The Duality of Testing: The Effect of Interim Testing on the Misinformation Effect and Its Underlying Mechanism

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

Date: 2023-03-18T00:00:00+00:00

Abstract

The effects of interim testing on the misinformation effect yield two distinct outcomes. Among them, retrieval enhanced suggestibility (RES) refers to participants who receive interim testing exhibiting lower accuracy on the final memory test and being more likely to report misinformation; the protective effect of testing (PET) refers to interim testing attenuating the misinformation effect and improving participants' memory performance. A systematic review of existing research indicates: First, these two phenomena can be explained respectively by memory reconsolidation theory, attentional capture hypothesis, retrieval fluency hypothesis (for RES), as well as memory strength theory, retrieval effort theory, and discrepancy detection theory (for PET). The relevant theories differ in terms of their stages of action and explanatory perspectives, and have been integrated into a new 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, among others. Finally, future research can continue to explore from two aspects: model validation and research expansion.

Full Text

The Dual Nature of Testing: The Influence of Interim Tests 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 an interim test on the misinformation effect has yielded two distinct outcomes. Retrieval enhanced suggestibility (RES) refers to the phenomenon where participants who receive an interim test exhibit lower accuracy on final memory tests and are more likely to report misinformation. The

protective effect of testing (PET) refers to the phenomenon where an interim test weakens the misinformation effect and improves memory performance. A systematic review of existing research reveals that these two phenomena can be explained by the reconsolidation account, attention 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 their stage of action and explanatory perspective, and have been integrated into a new theoretical model. Furthermore, several potential factors influence the separation of RES and PET, including characteristics of the original information materials, type of interim test, and features of the misinformation itself. Future research should focus on model validation and appropriate expansion of investigations.

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

1 Introduction

Misinformation refers to information that is objectively inaccurate (Scheufele & Krause, 2019). Research has shown that misinformation presented after an event can alter individuals' memory of the original event, a phenomenon known as the misinformation effect (Loftus, 2005; Wang & Jia, 2021). 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 era of information explosion, such situations are ubiquitous. Indeed, misinformation poses substantial harm to our cognition, emotions, and daily decision-making (Ecker et al., 2022), and research on this topic has grown rapidly in recent years (Wen & Zhang, 2018). Therefore, misinformation research holds both academic value and practical significance.

As researchers have continued to explore the misinformation effect, a standard paradigm has emerged. This paradigm consists of three phases: first, an original information phase where participants read or watch textual or video materials; second, a post-event information or narrative phase where participants are given additional materials that may be consistent with, irrelevant to, or contradictory to the original information; and third, a final test phase where participants answer questions about the original information to assess changes in memory accuracy (Loftus et al., 1978; Szpitalak et al., 2016; Wang & Jia, 2021). 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 crime—an immediate retrieval that occurs before encountering misinformation. Based on this, researchers have begun to focus on this first test, distinct from the final test, and have termed it the "initial test" or "interim test" (Chan et al., 2009; Gordon et al., 2020).

The testing effect demonstrates that testing between learning and a final assessment improves performance on that final test (Roediger & Karpicke, 2006b; Wang & Zhang, 2017; Yang et al., 2021). Consequently, researchers hypothesized that an interim test 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 confirmed that participants who received an interim test showed a more severe misinformation effect on the final test (Chan et al., 2009; Chan et al., 2017; Gordon et al., 2020), a phenomenon termed "retrieval enhanced suggestibility" (RES). Here, suggestibility refers to individuals' susceptibility to misinformation—the likelihood of being misled by inaccurate information (Wang et al., 2008; Cao et al., 2015). On the other hand, when conditions such as cognitive interviews, question introductions, or warnings were employed, the interim test reduced suggestibility (Memon et al., 2010; LaPaglia & Chan, 2013; Szpitalak et al., 2021; Chan et al., 2022), a phenomenon termed the "protective effect of testing" (PET).

As two contradictory phenomena, the emergence of RES and PET indicates 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. In what follows, this paper will review and analyze existing theories and relevant evidence, using research findings from the misinformation field as the main thread and supplementing them with findings from the learning domain, to clarify the nature of interim testing' s influence under different conditions and its underlying mechanisms.

2 How Interim Tests Enhance Suggestibility

RES was first discovered by Chan et al. (2009). In their experiment, participants first watched a video of a hijacking, 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 syringe"), subsequently read misinformation (e.g., "the terrorist used a chloroform-soaked rag to knock out the flight attendant"), and finally took a final test with identical questions. The results showed that participants had difficulty retrieving the original memory, leading to decreased accuracy on the final test, while simultaneously showing increased reporting of misinformation, indicating they had been influenced by the presented information. Existing explanations for RES fall into two main categories: first, original memory impairment, where the interim test either damages participants'original memory or blocks access to it, preventing correct reporting of original information; and second, enhanced learning of misinformation, where the interim test increases participants' attention to the narrative information, strengthening encoding and retrieval of misinformation, leading to greater reporting of inaccurate information on the final test.

