

《心理科学进展》审稿意见与作者回应

题目：测试的两面性：中期测试对错误信息效应的影响及其作用机制

作者：何宁 李梦 康彬 王梦云 岳云帆

第一轮

审稿人 1 意见：

该文章对测试增强个体对错误信息的易感性进行了详细综述，该综述系统讨论了相关的研究文献，有架构的整理了以往的研究内容，并提出了自己的一些理论思考。同时，综述问题具有重要的应用价值，国内的确需要更多的研究者注意到这个研究话题，因此，本人推荐该文章的发表。但是，在发表之前，仍需要对该文章做出必要的修改。

意见 1：

文章还存在一些错别字，如“针对错误信息研究的基本范式业已形成”。

回应：

非常感谢审稿专家对本研究价值的肯定和恳切建议，我们对照您的建议做了逐一修改，修改的部分在原文中用红色字体标出。

经过专家提醒，我们将“针对错误信息研究的基本范式业已形成”修改为“针对错误信息研究的基本范式**已经**形成”。此外，我们对全文文字进行了多次批判性阅读，对文中的一些表述进行了修改和完善，并添加了相应的参考文献。

意见 2：

将“the reconsolidation account”翻译成“记忆再巩固理论”是否更合适。

回应：

非常感谢审稿专家的提醒。根据建议，我们查阅了相关文献发现，国内学者均将“the reconsolidation account”翻译为“记忆再巩固理论”。因此，我们将“再整合理论”修改为“**记忆再巩固理论**”，并将相关概念中的“整合”均修改为“**巩固**”，请见“2.1 测试损害原始记忆”第二段。

相关参考文献：

郭滢, 龚先旻, 王大华. (2021). 错误记忆产生的认知与神经机制: 信息加工视角. *心理科学进展*, 29(01), 79 - 92.

李俊娇, 陈伟, 石佩, 董媛媛, 郑希付. (2022). 预期错误在恐惧记忆更新中的作用与机制. *心理科学进展*, 30(04), 834 - 850.

庄楚群, 王文清, 胡静初, 张蔚欣, 王鹏贵, 郑希付. (2017). 提取-消退范式中复合刺激对恐惧消退的影响. *心理学报*, 49(3), 329 - 335.

意见 3：

MMRF 指的是不是 modified modified free recall 测试？是不是应该缩写成 MMFR？

回应：

非常感谢审稿专家细致的建议。原论文中的“MMRF”的确指的是“modified modified free recall 测试”。根据建议，我们认真查阅了相关文献，发现此处缩写有误，应该缩写为“MMFR”。

因此，我们将“MMRF”修改为“MMFR”，并放入了英文名词“the modified modified free recall test”，请见“2.1 测试损害原始记忆”第三段。

相关参考文献：

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.

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

Sheaffer, R., Gal, R., & Pansky, A. (2021). Resisting misinformation via discrepancy detection: effects of an unaware suspicion cue. *Memory*, 1–11.

意见 4：

The forward testing effect 是不是翻译成“前向测试效应”更合适？

回应：

非常感谢审稿专家的意见。根据建议，我们查阅了相关文献发现，国内学者更多将“the forward testing effect”翻译为前向测试效应，因此我们作出了相应修改，请见“2.2 测试增强后续学习”第一段。

相关参考文献：

王堂生, 杨春亮, 钟年. (2020). 记忆的前向测试效应对老年人学习新事物的作用. *心理学报*, 52(11), 1266 – 1277.

马小凤, 李甜甜, 贾瑞红, 魏婕. (2022). 空间路线学习中的前向测试效应. *心理学报*, 54(12), 1433 – 1442.

意见 5：

“目前，这一解释得到了来自认知神经科学研究证据的支持（Pastötter et al., 2011）”，该研究并不是支持“干扰与分离”解释，而支持的是“学习与编码”解释。

回应：

非常感谢审稿专家的悉心建议。经过专家提醒，我们再次认真阅读了该文献。Pastötter 等人（2011）认为，测试促进了上下文分离，因此导致了列表之间编码重置，使得后续学习和先前学习一样有效。相应地，生理结果显示， α 能量会随着持续学习而不断增加，但在加入测试后， α 能量重新降至初始水平，这说明测试导致了编码重置。但是我们在进一步学习脑电领域相关研究后发现， α 波和前摄干扰无关（Kliegl et al., 2015）。因此，在“干扰与分离”部分引用 Pastötter 等人（2011）的研究并不合适。

因此，我们在“2.2.1 干扰与分离”部分将其替换为另一项生理学研究（Pajkossy et al., 2019）。该研究发现，与重新学习相比，中期测试减少了被试在最终测试中的瞳孔扩张。瞳孔大小反映了执行任务时的资源处理量，这说明中期测试减少了最终回忆时的处理负荷。

具体内容如下：

Pajkossy 等人（2019）的研究为这一解释提供了生理学证据。他们发现，与重新学习相比，中期测试减少了被试在最终测试中的瞳孔扩张。瞳孔大小反映了执行任务时的资源处理量，这说明中期测试减少了最终回忆时的处理负荷。

相关参考文献：

Kliegl, O., Pastötter, B., & Bäuml, K. H. T. (2015). The contribution of encoding and retrieval processes to proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(6), 1778–1789.

Pajkossy, P., Szöllösi, Á., & Racsmány, M. (2019). Retrieval practice decreases processing load of recall: Evidence revealed by pupillometry. *International Journal of Psychophysiology*, 143, 88–95.

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.

意见 6:

文章较少对“warning”降低 RES 的研究进行介绍，可以从干预或理论的角度对相关研究进行整理报告。

回应:

非常感谢审稿专家恳切的意见。根据专家建议，我们查阅了警告与错误信息的相关文献后，认为警告的确可以作为降低错误信息暗示性的一种干预手段，并对警告的作用机制进行了总结梳理，并在“5 总结与展望”第五段中有关干预的部分进行了补充，具体如下：

已有研究发现，警告可能是种有效的干预手段。在错误信息及中期测试的研究中，当被试被告知“叙述信息来源不确定，因此无法验证其准确性”时，他们对错误信息的敏感性会降低（Thomas et al., 2010; Oeberst & Blank, 2012; Blank & Launay, 2014; Polczyk, 2017; Higham et al., 2017; Szpitalak et al., 2021; Chan et al., 2022; Bulevich et al., 2022）。这可能是因为警告鼓励被试忽略错误信息的检索流畅性，从而进行更努力的检索（Thomas et al., 2010），也更能发现原始信息和错误信息之间的差异（Polczyk, 2017; Higham et al., 2017）。Chan 等人（2022）认为，警告可以让被试选择“遗忘”叙述信息，减少错误信息的进一步加工，类似于“定向遗忘”。然而，如果被试实际上并没有遇到错误信息，警告可能会使被试对正确的信息也产生怀疑，这被称作“污染真相相应”（Szpitalak et al., 2021）。因此，警告在现实情境中是否同样有效，还需进一步的验证。未来研究可以从已有理论证据入手，进一步发展出兼具科学性、有效性的干预手段。

相关参考文献:

- 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 & Cognition*, 50(1), 45–58.
- 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*. Advance online publication. <https://doi.org/10.1037/xap0000394>
- 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.
- Oeberst, A., & Blank, H. (2012). Undoing suggestive influence on memory: The reversibility of the eyewitness misinformation effect. *Cognition*, 125(2), 141–159.
- 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.
- Szpitalak, M., Woltmann, A., Polczyk, R. & Kekus, M. (2021). Memory Training as a Method for Reducing the Misinformation Effect. *Current Psychology*, 40, 5410–5419.
- 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, 149–157.

意见 7:

作者用“分散注意力”相关的研究来支持“检索流畅性”解释，该观点是值得商榷的，“分散注意力”相关的研究更多的是用来支持“注意力捕获假说”。

回应：

非常感谢审稿专家提出的意见。根据专家建议，我们再次回顾了 Gordon 和 Thomas (2017) 及 Gordon 等人 (2020) 的两篇研究。他们的研究均通过在叙述信息阶段加入次要任务以分散注意力，结果发现 RES 被消除了。经过专家提醒，我们认为该结果可以采用检索流畅性来解释，即注意力影响了错误信息的检索流畅性，但无法为检索流畅性提供直接有力的证据。因此，考虑到实验证据与论述内容的相关性，我们删除了“最新研究中，实验者要求被试在编码叙述信息时同时开展其他任务以分散注意力，结果表明，尽管中期测试依然增强了被试对错误信息的学习，但并未出现 RES 效应 (Gordon & Thomas, 2017; Gordon et al., 2020)”部分内容。

意见 8：

PET 现象需要给出一个具体的研究案例加以说明，否则普通读者可能并不明白该效应具体指的是何种现象。

回应：

非常感谢审稿专家的建议，我们在“3 中期测试如何发挥保护作用”第一段中添加了家庭场景范式的案例，向读者提供了能够生动解释该现象的生活实例，以增强论文的可读性。具体内容如下：

例如，在家庭场景范式中，首先给被试呈现一张家庭场景的照片（如“书桌”），场景中包含了一些生活用品，但缺少了最可能存在的物品（如“纸”和“笔”）。然后，被试会接触以音频或文本形式呈现的错误信息（如“他终于可以集中精神了，他咬了一口苹果，抓起笔，开始写文章”）。结果发现，中期测试组的被试在最终测试中更不可能报告本不存在的物品（如“笔”）。这意味着中期测试降低了错误信息的暗示性 (Pereverseff et al., 2020)。

相关参考文献：

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.

意见 9：

在作者建立的第一个模型中（见图 1），作者依然加入了“再巩固”解释，但该解释已经被大量研究证据所反驳，大量研究发现中部测试不仅没有降低原来信息的记忆，反而对原始信息记忆有促进作用。因此，请斟酌在你们最新的模型中，是否还加入该解释。

回应：

非常感谢审稿专家的意见。根据建议，我们认真考虑了中期测试和原始记忆的关系。鉴于广泛发现的测试效应，即中期测试可以增强个体对原始事件的记忆，这与“记忆再巩固”的解释相冲突。此外，错误信息相关研究中存在与“再巩固”矛盾的结果 (Chan et al., 2009; Chan & Langley, 2011; Gordon & Thomas, 2014)，例如，中期测试组的被试对原始信息和叙述信息的记忆都比对照组好。因此，我们在模型中删去了“记忆再巩固”部分内容，具体请见“2.3 RES 的作用路径”。

相关参考文献：

Chan, J. C. K., & LaPaglia, J. A. (2011). The dark side of testing: Repeated retrieval can enhance eyewitness suggestibility. *Journal of Experimental Psychology: Applied*, 17(4), 418–432.

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.

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

意见 10:

the bifurcation model theory 翻译成“歧化模型”是否更合适?

回应:

非常感谢审稿专家提出的意见。在我们目前所能查到的文献中，研究者提到“the bifurcation model theory”时有两种翻译：“二分模型假说”（王植洵，张锦坤，2017）和“基于分布的二分模型假说”（刘兆敏 等，2011；黄晓雪，张锦坤，2015）。因此，我们保留了最新用法，即“二分模型”。

相关参考文献:

黄晓雪，张锦坤. (2015). 提取练习效应的情境背景理论. *心理技术与应用*(12), 15 - 20.

刘兆敏，罗良，张玮. (2011). 记忆提取促进学习: 实验证据与机制解释. *心理发展与教育*(04), 441 - 447.

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(09), 1512 - 1520.

