The Dual Nature of Testing: The Effect of Interim Testing on the Misinformation Effect and Its Mechanism (Postprint)

Authors: He Ning, Li Meng, Kang Bin, Wang Mengyun, Yue Yunfan

Date: 2023-10-09T00:00:00+00:00

Abstract

Interim testing exerts two distinct effects on the misinformation effect. Specifically, retrieval enhanced suggestibility (RES) denotes that participants who undergo interim testing exhibit lower accuracy on the final memory test and are more likely to report misinformation; conversely, the protective effect of testing (PET) indicates that interim testing attenuates the misinformation effect and enhances participants' memory performance. A systematic review of extant research demonstrates: First, these two phenomena can be respectively explained by memory reconsolidation theory, attentional capture hypothesis, and retrieval fluency hypothesis (for RES), as well as memory strength theory, retrieval effort theory, and discrepancy detection theory (for PET). These theories differ in both the stage of operation and explanatory perspective, and have been integrated into a novel theoretical model. Second, the dissociation between RES and PET is subject to several potential influencing factors, including original information materials, types of interim testing, and characteristics of misinformation. Finally, future research may continue to explore from two aspects: model validation and research expansion.

Full Text

Preamble

ChinaXiv Cooperative Journal: Advances in Psychological Science
2023, Vol. 31, No. 9, 1626-1641

https://doi.org/10.3724/SP.J.1042.2023.01626

Research Method

The Dual Nature of Testing: The Influence of Interim Testing on the Misinformation Effect and Its Mechanism

HE Ning, LI Meng, KANG Bin, WANG Mengyun, YUE Yunfan
(School of Psychology, Shaanxi Normal University, Xi' an 710062, China)

Abstract

The effect of interim testing on the misinformation effect yields two distinct outcomes. "Retrieval enhanced suggestibility" (RES) refers to the phenomenon where participants who receive an interim test show lower accuracy on final memory tests and are more likely to report misinformation. Conversely, the "protective effect of testing" (PET) describes how interim testing weakens the misinformation effect and improves memory performance. A systematic review of existing research reveals that these two phenomena can be explained by distinct theoretical frameworks: the reconsolidation account, attention capture hypothesis, and retrieval fluency hypothesis (for RES), and the memory strength theory, retrieval effort theory, and discrepancy detection theory (for PET). These theories differ in their stage of action and explanatory perspective, and have been integrated into a new theoretical model. Furthermore, the divergence between RES and PET is influenced by several potential factors, including characteristics of the original information material, type of interim test, and features of the misinformation itself. Future research should explore model validation and research expansion.

Keywords: misinformation, interim test, retrieval enhanced suggestibility, protective effect of testing
Classification Code: B842

Introduction

Misinformation refers to information that is objectively inaccurate [**?**, **?**]. Research has shown that misinformation presented after an event can alter individuals' memory of the original event, a phenomenon known as the misinformation effect [**?**, **?**, **?**, **?**]. In criminal investigations, for instance, eyewitnesses may encounter ambiguous media reports or hear inaccurate discussions about the case after leaving the scene, leading them to doubt their original memory and alter their recollection of the incident. This significantly interferes with case resolution and increases social instability. In today' s information explosion era, such situations are ubiquitous. Indeed, misinformation poses substantial harm to our cognition, emotions, and daily decision-making [**?**, **?**], and research on this topic has grown rapidly in recent years [**?**, **?**, **?**]. Therefore, misinformation research holds both academic value and practical significance.

As researchers continue to explore the misinformation effect, a standard paradigm has emerged, consisting of three phases: an original information phase where participants read or watch textual or video material; a post-event

information or narrative phase where participants receive additional material that may be consistent with, irrelevant to, or contradictory with the original information; and a final test phase where participants answer questions about the original information to assess changes in memory accuracy [?, ?, ?, ?, ?]. However, this paradigm does not adequately simulate real-world situations. Returning to the criminal investigation example, eyewitnesses may have already undergone simple questioning by police when reporting the incident—an immediate retrieval that occurs before encountering misinformation. Based on this observation, researchers have begun to examine this first test, distinct from the final test, and have termed it the "initial test" or "interim test" [?, ?, ?].

The testing effect demonstrates that testing between learning and final assessment improves final test performance [?, ?, ?, ?, ?]. Consequently, researchers hypothesized that interim testing would strengthen memory for the original event and thereby reduce the impact of misinformation. However, early studies yielded contradictory results. On one hand, numerous studies have shown that participants who received interim tests exhibited more severe misinformation effects on final tests [?, ?, ?, ?], a phenomenon termed "retrieval enhanced suggestibility" (RES). Here, suggestibility refers to individuals' susceptibility to misinformation—the likelihood of being misled [?, ?, ?]. On the other hand, under conditions involving cognitive interviews, question framing, or warnings, interim testing has been found to reduce suggestibility [?, ?, ?, ?, ?], a phenomenon called the "protective effect of testing" (PET).

As two contradictory phenomena, RES and PET indicate that the effect of interim testing is not stable. Why does testing enhance suggestibility, and under what conditions does it protect original information? Scholars have yet to reach consensus on this question. Drawing on research findings from the misinformation domain and supplemented by studies from the learning domain, this paper reviews existing theories and related evidence to clarify the nature of interim testing' s influence under different conditions and its underlying mechanisms.

2. How Interim Testing Enhances Suggestibility

RES was first discovered by Chan et al. (2009). In their experiment, participants first watched a hijacking video, then took an interim test (e.g., "What did the terrorist use to knock down the flight attendant?" with the correct answer being "a hypodermic needle"), subsequently read misinformation (the terrorist used a chloroform-soaked cloth to subdue the flight attendant), and finally took a final test with the same questions. Results showed that participants had difficulty retrieving original memories, leading to decreased accuracy on the final test, while simultaneously showing increased reporting of misinformation. Current explanations for RES fall into two categories: first, original memory impairment, where interim testing either damages participants' original memories or blocks access to them, preventing correct reporting; second, enhanced learning of misinformation, where interim testing increases attention to narrative infor-

mation, strengthening encoding and retrieval of misinformation, leading to more misinformation reports on the final test.

2.1 Testing Impedes Access to Original Memory

Participants who receive interim tests report original information less frequently on final tests, possibly because the accessibility of original information is hindered. Indeed, information is highly fragile immediately after encoding and vulnerable to interference, only gradually consolidating over time to form long-term memory. The reconsolidation account [**?**, **?**, **?**] posits that this consolidation process is not irreversible—when reactivated, memories return to a fragile state and require reconsolidation to restabilize. When misinformation is presented during this reconsolidation window, original memories remain unstable and are easily disrupted. In other words, original memories undergo a "consolidation-activation-reconsolidation-new information intrusion" process. Interim testing serves as a strong activation method that allows new information to invade original memories extensively. Particularly when new information contradicts original information, this strong reactivation may trigger overcorrection, leading to forgetting of original information and updating with new information [**?**, **?**].

After interim testing activates original information, the presentation of misinformation can have two possible effects: first, misinformation may overwrite original information, causing complete loss of original memory (the updating hypothesis); second, misinformation may not destroy original information but merely impair its accessibility, preventing successful retrieval on the final test (the inhibition hypothesis, [**?**, **?**]). Currently, the inhibition hypothesis has received broad support. Studies have found that when participants are allowed to provide more than one answer on the final test (the modified modified free recall test, MMFR), the likelihood of correctly reporting original information does not differ between interim test and control groups, and may even be higher for the test group [**?**, **?**, **?**]. McCloskey and Zaragoza (1985) found that when the final test required forced choice between the correct answer and other options (but not misinformation), interim testing had no effect on accuracy. Based on this finding, they argued that memory for original details is not impaired by misinformation. In other words, interim testing merely reduces the accessibility of original information, preventing its successful retrieval on the final test. Consequently, improving accessibility of original information should enhance final test performance. For example, Gordon and Shapiro (2012) used semantic network models to prime key details in original memory by activating related concepts, thereby reducing misinformation effects and improving final test accuracy. Although their study did not involve interim testing of original information, it demonstrated that original and misinformation coexist in memory, and that increasing accessibility of original information reduces suggestibility. To directly demonstrate reduced accessibility, future research could compare differences between recognition and recall. Scully et al. (2017) argued that if participants can recognize but not recall information, this indicates reduced accessibility.