2.1 Testing Hinders Access to Original Memory

The lower likelihood of reporting original information on the final test among participants who received an interim test may be due to blocked accessibility of that information. Indeed, information is highly vulnerable immediately after encoding and can be easily disrupted, only gradually integrating and consolidating over time to form long-term memory. The reconsolidation account (Scully et al., 2017; Carneiro et al., 2021) posits that this consolidation process is not irreversible—when reactivated, memories return to a vulnerable state and must be reconsolidated to restabilize. When misinformation is presented during this reconsolidation window, the original memory is in an unstable state and easily disrupted and altered by the inaccurate information. In other words, original memories undergo a process of "consolidation—reactivation—reconsolidation—new information intrusion." As a strong reactivation method, interim testing leads to substantial intrusion of new information into original memories. Particularly when the new information contradicts the original, this strong reactivation may cause overcorrection, resulting in forgetting of original information and updating with new information (Scully et al., 2017).

After the interim test reactivates original memory, the presentation of misinformation may have two possible effects: first, the interim test may cause misinformation to overwrite original memory, completely destroying the original memory trace, known as the updating hypothesis; or second, the misinformation may not destroy the original memory but merely impair its accessibility, preventing successful retrieval of original information on the final test, known as the inhibition hypothesis (Chan & Langley, 2011).

Currently, the inhibition hypothesis has received widespread 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, or MMFR test), the likelihood of correctly reporting original information does not differ from, or is even higher than, that of control groups (Chan et al., 2009; Gordon & Thomas, 2014). McCloskey and Zaragoza (1985) found that when the final test required participants to choose between the correct answer and other options (but not misinformation), the interim test 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 the accessibility of original information should enhance final test performance. For example, Gordon and Shapiro (2012), drawing on semantic network models, primed key details from original memory by activating related concepts, thereby reducing the misinformation effect 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 suggestibility decreases when original information accessibility is enhanced. To directly demonstrate reduced accessibility of original information, future research could compare differences between recognition and

recall. Scully et al. (2017) suggested that if participants can recognize but not recall information, this would indicate reduced accessibility.

2.2 Testing Enhances Subsequent Learning

Another prominent manifestation of RES is the increased reporting rate of misinformation, indicating that interim testing enhances participants' attention to, learning of, or encoding of inaccurate information. As early as the 1970s, Tulving and Watkins (1974) discovered that testing previously learned material could enhance learning of new information, a phenomenon known as the "forward testing effect" or "test-potentiated learning" (Yang et al., 2018; Pastötter & Frings, 2019). This effect has been extensively validated in the learning domain (Wissman et al., 2011; Chan et al., 2018; Yang et al., 2018; Choi & Lee, 2020) and is now emerging in the misinformation field. Gordon and Thomas (2014) found that participants who received an interim test recalled more details from the post-event narrative compared to those without an interim test. When this information was inaccurate or inconsistent with original information, RES emerged.

Three possible explanations exist for how interim testing enhances learning of misinformation. First, interim testing helps participants better distinguish between original and new information, thereby reducing the impact of original memory on new learning—that is, reducing proactive interference. Second, interim testing increases participants' attention to details in the narrative that are relevant to the test, altering encoding patterns and thereby enhancing learning of new information. Third, interim testing increases the retrieval fluency of misinformation, making it easier for participants to retrieve and report misinformation rather 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 distinguishes between two encoding events (original and narrative information), thereby reducing the influence of original memory on new information learning. The phenomenon where testing promotes contextual separation and reduces proactive interference has been widely documented (Szpunar et al., 2008; Wahlheim, 2015; Bufe & Aslan, 2018; Dang et al., 2021; Yang et al., 2022).

In Szpunar et al.'s (2008) classic study, as the amount of previously learned material increased, so did the interference from that material on subsequent learning; testing previous material improved recall accuracy for subsequent material and generated less proactive interference. Recent research has further found that interference from previous material mediates the relationship between interim testing and subsequent learning performance (Dang et al., 2021; Yang et al., 2022). Szpunar et al. (2008) explained this using source monitoring and information overload: if previous lists were not tested, participants would need to extract all previously presented information when tested on the final material, creating excessive information load that could cause confusion between materi-

als. Adding tests helps participants separate lists from one another, enabling better source monitoring on one hand and reducing information overload and interference from previous material on the other, thereby facilitating recall of new information.

In the misinformation domain, interim testing may separate original from narrative information, reducing interference from the former on learning of the latter and increasing recall of misinformation on the final test. However, it should be noted that the relationship between old and new materials in the misinformation domain is more clearly correlated and conflicting, and involves more examination of episodic memory. Whether interim testing can still promote contextual separation and reduce interference therefore requires further verification.

2.2.2 Learning and Encoding In testing effect research, the finding that testing can enhance learning of new material has been repeatedly confirmed (Roediger & Karpicke, 2006a; Richland et al., 2009; Carpenter, 2011; Wissman et al., 2011). Gordon and Thomas (2017) examined participants' memory for narrative information on the final test and found that the interim test group recalled more post-event information, providing direct evidence that interim testing enhances learning of narrative information. Three different theoretical explanations currently exist for this phenomenon.

The attention capture hypothesis. The attention capture hypothesis (Thomas et al., 2010; Gordon & Thomas, 2014; Gordon et al., 2015) posits that interim testing increases the salience of test-relevant information, thereby strengthening learning of misinformation. Gordon and Thomas (2014) found that participants who received an interim test spent significantly longer reading narrative information than those without an interim test, suggesting that interim testing may increase attention to narrative information. When a secondary task was added during the narrative information phase to distract participants, the difference in misinformation rates between test and control groups disappeared, effectively eliminating RES (Gordon & Thomas, 2017; Gordon et al., 2020). This evidence suggests that additional attention to misinformation can explain the emergence of RES.