意见 11:

RES 和 PET 效应的区别的产生机制还需要进一步的完善，已有研究发现，信息呈现背景的一致性是一个重要的影响因素，当原始信息是以一个完整故事呈现，具有上下文关系时，更容易产生 RES；当信息是以情景图片呈现，而非故事性成现时，反而出现了 PET 效应。相关的研究结果和理论解释需要讨论。

回应:

(1) 非常感谢审稿专家的恳切建议。经过专家的提醒，我们查阅了“背景一致性”的相关文献。背景一致性是指当记忆提取时的背景与存储时的背景一致，人们可以回忆起更多的事件细节（Tulving & Thomson, 1973; Smith & Vela, 2001）。背景一致性可以影响错误信息效应，如在最终测试时恢复错误信息编码时的情景，导致了更大的暗示性（Roebbers & McConkey, 2003）。然而，我们未能从背景一致性的角度找到解释“故事呈现导致 RES，图片呈现导致 PET”的直接证据。一方面，背景一致性强调背景相同可以促进回忆，而在 RES 和 PET 中，测试时和编码时的背景都是不同的，且并未尝试恢复编码时的背景。另一方面，在我们目前了解的相关文献中，背景一致性研究并未涉及视频和图片两种呈现方式的区分。

尽管在材料类型的影响中并未找到“背景一致性”的相关研究证据，但在学习过程中，我们发现它可以为中期测试类型的影响提供解释。具体来说，认知访谈和自我管理访谈能够降低暗示性的原因可能是恢复了原始信息编码时的情景，促进了原始记忆与新的测试背景的融合，从而提高了在最终测试中的回忆率。我们在“4.2 中期测试类型”部分补充了相关内容，具体如下：

认知访谈和自我管理访谈可以降低错误信息暗示性的另一种解释来自背景一致性效应的相关研究。背景一致性是指当记忆提取时的背景与存储时的背景一致，人们可以回忆起更多的事件细节（Tulving & Thomson, 1973; Smith & Vela, 2001）。研究发现，背景一致性可以影响错误信息的暗示性（Roebbers & McConkey, 2003; Drohan-Jennings et al., 2010）。而中期测试采取认知访谈和自我管理访谈（包括恢复上下文、场景草图等手段）时，可以有效提取原始记忆的编码背景，从而提高原始信息回忆率（Hope et al., 2014; Pinto & Stein, 2015; Jack et al., 2015）。这一过程可以促进原始信息和新的测试背景相结合，因此被试在最终测试中（与中期测试相同的测试背景）更容易提取到原始信息（Karpicke et al., 2014），从而产生了 PET。

(2) 针对不同的原始信息材料分别产生了 RES 和 PET 效应的现象, 我们在“4.1 原始信息材料”部分补充了心理模型理论, 具体内容如下:

根据心理模型理论, 当原始事件具有叙述的连贯性、逻辑性和完整性时, 人们会形成关于事件如何展开的心理模型; 并且人们更喜欢坚持连贯和完整的心理模型, 即使它是不正确的 (Johnson & Seifert, 1994)。因此, 如果模型的一个组成部分缺失, 出现空白, 被试可能会接受不正确的信息以维持心理模型的完整性 (Lewandowsky et al., 2012)。相对应地, 在采用视频材料的研究中, 个体更关注故事情节的动态发展, 因而会构建一个包含丰富上下文信息的心理模型。如果原始信息记忆强度不足或无法访问, 被试便更可能接受错误信息以填补心理模型的缺失。然而, 在采用图片材料的研究中, 原始记忆和叙述信息都是由孤立的信息组成, 不存在上下文关系 (Huff et al., 2016; Pereverseff, 2020)。这时个体就难以构建连贯、完整的心理模型, 因而很难把错误信息整合到原来的记忆中, 错误信息的暗示性降低, PET 产生。

相关参考文献:

- 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.
- 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., 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.
- 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.
- 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.
- Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological science in the public interest*, 13(3), 106–131.
- 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.
- Pinto, L. H., & Stein, L. M. (2015). The theoretical basis of the technique of context reinstatement in the cognitive interview. *Avances en Psicología Latinoamericana*, 33(2), 285–302.
- Roebbers, 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.
- Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic bulletin & review*, 8(2), 203–220.
- Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80(5), 352–373.
-

审稿人 2 意见：

《测试是好是坏：中期测试对错误信息的影响》一文对文献资料把握较全面，涉及内容较详实，作者在分析已有文献的基础上尝试提出了 RES 产生的双路径模型以及 PET 产生的理论模型，表明作者对这一领域的研究问题进行了深入思考。但论文存在一些关键问题需认真考虑。

意见 1：

文章的理论价值不足。作者在引言中提出本文的核心问题是中期测试究竟在什么情况下增强暗示性，又在什么条件下起保护作用？其中的机制到底如何？作者试图将 RES 与 PET 作为两种对立的结果来分析，但实际上，已有研究已经指出测试的保护作用是在限制性条件下削弱错误信息的暗示性而产生的，那就表明 RES 和 PET 本身不是矛盾的现象。已有研究基本上已经对作者提出的问题进行了回答，本文没有新的理论贡献，如果作者将重点放在 RES 的双路径模型的提出上，可能会更好。

回应：

非常感谢审稿专家悉心提出的每一条意见和建议，我们根据建议做了逐一修改，修改的部分在原文中用红色字体标出。

经过专家的提醒，我们发现本文的理论价值和研究意义确实需要认真梳理从而进一步凸显。因此，我们从以下三方面进行了修改：

(1) RES 和 PET 是中期测试增强或减弱错误信息暗示性的不同表现，实际上回答了同一个问题，两者并不矛盾。但对 PET 进行理论探索依然是有价值的。对此，我们在 PET 的引入中简述了其研究与 RES 研究的互补性，并在“3 中期测试如何发挥保护作用”第二段补充了以下内容：

事实上，RES 和 PET 针对的是同一个问题，即“中期测试是增强抑或降低错误信息的暗示性？”故它们似乎是一种现象的两种矛盾表现。因此，这意味着对 PET 的研究可以补充对 RES 的讨论。首先，根据 RES 的现有理论及研究，只能解释如何将中期测试的负面影响降低至没有中期测试时的基线水平，而无法发挥出中期测试应有的优势。换句话说，RES 相关理论只能解释如何消除 RES，却无法说明何时产生 PET (Chan et al., 2009; Thomas et al., 2010; Gordon & Thomas, 2014, 2017; Gordon et al., 2020)。其次，探究如何发挥中期测试的正面影响并减少错误信息的暗示性，具有更广泛的现实意义。然而，目前却未形成解释 PET 现象的成熟理论框架，造成中期测试作用机制的部分空白。因此，我们梳理了相关研究和理论，包括记忆强度、提取努力理论、差异检测理论和迁移适当加工理论，以期为 PET 现象提供初步探索。

(2) 根据专家意见，我们调整了“RES 双路径模型”内容的比重，将“2.3 RES 的作用路径”作为单独一部分进行详细阐述。同时，考虑到“干扰与分离”作为“测试增强后续学习”中的其中一个机制，且文中并不涉及“干扰与分离”与“测试前学习”的相关表述，因此，我们在“RES 双路径模型”中删去了“干扰与分离”。具体改动如下：

2.3 RES 的作用路径

综上，可以看到，尽管研究者们为 RES 现象提供了丰富多样的理论解释，但忽视了理论间的相关性和互补性，因而无法把握 RES 产生机制的全貌。基于此，本文提出了一个针对原始信息和叙述信息学习的双路径模型（如图 1 所示）。

根据该模型，中期测试同时影响个体对原始信息和错误信息的记忆，并最终导致了 RES。在第一条路径中，根据测试效应 (Roediger & Karpicke, 2006b; 王植洵, 张锦坤, 2017; Yang et al., 2021)，中期测试可以增强个体对原始事件的记忆。然而，错误信息的出现扭转了这一作用。尽管已有研究证明原始记忆并没有被完全破坏 (McCloskey & Zaragoza, 1985; Chan et

al., 2009; Gordon & Thomas, 2014)，但是错误信息的出现仍然降低了原始信息的可访问性，进而降低了最终测试正确率。

在第二条路径中，中期测试通过多种机制增强了对错误信息的学习。首先，中期测试隔离了两次信息材料，这一方面减少了原始信息的记忆对错误信息学习的干扰（Szpunar et al., 2008），另一方面减少了最终测试中的处理负荷，从而提高了错误信息的回忆成绩（Pajkossy et al., 2019）。其次，检索失败、测试期望等会激发被试对叙述信息努力编码的动机（Richland et al., 2009; Weinstein et al., 2014; Cho et al., 2017; Yang et al., 2018）。最后，中期测试会提高被试的注意力（Thomas et al., 2010; Gordon & Thomas, 2014; Gordon et al., 2015），增加被试对叙述信息的阅读时间同时增强对错误信息的编码（Pastötter et al., 2011; Cho et al., 2017）。同时，当错误信息的编码被增强后，被试更易在最终测试中提取到错误信息，即检索流畅性提高（Thomas et al., 2010; Gordon & Thomas, 2014）。

值得注意的是，RES 的产生最终取决于原始信息可访问性和错误信息检索流畅性的相对强度，即不能孤立地看待任何一条作用路径。Thomas 等人（2010）认为，RES 是由于个体在检索到具有更高流畅性的错误信息之后，过早地终止了回忆原始信息所需要的进一步检索。这说明正是原始信息在“流畅性”的竞争中失败，导致了它被抛弃的命运。从这个角度，提高原始信息的可访问性或降低错误信息的检索流畅性，或者提供进一步检索到原始信息的机会即可减少 RES（Chan et al., 2009; Gordon & Shapiro, 2012; Gordon & Thomas, 2014; Gordon & Thomas, 2017; Gordon et al., 2020）。

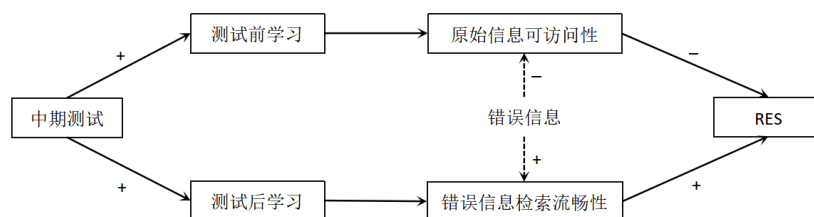


图 1 RES 产生的双路径模型图

(3) 鉴于专家指出，本文理论价值不足，并考虑到 RES 和 PET 实际上解决了同一个问题，即“中期测试是增强或是降低错误信息的暗示性”，我们在 RES 和 PET 相关论证结束后整合了两者的作用机制，并提出了一个总模型。我们补充了“3.6 统合 RES 和 PET 的总模型”部分，内容如下：

3.6 统合 RES 和 PET 的总模型

上述分析表明，RES 和 PET 的理论解释角度多样、内容丰富，为中期测试作用的产生提供了多维解读。然而，作为两种矛盾的现象，RES 和 PET 的研究难以孤立发展，解释间的交融互通必不可少。为此，如图 3 所示，我们根据作用阶段与作用关系统合了两种现象的相关理论，并提出了一个涵盖各大主流解释的上位模型，从而为中期测试对错误信息暗示性的影响机制提供更加深入全面地考察与探究。

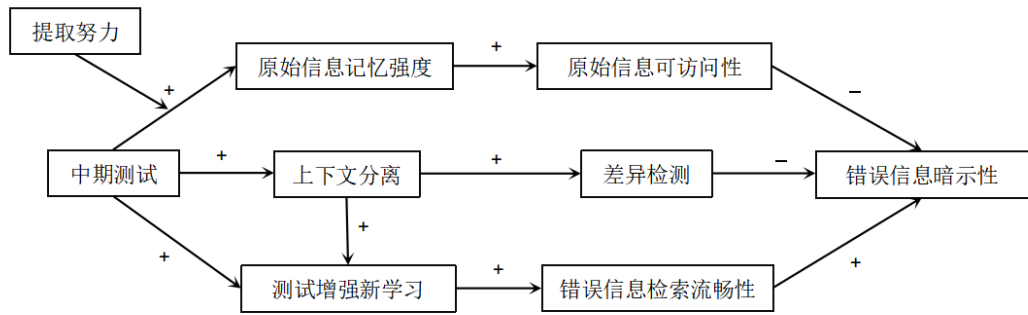


图3 中期测试对错误信息暗示性的作用路径图

我们认为，中期测试主要通过三条路径影响错误信息的暗示性。第一条路径解释了中期测试对原始信息的影响。首先，鉴于广泛存在的测试效应（Roediger & Karpicke, 2006b; 王植洵, 张锦坤, 2017; Yang et al., 2021），中期测试可以增强原始信息的记忆。根据提取努力理论，中期测试难度更大时，更能增强原始信息的储存强度（Karpicke & Roediger, 2007; Stenlund et al., 2014）。但这一过程取决于中期测试是否提取成功（Bjork & Bjork, 1992, 2011; Kornell et al., 2011; 王植洵, 张锦坤, 2017）。当原始信息被提取成功且记忆强度增加时，被试在最终测试时可以更轻松地访问原始信息。