2.2 Testing Enhances Subsequent Learning

Another prominent manifestation of RES is increased reporting of misinformation, indicating that interim testing enhances participants' attention to, learning of, or encoding of misinformation. As early as the 1970s, Tulving and Watkins (1974) found that testing prior learning could enhance learning of new information, a phenomenon known as the "forward testing effect" or "test-potentiated learning" [**?**, **?**, **?**]. This effect has been widely validated in the learning domain [**?**, **?**, **?**, **?**, **?**] and is now emerging in the misinformation domain. Gordon and Thomas (2014) found that compared to participants without interim tests, those who received interim tests could better recall post-event narrative details. When this information was false or inconsistent with original information, RES emerged.

Interim testing may strengthen learning of misinformation through three possible explanations. First, interim testing helps participants better distinguish between original and new information, thereby reducing proactive interference from original memory. Second, interim testing increases attention to details in misinformation related to the test, changing encoding patterns and enhancing learning of new information. Third, interim testing increases retrieval fluency of misinformation, making it more easily retrieved and reported than original information on the final test.

2.2.1 Interference and Separation Chan et al. (2009) proposed that one reason interim testing enhances learning of misinformation is that testing separates two encoding events (original and narrative information), thereby reducing the influence of original memory on new information learning. The phenomenon of testing promoting contextual separation and reducing proactive interference has been widely documented [**?**, **?**, **?**, **?**, **?**, **?**]. In Szpunar et al.' s (2008) classic study, as the amount of prior learning material increased, interference with subsequent learning also increased; testing prior material improved recall accuracy for subsequent material and produced less proactive interference. Recent research further found that interference from prior materials mediates the relationship between interim testing and subsequent learning performance [**?**, **?**, **?**]. Szpunar et al. (2008) explained this using source monitoring and information overload: if prior lists were not tested, participants had to retrieve all previously presented information during testing of the final material, creating excessive information load that could cause confusion between materials. Adding testing helped participants separate lists and distinguish them, which on one hand facilitated better source monitoring, and on the other hand reduced information overload and proactive interference from prior materials, benefiting recall of new information.

In the misinformation domain, interim testing may separate original and narrative information, reducing interference from the former on the latter and increasing recall of misinformation on the final test. However, it should be noted that in the misinformation domain, the relationship between old and new materials

is more clearly correlated and conflicting, and more emphasis is placed on examining episodic memory. Whether interim testing can still promote contextual separation and reduce interference requires further verification.

2.2.2 Learning and Encoding In testing effect research, testing has been repeatedly shown to enhance learning of new material [**?**, **?**, **?**, **?**, **?**]. Gordon and Thomas (2017) examined participants' memory for narrative information on the final test and found that the interim test group recalled more post-test information, providing direct evidence that interim testing enhances narrative information learning. Currently, there are three different theoretical explanations: the attention capture hypothesis, encoding reset theory, and encoding strategy theory.

The attention capture hypothesis [**?**, **?**, **?**, **?**] posits that interim testing increases the attractiveness of test-related information, thereby strengthening participants' learning of misinformation. Gordon and Thomas (2014) found that participants who received interim tests spent significantly longer reading narrative information than those without interim tests, suggesting that interim testing may increase attention to narrative information. When a secondary task was added during the misinformation phase to distract participants, the difference in misinformation rates between test and control groups disappeared, effectively canceling out RES [**?**, **?**, **?**]. This evidence suggests that additional attention to misinformation can explain RES.

However, this explanation is incomplete. For example, when an immediate final test was administered, both the interim test group and a detail-emphasis group (where attention was drawn through red underlining) showed similar RES; but after a 48-hour delay, the interim test group performed better than both the detail-emphasis and control groups [**?**, **?**]. This indicates that capturing attention alone cannot produce the same effect as interim testing, suggesting that more complex mechanisms underlie RES.

Encoding reset theory posits that encoding efficiency continuously declines during sustained learning, while testing can separate contexts and reduce memory load and inefficient encoding, making learning of new material as effective as early encoding [**?**, **?**, **?**, **?**]. Cognitive neuroscience research has found that alpha power continuously increases with sustained learning but returns to initial levels after testing, showing better learning effects [**?**, **?**], providing strong physiological evidence for this theory. Currently, this theory has not been applied to misinformation research, but future studies could combine cognitive neuroscience approaches to investigate RES.

Cho et al. (2017) proposed encoding strategy theory, suggesting that testing enhances new learning through changes in encoding strategies. Through interim testing, participants recognize deficiencies in their memory and consequently exert more effort in subsequent learning, seeking better memory methods [**?**, **?**, **?**, **?**]. Research has shown that interim testing can prompt participants to use

more effective encoding strategies [?, ?]. For example, compared to no testing or restudy, interim testing enhances semantic clustering in subsequent learning [?, ?, ?]. Yang et al. (2022) also found that interim testing enhances learning of new material by improving temporal clustering in subsequent learning. Research in the misinformation domain has found that different encoding strategies can affect participants' suggestibility [?, ?]. Therefore, interim testing may enhance participants' learning of misinformation through changes in encoding strategies.

Motivation theory. The above explanations primarily adopt a cognitive perspective, clarifying testing' s influence on underlying learning mechanisms but neglecting learners as active agents with their own motivations and needs. Motivation theory addresses this gap from perspectives such as retrieval failure and test expectancy.

When participants fail to recall event details (i.e., retrieval failure), they become aware of their memory' s fragility, thereby increasing their attention to and learning of subsequent material [?, ?, ?]. Richland et al. (2009) argued that retrieval failure strengthens retrieval pathways between questions and answers and encourages more elaborate deep processing. Grimaldi and Karpicke (2012) suggested that retrieval attempts after test failure activate a set of possible answers, and when restudied material overlaps with this set, learning of corresponding items is enhanced. If misinformation happens to appear in this set, RES emerges. Additionally, interim testing may activate higher test expectancy —the prediction that testing will continue—which is termed "test expectancy theory"[?, ?, ?]. Research has found that test expectancy, especially expecting more difficult subsequent tests, enhances participants' learning of material [?, ?, ?]. In the misinformation paradigm, participants have no test expectancy when encoding original information but do have it (derived from interim testing) when encoding misinformation, leading them to learn original information less well than misinformation, thereby producing RES.

2.2.3 Retrieval Fluency The enhancement of new information learning by interim testing cannot fully explain RES because testing also enhances learning of tested information (original information) [?, ?, ?, ?, ?]. So how does misinformation win in the competition between two similarly strengthened memories? The retrieval fluency hypothesis proposed by Thomas et al. (2010) may provide insight. Retrieval fluency refers to the ease with which information can be retrieved from memory. This hypothesis suggests that interim testing enhances attention to and encoding of post-test material, thereby increasing its retrieval fluency on the final test, making it more likely to be reported.