It is worth noting that this explanation is incomplete. For example, when the final test was administered immediately, the interim test group and a detail-emphasis group (where attention was drawn through red underlining) showed similar RES; however, when the final test was delayed by 48 hours, the interim test group performed better than both the detail-emphasis and control groups (Thomas et al., 2017). This indicates that capturing attention does not produce the same effect as interim testing. Therefore, more complex mechanisms may underlie RES.

Encoding patterns. Interim testing enhances learning of new information—that is, it enhances encoding of misinformation—thereby strengthening suggestibility. In the learning domain, researchers have proposed the encoding reset theory

(Pastötter et al., 2011) and the encoding strategy theory (Cho et al., 2017).

The encoding reset theory suggests that encoding efficiency continuously declines during sustained learning, but testing can separate contexts and reduce memory load and inefficient encoding, enabling individuals to learn new material as effectively as early encoding (Pastötter et al., 2018; Yang et al., 2018). Cognitive neuroscience research has found that alpha power increases with sustained learning but returns to initial levels when testing is introduced, accompanied by better learning outcomes (Pastötter et al., 2011), 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 validate RES.

Cho et al. (2017) proposed the encoding strategy theory, which posits that testing enhances new learning through changes in encoding strategies. Through interim testing, participants recognize deficiencies in their memory and consequently invest more effort in subsequent learning and seek better memory methods (Pyc & Rawson, 2012; Soderstrom & Bjork, 2014; Cho et al., 2017). Studies have found that interim testing can prompt participants to use more effective encoding strategies (Soderstrom & Bjork, 2014). For example, compared to no testing or restudying, interim testing enhanced semantic clustering in subsequent learning (Chan et al., 2018; Dang et al., 2021). Yang et al. (2022) also found that interim testing enhanced 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 (LaPaglia, 2013). 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 the impact of testing on underlying learning mechanisms, but neglect that learners, as active agents, have their own motivations and needs. Motivation theory therefore provides supplementary explanations from perspectives such as retrieval failure and test expectancy.

When participants fail to recall event details—that is, when retrieval failure occurs—they become aware of their fuzzy memory, thereby increasing their attention to and learning of subsequent material (Cho et al., 2017; Yang et al., 2018). Richland et al. (2009) argued that this is because 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 will emerge.

Additionally, interim testing may activate higher test expectancy—that is, the prediction that testing will continue—which is known as the test expectancy theory (Weinstein et al., 2014; Yang et al., 2018). Research has found that test expectancy, especially expecting more difficult subsequent tests, enhances

participants' learning of material (Agarwal & Roediger, 2011; Weinstein et al., 2014). In the misinformation paradigm, participants have no test expectancy when encoding original information but do have it when encoding misinformation (derived from the interim test), leading them to learn original information less well than misinformation and consequently producing RES.

2.2.3 Retrieval Fluency Another reason why the explanation that interim testing enhances new information learning cannot fully account for RES is that the enhancing effect of testing on learning of tested information (original information) has also been widely validated (Roediger & Karpicke, 2006b; Wang & Zhang, 2017; Yang et al., 2021). So in the competition between two equally enhanced memories, how can misinformation win? 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 the retrieval fluency of this information on the final test, making it more likely to be reported.

Thomas and Gordon conducted a series of studies to validate this theory. In their research, Thomas et al. (2010) used confidence and response time as indicators of retrieval fluency. They argued that confidence is affected by the ease of memory retrieval—the more fluent, the more confident—while response latency more directly assesses retrieval fluency, with faster responses indicating greater fluency. The results showed that participants in the interim test group responded more quickly when selecting misinformation and did so with high confidence, indicating that interim testing enhanced retrieval fluency of misinformation. However, when they suppressed retrieval fluency by warning participants that the source of narrative information was uncertain and its authenticity unknown, participants took longer to make choices and their inflated confidence disappeared, weakening RES. Another experiment found that when participants were asked to provide two answers on the final test, the interim test group showed better memory for both the original video and the misinformation, demonstrating that testing enhanced learning of both pre-test and post-test information. However, because misinformation was more easily retrieved, RES emerged (Gordon & Thomas, 2014). 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 the retrieval fluency of misinformation dissipated after 48 hours.

Overall, existing theoretical explanations for RES can largely be categorized into "encoding" and "retrieval" perspectives. "Encoding" explanations suggest that interim testing improves participants' 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 invest more effort during the narrative information phase;

encoding reset and attention capture provide important preconditions for enhanced subsequent learning, enabling participants to "be able" to elaborately encode misinformation during the narrative information phase; and changed encoding strategies provide methodological convenience for enhanced subsequent learning, making it "easy" for participants to engage in efficient encoding during the narrative information phase. "Retrieval" explanations, conversely, suggest that the retrieval process of interim testing influences the emergence of RES. First, according to the reconsolidation account, retrieving information destabilizes original memory, making it more vulnerable to misinformation. Second, retrieval itself promotes contextual separation and reduces interference from previous material. Finally, the retrieval fluency hypothesis suggests that RES emerges because misinformation is more easily retrieved on the final test, causing participants to respond more quickly and confidently when faced with incorrect options.

