Seifert, C.M.: 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 (1994)
19. Kahneman, D.: Thinking fast and slow. New York: Farrar, Straus and Giroux (2011)
20. Kahneman, D., Klein, G.: Conditions for intuitive expertise: a failure to disagree. American psychologist **64**(6), 515 (2009)
21. Kendeou, P., Butterfuss, R., Kim, J., Van Boekel, M.: Knowledge revision through the lenses of the three-pronged approach. Memory & Cognition **47**, 33–46 (2019)
22. Lewandowsky, S., Ecker, U.K., Seifert, C.M., Schwarz, N., Cook, J.: Misinformation and its correction: Continued influence and successful debiasing. Psychological Science in the Public Interest **13**(3), 106–131 (2012)
23. Martel, C., Mosleh, M., Rand, D.G.: You're definitely wrong, maybe: Correction style has minimal effect on corrections of misinformation online. Media and Communication **9**(1), 120–133 (2021)
24. McIlhiney, P., Gignac, G.E., Weinborn, M., Ecker, U.K.: Sensitivity to misinformation retractions in the continued influence paradigm: Evidence for stability. Quarterly Journal of Experimental Psychology **75**(7), 1259–1271 (2022)
25. Metzger, M.J., Flanagin, A.J.: Credibility and trust of information in online environments: The use of cognitive heuristics. Journal of pragmatics **59**, 210–220 (2013)
26. Miller, A.L., Wissman, K.T., Peterson, D.J.: The continued influence effect: Examining how age, retraction, and delay impact inferential reasoning. Applied Cognitive Psychology **36**(3), 708–723 (2022)
27. Pachur, T., Hertwig, R., Steinmann, F.: How do people judge risks: Availability heuristic, affect heuristic, or both? Journal of Experimental Psychology: Applied **18**(3), 314 (2012)

28. Pennycook, G., Cannon, T.D., Rand, D.G.: Prior exposure increases perceived accuracy of fake news. Journal of Experimental Psychology: General **147**(12), 1865–1880 (2018)
29. Prike, T., Blackley, P., Swire-Thompson, B., Ecker, U.K.: Examining the replicability of backfire effects after standalone corrections. Cognitive Research: Principles and Implications **8**(1), 39 (2023)
30. Reinero, D.A., Harris, E.A., Rathje, S., Duke, A., Van Bavel, J.J.: Partisans are more likely to entrench their beliefs in misinformation when political outgroup members fact-check claims (2023)
31. Roets, A., et al.: 'fake news': Incorrect, but hard to correct. the role of cognitive ability on the impact of false information on social impressions. Intelligence **65**, 107–110 (2017)
32. Roozenbeek, J., Schneider, C.R., Dryhurst, S., Kerr, J., Freeman, A.L., Recchia, G., Van Der Bles, A.M., Van Der Linden, S.: Susceptibility to misinformation about covid-19 around the world. Royal Society open science **7**(10), 201199 (2020)
33. Sanderson, J.A., Farrell, S., Ecker, U.K.: Examining the role of information integration in the continued influence effect using an event segmentation approach. PloS one **17**(7), e0271566 (2022)
34. Sanderson, J.A., Gignac, G.E., Ecker, U.K.: Working memory capacity, removal efficiency and event specific memory as predictors of misinformation reliance. Journal of Cognitive Psychology **33**(5), 518–532 (2021)
35. Stanovich, K.E.: Miserliness in human cognition: The interaction of detection, override and mindware. Thinking & Reasoning **24**(4), 423–444 (2018)
36. Susmann, M.W., Wegener, D.T.: The role of discomfort in the continued influence effect of misinformation. Memory & Cognition **50**(2), 435–448 (2022)
37. Susmann, M.W., Wegener, D.T.: How attitudes impact the continued influence effect of misinformation: The mediating role of discomfort. Personality and Social Psychology Bulletin **49**(5), 744–757 (2023)
38. Susmann Mark W., Wegener, D.T.: The role of discomfort in the continued influence effect of misinformation. Memory Cognition **50**, 435–448 (2022)
39. Swire, B., Ecker, U.K., Lewandowsky, S.: The role of familiarity in correcting inaccurate information. Journal of Experimental Psychology: Learning, Memory, and Cognition **43**(12), 1948–1961 (2017)
40. Swire-Thompson, B., Cook, J., Butler, L.H., Sanderson, J.A., Lewandowsky, S., Ecker, U.K.: Correction format has a limited role when debunking misinformation. Cognitive Research: Principles and Implications **6**(1), 83 (2021)
41. Tannenbaum Melanie B., Hepler Justin, Z.R.S.S.L.J.S.W.K.A.D.: Appealing to fear: A meta-analysis of fear appeal effectiveness and theories. Psychological Bulletin **141**(6), 1178–1204 (2015)
42. Watts, J., Moyer-Gusé, E.: Countering narrative misinformation: Investigating the effects of narrative corrections and character trust on story-related knowledge of hpv. Journal of Health Communication **27**(8), 593–602 (2022)
43. Wilkes, A., Leatherbarrow, M.: Editing episodic memory following the identification of error. The Quarterly Journal of Experimental Psychology **40**(2), 361–387 (1988)
44. Williamson, J.B., Drago, V., Harciarek, M., Falchook, A.D., Wargovich, B.A., Heilman, K.M.: Chronological effects of emotional valence on the self-selected retrieval of autobiographical memories. Cognitive and Behavioral Neurology **32**(1), 11–15 (2019)
45. Yonelinas, A.P., Ritchey, M.: The slow forgetting of emotional episodic memories: An emotional binding account. Trends in cognitive sciences **19**(5), 259–267 (2015)