Thomas and Gordon conducted a series of studies to test this theory. In their research, Thomas et al. (2010) used confidence and response time as indicators of retrieval fluency. They argued that confidence is influenced by retrieval ease—the more fluent, the more confident; response delay provides a more direct assessment of retrieval fluency, with faster responses indicating greater fluency. Results showed that participants in the interim test group responded

faster when selecting misinformation and showed high confidence, indicating that interim testing enhanced misinformation retrieval fluency. However, when they suppressed retrieval fluency by warning participants that narrative information sources were uncertain and unverifiable, participants took longer to make choices, confidence inflation disappeared, and RES was weakened. Another experiment found that when participants were required to provide two answers on the final test, the interim test group showed better memory for both the original video and misinformation, suggesting that testing enhanced learning of both pre-test and post-test information. However, because misinformation was more easily retrieved, RES occurred [**?**, **?**]. Thomas et al. (2017) found that after a 48-hour delay, the interim test group was even more accurate than the control group, with no significant difference in misinformation reporting rates, possibly because misinformation retrieval fluency diminished after 48 hours.

Overall, existing theoretical explanations for RES can be categorized from two perspectives: "encoding" and "retrieval." The "encoding" perspective suggests that interim testing improves learning of subsequent material, including motivation theory, attention capture hypothesis, encoding reset theory, and encoding strategy theory. Specifically, retrieval failure and test expectancy provide motivational factors for enhanced subsequent encoding, making participants "want" to expend more effort during the narrative information phase; encoding reset and attention capture provide important preconditions for enhanced learning, enabling participants to "be able to" engage in detailed encoding of misinformation; and changed encoding strategies facilitate enhanced learning, making participants "find it easy" to encode efficiently during the narrative information phase. The "retrieval" perspective suggests that the retrieval process in interim testing influences RES. First, according to the reconsolidation account, retrieval destabilizes original memories, making them more vulnerable to misinformation. Second, retrieval itself promotes contextual separation and reduces interference from prior material. Finally, the retrieval fluency hypothesis posits that RES occurs because misinformation is more easily retrieved on the final test, leading participants to respond faster and with greater confidence when faced with misinformation options.

2.3 The Pathway of RES

As described above, although researchers have provided diverse theoretical explanations for RES, they have overlooked relationships between theories and cannot capture the full picture of RES mechanisms. Based on this, we propose a dual-pathway model for learning of original and narrative information (see Figure 1 [Figure 1: see original paper]).

According to this model, interim testing simultaneously influences memory for both original and misinformation, ultimately leading to RES. In the first pathway, based on the testing effect [**?**, **?**, **?**, **?**, **?**], interim testing should enhance memory for original information. However, the emergence of misinformation reverses this effect. Although research has shown that original memory is not

completely destroyed [?, ?, ?, ?], misinformation still reduces accessibility of original information, thereby decreasing final test accuracy.

In the second pathway, interim testing enhances learning of misinformation through multiple mechanisms. First, interim testing separates the two information materials, reducing interference from original memory on misinformation learning [?, ?]. Second, retrieval failure and test expectancy motivate participants to encode narrative information effortfully [?, ?, ?, ?, ?]. Finally, interim testing increases participants' attention [?, ?, ?, ?], increasing reading time for narrative information while simultaneously enhancing encoding of misinformation [?, ?, ?]. When encoding of misinformation is enhanced, participants can more easily retrieve it on the final test, i.e., retrieval fluency increases [?, ?, ?].

Overall, the emergence of RES ultimately depends on the relative strength of original information accessibility versus misinformation retrieval fluency. Neither pathway can be viewed in isolation. Thomas et al. (2010) argued that RES occurs because individuals terminate further retrieval needed to recall original information after retrieving more fluent misinformation. This suggests that it is original information' s failure in the "fluency competition" that leads to its abandonment. From this perspective, improving original information accessibility or reducing misinformation retrieval fluency, or providing opportunities for further retrieval of original information, could reduce RES [?, ?, ?, ?, ?, ?].

3. How Interim Testing Provides Protection

Although RES has been widely validated, researchers have also discovered PET under certain conditions. For example, in the household scene paradigm, participants first view a photo of a household scene (e.g., a "desk") containing various items but missing the most likely objects (e.g., "paper" and "pen"). They then encounter misinformation presented in audio or text form (e.g., "He could finally concentrate; he took a bite of an apple, grabbed a pen, and began writing"). Results showed that participants in the interim test group were less likely to report non-existent items (e.g., "pen") on the final test, meaning interim testing reduced suggestibility [?, ?].

In fact, RES and PET address the same question—"Does interim testing enhance or reduce suggestibility?" —and thus appear to be two contradictory manifestations of the same phenomenon. This means PET research can supplement RES discussions. First, existing RES theories and research can only explain how to reduce interim testing' s negative effects to baseline levels observed without testing, but cannot explain how to leverage testing' s positive effects. In other words, RES theories can explain how to eliminate RES but not when PET occurs [?, ?, ?, ?, ?, ?]. Second, exploring how to harness the positive effects of interim testing to reduce suggestibility has broader practical significance. However, no mature theoretical framework currently exists to explain PET, leaving a gap in discussions of interim testing' s mechanisms. Therefore, we review relevant research and theories, including memory strength, retrieval effort the-

ory, discrepancy detection theory, and transfer-appropriate processing theory, to provide preliminary exploration of PET.

3.1 Memory Strength

Some researchers argue that individuals are susceptible to misinformation because their memory strength is insufficient, leaving them uncertain about their judgments and more likely to accept misinformation. Therefore, "strong" memory for original events enables people to more easily detect discrepancies between original events and misinformation, leading them to reject misinformation and make correct choices [**?**, **?**, **?**, **?**, **?**]. In the misinformation paradigm, the standard for measuring whether original memory is strong (memory strength) is whether the interim test was answered correctly [**?**, **?**]. Chan and Langley (2011) found that when interim test responses were correct, the probability of reporting misinformation did not differ significantly from participants who received no interim test, suggesting that strong memory counteracted RES. In LaPaglia and Chan' s (2013) study, when misinformation was presented in question stem format, participants who answered interim test questions correctly showed higher final test accuracy than those who answered incorrectly, even resisting misinformation effects. Gabbert et al. (2012) also argued that PET occurs because testing strengthens original memory strength, enabling better discrepancy detection. Therefore, when individuals' memory for original information is sufficiently accurate and robust, they can resist misinformation effects, which may be one reason PET emerges. We hypothesize that when individuals have strong enough memory for original information, interim testing will exert its protective effect and reduce suggestibility; when original memory strength is insufficient, individuals are more easily misled by misinformation, resulting in RES.

3.2 Retrieval Effort Theory

In studies that found PET, when participants expended more effort to retrieve information during interim testing, their memory for original information was enhanced, making it easier to identify misinformation and reduce suggestibility. The most direct evidence comes from Pansky and Tenenboim (2011), who categorized information into basic-level (e.g., "chair") and subordinate-level (e.g., "wooden chair") and conducted gist tests and verbatim tests at two different depths. Gist tests required participants to answer at the basic level, while verbatim tests required answers at the subordinate level. Results showed that both test types produced testing effects, but only verbatim tests reduced suggestibility.

Regarding this, retrieval effort theory posits that the testing effect results from cognitive effort expended during the retrieval learning process [**?**, **?**, **?**, **?**, **?**]. This theory has received extensive support in the testing effect domain. For example, participants who engaged in deep retrieval during initial tests showed better subsequent recognition performance [**?**, **?**]. To further explain this phe-

nomenon, researchers have developed the desirable difficulty theory [?, ?, ?] and the bifurcation model theory [?, ?]. The desirable difficulty theory distinguishes between storage strength and retrieval strength. The former reflects the relative durability of memory traces or persistence of learning; the latter reflects the momentary accessibility of memory traces—the ease with which information can be retrieved, similar to retrieval fluency. Meanwhile, the theory assumes an inverse relationship between retrieval strength and storage strength: only when retrieval practice involves high effort (low retrieval strength) and successful retrieval can it enhance memory strength (high storage strength) and promote long-term learning [?, ?]. In the bifurcation model hypothesis, all items initially show a normal distribution of retrieval strength. During repeated study, retrieval strength increases modestly, maintaining a normal distribution; during retrieval practice, successfully retrieved items increase in retrieval strength while failed items decrease, with the magnitude of increase depending on retrieval difficulty—greater difficulty produces larger increases [?, ?, ?].