2.3 The Pathway of RES

In summary, although researchers have provided rich and diverse theoretical explanations for RES, they have neglected the correlations and complementarities among theories, and thus cannot grasp the full picture of RES' s underlying mechanisms. Based on this, this paper proposes a dual-pathway model for learning of both original and narrative information (as shown in Figure 1 [Figure 1: see original paper]).

According to this model, interim testing simultaneously affects individuals' memory for both original and misinformation, ultimately leading to RES. In the first pathway, based on the testing effect (Roediger & Karpicke, 2006b; Wang & Zhang, 2017; Yang et al., 2021), interim testing can enhance memory for the original event. However, the appearance of misinformation reverses this effect. Although research has proven that original memory is not completely destroyed (McCloskey & Zaragoza, 1985; Chan et al., 2009; Gordon & Thomas, 2014), the presence of misinformation still reduces the 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 isolates the two information materials, thereby reducing interference from original memory on misinformation learning (Szpunar et al., 2008). Second, retrieval failure and test expectancy motivate participants to encode narrative information effortfully (Richland et al., 2009; Weinstein et al., 2014; Cho et al., 2017; Yang et al., 2018). Finally, interim testing increases participants' attention (Thomas et al., 2010; Gordon & Thomas, 2014; Gordon et al., 2015), increasing reading time for narrative information and simultaneously enhancing encoding of misinformation (Pastötter et al., 2011; Cho et al., 2017). When encoding of misinformation is enhanced, participants can more easily retrieve it on the final test—that is, retrieval fluency increases (Thomas et al., 2010; Gordon & Thomas, 2014).

Taken together, the emergence of RES ultimately depends on the relative strength of original information accessibility versus misinformation retrieval fluency—that is, neither pathway can be viewed in isolation. Thomas et al. (2010) argued that RES occurs because individuals prematurely terminate further retrieval needed to recall original information after retrieving more fluent misinformation. This suggests that it is precisely original information' s "failure" in the fluency competition that leads to its abandonment. From this perspective, increasing original information accessibility or decreasing misinformation retrieval fluency, or providing opportunities for further retrieval of original information, can reduce RES (Chan et al., 2009; Gordon & Shapiro, 2012; Gordon & Thomas, 2014, 2017; Gordon et al., 2020).

3 How Interim Tests Exert a Protective Effect

Although RES has been widely validated, researchers have also discovered the opposite PET under certain conditions. For example, in the household scene paradigm, participants are first shown a photo of a household scene (e.g., a "desk") containing various daily 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 the pen, and began writing"). The results showed that participants in the interim test group were less likely to report nonexistent items (e.g., "pen") on the final test, indicating that interim testing reduced suggestibility (Pereverseff et al., 2020).

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 a single phenomenon. This means that research on PET can supplement the discussion of RES. First, existing theories and research on RES can only explain how to reduce the negative impact of interim testing to the baseline level observed without interim testing, but cannot leverage the inherent advantages of interim testing. In other words, RES-related theories can explain how to eliminate RES but cannot specify when PET will occur (Chan et al., 2009; Thomas et al., 2010; Gordon & Thomas, 2014, 2017; Gordon et al., 2020). Second, investigating how to exert the positive influence of interim testing and reduce suggestibility has broader practical significance. However, no mature theoretical framework for explaining PET has yet been developed, leaving a gap in discussions of interim testing' s mechanisms. Therefore, we have reviewed relevant research and theories, including memory strength, retrieval effort theory, discrepancy detection theory, and transfer-appropriate processing theory, to provide preliminary exploration for PET.

3.1 Memory Strength

Some researchers argue that individuals are susceptible to misinformation because their memory strength is insufficient, leaving them unconfident in their judgments and more likely to accept inaccurate information. Therefore, a

"strong" memory for the original event enables people to more easily detect discrepancies between the original event and misinformation, thereby rejecting the inaccurate information and making correct choices (Peterson et al., 2004; Loftus, 2005; Guo & Li, 2007).

In the misinformation paradigm, the standard for measuring whether original memory is strong (memory strength) is whether the interim test was answered correctly (Wang & Yang, 2021). Chan and Langley (2011) found that when the interim test was answered correctly, the probability of participants reporting misinformation did not differ significantly from that of participants without an interim test, indicating that strong memory offset RES. In LaPaglia and Chan' s (2013) study, when misinformation was presented in question stem format, participants who answered the interim test correctly showed higher final test accuracy than those who answered incorrectly, and were even unaffected by misinformation. Gabbert et al. (2012) also argued that PET occurs because testing strengthens participants' original memory, enabling them to better detect discrepancies. Therefore, participants' memory for the original event being sufficiently accurate and robust to resist misinformation may be one reason for PET. We hypothesize that when individuals' memory for original information is strong enough, interim testing will exert its protective effect and reduce suggestivity; when original memory strength is insufficient, individuals are more easily misled by misinformation, and RES emerges.

3.2 Retrieval Effort Theory

In studies that have found PET, when participants expend more effort to retrieve information during the interim test phase, their memory for original information is enhanced, making it easier for them to identify misinformation and reduce suggestibility. The most direct evidence comes from Pansky and Tenenboim (2011), who divided information into basic-level (e.g., "chair") and subordinate-level (e.g., "wooden chair") categories and conducted both gist 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. The results showed that both test types produced testing effects, but only verbatim tests reduced suggestibility.