第二条路径涉及错误信息的影响。与 RES 模型一致，在此不再赘述。第三条路径同时作用于原始信息和错误信息。首先，中期测试使上下文产生变化，导致编码重置，从而区分了原始信息和错误信息两次编码事件（Szpunar et al., 2008; Pastötter et al., 2008, 2011, 2018; Chan et al., 2009; Wahlheim, 2015; Bufo & Aslan, 2018; Dang et al., 2021）。其次，这种更明确的区分有利于被试察觉信息之间的差异，从而降低错误信息的暗示性（Tousignant et al., 1986; LaPaglia, 2013; Mullet & Marsh, 2016; Polczyk, 2017; Putnam et al., 2017; Sheaffer et al., 2021; Bailey et al., 2021）。

这一模型整合了 RES 和 PET 的相关理论研究，对中期测试影响错误信息暗示性的机制进行了更详细的阐述。但该模型是在现有相关理论上初步搭建的，各理论之间存在何种联系、路径之间如何相互作用，以及是否存在其它关键变量和影响因素，仍需更多的理论和实证研究加以回应和补充。

意见 2:

综述文章写作的核心线索不突出。本文在阐述中期测试如何增强错误信息的暗示性时，从原始记忆受损和错误信息学习程度增强两个方面来阐述。但是在错误信息学习程度增强这个方面，用较长的篇幅列举了中期测试减少了前摄性干扰，加强了新信息学习的机制解释与实验证据。但是，只有检索流畅性假说是比较直接的支持性假说，其余的内容占据大量篇幅却都在讲测试如何增强学习，而不是如何增强错误记忆，写作偏离了主题。

回应:

非常感谢审稿专家的意见。经过专家提醒，我们发现原文在“测试增强后续学习”部分确实花费过多笔墨在与错误信息无关的细节上，偏离了主线问题。其中，“2.2.1 干扰与分离”和“2.2.2 学习与编码”两节对学习领域的研究讲解过于冗长，且未能很好地将不同领域的研究联系起来以佐证有关测试作用的论点。因此，为了强调本文的核心线索，我们调整了原文中这两节的论证方式，改为“以研究引问题，以问题带研究”。一方面在结构上凸出测试作用的中心性，将错误信息领域的观点和研究当作重点，学习领域的理论和证据用作补充；另一方面在内容上对学习领域的研究证据进行了大量删减，增加了中期测试与错误信息

相关内容，以突出两者的联系。具体修改内容如下：

(1) “2.2.1 干扰与分离”部分删去了部分学习领域的研究，并缩减了字数，具体改动如下：

原文：

Chan 等人（2009）认为，中期测试之所以能够增强对事后错误信息的学习，其中一个原因是测试区分了两次不同的编码事件（原始信息和叙述信息），从而减少了原始记忆对新信息学习的影响。在正向测试效应的相关研究中，这种促进上下文分离并减少前摄干扰的现象已被广泛发现（Szpunar et al., 2008; Wahlheim, 2015; Bufe & Aslan, 2018）。如在 Szpunar 等人（2008）的经典研究中，当学习 5 个列表的单词时，如果分别对前四个列表进行测试，可以提高对第 5 个列表回忆的正确率，且先前列表中单词的入侵也更少。他们采用来源监控和信息过载进行解释：如果先前的列表没有进行测试，那么在对最后材料进行测试时，被试需要对之前出现的所有信息进行检索，信息负荷过重，可能会导致材料之间的混淆以及先前材料的干扰。加入测试有利于被试将列表彼此分开并加以区分，一方面可以更好地进行来源监测，另一方面也减少了信息过载，有利于对新信息的学习。

最新研究发现，测试可以通过减少旧材料的干扰来增强对新信息的学习，但若被试没有意识到二者之间的区别，则干扰依然存在（Wahlheim, 2015; Dang et al., 2021）。这说明测试可能是通过上下文的变化来帮助个体区分测试前后的学习，以减少信息间的干扰，提高学习成绩。目前，这一解释得到了来自认知神经科学研究证据的支持（Pastötter et al., 2011）。尽管如此，测试是通过减少干扰来增强学习的这一解释仍然受到了挑战。如在 Wissman 等人（2011）的研究中，他们加入了一个没有先前材料的组（不存在干扰），结果发现学习并进行中期测试的被试对新材料的回忆率竟然高于无干扰组。这可能意味着测试的“好处”并不止于减少干扰。

就错误信息领域而言，中期测试可能分离了原始信息与叙述信息，减少了前者对后者的干扰，提高了最终测试中错误信息的回忆率。但需要注意的是，一方面，学习领域对测试效应的研究大多采用单词对、词表或文本材料，而错误信息领域则更多考察情景记忆；另一方面，由于研究目的的不同，错误信息领域的新旧材料在关系上具有更加明显的相关性和冲突性。那么，中期测试是否仍能在学习叙述信息时促进分离、减少干扰？因此，将这一解释应用于错误信息范式时还需进一步的验证。

修改后：

Chan 等人（2009）认为，中期测试之所以能够增强对事后错误信息的学习，其中一个原因是测试区分了两次不同的编码事件（原始信息和叙述信息），从而减少了原始记忆对新信息学习的影响。**测试**促进上下文分离并减少前摄干扰的现象已被广泛发现（Szpunar et al., 2008; Wahlheim, 2015; Bufe & Aslan, 2018; Dang et al., 2021）。如在 Szpunar 等人（2008）的经典研究中，当学习 5 个列表的单词时，如果分别对前四个列表进行测试，可以提高对第 5 个列表回忆的正确率，且先前列表中单词的入侵也更少。

测试可以增强新材料的学习，可能是由于测试导致上下文产生变化，被试在最终测试时只需检索最新编码的材料，减少了信息过载及材料之间的混淆和干扰，从而提高了新材料的回忆成绩（Szpunar et al., 2008）。Pajkossy 等人（2019）的研究为这一解释提供了生理学证据。他们发现，与重新学习相比，中期测试减少了被试在最终测试中的瞳孔扩张。瞳孔大小反映了执行任务时的资源处理量，这说明中期测试减少了最终回忆时的处理负荷。

就错误信息领域而言，中期测试可能分离了原始信息与叙述信息，减少了前者对后者的干扰，提高了最终测试中错误信息的回忆率。但需要注意的是，**错误信息领域的新旧材料在关系上具有更加明显的相关性和冲突性，且更多考察情景记忆。那么，中期测试是否仍能促进上下文分离并减少干扰还需进一步的验证。**

(2) “2.2.2 学习与编码”部分第一段中,删去了“Wissman 等人(2011)在考虑了来源混淆……学习和编码”研究内容的阐述,具体如下:

原文:

在测试效应的相关研究中,测试可以增加对新材料的学习效果已经被多次证实(Roediger & Karpicke, 2006a; Richland et al., 2009; Carpenter, 2011)。Wissman 等人(2011)在考虑了来源混淆、干预效果、测试期望等可能性后,依然发现了测试对于新材料学习的稳定的增强效果。因此,他们认为,对先前材料的测试(尤其当测试失败时)会增强对测试后材料的学习和编码。Gordon 和 Thomas (2017)在最终测试时测试了被试对叙述信息的记忆,结果发现中期测试组的被试回忆起更多的测试后信息,为中期测试增强叙述信息的学习提供了直接证据。对此,目前存在三种不同的理论解释。

修改后:

在测试效应的相关研究中,测试可以增加对新材料的学习效果已经被多次证实(Roediger & Karpicke, 2006a; Richland et al., 2009; Carpenter, 2011; **Wissman et al., 2011**)。Gordon 和 Thomas (2017)在最终测试时测试了被试对叙述信息的记忆,结果发现中期测试组的被试回忆起更多的测试后信息,为中期测试增强叙述信息的学习提供了直接证据。对此,目前存在三种不同的理论解释。

(3) “注意捕捉假说”部分主要调整了文本结构,缩减了篇幅,具体改动如下:

原文:

Gordon 和 Thomas (2014)发现,接受过中期测试的被试对叙述信息的阅读时间要显著长于没有接受中期测试的被试。阅读时间是注意力的一个重要指标,中期测试可能增加了被试对叙述信息的注意力。更进一步,在他们的研究中,当中期测试的内容为错误信息相关细节而非其本身时,被试同样倾向于报告该信息。这说明中期测试可能增强了个人对整体而非细节的学习。据此,研究者提出了注意捕捉假说(the attention capture hypothesis, Thomas et al, 2010; Gordon & Thomas, 2014; Gordon et al, 2015)。该假说认为在中期测试后,与测试相关的信息更容易吸引被试的注意力,从而加强对错误信息的学习。

研究发现,测试可以增强个体对后续学习材料的注意,增加对材料的阅读时间,并减少心不在焉的情况(Szpunar et al, 2013)。而在中期测试后加入次要任务,以分散被试对叙述信息的额外注意力时,中期测试组和对照组产生的错误信息没有差异,即 RES 效应被抵消了(Gordon & Thomas, 2017)。这些证据说明 RES 产生的部分原因是因为增强了对测试后信息的额外注意。

但应该说明的是,注意力分配并不能完全解释 RES 效应。例如,与中期测试直接询问错误信息相比,询问相关细节尽管会产生相似水平的 RES,但阅读时间却明显更短(Gordon & Thomas, 2014)。除此之外,通过加粗或者斜体、红色闪烁字体等方式吸引注意力后,尽管被试花费更多时间学习所强调的内容,但是学习效果仍然不如测试增强的学习(Richland et al, 2009; Shapiro & Gordon, 2012)。Thomas 等人(2017)则发现学习效果的差异与最终测试的时间有关:当立即进行最终测试时,中期测试组与强调细节组(标有红色下划线)表现出了相似的 RES;但当测试延迟 48 小时时,中期测试组的表现要比强调细节组和对照组更好。因此,RES 背后可能有更复杂的机制。

修改后:

注意捕捉假说 (the attention capture hypothesis, Thomas et al, 2010; Gordon & Thomas, 2014; Gordon et al, 2015) 认为,中期测试提高了测试相关信息的吸引力,从而加强了被试对错误信息的学习。Gordon 和 Thomas (2014) 发现,接受过中期测试的被试对叙述信息的阅读时间要显著长于没有接受中期测试的被试,这意味着中期测试可能提高了被试对叙述信息的注意。对此,当中期测试后加入次要任务以分散被试注意力时,测试组与对照组的错误

信息率便没有明显差异,即 RES 效应被抵消了(Gordon & Thomas, 2017; Gordon et al., 2020)。这些证据说明,对错误信息的额外注意可以解释 RES 的产生。

值得一提的是,这种解释并不完全。例如,当立即进行最终测试时,中期测试组与强调细节组(通过红色下划线吸引注意力)表现出了相似的 RES;但当测试延迟 48 小时时,中期测试组的表现要比强调细节组和对照组更好(Thomas et al., 2017)。这说明,吸引注意力并不能产生与中期测试相同的效果。因此,RES 背后可能具有更为复杂的机制。

(4)“编码方式”部分删减了不太重要的部分内容,并调整了语言表述,具体改动如下:

原文:

测试增强学习,也可以说是增强了对新信息的编码。为此,研究者提出了编码重置理论(the encoding reset theory, Pastötter et al., 2011)和编码策略理论(the encoding strategy theory, Cho et al., 2017)。

编码重置理论认为,测试使得上下文分离并减少了记忆负荷以及低效率的编码,从而使得新材料的学习和早期编码一样有效(Yang et al., 2018)。Pastötter 等人(2018)对同一列表中不同序列位置的回忆曲线的研究,证明了在连续的学习中,早期位置的材料学习效果更好,这说明在持续的学习中学习效率会降低。测试使测试前和测试后材料获得了相似的序列位置的曲线,说明测试使编码重置了,且增强了测试后材料的学习。该研究为编码重置理论提供了证据。在神经认知领域, α 能量会随着持续学习而不断增加,但在加入测试后, α 能量与开始学习时没有差异,并且产生了更好的测试后学习效果(Pastötter et al., 2011)。

Cho 等人(2017)提出了编码策略理论,即正向测试效应源于编码策略的改变。他们认为,测试使被试认识到自己的记忆存在不足,并在之后的学习中花费更多时间学习后续材料。Soderstrom 和 Bjork(2014)的研究提供了更有力的证据——与重新研究相比,中期测试组在第二次学习期间花费时间更久,且多花费在中期测试未能成功回忆的项目以及较难的项目上;同时,测试组在后期学习中会使用更多有效的学习策略(如语义参考:将单词对与生活现实情况联系起来),使其在最终测试的回忆率更好。Dang 等人(2021)也发现,测试组的被试会使用更好的语义加工策略,从而提高最终成绩。此外,有实验表明也发现,不同的编码策略影响了被试对错误信息的敏感性(LaPaglia, 2013)。

除了以上两种理论,中期测试还可能通过无意识地激活相关信息来增强后续编码(Carpenter, 2011; Grimaldi & Karpicke, 2012)。例如,Carpenter(2011)的研究说明,测试可能激活了一个与线索和目标都相关的中介信息,并通过中介信息来增强对目标的记忆。但当错误信息出现时,中介信息反而会同时与矛盾的信息产生联系,从而干扰或混淆原始记忆,产生 RES 效应。

修改后:

中期测试增强对新信息的学习,即增强了对错误信息的编码,因此提高了它们的暗示性。在学习领域,研究者提出了编码重置理论(the encoding reset theory, Pastötter et al., 2011)和编码策略理论(the encoding strategy theory, Cho et al., 2017)。

编码重置理论认为,在持续的学习中编码效率会不断降低,而测试能够分离上下文并减少记忆负荷与低效率编码,从而使得个体对新材料的学习和早期编码一样有效(Pastötter et al., 2018; Yang et al., 2018)。认知神经领域的研究发现, α 能量会随着持续学习而不断增加,但在加入测试后, α 能量重新降至初始水平,并且表现出更好的学习效果(Pastötter et al., 2011),这为该理论提供了强有力的生理证据。目前,该理论尚未应用于错误信息研究,之后可结合认知神经研究对 RES 加以验证。

Cho 等人(2017)提出了编码策略理论,即测试增强新学习源于编码策略的改变。通过中期测试,被试认识到自己的记忆存在不足,因而会在后续学习中更加努力,并寻求更好的

记忆方法 (Soderstrom & Bjork, 2014; Cho et al., 2017)。换言之, 中期测试能够促使被试使用更有效的编码策略, 从而提高记忆成绩 (Dang et al., 2021)。错误信息领域的研究发现, 不同的编码策略可以影响被试对错误信息的敏感性 (LaPaglia, 2013)。因此, 中期测试可能通过编码策略的改变增强对错误信息的学习, 最终导致出现 RES 效应。

(5) “动机理论”部分删减了不太重要的细节, 增加了错误信息相关表述, 具体改动如下:

原文:

上述解释主要从认知角度出发, 阐明了测试对底层学习机制的影响, 但忽视了学习者作为一个能动的主体, 有着自身的动机与需要。为此, 可以参考学习领域的相关研究。Yang 等人 (2021) 对课堂教学的测试效果进行了元分析, 发现测试效果和重复测试次数、时间以及周期正相关, 即测试的次数越多, 不断“学习-测试”的持续时间越长、周期越多, 测试效果越好。动机理论 (the motivation theory) 从检索失败、测试期望等角度进行了解释。

测试增强学习的一个典型现象是中期测试检索失败, 即当被试未能成功回忆起事件细节时, 会意识到自己记忆模糊, 从而增强被试对后续材料的学习和关注 (Cho et al., 2017; Yang et al., 2018)。Richland 等人 (2009) 发现, 与延长学习时间、强调细节内容、牢记测试问题等方式相比, 被试对初次测试回答错误的项目记忆更好, 且该优势至少能够保持一周。他们认为, 这是由于检索失败增强了问题和答案之间的检索路径, 或鼓励被试进行更精细的深度处理。

类似地, 搜索集理论 (the search set theory) 认为, 测试失败后的检索尝试会激活一个可能答案的集合, 当再次学习的材料与之重叠时, 便会增强相应项目的学习 (Grimaldi & Karpicke, 2012)。在研究中, 如果测试或材料引导被试关注另外一个相关但不正确的选项 (类似于错误信息), 被试的错误率便会高于未测试组, 且表现出更多的前摄干扰与材料入侵。这在某种程度上解释了 RES 效应的产生。当被试在中期测试回答错误或者不自信时, 会自发产生一个搜索集。如果叙述信息中刚好出现集合中的选项 (可能是错误信息), 便会增强该选项的记忆。这是中期测试促进测试后信息的学习和编码的一种解释。

另外, 参加中期测试可能会激活更高的测试预期, 即预测之后仍要进行测试, 这被称为测试期望理论 (the test expectancy theory, Weinstein et al., 2014; Yang et al., 2018)。Weinstein 等人 (2014) 发现, 告知被试之后会进行测试 (即测试期望), 使未测试组和测试组一样增强了个体对新材料的学习, 并减少了先前材料的干扰。Agarwal 和 Roediger (2011) 采用闭卷考试和开卷考试两种测试期望来测量被试两天后的测试表现, 结果发现, 预期最终会进行闭卷考试的被试比预期开卷的被试花费更多时间进行学习, 且两天后最终成绩更好。这说明测试期望, 尤其是预期之后进行更具难度的测试会增强被试对材料的学习。但它无法完全解释正向测试效应。如 Geller 和 Peterson (2021) 发现, 相比于低测试期望组, 高测试期望的被试尽管在识别回忆和提示性回忆中表现更好, 但并不总会花费更多的学习时间。又如在学习材料前让被试了解之后可能进行的测试类型, 并不能产生正向测试效应 (Wissman et al., 2011; Choi & Lee, 2020)。

修改后:

上述解释主要从认知角度出发, 阐明了测试对底层学习机制的影响, 但忽视了学习者作为一个能动的主体, 有着自身的动机与需要。为此, 动机理论 (the motivation theory) 从检索失败、测试期望等角度进行了补充解释。

当被试未能成功回忆起事件细节, 即检索失败时, 会意识到自己记忆模糊, 从而增强他 (她) 对后续材料的关注和学习 (Cho et al., 2017; Yang et al., 2018)。Richland 等人 (2009) 认为, 这是由于检索失败增强了问题和答案之间的检索路径, 并鼓励被试进行更精细的深度处理。Grimaldi 和 Karpicke (2012) 则认为, 测试失败后的检索尝试会激活一个可能答案的

集合，当再次学习的材料与之重叠时，便会增强相应项目的学习。如果错误信息刚好出现在这个集合中，便会产生 RES 效应。

另外，参加中期测试可能会激活更高的测试预期，即预测之后仍要进行测试，这被称为测试期望理论（the test expectancy theory, Weinstein et al., 2014; Yang et al., 2018）。研究发现，测试期望，尤其是预期之后进行更具难度的测试会增强被试对材料的学习（Agarwal & Roediger, 2011; Weinstein et al., 2014）。在错误信息范式中，被试在接触原始信息时并无测试预期，但在编码错误信息时却有（源于中期测试），这会导致他们对原始信息的学习不如对错误信息的学习，进而产生 RES。

意见 3:

对 PET 的理论解释多借鉴测试效应的解释理论，但是没有对这些理论如何解释测试降低对错误信息的敏感性进行重点分析。

回应:

非常感谢审稿专家指出该问题。我们调整了“3 中期测试如何发挥保护作用”部分的行文逻辑和结构，并增加了错误信息敏感性的相关分析，具体修改内容如下：

(1) “3.1 记忆强度”部分删去了郭秀艳和李荆广（2007）的研究的详细表述，并在第一段和第二段分别加入了错误信息的观点阐释，另对部分语句进行了调整。具体改动如下：

原文:

有研究者认为，对原始事件的“强大”记忆能使人们更容易发现原始事件与错误信息之间的差异，从而拒绝错误信息，做出正确的选择（Peterson et al., 2004; Loftus, 2005）。该解释得到了包括错误信息在内的多个领域研究支持。

在错误信息范式中，原始记忆是否强大（记忆强度）是以中期测试正确率为指标（Wang & Yang, 2021）。Chan 和 Langley（2011）发现，当中期测试回答正确时，被试报告错误信息的概率与未接受中期测试的被试差异不显著，这说明强大的记忆抵消了 RES 效应。在 LaPaglia 和 Chan（2013）的研究中，以问题题干的形式呈现错误信息时，被试在中期测试回答正确时的最终测试正确率比错误时更高，甚至不受错误信息影响，错误信息效应被缓解了。Gabbert 等人（2012）也认为，出现 PET 的原因是测试增强了被试的原始记忆强度，从而更能进行差异检测。在该研究中，被试的正确率越高，即记忆越好，错误信息的报告率越低，两者呈负相关。因此，被试对原始事件的记忆足够准确、牢固，可能是出现 PET 的原因之一。

在非标准的错误信息范式中，郭秀艳和李荆广（2007）发现，原始信息呈现时间越长，真实记忆越强，错误记忆越弱；两周后，错误记忆程度增强，这可能与真实记忆强度下降有关。他们认为，被试的受暗示性正是由于自身的记忆强度不足，对自己的判断没有自信，从而更易被误导。

修改后:

有研究者认为，被试的受暗示性是由于自身的记忆强度不足，对自己的判断没有自信，从而更易被误导。因此，对原始事件的“强大”记忆能使人们更容易发现原始事件与错误信息之间的差异，从而拒绝错误信息，做出正确的选择（Peterson et al., 2004; Loftus, 2005; 郭秀艳, 李荆广, 2017）。该解释得到了包括错误信息在内的多个领域研究支持。

在错误信息范式中，原始记忆是否强大（记忆强度）是以中期测试正确率为指标（Wang & Yang, 2021）。Chan 和 Langley（2011）发现，当中期测试回答正确时，被试报告错误信息的概率与未接受中期测试的被试差异不显著，这说明强大的记忆抵消了 RES 效应。在 LaPaglia 和 Chan（2013）的研究中，以问题题干的形式呈现错误信息时，被试在中期测试

回答正确时的最终测试正确率比错误时更高，甚至不受错误信息影响。Gabbert 等人（2012）也认为，出现 PET 的原因是测试增强了被试的原始记忆强度，从而更能进行差异检测。在该研究中，被试的正确率越高，即记忆越好，错误信息的报告率越低，两者呈负相关。因此，被试对原始事件的记忆足够准确、牢固，并以此抵御了错误信息的影响，可能是出现 PET 的原因之一。我们假设，当个体对原始信息的记忆足够强大时，中期测试会发挥其应有的保护作用，降低错误信息的暗示性；当原始记忆强度不足时，个体更容易被错误信息误导，出现 RES。这一推论需要未来研究的进一步验证。