According to this theory, when participants expend cognitive effort to retrieve original information during interim testing, it can enhance memory strength for original information. Especially for information with insufficient memory strength, when more effort is expended and retrieval succeeds, memory strength improves substantially; if retrieval fails during interim testing, participants may lose confidence and abandon further retrieval attempts on subsequent tests. When original memory is strong enough, participants may retrieve it easily without much effort, but correspondingly, these memories cannot be strengthened. Thus, an interactive relationship exists between retrieval effort and memory strength. To harness the positive effects of interim testing, research should increase interim test difficulty or retrieval effort.

3.3 Discrepancy Detection Theory

Discrepancy detection theory posits that detecting discrepancies between original and misinformation can reduce the misinformation effect—participants who notice discrepancies are more likely to identify and reject misinformation, thereby improving final test accuracy [?, ?, ?, ?, ?, ?]. Research shows that when participants conduct discrepancy detection on each piece of narrative information while reading, suggestibility can be effectively reduced [?, ?, ?].

Studies on contradictory and additive misinformation can provide supplementary evidence for discrepancy detection theory. Contradictory misinformation conflicts with original events (i.e., “calling a deer a horse”), while additive misinformation includes information present in the narrative but absent from the original event (i.e., “creating something from nothing”). The key difference is that contradictory misinformation conflicts with original memory, making it more suitable for discrepancy detection. Research has found that compared to additive misinformation, participants report less contradictory misinformation [?, ?, ?], and this difference is more pronounced in the interim test group [?, ?]. This suggests that interim testing further enhances discrepancy detection.

Studies have shown that compared to restudy, testing helps participants detect changes in learning material [**?**, **?**]. Additionally, in Gordon et al.' s (2015) study, when participants answered interim test questions correctly, they spent more time reading misinformation, indicating they detected inconsistencies between original memory and narrative information [**?**, **?**, **?**]. Therefore, when discrepancy recall tasks were added to the final test, participants in the interim test group were better able to detect differences between original and false narrative information [**?**, **?**]. In summary, interim testing can reduce misinformation effects by promoting discrepancy detection.

3.4 Transfer-Appropriate Processing Theory

LaPaglia and Chan (2012) discovered PET in research on face recognition. They argued that interim testing "sealed" correct memories containing almost no misinformation, and that final testing might activate these memories to resist misinformation. In other words, testing may "transfer" some original memory, but how does this transfer occur, and what factors influence its effectiveness? Transfer-appropriate processing theory from the learning domain may provide empirical support.

Transfer-appropriate processing theory posits that testing allows participants to practice retrieval operations, thereby performing better on subsequent retrieval —in other words, transferring "testing skills" to later tests [**?**, **?**, **?**, **?**]. From this, we can hypothesize that testing works best when interim and final tests are identical. This theory has received some empirical support. On one hand, whether test types (e.g., short answer or multiple choice) match does affect test performance. Meta-analyses by Adesope et al. (2017) and Yang et al. (2021) found larger testing effects when formats matched. On the other hand, the degree of cue matching between the two tests also influences testing effects. For instance, Veltre et al. (2015) found better memory performance when interim and final tests used the same cue type (semantic vs. spelling cues). This suggests that matching interim and final tests in type or other conditions is effective and may influence the "relay transfer" of testing to original information, thereby affecting PET.

However, some questions remain about this theory. For example, Kang et al. (2007) used two test formats (short answer and multiple choice) as initial tests to examine performance under matched and mismatched test formats, finding that short-answer testing promoted better final test performance regardless of format. Rowland' s (2014) meta-analysis obtained similar results—testing effects were larger when interim tests used free recall rather than recognition tests. This seems to suggest that testing effects depend on retrieval effort rather than format matching. Therefore, whether test format matching can enhance PET in the misinformation domain requires further verification.

Correspondingly, theoretical explanations for PET can also be summarized from "encoding"and "retrieval"perspectives. From the "encoding"perspective, memory

strength theory posits that "strong" original memory can reduce suggestibility. During interim testing, participants have opportunities for additional encoding of original information, which enhances original memory [?, ?] and promotes PET. Additionally, interim testing reduces suggestibility by promoting discrepancy detection, manifested as spending more time reading misinformation. This indicates that discrepancy detection enhances some form of critical encoding of misinformation.

From the "retrieval" perspective, retrieval effort theory focuses on direct benefits of retrieval—examining the degree of effort in the retrieval process and assuming that more cognitive effort leads to greater memory strength enhancement. Transfer-appropriate processing theory, in contrast, views "retrieval" as a special skill that teaches participants "how to test," thereby improving performance on subsequent tests. Therefore, transfer-appropriate processing theory emphasizes whether interim and final test formats are consistent.

3.5 The Pathway of PET

Although no mature explanation for PET exists in the misinformation domain, the above theories clearly provide preliminary exploration from different perspectives. To further integrate and develop explanations for PET, we incorporate existing mainstream research into a unified theoretical framework (see Figure 2 [Figure 2: see original paper]). First, higher participant memory strength and retrieval effort required by testing both help exert testing' s protective effect, while more effortful retrieval can also improve memory performance [?, ?, ?, ?, ?, ?]. Second, beyond direct effects, strong memory and difficult retrieval enhance participants' discrepancy detection ability, making them more likely to detect misinformation and thereby reduce suggestibility [?, ?, ?, ?]. Finally, according to transfer-appropriate processing theory, the influence of retrieval effort on PET may be moderated by test format matching. If interim and final test formats are consistent, retrieval effort may be more helpful in enhancing testing' s protective effect [?, ?, ?, ?].

3.6 Integrated Model of RES and PET

The above analysis shows that theories related to RES and PET are not only rich in content but also provide diverse interpretations of interim testing' s influence on the misinformation effect and its mechanisms from different perspectives. Overall, relationships among these theories can be summarized in two aspects: first, the stage at which interim testing effects appear in the misinformation paradigm differs; second, they explain the mechanisms of interim testing on misinformation from different "encoding" or "retrieval" perspectives. For clarity, we compare and integrate relevant theories along two dimensions— "stage of action" and "explanatory perspective" —as shown in Table 1 .

Table 1 Comparison of Theories for RES and PET

Phenomenon	Theory	Stage of Action	Explanatory Perspective
RES	Reconsolidation Account	Reactivation	Encoding/Retrieval
RES	Interference & Separation	Encoding	Encoding
RES	Attention Capture Hypothesis	Encoding	Encoding
RES	Encoding Reset Theory	Encoding	Encoding
RES	Encoding Strategy Theory	Encoding	Encoding
RES	Retrieval Fluency Hypothesis	Final Test	Retrieval
PET	Memory Strength	Interim Test	Encoding
PET	Retrieval Effort Theory	Interim Test	Retrieval
PET	Discrepancy Detection Theory	Final Test	Encoding/Retrieval
PET	Transfer-Appropriate Processing Theory	Final Test	Retrieval

However, as two contradictory phenomena, RES and PET cannot be studied in isolation, and integration and communication between explanations are essential. Therefore, as shown in Figure 3 [Figure 3: see original paper], we have unified theories for both phenomena based on the logical sequence and possible causal relationships among different theories, proposing an overarching model that encompasses major mainstream explanations to provide more in-depth and comprehensive investigation of the mechanisms through which interim testing influences the misinformation effect.