Retrieval effort theory posits that the testing effect results from the cognitive effort expended during the retrieval of learned information (Bjork, 1975; Karpicke & Roediger, 2007; Rowland, 2014; Stenlund et al., 2016). This theory has received extensive support in the testing effect literature. For example, participants who engaged in deep retrieval during initial testing showed better recognition performance on subsequent tests (Jacoby et al., 2005). To further explain this phenomenon, researchers developed the desirable difficulty theory (Bjork & Bjork, 1992; Bjork & Bjork, 2011) and the bifurcation model theory (Kornell et al., 2011).

The desirable difficulty theory distinguishes between storage strength and re-

trieval strength. The former reflects the relative permanence of memory traces or the durability of learning; the latter reflects the momentary accessibility of memory traces—the ease with which information can be recalled, similar to retrieval fluency. Meanwhile, this theory assumes that retrieval strength is negatively correlated with increases in storage strength: only when successful retrieval requires substantial effort (low retrieval strength) can retrieval practice enhance memory strength (high storage strength) and promote long-term learning (Zhang et al., 2008). In the bifurcation model hypothesis, all items initially follow a normal distribution of retrieval strength. When restudied, items show small increases in retrieval strength and the overall distribution remains normal; when retrieval practice occurs, 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 (Wang & Zhang, 2017).

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

3.3 Discrepancy Detection Theory

Discrepancy detection theory posits that detecting discrepancies between original and misinformation can reduce the misinformation effect—that is, participants who notice discrepancies are more likely to identify and reject misinformation than those who do not, thereby improving final test accuracy (Tousignant et al., 1986; Mullet & Marsh, 2016; Polczyk, 2017; Putnam et al., 2017; Sheaffer et al., 2022). Studies have shown that when participants engage in discrepancy detection for each piece of narrative information while reading it, suggestibility can be effectively reduced (LaPaglia, 2013; Bailey et al., 2021).

Research on contradictory and additive misinformation can provide evidence for discrepancy detection theory. Contradictory misinformation refers to information that conflicts with the original event—that is, "calling a deer a horse." Additive misinformation refers to information that appears in the narrative but was not present in the original event—that is, "creating something from nothing." The difference is that contradictory misinformation conflicts with original memory, making it more suitable for discrepancy detection. Studies have found that participants report less contradictory than additive misinformation (Huff &

Umanath, 2018; Umanath et al., 2019), and this difference is more pronounced in the interim test group (Gordon et al., 2015). This suggests that interim testing further enhances discrepancy detection.

Research has shown that compared to restudying, tested participants are better at detecting changes in learning materials (Wahlheim, 2015). Additionally, in Gordon et al.' s (2015) study, when participants answered the interim test correctly, they spent more time reading misinformation, implying they detected inconsistency between original memory and narrative information (Tousignant et al., 1986; Loftus, 2005). Therefore, when a discrepancy recall task was added to the final test, the interim test group was better able to detect differences between original and misinformation (Gordon & Thomas, 2017). In summary, interim testing can reduce the misinformation effect by promoting discrepancy detection.

3.4 Transfer-Appropriate Processing Theory

LaPaglia and Chan (2012) discovered PET in a study on face recognition. They argued that interim testing "seals" correct memory containing almost no misinformation, which may be activated during the final test to resist misinformation. In other words, testing may "transfer" part of original memory, but how does this "transfer" occur, and what factors affect its effectiveness? The transfer-appropriate processing theory from the learning domain can provide some empirical evidence.

Transfer-appropriate processing theory posits that testing allows participants to practice retrieval operations, thereby performing better on subsequent retrieval attempts—in other words, transferring "testing skills" to later tests (Morris et al., 1977; Roediger & Karpicke, 2006b; Yang et al., 2021). From this, we can hypothesize that testing is most effective 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 that testing effects are larger when formats are consistent. On the other hand, the degree of cue matching between the two tests also affects testing' s impact. For example, Veltre et al. (2015) used two types of tests providing semantic cues and spelling cues, finding that memory performance was better when interim and final test cue types matched. This suggests that matching interim and final tests in type or other conditions is effective and may affect the "relay transfer" of testing' s effect on original memory, thereby influencing PET emergence.

However, some questions remain about this theory. For instance, Kang et al. (2007) used two test formats (short answer and multiple choice) as initial tests to examine participants' performance when final test formats matched or mismatched, finding that short answer tests promoted final test performance better than multiple choice tests regardless of format. Rowland' s (2014) meta-

analysis obtained similar results—testing effects were larger when interim tests used free recall compared to 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 still 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 their original memory (Yang et al., 2021) and promotes PET emergence. Additionally, interim testing reduces suggestibility by promoting discrepancy detection, manifested as spending more time reading misinformation. This suggests that discrepancy detection enhances participants' critical encoding of misinformation. From the "retrieval" perspective, retrieval effort theory focuses on the direct benefits of retrieval—examining the degree of effort in the retrieval process and positing that more cognitive effort leads to greater memory strength enhancement. Transfer-appropriate processing theory, conversely, views "retrieval" as a special skill that teaches participants "how to test," thereby improving performance on subsequent tests. Therefore, transfer-appropriate processing theory places greater emphasis on 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 achieve further integration and development of PET explanations, this paper incorporates existing mainstream research into a unified theoretical framework (as shown in Figure 2 [Figure 2: see original paper]). First, higher participant memory strength and the retrieval effort required by testing both help exert testing' s protective effect, while more effortful retrieval can also improve individuals' memory performance (Chan & Langley, 2011; Pansky & Tenenboim, 2011; LaPaglia & Chan, 2013; Wang & Zhang, 2017). Second, beyond direct effects, strong memory and difficult retrieval also enhance participants' discrepancy detection ability, making them more likely to detect misinformation and thereby reduce suggestibility (LaPaglia, 2013; Gordon & Thomas, 2017; Bailey et al., 2021). Finally, according to transfer-appropriate processing theory, the impact 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 (Morris et al., 1977; Roediger & Karpicke, 2006b; Yang et al., 2021).