（2）“3.2 提取努力理论”部分调整了段落结构，删去了偏离核心线索的研究，并增加了对错误信息敏感性的论述。具体改动如下：

原文：

提取努力理论（the retrieval effort theory）认为，测试效应缘于从记忆中检索学习信息的过程中所花费的认知努力。Bjork（1975）认为，深度的、努力的提取可以增强测试效果。与容易的检索相比，困难检索将更有利于长期保留（Karpicke & Roediger, 2007; Stenlund et al., 2014）。该理论在测试效应领域得到了广泛的支持。例如，相比而言，那些在初始测试中进行深度提取的被试在随后的测试中再认成绩更好（Jacoby et al., 2005）；增加初始测验的间隔会使得提取更困难，进而提高被试的记忆效果（Karpicke & Roediger, 2007）。Rowland（2014）的元分析也发现，与识别测试相比，难度更大的回忆测试产生了更明显的测试效应。但是元分析的结论并不一致，如 Adesope 等人（2017）的元分析发现，多项选择测试比线索回忆、自由回忆和短回答测试产生了更大的记忆收益；Yang 等人（2021）的元分析则发现，识别测试和回忆测试产生的测试效应没有差异。

为了进一步解释该现象，研究者还发展出了必要难度理论（the desirable difficulty theory, Bjork, 1992, 2011）和二分模型理论（the bifurcation model theory; Kornell et al., 2011）。必要难度理论区分了储存强度（storage strength）和提取强度（retrieval strength）。前者反映了记忆痕迹的相对持久性或学习的持久性；后者反映了记忆痕迹的瞬时可及性，即记忆能够被想起的容易程度，类似于检索流畅性。与此同时，该理论假设提取强度与储存强度的增加负相关：只有当花费较多努力提取成功时（低提取强度），提取练习才能增强记忆强度，促进长期学习（高储存强度）。在二分模型假说中，所有项目原有的提取强度呈正态分布。当重复学习时，项目的提取强度增加幅度较小，整体仍呈现正态分布；当提取练习时，提取成功的项目提取强度增加，提取失败的项目提取强度减弱，其中增加幅度取决于提取难度——难度越高增幅越大（王植洵, 张锦坤, 2017）。

具体到错误信息领域，该理论也为 PET 现象的解释提供了一些借鉴。当被试完成中期测试，即花费更多认知努力进行提取后，他们对原始信息的记忆会得到增强，从而更容易甄别错误信息，以降低其暗示性。最直观的证据来自 Pansky 和 Tenenboim（2011），他们将信息分为基本层（如“椅子”）和从属层（如“木椅子”），并在两种不同的深度进行要点测试和逐字测试。其中，要点测试要求被试在基本层回答问题，而逐字测试要求被试必须在从属层回答问题。结果发现，两种测试均可以产生测试效应，但只有逐字测试可以降低暗示性。

修改后：

在发现 PET 现象的研究中，当被试在中期测试时花费更多努力以提取信息时，他们对原始信息的记忆会得到增强，从而更容易甄别错误信息，以降低其暗示性。最直观的证据来自 Pansky 和 Tenenboim（2011），他们将信息分为基本层（如“椅子”）和从属层（如“木椅子”），并在两种不同的深度进行要点测试和逐字测试。其中，要点测试要求被试在基本层回答问题，而逐字测试要求被试必须在从属层回答问题。结果发现，两种测试均可以产生测试效应，但只有逐字测试可以降低暗示性。

对此，提取努力理论（the retrieval effort theory）认为，测试效应缘于从记忆中检索学习信息的过程中所花费的认知努力（Bjork, 1975; Karpicke & Roediger, 2007; Stenlund et al., 2014; Rowland, 2014）。该理论在测试效应领域得到了广泛的支持。例如，相比而言，那些在初始测试中进行深度提取的被试在随后的测试中再认成绩更好（Jacoby et al., 2005）。为了进一步解释该现象，研究者发展出了必要难度理论（the desirable difficulty theory, Bjork, 1992, 2011）和二分模型理论（the bifurcation model theory; Kornell et al., 2011）。必要难度理论区分了储存强度（storage strength）和提取强度（retrieval strength）。前者反映了记忆痕迹的相对持久性或学习的持久性；后者反映了记忆痕迹的瞬时可及性，即记忆能够被想起的容易程度，类似于检索流畅性。与此同时，该理论假设提取强度与储存强度的增加负相关：只有当花费较多努力提取成功时（低提取强度），提取练习才能增强记忆强度，促进长期学习（高储存强度）。在二分模型假说中，所有项目原有的提取强度呈正态分布。当重复学习时，项目的提取强度增加幅度较小，整体仍呈现正态分布；当提取练习时，提取成功的项目提取强度增加，提取失败的项目提取强度减弱，其中增加幅度取决于提取难度——难度越高增幅越大（王植洵, 张锦坤, 2017）。

根据该理论，被试在中期测试中花费认知努力检索原始信息，可以增强原始信息的记忆强度。尤其对于记忆强度不足的信息，当花费更多努力并成功提取时，该信息的记忆强度会得到大幅提升；如果在中期测试阶段提取失败，被试可能失去信心，从而放弃在之后的测试中进行更进一步的检索。原始记忆足够强时，中期测试中可能不需花费太多努力便可轻松检索到，但相应的，这些记忆强度也无法得到进一步增强。可见，提取努力和记忆强度之间存在交互关系。如果想要促进中期测试的正面作用，应加大中期测试的难度或检索深度。

（3）“3.3 差异检测理论”部分进行了两处修改。首先，将第一段中“如在 LaPaglia（2013）的实验中……更不容易被暗示”修改为“研究表明，被试在阅读叙述信息时，如果对每一条叙述信息进行差异检测，可以有效降低错误信息的暗示性（LaPaglia, 2013; Bailey et al., 2021）”。

其次，结合审稿人1的意见6，我们将“警告”的相关研究统一放在了“5 总结与展望”第五段。因此，为了突出差异检测在错误信息效应中的作用，我们将原论文中“4.1 原始信息”部分中关于“矛盾型错误信息和附加型错误信息”的内容替换了“警告”相关内容，放在了“3.3 差异检测理论”部分第二段。具体内容如下：

矛盾型和附加型错误信息的相关研究可以为差异检测理论补充证据。其中，矛盾型错误信息是指与原始事件相冲突的信息，即“指鹿为马”；附加型错误信息是指原始事件中没有出现但在叙述信息中出现的消息，即“无中生有”。两者的区别在于，矛盾型错误信息和原始记忆存在冲突，更适合进行差异检测。研究发现，与附加型错误信息相比，被试报告更少的矛盾型错误信息（Huff & Umanath, 2018; Umanath et al., 2019），且这种差异在中期测试组更明显（Gordon et al., 2015）。这说明中期测试进一步增强了差异检测。

（4）“3.4 迁移适当加工理论”部分第二段末尾增加了“这意味着中期测试和最终测试在类型或其他条件上的匹配是有效的，可能影响测试对原始信息的“接力传递”，进而影响PET的产生”的描述。

对第三段内容也进行了修改，具体如下：

原文：

然而，对于该理论也存在一些质疑。如Kang等（2007）采用两种测试形式（简答题和选择题）作为初始测试，考察被试在初始测试形式与最终测试形式相同和不同时的成绩，结果发现，相对于无测试组，两种测试形式都会产生测试效应，且简答式测试比选择题测试更好地促进了被试最终测试（不论形式如何）的成绩。在Rowland（2014）的元分析中，格式匹配并没有调节测试效应——与识别测试相比，不论最终测试的格式是否一致，初始测试使

用自由回忆测试时的测试效应更大。

修改后：

然而，对于该理论也存在一些质疑。如 Kang 等（2007）采用两种测试形式（简答题和选择题）作为初始测试，考察被试在**测试格式匹配和不匹配**时的成绩，结果发现，简答式测试比选择题测试更好地促进了被试最终测试（不论形式如何）的成绩。**Rowland（2014）的元分析也得到了类似结果——与识别测试相比，中期测试使用自由回忆时得到的测试效应更大。这似乎意味着测试效应取决于提取努力而非格式匹配。因此，在错误信息领域，测试格式的匹配是否可以增强 PET 效应还需进一步验证。**

意见 4：

文章中的小标题不够严谨。就第四部分而言，4.2 和 4.3 非常容易引起歧义，第四部分本来在谈中期测试产生两种分离结果的潜在影响因素，那 4.2 的标题是中期测试，中期测试本身如何影响中期测试呢？应该写成中期测试的类型。同理，4.3 应该写成错误信息的特点。

回应：

非常感谢审稿专家的提醒，我们将“4 RES 与 PET 分离的潜在影响因素”部分的小标题分别修改为“4.1 原始信息**材料**”、“4.2 中期测试**类型**”、“4.3 错误信息**特点**”。

第二轮

审稿人 1 意见：

经过上一轮的修改，该论文已有较大提高，基本达到可被发表水平。但仍有一些不足之处需要进一步修改。

意见 1：

摘要的第一句话存在较大问题“其中测试增强暗示性（retrieval enhanced suggestibility, RES）是指中期测试破坏了原始记忆中期测试并没有破坏原始记忆”，大量 MMFR 和选择再认研究表明，中期测试并没有破坏原始信息的记忆，RES 的产生可能是其他原因导致的。

回应：

非常感谢审稿专家第二轮细致而深入的意见，据此我们在原文中进行了相应的补充，第二轮修改的部分在原文中用绿色字体标出。

经过专家提醒，我们发现摘要中对 RES 的定义确实存在问题。因此，我们参照已有研究（Thomas et al., 2010; Gordon et al., 2015），将“其中测试增强暗示性（retrieval enhanced suggestibility, RES）是指中期测试破坏了原始记忆，导致被试更易接受或报告错误信息”修改为“其中测试增强暗示感受性（retrieval enhanced suggestibility, RES）是指接受中期测试的被试在最终记忆测试中的正确率较低，且更可能报告错误信息”。

相关参考文献：

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, 149–157.

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.

意见 2:

“但当测试激活一周后再学习错误信息时，RES 现象却消失了 (Chan et al., 2017)”，Chan 等人的研究发现一周后 RES 结果并不显著，但这并不意味着 RES 消失了，不显著的结果不能说明效应不存在。如果我没记错的话，Chan 等人并不认为他们的研究结果支持记忆在巩固理论。更重要的是，Chan 等人提出了其他理论来进行解释，他们自己都不支持再巩固理论。再巩固理论被大量研究所证实，建议作者在讨论再巩固理论时，可以弱化讨论，现在更多证据反驳而不是支持该理论。

回应:

(1) 非常感谢审稿专家悉心提出的建议。我们重新梳理了 Chan 等人的已发表文章并发现，Chan 等人 (2017) 认为已有研究并不能严格检验记忆再巩固理论，它只能为 RES 提供了一个补充性的但不充分的解释。且 Chan 等人 (2009) 的研究并没有直接证实其再巩固的假设。而在他们 2011 年的两个研究中，当错误信息推迟一周呈现时依然产生了 RES 现象 (Chan & Langley, 2011; Chan & LaPaglia, 2011)，这一发现反驳了再巩固理论。因此，使用 Chan 等人 (2017) 的研究来支持再巩固理论并不合适。

相关参考文献:

- Chan, J. C. K., & Langley, 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., & LaPaglia, J. A. (2011). The dark side of testing: Repeated retrieval can enhance eyewitness suggestibility. *Journal of Experimental Psychology: Applied*, 17(4), 418–432.
- Chan, J. C. K., Manley, K. D., & Lang, K. (2017). Retrieval enhanced suggestibility: A retrospective and new investigation. *Journal of Applied Research in Memory and Cognition*, 6(3), 213–229.
- 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.