We propose that interim testing primarily influences the misinformation effect

through three pathways. The first pathway explains interim testing' s influence on original information. First, given the widely established testing effect [?, ?, ?, ?, ?], interim testing should enhance memory for original information. According to retrieval effort theory, when interim testing is more difficult, it better enhances storage strength of original information [?, ?, ?]. However, this process depends on whether retrieval succeeds during interim testing [?, ?, ?, ?, ?, ?]. When original information is successfully retrieved and memory strength increases, participants can more easily access original information on the final test.

The second pathway concerns the influence of misinformation. Consistent with the RES model, we will not elaborate further here. The third pathway simultaneously affects both original and misinformation. First, interim testing creates contextual changes, leading to encoding reset, thereby separating original and misinformation encoding events [?, ?, ?, ?, ?, ?, ?, ?]. Second, this clearer distinction helps participants detect discrepancies between information sources, thereby reducing suggestibility [?, ?, ?, ?, ?, ?, ?, ?].

This model integrates theoretical research on RES and PET, providing a more detailed exposition of the mechanisms through which interim testing influences the misinformation effect. It offers several theoretical values: First, it places RES and PET research within the same framework, helping to focus research directions and highlight key issues. Previous studies treated RES and PET as contradictory phenomena, often biased in their explanations and insufficiently comprehensive in considering influencing factors, making it difficult to effectively accumulate research findings. A unified research framework can provide clues for researchers to explain phenomena, help the misinformation domain converge research directions, and better uncover theoretical value behind studies. Second, the model connects theories from learning and misinformation domains, providing a bridge for interdisciplinary integration. The learning domain has mature theories, while the misinformation domain has diverse findings, making this connection valuable in two ways: it fills theoretical gaps in the misinformation domain, providing a solid theoretical foundation for related research; and it expands the application scope of testing effect theories, enriching their theoretical connotations.

4. Potential Influencing Factors for the Separation of RES and PET

Based on the above explanatory framework, we can see that under different stages of the misinformation paradigm, the nature of interim testing' s effect depends on the types or characteristics of original information, interim testing, and misinformation. Thus, potential influencing factors for the separation of RES and PET can be broadly divided into three categories:

4.1 Original Information Material

First, for different types of original material, the nature of interim testing' s effect differs. Under otherwise consistent conditions, studies using video materials have shown RES [?, ?, ?, ?], while those using picture materials have shown PET [?, ?, ?]. According to mental model theory, when original events have narrative coherence, logic, and completeness, people form mental models about how events unfold; and people prefer to maintain coherent and complete mental models even if they are incorrect [?, ?]. Therefore, if a component of the model is missing, creating a gap, participants may accept incorrect information to maintain mental model completeness [?, ?]. Correspondingly, in studies using video materials, individuals focus more on the dynamic development of the storyline and thus construct mental models rich in contextual information. If original memory strength is insufficient or inaccessible, participants are more likely to accept misinformation to fill gaps in their mental models. However, in studies using picture materials, original memory and narrative information consist of isolated pieces without contextual relationships [?, ?, ?]. In this case, individuals cannot construct coherent, complete mental models and thus find it difficult to integrate misinformation into original memory, reducing suggestibility and producing PET.

Additionally, the contextual information content of misinformation also affects suggestibility. Research has found that RES occurs when misinformation is presented in narrative form, while PET emerges when misinformation is introduced in question stem format [?, ?]. LaPaglia (2013) argued that this is because narrative misinformation contains more contextual details, while question stems provide isolated information. Correspondingly, when misinformation is presented sentence-by-sentence rather than in paragraph form, the interim test group recalled more correct information [?, ?]. Therefore, the richness of misinformation content, or its contextual information load, influences suggestibility. Specifically, richer content leads to RES, while isolated information leads to PET [?, ?]. This may be because when narratives contain excessive information, individuals have difficulty conducting discrepancy detection, but when information is presented in isolation, discrepancies become more salient.

4.2 Interim Test Type

The type of interim test influences its effectiveness. Memon et al. (2010) used cognitive interviews (a questioning method involving extensive and detailed memory retrieval, including context reinstatement, reporting everything, changing perspective, and reversing order) and found that interviews increased reporting of correct details and reduced false agreement with fabricated items one week later. Gabbert et al. (2012) found that completing self-administered interviews (including details such as event sequence, person appearance, and even scene sketches) improved accuracy and made individuals more resistant to narrative misinformation and misleading questions. However, when cued recall was used, interim testing often led to stronger suggestibility [?, ?, ?]. According to re-

trieval effort theory, this is because cognitive interviews and self-administered interviews require more effortful and deeper retrieval than cued recall, thereby strengthening original memory and enabling better rejection of misinformation.

Another explanation for why cognitive interviews and self-administered interviews reduce suggestibility comes from research on context reinstatement effects. Context reinstatement refers to the phenomenon where people recall more event details when the retrieval context matches the encoding context [**?**, **?**, **?**]. Research has found that context reinstatement can influence individuals' suggestibility [**?**, **?**, **?**]. When cognitive interviews and self-administered interviews (including context reinstatement, scene sketches, and other methods) are used, interim testing can help participants effectively retrieve the encoding context of original memory, thereby improving original information recall rates [**?**, **?**, **?**, **?**]. This process facilitates integration of original information with new test contexts, making it easier for participants to retrieve original information on final tests (with the same test context as interim tests), thus producing PET.

4.3 Misinformation Characteristics

The misleading nature of misinformation is inseparable from its own characteristics. For example, compared to central details (more vivid, attention-grabbing key content), memory for peripheral details (less important information in events) is more easily misled. Wilford et al. (2014) found that for central details, test and control groups performed similarly, with no RES observed; for peripheral details, RES was only canceled when interim test responses were correct. In fact, in their study, central details were those reported more frequently by most participants, indicating better memory strength that could better resist misinformation interference. Other studies have also found that memory for central details is generally less susceptible to misinformation than peripheral details [**?**, **?**]. According to memory strength theory, central details, due to their inherently better memory strength, are less vulnerable to misinformation interference; for peripheral details that are easily overlooked in events, individuals' own memories are less clear and robust, making them more susceptible to misinformation.

5. Summary and Outlook

This paper provides a theoretical discussion of the two different effects of interim testing on the misinformation effect. RES occurs because interim testing impedes access to original information and promotes separation, increases attention, and motivates participants to learn misinformation. However, when original memory is sufficiently strong, or when participants are encouraged to engage in more detailed retrieval, reminded to attend to information discrepancies, or helped to transfer strategies or memories, testing can exert its protective effect. Based on our review of relevant theories, we propose pathway diagrams for both RES and PET and construct an integrated model of interim testing'

s influence on the misinformation effect, while analyzing potential factors influencing the separation of RES and PET. Building on our theoretical model, future research can be developed in two directions.

5.1 Model Validation

The model we propose integrates relatively independent research findings and theoretical foundations from different domains, providing a relatively comprehensive explanation for phenomena discovered in existing research. However, the model is still in its initial development stage. The relationships among theories, interactions between pathways, and whether other key variables and influencing factors exist all require more targeted theoretical and empirical research.

First, future research needs to directly test different pathways. Regarding the original information and misinformation pathways, the relationship from "original information memory strength" to "original information accessibility" and from "testing enhances new learning" to "misinformation retrieval fluency" is essentially an "encoding-retrieval" relationship. Future research could manipulate depth of processing [**?**, **?**] or add distraction tasks [**?**, **?**, **?**] to alter learning levels of information, and measure retrieval fluency and accessibility through response time and confidence level on final tests [**?**, **?**] to examine changes in the misinformation effect. On this basis, future research could use mediation analysis to test whether interim testing affects the misinformation effect by influencing encoding levels of original and misinformation, changing accessibility of original information and retrieval fluency of misinformation.