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, the 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's effect on misinformation from different "encoding"or "retrieval"perspectives. To facilitate understanding, we have compared and integrated relevant theories along two dimensions– "stage of action" and "explanatory perspective" –as shown in Table 1 .

However, as two contradictory phenomena, research on RES and PET cannot develop 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 related to both phenomena according to their logical sequences and possible causal connections, proposing an overarching model encompassing major mainstream explanations to provide more in-depth and comprehensive investigation of the mechanisms underlying interim testing' s influence on the misinformation effect.

We propose that interim testing primarily affects the misinformation effect through three pathways. The first pathway explains interim testing' s influence on original information. First, given the widely documented testing effect (Roediger & Karpicke, 2006b; Wang & Zhang, 2017; Yang et al., 2021), interim testing can enhance memory for original information. According to retrieval effort theory, when interim tests are more difficult, they better enhance the storage strength of original information (Karpicke & Roediger, 2007; Stenlund et al., 2016). However, this process depends on whether retrieval during interim testing is successful (Bjork & Bjork, 1992; Bjork & Bjork, 2011; Kornell et al., 2011; Wang & Zhang, 2017). 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. This is consistent with the RES model and will not be reiterated here. The third pathway acts on both original and misinformation simultaneously. First, interim testing changes the context, leading to encoding reset that distinguishes the two encoding events of original and misinformation (Szpunar et al., 2008; Chan et al., 2009; Pastötter et al., 2011; Wahlheim, 2015; Bufe & Aslan, 2018; Pastötter et al., 2018; Dang et al., 2021). Second, this clearer distinction facilitates participants' detection of discrepancies between information sources, thereby reducing suggestibility (Tousignant et al., 1986; LaPaglia, 2013; Mullet & Marsh, 2016; Polczyk, 2017; Putnam et al., 2017; Bailey et al., 2021; Sheaffer et al., 2022).

This model integrates theoretical research on both RES and PET, providing a more detailed elaboration of the mechanisms by which interim testing influences the misinformation effect, and holds certain theoretical value. First, the model places research on RES and PET within the same framework, which helps focus research directions and highlight key issues. Previous research treated RES and PET as contradictory phenomena, often biasing explanations and inadequately

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 excavate theoretical value. Second, the model connects theories from the learning and misinformation domains, providing a bridge for interdisciplinary integration. The learning domain has mature theories, while the misinformation domain has diverse findings, so this connection has dual value: on one hand, it fills theoretical gaps in the misinformation domain, providing a solid theoretical foundation for related research; on the other hand, it expands the application scope of testing effect theories and enriches their theoretical connotations.

4 Potential Influencing Factors for the Separation of RES and PET

Based on the existing explanatory framework, we can see that across different stages of the misinformation paradigm, the nature of interim testing' s effect depends on the type 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 Materials

First, for different types of original materials, the nature of interim testing' s effect differs. Under otherwise consistent conditions, studies using video materials have shown RES (Chan et al., 2017; Gordon & Thomas, 2017; Gordon et al., 2020), while those using picture materials have shown PET (Huff et al., 2016; Pereverseff et al., 2020). According to mental model theory, when original events have narrative coherence, logic, and completeness, people form mental models of how events unfold; and people prefer to maintain coherent and complete mental models even when they are incorrect (Johnson & Seifert, 1994). Therefore, if a component of the model is missing, creating a gap, participants may accept incorrect information to maintain the mental model' s integrity (Lewandowsky et al., 2012). Correspondingly, in studies using video materials, individuals focus more on the dynamic development of storylines and thus construct mental models rich in contextual information. If original memory is insufficiently strong or inaccessible, participants are more likely to accept misinformation to fill gaps in their mental models. However, in studies using picture materials, both original memory and narrative information consist of isolated items without contextual relationships (Huff et al., 2016; Pereverseff et al., 2020). In this case, individuals have difficulty constructing coherent, complete mental models and thus cannot easily integrate misinformation into original memory, reducing suggestibility and producing PET.

4.2 Type of Interim Test

The type of interim test affects its effectiveness. Memon et al. (2010) used cognitive interviews (a questioning method involving extensive and detailed memory retrieval, including context reinstatement, report everything, change perspective, and reverse order) and found that interviews increased reporting of correct details and reduced false agreement with fabricated items after one week. Gabbert et al. (2012) found that completing a self-administered interview (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 (Gordon & Thomas, 2017; Gordon et al., 2020). According to retrieval effort theory, this is because cognitive interviews and self-administered interviews, compared to cued recall, require more effortful and deeper retrieval, thereby strengthening original memory and enabling better rejection of misinformation.