(2) 根据审稿专家的意见，我们对“2.1 测试损害原始记忆”部分的结构进行了调整，将叙述重点放在“原始信息是否被错误信息完全破坏”这一问题。具体修改如下：

原文:

2.1 测试损害原始记忆

接受中期测试的被试正确报告原始信息的可能性较低，针对这一现象的一种解释是被试的原始记忆受损，即中期测试导致了错误信息对原始信息的覆盖，这被称为更新假说

(updating hypothesis)；另一种解释是受损的并非原始记忆本身，而是其可访问性，即无法成功提取原始信息，这被称为抑制假说 (inhibition hypothesis; Chan & Langley, 2011)。

支持更新假说的研究者认为，当原始记忆不稳定时，呈现错误信息会使它与原始信息相结合，从而取代了原始记忆。事实上，信息在刚刚被编码之后是非常脆弱的，极易受到干扰，只能随着时间推移慢慢整合并巩固下来，从而形成长期记忆。记忆再巩固理论 (the reconsolidation account, Scully et al., 2017; Paula et al., 2021; Carneiro et al., 2021) 认为，这种“巩固过程”并非不可逆，当记忆被重新激活，将再次回到脆弱状态，为了使这些记忆重新稳定下来，会出现一个临时窗口进行再次巩固。在再巩固期间呈现错误信息时，原始记忆处于不稳定的状态，很容易受到错误信息的干扰而发生变化。换言之，原始记忆要经历一个“巩固——激活——再巩固——新信息入侵”的过程。而中期测试正是一种强烈的激活手段，会导致新信息大量入侵原始记忆。尤其当新信息与原始信息相矛盾时，这种强烈的再激活可能会引起过度矫正，从而导致原始信息的遗忘和新信息的更新 (Scully et al., 2017)。相关证据表明，在呈现原始事件一周后，通过中期测试激活原始信息并立即呈现错误信息时，出现了 RES 现象；但当测试激活一周后再学习错误信息时，RES 现象却消失了 (Chan et al., 2017)。

这说明，中期测试只有在记忆“脆弱”时才能增强错误信息的暗示性，证实了记忆再巩固理论。这在一定程度上解释了 RES 效应的结果，但仍然受到了一些质疑。例如，有研究发现，当错误信息在再巩固阶段结束以后呈现时，仍然出现了 RES 现象（Chan & Langley, 2011; Chan & LaPaglia, 2011）。

抑制假说也得到了不少实证研究的支持。如允许被试在最终测试时提供不止一个答案（the modified modified free recall test, MMFR 测试）时，与对照组相比，中期测试组正确报告原始信息的可能性没有差异甚至更高（Chan et al., 2009; Gordon & Thomas, 2014）。McCloskey 和 Zaragoza（1985）发现，当最终测试要求被试在正确答案和其他选项（但不是错误信息）间进行强迫选择时，中期测试对正确率没有影响。基于这一发现，他们认为，对原始细节的记忆不会因错误信息而受损。也就是说，中期测试只是阻碍了被试对原始信息的访问，导致其在最终测试时无法成功提取。在此基础上，如果增强原始信息的可访问性，就可以提高最终测试的正确率。例如，Gordon 和 Shapiro（2012）参照语义网络模型，通过激活相关概念来启动原始记忆中的关键细节，从而降低了错误信息的影响，提高了最终测试的正确率。尽管他们的研究并没有对原始信息进行中期测试，但也说明原始信息和错误信息在记忆中是共存的，且当增加了原始信息的可访问性之后，错误信息的暗示性就降低了。为了直接证明原始信息可访问性被阻碍，未来研究可以尝试比较再认和回忆之间的差异。Scully 等人（2017）认为，如果能再认但不能回忆，便说明可访问性受损。

修改后：

2.1 测试阻碍原始信息的访问

接受中期测试的被试在最终测试中报告原始信息的可能性较低，可能是因为原始信息的可访问性受到了阻碍。事实上，信息在刚刚被编码之后是非常脆弱的，极易受到干扰，只能随着时间推移慢慢整合并巩固下来，从而形成长期记忆。记忆再巩固理论（the reconsolidation account, Scully et al., 2017; Paula et al., 2021; Carneiro et al., 2021）认为，这种“巩固过程”并非不可逆，当记忆被重新激活，将再次回到脆弱状态，为了使这些记忆重新稳定下来，会出现一个临时窗口进行再次巩固。在再巩固期间呈现错误信息时，原始记忆处于不稳定的状态，很容易受到错误信息的干扰而发生变化。换言之，原始记忆要经历一个“巩固——激活——再巩固——新信息入侵”的过程。而中期测试正是一种强烈的激活手段，会导致新信息大量入侵原始记忆。尤其当新信息与原始信息相矛盾时，这种强烈的再激活可能会引起过度矫正，从而导致原始信息的遗忘和新信息的更新（Scully et al., 2017）。

中期测试激活原始信息之后，错误信息的呈现存在两种可能的影响：一是中期测试导致了错误信息对原始信息的覆盖，使被试的原始记忆完全丢失，这被称为更新假说（updating hypothesis）；二是错误信息并未破坏原始信息，只是损害了其可访问性，导致在最终测试中无法成功提取原始信息，这被称为抑制假说（inhibition hypothesis; Chan & Langley, 2011）。

目前，抑制假说得到了广泛的支持。研究发现，当允许被试在最终测试时提供不止一个答案（the modified modified free recall test, MMFR 测试）时，与对照组相比，中期测试组正确报告原始信息的可能性没有差异甚至更高（Chan et al., 2009; Gordon & Thomas, 2014）。McCloskey 和 Zaragoza（1985）发现，当最终测试要求被试在正确答案和其他选项（但不是错误信息）间进行强迫选择时，中期测试对正确率没有影响。基于这一发现，他们认为，对原始细节的记忆不会因错误信息而受损。也就是说，中期测试只是阻碍了被试对原始信息的访问，导致其在最终测试时无法成功提取。在此基础上，如果增强原始信息的可访问性，就可以提高最终测试的正确率。例如，Gordon 和 Shapiro（2012）参照语义网络模型，通过激活相关概念来启动原始记忆中的关键细节，从而降低了错误信息的影响，提高了最终测试的正确率。尽管他们的研究并没有对原始信息进行中期测试，但也说明原始信息和错误信息在记忆中是共存的，且当增加了原始信息的可访问性之后，暗示感受性就降低了。为了直接证

明原始信息可访问性被阻碍，未来研究可以尝试比较再认和回忆之间的差异。Scully 等人（2017）认为，如果能再认但不能回忆，便说明可访问性受损。

意见 3:

“Pajkossy 等人（2019）的研究为这一解释提供了生理学证据。他们发现，与重新学习相比，中期测试减少了被试在最终测试中的瞳孔扩张。瞳孔大小反映了执行任务时的资源处理量，这说明中期测试减少了最终回忆时的处理负荷”。Pajkossy 等人（2019）的研究结果如果支持了“干扰与分离”理论？这一点却是令人费解，为何更多的认知努力就是支持“干扰与分离”的支持证据？比较牵强。

回应:

非常感谢审稿专家的意见。Szpunar 等人（2008）认为，测试增强新学习的原因是通过促进上下文分离，从而减少了最终测试时的信息过载。基于此，我们试图引用 Pajkossy 等人（2019）的研究为此解释提供支持，即中期测试确实减少了最终测试时的处理负荷。但经过专家的提醒，我们发现该研究与“干扰与分离”这一机制并无直接联系，无法提供客观、有力的证据。因此，我们删除了这一研究，并对“2.2.1 干扰与分离”部分进行以下修改：

原文:

2.2.1 干扰与分离

Chan 等人（2009）认为，中期测试之所以能够增强对事后错误信息的学习，其中一个原因是测试区分了两次不同的编码事件（原始信息和叙述信息），从而减少了原始记忆对新信息学习的影响。测试促进上下文分离并减少前摄干扰的现象已被广泛发现（Szpunar et al., 2008; Wahlheim, 2015; Bufe & Aslan, 2018; Dang et al., 2021）。如在 Szpunar 等人（2008）的经典研究中，当学习 5 个列表的单词时，如果分别对前四个列表进行测试，可以提高对第 5 个列表回忆的正确率，且先前列表中单词的入侵也更少。

测试可以增强新材料的学习，可能是由于测试导致上下文产生变化，被试在最终测试时只需检索最新编码的材料，减少了信息过载及材料之间的混淆和干扰，从而提高了新材料的回忆成绩（Szpunar et al., 2008）。Pajkossy 等人（2019）的研究为这一解释提供了生理学证据。他们发现，与重新学习相比，中期测试减少了被试在最终测试中的瞳孔扩张。瞳孔大小反映了执行任务时的资源处理量，这说明中期测试减少了最终回忆时的处理负荷。

就错误信息领域而言，中期测试可能分离了原始信息与叙述信息，减少了前者对后者的干扰，提高了最终测试中错误信息的回忆率。但需要注意的是，错误信息领域的新旧材料在关系上具有更加明显的相关性和冲突性，且更多考察情景记忆。那么，中期测试是否仍能促进上下文分离并减少干扰还需进一步的验证。

修改后:

2.2.1 干扰与分离

Chan 等人（2009）认为，中期测试之所以能够增强对事后错误信息的学习，其中一个原因是测试区分了两次不同的编码事件（原始信息和叙述信息），从而减少了原始记忆对新信息学习的影响。测试促进上下文分离并减少前摄干扰的现象已被广泛发现（Szpunar et al., 2008; Wahlheim, 2015; Bufe & Aslan, 2018; Dang et al., 2021; Yang et al., 2022）。

在 Szpunar 等人（2008）的经典研究中，随着先前学习的材料增多，后续学习中受到先前材料的干扰程度也随之增大；而对先前材料进行测试可以提高后续材料的回忆正确率，且产生更少的前摄干扰。最新研究进一步发现，先前材料的干扰在中期测试和后续学习的记忆表现之间发挥中介作用（Dang et al., 2021; Yang et al., 2022）。Szpunar 等人（2008）采用来源监控和信息过载进行解释：如果先前的列表没有进行测试，那么在对最后的材料进行测试时，被试需要对之前出现的所有信息进行提取，信息负荷过重，可能会导致材料之间的混淆。

加入测试有利于被试将列表彼此分开并加以区分，一方面可以更好地进行来源监测，另一方面也减少了信息过载和先前材料的干扰，有利于对新信息的回忆。

就错误信息领域而言，中期测试可能分离了原始信息与叙述信息，减少了前者对后者的干扰，提高了最终测试中错误信息的回忆率。但需要注意的是，错误信息领域的新旧材料在关系上具有更加明显的相关性和冲突性，且更多考察情景记忆。那么，中期测试是否仍能促进上下文分离并减少干扰还需进一步的验证。

相关参考文献：

- 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.
- Szpunar, K. K., McDermott, K. B., & Roediger, H. L. (2008). Testing during study insulates against the buildup of proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1392–1399.
- Wahlheim, C. N. (2015). Testing can counteract proactive interference by integrating competing information. *Memory & Cognition*, 43(1), 27–38.
- Bufe, J., & Aslan, A. (2018). Desirable difficulties in spatial learning: testing enhances subsequent learning of spatial information. *Frontiers In Psychology*, 9,1701.
- 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.
- 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.