Regarding the third pathway, the relationship between contextual separation and discrepancy detection also warrants further examination. Previous research found that interference from prior material only decreases when participants detect discrepancies between information sources; otherwise, interference remains [**?**, **?**]. However, in that study, discrepancy detection was conducted on whole materials rather than individual items, so whether the mechanisms of interference and separation must operate through discrepancy detection requires more evidence. Future research could examine participants' memory for misinformation on final tests and measure proactive interference by counting the number of original information items reported, while adding discrepancy detection tests after final tests (asking participants whether they noticed discrepancies in each piece of information, [**?**, **?**]) to validate this pathway.

Second, future research needs to combine different pathways to comprehensively explore interim testing' s mechanisms. On one hand, different pathways are not independent and may interact. For example, researchers have proposed that strong original memory promotes discrepancy detection [**?**, **?**, **?**, **?**], but their studies did not provide direct evidence for this assumption. Gabbert et al. (2012) found that higher accuracy (i.e., better memory) correlated negatively with misinformation reporting rates, but this alone is insufficiently direct and complete

evidence. Moreover, whether the encoding level of misinformation also promotes discrepancy detection requires further investigation. On the other hand, different pathways operate simultaneously, but their relative contributions remain unknown. Yang et al. (2022) used mediation analysis to examine both independent and joint effects of encoding strategies and proactive interference on the forward testing effect, finding that proactive interference played a more important role than strategy change. Therefore, future research could adopt Yang et al.' s (2022) approach to quantitatively analyze the nature and magnitude of interim testing' s influence on the misinformation effect across the three pathways under different conditions, thereby providing focus for subsequent research or intervention development.

5.2 Research Expansion

While researchers continuously refine objects of study and enrich experimental evidence, they have also provided preliminary exploration for the intersection of learning and misinformation domains. However, existing research retains characteristics of the learning domain in theoretical foundation and variable selection, making it difficult to resolve contradictions between different phenomena and limiting further extension of theoretical and practical value. Therefore, future research should broaden perspectives and develop subsequent studies from two aspects: introducing new variables and developing intervention methods.

First, the current research scope is relatively narrow and needs to further examine individual differences and social factors. On one hand, individual differences influence how people process misinformation. For example, the effect of interim testing may be influenced by individuals' need for cognition. Research has found that individuals with high need for cognition are less susceptible to misinformation [**?**, **?**]. Need for cognition refers to a trait reflecting individuals' willingness to invest effort in deep cognitive processing [**?**, **?**]. High-need-for-cognition individuals tend to engage in more elaborate thinking and recall more information in memory tests compared to low-need-for-cognition individuals [**?**, **?**, **?**]. This suggests that high-need-for-cognition individuals may expend more effort retrieving information during interim testing, thereby strengthening original memory. Additionally, Leding and Antonio (2019) argued that high-need-for-cognition individuals are better able to reject misinformation because they expend greater effort on discrepancy detection. Therefore, we predict that need for cognition influences the effect of interim testing on the misinformation effect through retrieval effort and discrepancy detection. Future research could further explore how individual factors such as cognition, emotion, and attitude influence RES or PET.

On the other hand, misinformation in real life often carries "social" characteristics. For example, when misinformation is introduced socially, individuals are influenced by misinformation from co-witnesses, a phenomenon known as "memory conformity" considered a form of misinformation effect [**?**, **?**, **?**]. Moreover, socially sourced misinformation causes greater harm than non-socially sourced

misinformation [?, ?]. Research has found that being aware of discrepancies between original information and socially sourced misinformation can reduce memory conformity, but even when participants detect discrepancies, they may still report misinformation because they lack confidence in their own memory [?, ?]. Interim testing, as an operation that can strengthen original memory and promote discrepancy detection, may help reduce memory conformity effects. Furthermore, future research could examine factors such as the social nature of information, its source, and individuals' social identities and social environments, thereby broadening the scope of interim testing research to better exert testing' s protective effect and achieve meaningful practical and theoretical depth.

Second, to expand research significance, effective intervention methods need to be developed. Based on existing research, we predict that increasing participants' retrieval effort, encouraging discrepancy detection, and guiding participants to recall their initial responses can exert testing' s protective effect and reduce suggestibility. However, these are only short-term experimental manipulations. To improve individuals' ability to resist misinformation, long-term effective interventions need to be developed. For example, training that reminds participants "human memory is unreliable" can reduce their suggestibility [?, ?]. Researchers believe one explanation is that this memory training promotes discrepancy detection. Unfortunately, in their study, interim testing did not affect the misinformation effect, so whether memory training can influence RES or PET by promoting discrepancy detection requires further research.

Previous studies have found that warnings may be an effective intervention. In misinformation and interim testing research, when participants are told "narrative information sources are uncertain and therefore cannot be verified for accuracy," their suggestibility decreases [?, ?, ?, ?, ?, ?, ?, ?]. This may be because warnings encourage participants to ignore misinformation retrieval fluency, thereby engaging in more effortful retrieval [?, ?] and better detecting discrepancies between original and misinformation [?, ?, ?]. Chan et al. (2022) argued that warnings allow participants to "forget" narrative information, reducing further processing of misinformation, similar to "directed forgetting." However, if participants do not actually encounter misinformation, warnings may cause them to doubt correct information as well, an effect called "tainting the truth" [?, ?]. Therefore, whether warnings are equally effective in real-world situations requires further verification. Future research could build on existing theoretical evidence to develop intervention methods that are both scientifically sound and effective.

References

曹晓君, 涂燊, 陈旭. (2015). 记忆暗示感受性的重复测试效应研究. 中国临床心理学杂志, 23(6), 980−984+1008.

郭秀艳, 李荆广. (2007). 误导信息干扰引发的错误记忆研究. 心理科学, 30(4), 814−819.

王红椿, 刘鸣, 张积家, 向祖强. (2008). 目击证人暗示感受性的研究及进展. 心理科学, 31(3), 744−747.

王析蕾, 贾宁. (2021). 错误信息效应的主要理论与影响因素的述评与展望. 心理技术与应用, 9(1), 52−64.

王植洵, 张锦坤. (2017). 提取练习效应的产生机制: 编码、提取的整合演变. 心理科学进展, 25(9), 1512−1520.

温家林, 张增一. (2018). 错误信息的产生、传播及识别和控制——错误信息已有研究评述. 科学与社会, 8(3), 108−122.

徐洁, 周宁. (2010). 认知需求对个体信息加工倾向性的影响. 心理科学进展, 18(4), 685−690.

张锦坤, 白学军, 杨丽娴. (2008). 国外关于测试效应的研究概述. 心理科学进展, 16(4), 661−670.

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research*, 87(3), 659−701.

Agarwal, P. K., & Roediger, H. L. III. (2011). Expectancy of an open-book test decreases performance on a delayed closed-book test. *Memory*, 19(8), 836−852.

Bailey, N. A., Olaguez, A. P., Klemfuss, J. Z., & Loftus, E. F. (2021). Tactics for increasing resistance to misinformation. *Applied Cognitive Psychology*, 35(4), 863−872.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56−64). New York, NY: Worth.

Bjork, R. A. (1975). Retrieval as a memory modifier: An interpretation of negative recency and related phenomena. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 123−144). Hillsdale, NJ: Erlbaum.

Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. Healy, S. Kosslyn, & R. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes* (Vol. 2, pp. 35−67). Hillsdale, NJ: Erlbaum.