Another explanation for why cognitive interviews and self-administered interviews can 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 (Tulving & Thomson, 1973; Smith & Vela, 2001). Research has found that context reinstatement can affect individuals' suggestibility (Roebers & McConkey, 2003; Drohan-Jennings et al., 2010). When cognitive interviews and self-administered interviews are used (including methods such as context reinstatement and scene sketches), interim testing can help participants effectively retrieve the encoding context of original memory, thereby improving recall of original information (Hope et al., 2014; Jack et al., 2015; Pinto & Stein, 2015). This process facilitates integration of original information with the new test context, so participants can more easily retrieve original information on the final test (which shares the same test context as the interim test), and PET emerges.

4.3 Characteristics of Misinformation

The misleading nature of misinformation is inseparable from its own characteristics. For example, compared to central details (i.e., key content that is more vivid and attention-grabbing), 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; but for peripheral details, RES was only offset when the interim test was answered correctly. In fact, in their study, central details were those reported more frequently by most participants, indicating inherently better memory strength and greater resistance to misinformation interference. Other studies have also found that memory for central details is generally less susceptible to misinformation than memory for peripheral details (Paz-Alonso & Goodman, 2008). According to memory strength theory, central details, due to their inherently better memory strength, are less vulnerable to misinformation interference;

whereas peripheral details, which are easily overlooked in events, are less clear and stable in memory and thus more susceptible to misinformation.

Additionally, the contextual information load of misinformation affects suggestibility. Research has found that RES emerged when misinformation was presented in narrative form, while PET emerged when misinformation was introduced in question stem format (LaPaglia & Chan, 2013). LaPaglia (2013) argued that this is because narrative misinformation contains more contextual details, while question stems provide isolated information. Correspondingly, when misinformation was presented sentence-by-sentence rather than in paragraph form, the interim test group recalled more correct information (Gordon & Thomas, 2014). Therefore, the content richness or contextual information load of misinformation itself affects suggestibility. Specifically, richer content leads to RES, while poorer content leads to PET (LaPaglia, 2013). This may be because when narratives contain too much information, individuals have difficulty conducting discrepancy detection, but when information is presented in isolation, discrepancies become more salient.

5 Summary and Outlook

This paper provides a theoretical discussion of the two different mechanisms by which interim testing produces effects in the misinformation effect. RES arises because interim testing hinders participants' access to original information and promotes their learning of misinformation through separation, increased attention, and motivation; but when original memory is sufficiently strong, or when participants are encouraged to engage in more detailed retrieval, reminded to notice information discrepancies, or helped to transfer strategies or memories, testing can exert a protective effect. Based on our review of relevant theories, we have proposed pathway diagrams for both RES and PET, constructed an integrated model of interim testing' s influence on the misinformation effect, and analyzed potential influencing factors for the separation of RES and PET. Building on the theoretical models proposed in this paper, future research can be conducted in the following two areas.

5.1 Model Validation

The models proposed in this paper integrate relatively independent research findings and theoretical foundations from different domains, and can provide relatively comprehensive explanations for phenomena discovered in existing research. However, these models are still in their initial construction stage. What connections exist among various theories, how pathways interact, and whether other key variables and influencing factors exist all require more targeted theoretical and empirical research to address and supplement.

First, future research needs to directly test different pathways. Regarding the pathways for original and misinformation, 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 (Cao et al., 2015) or add distracting tasks (Gordon & Thomas, 2017; Gordon et al., 2020) to alter the degree of information learning, and measure retrieval fluency and accessibility through response time and confidence level on the final test (Thomas et al., 2010) to examine changes in the misinformation effect. On this basis, future research could use mediation analysis to test whether interim testing affects the degree of encoding of original and misinformation, changes the accessibility and retrieval fluency of original information, and ultimately influences the misinformation effect.

Regarding the third pathway, the relationship between contextual separation and discrepancy detection also warrants further examination. Research has found that interference from previous material only weakens when participants detect discrepancies between information sources; otherwise, interference persists (Wahlheim, 2015). However, in that study, discrepancy detection was of overall material rather than individual items, so whether the interference and separation mechanisms must operate through discrepancy detection requires more evidence. Future research could examine memory for tested misinformation on the final test, measure proactive interference by the number of original information items participants report, and add a discrepancy detection test after the final test (asking participants whether they detected discrepancies in each piece of information, Polczyk, 2017) to validate this pathway.

Second, future research needs to combine different pathways to comprehensively explore the mechanisms of interim testing. On one hand, different pathways are not independent and may interact. For example, researchers have argued that strong original memory promotes discrepancy detection (Peterson et al., 2004; Loftus, 2005; Gabbert et al., 2012), but their studies have not provided direct evidence for this hypothesis. Gabbert et al. (2012) found that higher participant accuracy (i.e., better memory) was negatively correlated with misinformation reporting rates, but this alone is insufficiently direct and adequate as evidence. Additionally, whether the degree of misinformation encoding 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 simultaneously examine the 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 have continuously refined their focus and enriched experimental evidence, they have also provided preliminary exploration for the intersection of learning and misinformation research. However, existing research still retains the imprint of the learning domain in its theoretical foundation and variable selection, making it difficult to handle contradictions between different phenomena and limiting further extension of theoretical and practical value. Therefore, future research should broaden its perspective and conduct subsequent studies from two aspects: variable introduction and intervention development.