意见 4:

关于“编码策略理论”，作者引用了部分其他研究证据，然而最直接的支持证据却并没有被引用。例如 Chan 等发表在 JML 上的研究发现中部测试提高后续的语义聚类编码，Yang 等（2021）发表在 JEPLMC 上的研究发现中部测试提高后续的时间聚类编码。

回应：

非常感谢审稿专家的意见和论文推荐。在认真阅读相关文献以后，我们对“编码方式”第三段内容进行了以下修改：

原文：

Cho 等人（2017）提出了编码策略理论，即测试增强新学习源于编码策略的改变。通过中期测试，被试认识到自己的记忆存在不足，因而会在后续学习中更加努力，并寻求更好的记忆方法（Soderstrom & Bjork, 2014; Cho et al., 2017）。换言之，中期测试能够促使被试使用更有效的编码策略，从而提高记忆成绩（Dang et al., 2021）。

修改后：

Cho 等人（2017）提出了编码策略理论，即测试增强新学习源于编码策略的改变。通过中期测试，被试认识到自己的记忆存在不足，因而会在后续学习中更加努力，并寻求更好的记忆方法（Pyc & Rawson, 2012; Soderstrom & Bjork, 2014; Cho et al., 2017）。研究发现，中期测试能够促使被试使用更有效的编码策略（Soderstrom & Bjork, 2014）。例如，与无测试或重新学习相比，中期测试增强了后续学习中的语义聚类（Chan et al., 2018; Dang et al., 2021）。Yang 等人（2022）也发现，中期测试通过提高后续学习的时间聚类水平，增强了对新材料的学习。

相关参考文献：

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

审稿人 3 意见:

该研究力图梳理 RES 和 PET 的理论解释及其内在机制，针对性地提出两个理论模型，同时分析了 RES 和 PET 分离的潜在影响因素，并从变量引入和干预手段两方面说明了该领域的未来研究方向。作者对两位审稿人的意见进行了修改，综述问题有一定的研究意义，作者也提出了自己的观点，但文章依然存在一些问题供作者参考：

意见 1:

测试究竟是好是坏？这是文章标题中重点关注点问题，但是通读下来，并没有找到明确的答案，这意味着题目不太恰当？重点在于关注测试的好坏吗？

回应:

非常感谢审稿专家第二轮恳切的意见。我们根据专家的建议进行了修改，本次修改内容在原文中用绿色字体标出。

经过专家的提醒，我们发现原题目确实存在不妥之处。一方面，中期测试对错误信息效应的影响并不稳定，即无法断言“好”或“坏”；另一方面，本文并非是对测试“好坏”的简单判断，而是关注中期测试产生影响的内在机制，即“为什么会出现 RES？”、“为什么会出现 PET？”以及“两者的边界条件是什么？”。因此，经过慎重考虑，我们将题目修改为“**测试的两面性：中期测试对错误信息效应的影响及其作用机制**”。

意见 2:

测试增强暗示性作为标题是否合适？增强什么的暗示性？何为暗示性？暗示了什么？从字面上看 suggestibility 翻译为暗示，有没有更好的解释？后文中提到敏感性等，术语应该一致。

回应:

非常感谢审稿专家的意见。首先，我们认真查阅了相关文献，发现国内已有研究对 suggestibility 的翻译并不一致，包括“暗示感受性”（王红椿 等, 2008; 龚梦园 等, 2009; 顾丽佳, 郭建友, 2014; 曹晓君 等, 2017）、“易受暗示性”（常河山, 刘新学, 2006; 曾守锤, 李其维, 2008）、“受暗示性”（郭秀艳, 李荆广, 2007; 李婷玉 等, 2018）等。在错误信息相关领域中，研究者多翻译为“暗示感受性”（王红椿 等, 2008; 曹晓君 等, 2015; 曹晓君 等, 2017），因此我们最终采用了“**暗示感受性**”这一翻译。

其次,暗示感受性是指个体对错误信息的易感性,即个体被错误信息误导的可能性(王红椿等,2008;曹晓君等,2015)。RES是指参与了中期测试的被试比没有经过测试的被试更可能受到错误信息的影响。换句话说,中期测试增强了个体的暗示感受性。“敏感性”的含义类似于易感性,因此文中“对错误信息的敏感性”指的就是暗示感受性。为了避免概念的混淆,我们将其统一表述为“暗示感受性”。

最后,我们在文中加入了暗示感受性的定义,具体请见“1 引言”第三段。

相关参考文献:

- 曹晓君,冉光明,陈旭.(2017).记忆暗示感受性的双加工机制研究.《心理学探新》,37(05),409-415.
- 曹晓君,涂桑,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(06),980-984+1008.
- 常河山,刘新学.(2006).压力与画对儿童错误记忆影响的实验研究.《心理科学》,29(03),624-627.
- 李婷玉,刘黎,李宜霖,朱莉琪.(2018).冲突情境下幼儿的选择性信任和信念修正.《心理学报》,50(12),1390-1399.
- 龚梦园,徐富明,方芳.(2009).事后聪明式偏差的理论模型及影响因素.《心理科学进展》,17(02),325-333.
- 顾丽佳,郭建友.(2014).安慰剂镇痛效应的个体差异及其影响因素.《中国临床心理学杂志》,22(06),994-998+993.
- 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(04),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(03),744-747.
- 曾守锤,李其维.(2008).学前儿童的现实性监测对易受暗示性的预测.《心理科学》,31(06),1313-1316.

意见 3:

the retrieval fluency hypothesis 翻译为检索流畅性,更常见的是翻译为“提取流畅性”。在后续的另外一个翻译中, retrieval 作者也把它翻译为“提取”。可以再斟酌。

回应:

非常感谢审稿专家指出该问题。我们查阅了相关文献发现,国内多个研究将其翻译为“提取流畅性”(贾宁,白学军,2007;刘宁等,2007;张锦坤等,2008;侯瑞鹤,俞国良,2008;王怀勇,刘永芳,2014;刘超等,2019),因此统一修改为“提取流畅性”。相对应地,鉴于心理学领域的相关研究一般将“retrieval”翻译为“提取”(聂爱情等,2018;陈伟等,2020;张环等,2020;周晨琛等,2020;李俊娇等,2021;王松雪等,2023),我们把文中的“检索”也统一修改为“提取”。

相关参考文献:

- 陈伟,李俊娇,林小裔,张晓霞,郑希付.(2020).行为干预情绪记忆再巩固:从实验室到临床转化.《心理科学进展》,28(02),240-251.
- 侯瑞鹤,俞国良.(2008).加工流畅性和提取流畅性与学习不良儿童学习判断的关系.《心理学报》,40(09),994-1001.
- 贾宁,白学军.(2007).学习判断的练习伴随低估效应.《心理科学进展》,15(03),429-435.
- 李俊娇,陈伟,胡琰健,曹杨婧文,郑希付.(2021).预期错误与急性应激对不同强度恐惧记忆提取消退的影响.《心理学报》,53(06),587-602.
- 刘超,陈功香,李开云.(2019).延迟学习判断的时间加工进程.《心理学探新》,39(03),232-237.
- 刘宁,张峰,张庆林,梁娜.(2007).重复学习判断中的低自信效应.《心理科学进展》,15(05),788-794.
- 聂爱情,李梦思,潘如.(2018).编码如何预测提取?以相继记忆效应为证.《心理科学进展》,26(10),1775-1786.
- 王怀勇,刘永芳.(2014).决策过程中的调节匹配效应及其机制.《心理科学》,37(01),182-189.
- 王松雪,程思,蒋挺,刘勋,张明霞.(2023).外在奖赏对陈述性记忆的影响.《心理科学进展》,31(01),78-86.

张环, 侯双, 王海曼, 廉宇煊, 杨海波. (2020). 他人在场条件下的社会分享型提取诱发遗忘. *心理学报*, 52(06), 716 - 729.

张锦坤, 白学军, 杨丽娟. (2008). 国外关于测试效应的研究概述. *心理科学进展*, 16(04), 661 - 670.

周晨琛, 姬鸣, 周圆, 徐泉, 游旭群. (2020). 不同注意状态下前瞻记忆意图后效的抑制效应. *心理科学*, 43(04), 777 - 784.

意见 4:

展望与前面的研究内容结合不够紧密, 尤其是第二部分, 关注干预, 这部分内容很多来自错误信息持续影响效应的研究里面的内容, 前文并未涉及, 应该关注的是中期测试。

回应:

非常感谢审稿专家指出这一问题。首先, 我们删除了错误信息持续影响效应研究中的有关内容。其次, 为了加强展望部分与上文的联系, 提升论文的探究深度与理论价值, 我们分别从个体、社会 and 干预手段的角度出发, 以认知需求、记忆从众、记忆训练、警告等影响因素为例, 通过机制探讨预测它们与中期测试的关系, 以期为未来研究提供思路。基于此, 我们对“5 总结与展望”部分的第二至四段进行了修改, 具体如下:

原文:

第一, 当前研究范围较狭窄, 需进一步考察个体差异和社会因素的影响。一方面, 个体间的差异会影响他们对错误信息的处理方式。例如, 认知方面, 高认知需求的个体更可能进行差异检测并拒绝错误信息 (Leding & Antonio, 2019), 工作记忆能力的差异也会影响错误信息的作用 (任芬 等, 2020); 情感方面, 强烈的情绪会使个体更容易相信错误信息 (Martel et al., 2020), 愤怒、恐惧和焦虑等情绪可以改变个体对错误信息的态度 (Weeks, 2015; Angeline et al., 2019)。又如, 个体对某一事件所持有的先前态度或信念往往会左右其对错误信息的接受程度 (Shen & Zhou, 2021)。研究表明, 个体更倾向于阅读与自己态度一致的错误信息, 且纠正这种错误甚至会适得其反 (Scheufele & Krause, 2019)。由此, 这些个体差异与中期测试的相互作用应该纳入未来研究的考虑范围。

另一方面, 现实生活中的错误信息往往带有“社会”色彩。首先, 信息的意义需要被重视。从已有研究不难看出, 错误信息效应的研究大多采用中性的实验材料, 考察记忆时也仅涉及“是”和“否”之类非此即彼的问题。这种将信息简单化的做法, 忽视了个体对意义的寻求。如有研究发现, 与不文明的信息相比, 人们更容易相信文明的信息 (Kim & Chen, 2020)。其次, 信息的来源需要被考虑。例如, 个体更容易相信来自可信来源的信息——与电视节目相比, 来自家人朋友的信息更有利于纠正错误信息 (Pluviano et al., 2020)。最后, 人际互动的影响需要被关注。已有研究发现, 被试在接受错误信息时与他人进行现实互动, 会受到来自他人的错误信息的强烈影响 (Numbers et al., 2014)。同时, 群体中认可错误信息的人越多, 个体越倾向于接受该信息 (Mojtahedi et al., 2018)。

第二, 为扩大研究的现实意义, 需开发有效的干预手段。根据已有研究, 我们预测通过增强被试的检索努力程度、鼓励被试进行差异检测、引导被试回忆自己最初的反应可以发挥测试的保护作用, 减少错误信息的暗示性。但这些都只是短期的实验操作, 若想要改善个体抵抗错误信息的能力, 还需开发出长期有效的干预方式。例如, 通过积极反馈增强个体对自己记忆能力的信息便可以减少错误信息的影响 (Szpitalak & Polczyk, 2019)。在信息传播领域, 当面对不确定信息时, 引导被试主动思考而非依赖直觉可以减轻反复看到假新闻的负面影响 (London & Raj, 2019); 通过让被试模拟编造假新闻的方式可以提高其发现并拒绝错误信息的能力 (Roozenbeek & Linden, 2019)。

修改后:

第一, 当前研究范围较狭窄, 需进一步考察个体差异和社会因素的影响。一方面, 个体

间的差异会影响他们对错误信息的处理方式。例如，中期测试的效果可能会受到个体认知需求的影响。已有研究发现，高认知需求的个体更不可能受到错误信息的影响（Leding & Antonio, 2019）。认知需求是指个体愿意投入精力进行深度认知加工的一种特质（Cacioppo & Petty, 1982）。高认知需求者倾向于进行更精细的思考，与低认知需求者相比，他们在记忆测试中回忆了更多信息（徐洁, 周宁, 2010）。这说明高认知需求的个体可能会在中期测试阶段花费更多努力提取信息，从而增强对原始信息的记忆。同时，Leding 和 Antonio (2019) 认为，高认知需求者是因为花费更大努力进行了差异检测，所以更能拒绝错误信息。因此，我们预测认知需求会通过提取努力、差异检测影响中期测试对错误信息效应的作用。未来研究可以进一步探索认知、情感、态度等个体因素对 RES 或 PET 现象的影响。