Blank, H., & Launay, C. (2014). How to protect eyewitness memory against the misinformation effect: A meta-analysis of post-warning studies. *Journal of Applied Research in Memory and Cognition*, 3(2), 77−88.

Bufe, J., & Aslan, A. (2018). Desirable difficulties in spatial learning: Testing enhances subsequent learning of spatial information. *Frontiers in Psychology*, 9, 1701.

Bulevich, J. B., Gordon, L. T., Hughes, G. I., & Thomas, A. K. (2022). Are witnesses able to avoid highly accessible misinformation? Examining the efficacy of different warnings for high and low accessibility postevent misinformation. *Memory and Cognition*, 50(1), 45−58.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42(1), 116−131.

Carpenter, S. K. (2011). Semantic information activated during retrieval contributes to later retention: Support for the mediator effectiveness hypothesis of the testing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(6), 1547−1552.

Carneiro, P., Lapa, A., & Finn, B. (2021). Memory updating after retrieval: When new information is false or correct. *Memory*, 29(9), 1156−1175.

Chan, J. C. K., & Langley, M. M. (2011). Paradoxical effects of testing: Retrieval enhances both accurate recall and suggestibility in eyewitnesses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(1), 248−255.

Chan, J. C. K., Manley, K. D., Davis, S. D., & Szpunar, K. K. (2018). Testing potentiates new learning across a retention interval and a lag: A strategy change perspective. *Journal of Memory and Language*, 102, 83−96.

Chan, J. C. K., Manley, K. D., & Lang, K. (2017). Retrieval enhanced suggestibility: A retrospective and a new investigation. *Journal of Applied Research in Memory and Cognition*, 6(3), 213−229.

Chan, J. C. K., O' Donnell, R., & Manley, K. D. (2022). Warning weakens retrieval-enhanced suggestibility only when it is given shortly after misinformation: The critical importance of timing. *Journal of Experimental Psychology: Applied*, 28(4), 694−716.

Chan, J. C. K., Thomas, A. K., & Bulevich, J. B. (2009). Recalling a witnessed event increases eyewitness suggestibility: The reversed testing effect. *Psychological Science*, 20(1), 66−73.

Cho, K. W., Neely, J. H., Crocco, S., & Vitrano, D. (2017). Testing enhances both encoding and retrieval for both tested and untested items. *Quarterly Journal of Experimental Psychology*, 70(7), 1211−1235.

Choi, H., & Lee, H. S. (2020). Knowing is not half the battle: The role of actual test experience in the forward testing effect. *Educational Psychology Review*, 32(3), 765−789.

Dang, X., Yang, C., & Chen, Y. (2021). Age difference in the forward testing effect: The roles of strategy change and release from proactive interference. *Cognitive Development*, 59, 101079.

Drohan-Jennings, D. M., Roberts, K. P., & Powell, M. B. (2010). Mental context reinstatement increases resistance to false suggestions after children have experienced a repeated event. *Psychiatry, Psychology and Law*, 17(4), 594−606.

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.

Gabbert, F., Hope, L., Fisher, R. P., & Jamieson, K. (2012). Protecting against misleading post-event information with self-administered interview. *Applied Cognitive Psychology*, 26(4), 568−575.

Gabbert, F., Memon, A., Allan, K., & Wright, D. B. (2004). Say it to my face: Examining the effects of socially encountered misinformation. *Legal and Criminological Psychology*, 9(2), 215−227.

Gordon, L. T., Bilolikar, V. K., Hodhod, T., & Thomas, A. K. (2020). How prior testing impacts misinformation processing: A dual-task approach. *Memory and Cognition*, 48(2), 314−324.

Gordon, L. T., & Shapiro, A. M. (2012). Priming correct information reduces the misinformation effect. *Memory and Cognition*, 40(5), 717−726.

Gordon, L. T., & Thomas, A. K. (2014). Testing potentiates new learning in the misinformation paradigm. *Memory and Cognition*, 42(2), 186−197.

Gordon, L. T., & Thomas, A. K. (2017). The forward effects of testing on eyewitness memory: The tension between suggestibility and learning. *Journal of Memory and Language*, 95, 190−199.

Gordon, L. T., Thomas, A. K., & Bulevich, J. B. (2015). Looking for answers in all the wrong places: How testing facilitates learning of misinformation. *Journal of Memory and Language*, 83, 140−151.

Grimaldi, P. J., & Karpicke, J. D. (2012). When and why do retrieval attempts enhance subsequent encoding? *Memory and Cognition*, 40(4), 505−513.

Higham, P. A., Blank, H., & Luna, K. (2017). Effects of postwarning specificity on memory performance and confidence in the eyewitness misinformation paradigm. *Journal of Experimental Psychology: Applied*, 23(4), 417−432.

Hope, L., Gabbert, F., Fisher, R. P., & Jamieson, K. (2014). Protecting and enhancing eyewitness memory: The impact of an initial recall attempt on performance in an investigative witness interview. *Applied Cognitive Psychology*, 28(3), 304−313.

Huff, M. J., & Umanath, S. (2018). Evaluating suggestibility to additive and contradictory misinformation following explicit error detection in younger and older adults. *Journal of Experimental Psychology: Applied*, 24(2), 180−195.

Huff, M. J., Weinsheimer, C. C., & Bodner, G. E. (2016). Reducing the misinformation effect through initial testing: Take two tests and recall me in the morning? *Applied Cognitive Psychology*, 30(1), 61−69.

Jack, F., Martyn, E., & Zajac, R. (2015). Getting the picture: Effects of sketch plans and photographs on children' s, adolescents' and adults' eyewitness recall.

Applied Cognitive Psychology, 29(5), 723−734.

Jacoby, L. L., Bishara, A. J., Hessels, S., & Toth, J. P. (2005). Aging, subjective experience, and cognitive control: Dramatic false remembering by older adults. *Journal of Experimental Psychology: General*, 134(2), 131−148.

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.

Kang, S. H. K., McDermott, K. B., & Roediger, H. L. III. (2007). Test format and corrective feedback modify the effect of testing on long term retention. *European Journal of Cognitive Psychology*, 19(4−5), 528−558.

Karpicke, J. D., Lehman, M., & Aue, R. W. (2014). Retrieval-based learning: An episodic context account. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 61, pp. 237−284). San Diego: Elsevier Academic Press.

Karpicke, J. D., & Roediger, H. L. III. (2007). Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(4), 704−719.

Kęku ś, M., Polczyk, R., Ito, H., Mori, K., & Barzykowski, K. (2021). Is your memory better than mine? Investigating the mechanisms and determinants of the memory conformity effect using a modified MORI technique. *Applied Cognitive Psychology*, 35(6), 1621−1630.

Kornell, N., Bjork, R. A., & Garcia, M. A. (2011). Why tests appear to prevent forgetting: A distribution-based bifurcation model. *Journal of Memory and Language*, 65(2), 85−97.

LaPaglia, J. A. (2013). Testing can enhance or reduce suggestibility: The importance of contextual detail during misinformation exposure (Unpublished doctoral dissertation). Iowa State University, Ames.

LaPaglia, J. A., & Chan, J. C. K. (2012). Retrieval does not always enhance suggestibility: Testing can improve identification performance. *Law and Human Behavior*, 36(6), 478−487.

LaPaglia, J. A., & Chan, J. C. K. (2013). Testing increases suggestibility for narrative-based misinformation but reduces suggestibility for question-based misinformation. *Behavioral Sciences and the Law*, 31(5), 593−606.

Leding, J. K., & Antonio, L. (2019). Need for cognition and discrepancy detection in the misinformation effect. *Journal of Cognitive Psychology*, 31(4), 409−415.

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.

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning and Memory*, 12(4), 361−366.

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.

McCloskey, M., & Zaragoza, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General*, 114(1), 1−16.