First, the current research scope is relatively narrow and needs to further examine individual differences and social factors. On one hand, individual differences affect 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 (Leding & Antonio, 2019). Need for cognition refers to a trait reflecting individuals' willingness to invest effort in deep cognitive processing (Cacioppo & Petty, 1982). 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 (Xu & Zhou, 2010). 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 will affect interim testing' s influence on the misinformation effect through retrieval effort and discrepancy detection. Future research could further explore the effects of cognitive, emotional, attitudinal, and other individual factors on RES or PET.

On the other hand, misinformation in real life often has a "social" dimension. For example, when misinformation is introduced socially, individuals are influenced by inaccurate information from co-witnesses, a phenomenon known as "memory conformity," considered a form of misinformation effect (Gabbert et al., 2004; Kękuś et al., 2021). Moreover, socially sourced misinformation causes greater harm than non-socially sourced misinformation (Gabbert et al., 2004). Research has found that detecting discrepancies between original and others'misinformation can reduce memory conformity, but participants may still report misinformation despite detecting discrepancies because they lack confidence in their own memory (Kękuś et al., 2021). As an operation that can strengthen original memory and promote discrepancy detection, interim testing may reduce memory conformity effects. Furthermore, future research could examine factors such as the social nature of information, its source, and individuals' social identity and social environment, thereby broadening the scope of interim testing research to better exert testing' s protective effect and achieve greater practical value and theoretical depth.

Second, to expand the practical significance of research, 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 that "human memory is unreliable" can reduce their suggestibility (Szpitalak et al., 2021). Researchers have argued that 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 through promoting discrepancy detection requires further research.

Existing research has found that warnings may be an effective intervention. In studies on misinformation and interim testing, when participants were told that "the source of narrative information is uncertain and therefore its accuracy cannot be verified," their suggestibility decreased (Thomas et al., 2010; Oeberst & Blank, 2012; Blank & Launay, 2014; Higham et al., 2017; Polczyk, 2017; Szpitalak et al., 2021; Bulevich et al., 2022; Chan et al., 2022). This may be because warnings encourage participants to ignore the retrieval fluency of misinformation and engage in more effortful retrieval (Thomas et al., 2010), and also help them better detect discrepancies between original and misinformation (Higham et al., 2017; Polczyk, 2017). Chan et al. (2022) argued that warnings may allow participants to choose 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, a phenomenon known as "tainting the truth response" (Szpitalak et al., 2021). Therefore, whether warnings are equally effective in real-world situations requires further verification. Future research could develop scientifically sound and effective intervention methods based on existing theoretical evidence.

References

Cao, X., Tu, S., & Chen, X. (2015). Repeated testing effects on memory suggestibility. *Chinese Journal of Clinical Psychology*, 23(6), 980-984+1008.

Guo, X., & Li, J. (2007). Research on false memory induced by misleading information interference. *Psychological Science*, 30(4), 814-819.

Wang, H., Liu, M., Zhang, J., & Xiang, Z. (2008). Research and progress on eyewitness suggestibility. *Psychological Science*, 31(3), 744-747.

Wang, X., & Jia, N. (2021). Review and prospects of major theories and influencing factors of the misinformation effect. *Psychology: Techniques and Applications*, 9(1), 52-64.

Wang, Z., & Zhang, J. (2017). The mechanism of retrieval practice effect: Inte-

grated evolution of encoding and retrieval. *Advances in Psychological Science*, 25(9),

Wen, J., & Zhang, Z. (2018). The generation, dissemination, identification, and control of misinformation: A review of existing research on misinformation. *Science and Society*, 8(3),

Xu, J., & Zhou, N. (2010). The influence of need for cognition on individual information processing tendency. *Advances in Psychological Science*, 18(4), 685–690.

Zhang, J., Bai, X., & Yang, L. (2008). An overview of foreign research on testing effects. *Advances in Psychological Science*, 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, R.A. (1975). Retrieval as a memory modifier: An interpretation of negative recency and related phenomena. In Solso R.L., (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 123–144). Hillsdale, NJ: Erlbaum.

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., & 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.

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.

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

9,1701.

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

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

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.

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., & 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., 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., 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),

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,

Gabbert, F., Hope, L., Fisher, R. P., & Jamieson, K. (2012). Protecting against misleading post-event information with a 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., & 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., 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., 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),

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 interview. *Applied Cognitive Psychology*, 28(3), 304–313.

Huff, M. J., & Umanath, S. (2018). Evaluating the 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.

Kękuś, 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 witness 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),

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.

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.

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),

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, 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),

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.

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.

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),

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). The 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, 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, 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),

Two sides of testing: The influence of interim tests 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 an interim test on the misinformation effect has been found to have two distinct results. "Retrieval enhanced suggestibility" (RES) refers to the observation that participants who received an interim test were less likely to respond correctly on the final memory test and more likely to report misinformation. The "protective effect of testing" (PET) refers to the observation that an interim test weakened the misinformation effect and improved participants'memory performance. A systematic review of existing studies shows that these two phenomena can be explained by the reconsolidation account, the attention capture hypothesis, and the retrieval fluency hypothesis (for RES), or by the memory strength theory, the retrieval effort theory, and the discrepancy detection theory (for PET). These related theories differ in both the stage of action and the perspective of explanation, and are integrated into a new theoretical model. In addition, there are some potential influences on the separation of RES and PET, including the original information material, the type of interim test, and the characteristics of the misinformation. Future research should begin with the testing of this theoretical model and expand it in appropriate directions.

Key words: misinformation, interim test, retrieval enhanced suggestibility, protective effect of testing

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