Memon, A., Zaragoza, M., Clifford, B. R., & Kidd, L. (2010). Inoculation or antidote? The effects of cognitive interview timing on false memory for forcibly fabricated events. *Law and Human Behavior*, 34(2), 105−117.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer-appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, 16(5), 519−533.

Mullet, H. G., & Marsh, E. J. (2016). Correcting false memories: Errors must be noticed and replaced. *Memory and Cognition*, 44(3), 403−412.

Oeberst, A., & Blank, H. (2012). Undoing suggestive influence on memory: The reversibility of the eyewitness misinformation effect. *Cognition*, 125(2), 141−159.

Pansky, A., & Tenenboim, E. (2011). Inoculating against eyewitness suggestibility via interpolated verbatim vs. gist testing. *Memory and Cognition*, 39(1), 155−170.

Pastötter, B., Engel, M., & Frings, C. (2018). The forward effect of testing: Behavioral evidence for the reset-of-encoding hypothesis using serial position analysis. *Frontiers in Psychology*, 9, 1197.

Pastötter, B., & Frings, C. (2019). The forward testing effect is reliable and independent of learners' working memory capacity. *Journal of Cognition*, 2(1), 37.

Pastötter, B., Schicker, S., Niedernhuber, J., & Bäuml, K. T. (2011). Retrieval during learning facilitates subsequent memory encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(2), 287−297.

Paz-Alonso, P. M., & Goodman, G. S. (2008). Trauma and memory: Effects of post-event misinformation, retrieval order, and retention interval. *Memory*, 16(1), 58−75.

Pereverseff, R. S., Bodner, G. E., & Huff, M. J. (2020). Protective effects of testing across misinformation formats in the household scene paradigm. *Quarterly Journal of Experimental Psychology*, 73(3), 425−441.

Peterson, C., Parsons, T., & Dean, M. (2004). Providing misleading and reinstatement information a year after it happened: Effects on long-term memory. *Memory*, 12(1), 1−13.

Pinto, L. H., & Stein, L. M. (2015). The theoretical basis of the technique of context reinstatement in the cognitive interview. *Avances en Psicologia Latinoamericana*, 33(2), 285−301.

Polczyk, R. (2017). The "memory" misinformation effect may not be caused by memory failures: Exploring memory states of misinformed subjects. *Polish Psychological Bulletin*, 48(3), 388−400.

Pyc, M. A., & Rawson, K. A. (2012). Why is test-restudy practice beneficial for memory? An evaluation of the mediator shift hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(3), 737−746.

Putnam, A. L., Sungkhasettee, V. W., & Roediger, H. L. III. (2017). When misinformation improves memory: The effects of recollecting change. *Psychological Science*, 28(1), 36−46.

Richland, L. E., Kornell, N., & Kao, L. S. (2009). The pretesting effect: Do unsuccessful retrieval attempts enhance learning? *Journal of Experimental Psychology: Applied*, 15(3), 243−257.

Roebers, C. M., & McConkey, K. M. (2003). Mental reinstatement of the misinformation context and the misinformation effect in children and adults. *Applied Cognitive Psychology*, 17(4), 477−493.

Roediger, H. L. III., & Karpicke, J. D. (2006a). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17(3), 249−255.

Roediger, H. L. III., & Karpicke, J. D. (2006b). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, 1(3), 181−210.

Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin*, 140(6), 1432−1463.

Scheufele, D. A., & Krause, N. M. (2019). Science audiences, misinformation, and fake news. *Proceedings of the National Academy of Sciences*, 116(16), 7662−7669.

Scully, I. D., Napper, L. E., & Hupbach, A. (2017). Does reactivation trigger episodic memory change? A meta-analysis. *Neurobiology of Learning and Memory*, 142, 99−107.

Sheaffer, R., Gal, R., & Pansky, A. (2022). Resisting misinformation via discrepancy detection: Effects of an unaware suspicion cue. *Memory*, 30(6), 695−705.

Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review*, 8(2), 203−220.

Soderstrom, N. C., & Bjork, R. A. (2014). Testing facilitates the regulation of subsequent study time. *Journal of Memory and Language*, 73, 99−115.

Stenlund, T., Sundström, A., & Jonsson, B. (2016). Effects of repeated testing on short- and long-term memory performance across different test formats. *Educational Psychology*, 36(10), 1710−1727.

Szpitalak, M., Polak, M., Polczyk, R., & Dukała, K. (2016). Influence of social, para-social, and nonsocial misleading post-event sources on memory performance. *European Journal of Social Psychology*, 46(2), 185−197.

Szpitalak, M., Woltmann, A., Polczyk, R., & Kękuś, M. (2021). Memory training as a method for reducing the misinformation effect. *Current Psychology*, 40(11), 5410−5419.

Szpunar, K. K., McDermott, K. B., & Roediger, H. L. III. (2008). Testing during study insulates against the buildup of proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(6), 1392−1399.

Thomas, A. K., Bulevich, J. B., & Chan, J. C. K. (2010). Testing promotes eyewitness accuracy with a warning: Implications for retrieval enhanced suggestibility. *Journal of Memory and Language*, 63(2), 149−157.

Thomas, A. K., Gordon, L. T., Cernasov, P. M., & Bulevich, J. B. (2017). The effect of testing can increase or decrease misinformation susceptibility depending on the retention interval. *Cognitive Research: Principles and Implications*, 2(1), 45.

Tousignant, J. P., Hall, D., & Loftus, E. F. (1986). Discrepancy detection and vulnerability to misleading postevent information. *Memory and Cognition*, 14(4), 329−338.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80(5), 352−373.

Tulving, E., & Watkins, M. J. (1974). On negative transfer: Effects of testing one list on the recall of another. *Journal of Verbal Learning and Verbal Behavior*, 13(2), 181−193.

Umanath, S., Ries, F., & Huff, M. J. (2019). Reducing suggestibility to additive versus contradictory misinformation in younger and older adults via divided attention and/or explicit error detection. *Applied Cognitive Psychology*, 33(5), 793−805.

Veltre, M. T., Cho, K. W., & Neely, J. H. (2015). Transfer-appropriate processing in the testing effect. *Memory*, 23(8), 1229−1237.

Wahlheim, C. N. (2015). Testing can counteract proactive interference by integrating competing information. *Memory and Cognition*, 43(1), 27−38.

Wang, L., & Yang, J. (2021). Effect of feedback type on enhancing subsequent

memory: Interaction with initial correctness and confidence level. *PsyCh Journal*, 10(5), 751−766.

Weinstein, Y., Gilmore, A. W., Szpunar, K. K., & McDermott, K. B. (2014). The role of test expectancy in the build-up of proactive interference in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(4), 1039−1048.

Wilford, M. M., Chan, J. C. K., & Tuhn, S. J. (2014). Retrieval enhances eyewitness suggestibility to misinformation in free and cued recall. *Journal of Experimental Psychology: Applied*, 20(1), 81−93.

Wissman, K. T., Rawson, K. A., & Pyc, M. A. (2011). The interim test effect: Testing prior material can facilitate the learning of new material. *Psychonomic Bulletin & Review*, 18(6), 1140−1147.

Yang, C., Luo, L., Vadillo, M. A., Yu, R., & Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin*, 147(4), 399−435.

Yang, C., Potts, R., & Shanks, D. R. (2018). Enhancing learning and retrieval of new information: A review of the forward testing effect. *NPJ Science of Learning*, 3(1), 8.

Yang, C., Zhao, W., Luo, L., Sun, B., Potts, R., & Shanks, D. R. (2022). Testing potential mechanisms underlying test-potentiated new learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 48(8), 1127−1143.

Note: Figure translations are in progress. See original paper for figures.