另一方面，现实生活中的错误信息往往带有“社会”色彩。例如，当错误信息以社会方式引入时，个体会受到来自共同目击者的错误信息的影响，这被称作“记忆从众（memory conformity）”，被认为是错误信息效应的一种形式（Gabbert et al., 2004; Kękuś et al., 2021）。并且，社会来源的错误信息比非社会来源的错误信息会产生更大的危害（Gabbert et al., 2004）。研究发现，意识到原始信息和来自他人的错误信息之间的差异可以降低记忆从众现象，但尽管被试意识到差异，仍然可能报告错误信息，这是因为他们对自己的记忆不自信（Kękuś et al., 2021）。而中期测试作为一种可以增强原始记忆并促进差异检测的操作手段，或许可以减少记忆从众效应。此外，未来的研究可从信息的社会性质、来源及个体的社会身份与所处的社会环境等因素入手，由此拓宽中期测试的研究范围，从而更好地发挥测试的保护作用，取得应有的现实价值与理论深度。

第二，为扩大研究的现实意义，需开发有效的干预手段。根据已有研究，我们预测通过增强被试的提取努力程度、鼓励被试进行差异检测、引导被试回忆自己最初的反应可以发挥测试的保护作用，降低暗示感受性。但这些都只是短期的实验操作，若想要改善个体抵抗错误信息的能力，还需开发出长期有效的干预方式。例如，关于人类记忆不可靠的训练可以降低被试的暗示感受性（Szpitalak et al., 2021）。研究者认为，一种解释是这种记忆训练促进了差异检测。遗憾的是，在他们的研究中，中期测试并没有影响错误信息效应，因此记忆训练是否可以通过促进差异检测影响 RES 或 PET，还需进一步的研究证实。

相关参考文献：

- 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(04), 685–690.
- Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42(1), 116–131.
- 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.
- 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.
- Leding, J. K., & Antonio, L. (2019). Need for cognition and discrepancy detection in the misinformation effect. *Journal of Cognitive Psychology*, 31(4), 409–415.
- Szpitalak, M., Woltmann, A., Polczyk, R. & Kekus, M. (2021). Memory Training as a Method for Reducing the Misinformation Effect. *Current Psychology*, 40, 5410–5419.

意见 5:

作者提出了模型图，看似有创新，但是只是提供了解释路径，其意义需要进一步挖掘，比如，有何理论上的贡献？对未来研究有何启发？

回应：

非常感谢审稿专家的恳切意见。根据专家建议，我们补充了该模型的理论价值和未来研究展望的内容。

(1) 为了更详细地阐明该模型的理论贡献，我们修改了“3.6 综合 RES 和 PET 的总模型”部分的第四段内容，具体修改如下：

原文：

这一模型整合了 RES 和 PET 的相关理论研究，对中期测试影响错误信息暗示性的机制进行了更详细的阐述。但该模型是在现有相关理论上初步搭建的，各理论之间存在何种联系、路径之间如何相互作用，以及是否存在其它关键变量和影响因素，仍需更多的理论和实证研究加以回应和补充。

修改后：

这一模型整合了 RES 和 PET 的相关理论研究，对中期测试影响错误信息效应的机制进行了更详细的阐述，具有一定的理论价值：第一，该模型将 RES 和 PET 的相关研究置于同一研究框架中，有利于聚焦研究方向、凸显研究重点。过往研究将 RES 与 PET 视为矛盾现象，在解释结果时往往有所偏重、影响因素考虑不够全面，难以有效积累研究成果；统一研究框架可以为研究者解释现象提供线索，帮助错误信息领域收束研究方向，从而更好地挖掘研究背后的理论价值。第二，该模型联结了学习领域和错误信息领域的相关理论，为领域间交融互通提供了桥梁。学习领域理论成熟，错误信息领域成果多样，因此这种联系的建立有两方面价值：一方面，它弥补了错误信息领域的理论空缺，为相关研究提供了坚实的理论基础；另一方面，它拓展了测试效应相关理论的应用范围，丰富了其理论内涵。

(2) 我们将“5 总结与展望”分为“5.1 模型验证”和“5.2 研究拓展”两个部分，并在“5.1 模型验证”中提出了针对模型的未来研究思路，具体内容如下：

5.1 模型验证

本文所提出的模型整合了相对独立的研究成果与不同领域的理论基础，能够为现有研究所发现的现象提供较全面的解释。但该模型仍处于搭建初期，各理论之间存在何种联系、路径之间如何相互作用，以及是否存在其他关键变量和影响因素，仍需更多针对性的理论和实证研究加以回应和补充。

第一，未来研究需要对不同路径进行直接验证。就原始信息和错误信息这两条路径而言，从“原始信息记忆强度”到“原始信息可访问性”与“测试增强新学习”到“错误信息提取流畅性”的关系本质上就是“编码——提取”的关系。未来研究可以通过操纵加工深度（曹晓君 等, 2015）或加入分心任务（Gordon & Thomas, 2017; Gordon et al., 2020）以改变信息的学习程度，并通过最终测试时的反应时间和信心水平测量提取流畅性和可访问性（Thomas et al., 2010），考察错误信息效应的变化。在此基础上，未来研究可以采用中介分析的方法，以检验中期测试是否可以通过影响原始信息和错误信息的编码程度，改变原始信息的可访问性和提取流畅性，最终影响错误信息效应。

就第三条路径来说，上下文分离与差异检测的关系也值得进一步验证。已有研究发现，只有当被试发现信息之间的差异时，先前材料的干扰才会减弱，否则干扰依然存在（Wahlheim, 2015）。但该研究中，差异检测是对整体材料而非独立项目的检测，因此这是否意味着干扰与分离的机制必须通过差异检测发挥作用还需要更多的证据支持。未来研究可通过在最终测试中考察测试错误信息的记忆，以被试报告原始信息的数量测量前摄干扰，并在最终测试后加入差异检测测试（询问被试在每条信息中是否发现了差异; Polczy, 2017），以验证该路径。

第二，未来研究需要结合不同路径以全面探讨中期测试的作用机制。一方面，不同路径间并非相互独立，可能存在交互影响。例如，研究者认为，强大的原始记忆会促进差异检测（Peterson et al., 2004; Loftus, 2005; Gabbert et al., 2012），但他们的研究未能为这一假设提

供直接证据。例如，Gabbert 等人（2012）的研究发现，被试的正确率越高（即记忆越好），错误信息的报告率越低，两者呈负相关。但仅以这一点作为证据显然是不够直接和充分的。此外，错误信息的编码程度是否同样会促进差异检测也需进一步探究。

另一方面，不同路径同时发挥作用，但相对贡献大小仍未可知。Yang 等人（2022）采用中介分析同时考察了编码策略和前摄干扰对前向测试效应的独立作用和共同作用，发现前摄干扰比策略改变发挥了更重要的作用。因此，未来研究可以借鉴 Yang 等人（2022）的方法，以量化分析不同条件下中期测试在三条路径上对错误信息效应的影响性质及大小，进而在开展后续研究或开发干预手段时有所侧重。

相关参考文献：

- 曹晓君, 涂桑, 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(06), 980 - 984+1008.
- Gabbert, F., Hope, L., Fisher, R., & Jamieson, K. (2012). Protecting against misleading post - event information with a Self - Administered Interview. *Applied Cognitive Psychology*, 26(4), 568-575.
- 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., Bilollikar, 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.
- 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.
- 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.
- 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.
- 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, 149-157.
- Wahlheim, C. N. (2015). Testing can counteract proactive interference by integrating competing information. *Memory & Cognition*, 43(1), 27-38.
- 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.
-

第三轮

审稿人 3 意见：

作者对论文的审稿意见做了非常细致的修改，但依然存在一些问题，供作者再斟酌：

意见 1：

论文第一段并未界定什么是错误信息。作为一个很关键的变量，应该给出界定。另外，什么是错误信息效应？在错误信息领域各种效应非常多，比如错误信息持续影响效应等，最好能够界定下错误信息效应。

回应：

非常感谢审稿专家第三轮的恳切建议，我们根据建议做了逐一修改，并对一些文字表述进行了修改和完善，本轮修改的内容在原文中用紫色字体标出。

经过专家的提醒，我们在论文第一段中加入了错误信息和错误信息效应的定义：“错误信息（misinformation）是指客观上不准确的信息（Scheufele & Krause, 2019）。研究发现，事件发生后呈现的错误信息会改变个体对原始事件的记忆，这被称为错误信息效应（misinformation effect, Loftus, 2005; 王析蕾, 贾宁, 2021）。”

相关参考文献：

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

意见 2：

论文第一段：当谈及“错误信息（misinformation）”时，人们常常会想到曲解既成事实或编造虚假事件以欺骗他人的谣言，却忽视了环境偶然干扰或他人无意分享的影响。这句话放这儿是什么意思呢？目的何在？

回应：

非常感谢审稿专家指出这一问题。这句话的本意是为了引出下文关于错误信息以及错误信息效应的概念。经过专家提醒，我们认为这句话和文章的主题并无直接联系，放在此处并不合适。因此，在原文中删去了这一表述。

意见 3：

（retrieval enhanced suggestibility, RES）翻译为“测试增强暗示感受性”是否合适？请再斟酌。

回应：

非常感谢审稿专家的意见。我们认真查阅了相关文献发现，国内已有研究者将 RES（retrieval enhanced suggestibility）翻译为“提取加强的暗示感受性”（曹晓君 等, 2015）。其它一些硕士论文将 RES 翻译为“提取增加暗示性”、“提取增加受暗示性效应”、“检索提高易感性效应”（韩珊, 2015; 钱慧昌, 2022; 王析蕾, 2021）。

我们参考了曹晓君等人（2015）的翻译，但考虑到“提取加强的暗示感受性”落脚于“暗示感受性”，似乎暗指 RES 本身是一种暗示感受性。而我们认为 RES 强调的是“加强暗示感受性”的这一过程或现象，它本身并不是暗示感受性。经过慎重考虑，我们最终决定去掉“的”，将 RES（retrieval enhanced suggestibility）翻译为“提取加强暗示感受性”，并在原文中进行了相应修改。

相关参考文献：

- 曹晓君, 涂桑, 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(06), 980 - 984+1008.
- 韩珊. (2015). 目击者记忆: 检索提高易感性效应的内部机制研究(硕士学位论文, 河北师范大学).
- 钱慧昌. (2022). 不同时间点的警告在记忆重复提取过程中对目击者记忆的保护作用的研究(硕士学位论文, 安徽医科大学).
- 王析蕾. (2021). 原始信息学习程度与叙述信息可信度对提取增强暗示性效应的影响(硕士学位论文, 河北师范大学).

意见 4：

关于中期测试存在诸多理论，作者也尝试构建了作用路径，但是对于不同理论之间的异

同是否可以做一个分析和整合？以便于读者理解诸多理论。

回应：

非常感谢审稿专家悉心提出的建议。经过认真思考和讨论，我们分别采用“编码”和“提取”两个视角对 RES 和 PET 的相关理论进行了比较和整合，并进行了以下修改：

(1) 就“2.1 测试阻碍原始记忆的访问”和“2.2 测试增强后续学习”两部分而言，我们在阐述所有 RES 相关理论之后，进一步从“编码”和“提取”两个角度对这些理论进行了比较和分析，具体修改内容如下：

总的来说，已有关于 RES 现象的理论解释大多可从“编码”和“提取”两个角度加以归类。“编码”解释认为，中期测试改善了被试对后续材料的学习，包括动机理论、注意捕捉假说、编码重置理论和编码策略理论。具体而言，检索失败、测试期望等为增强后续编码提供了动机因素，使被试在叙述信息阶段“想要”花费更多努力；编码重置、注意捕捉为增强后续学习提供了重要的前提条件，使被试在叙述信息阶段“能够”对错误信息进行精细编码；改变编码策略为增强后续学习提供了手段上的便利，使被试在叙述信息阶段“易于”进行高效编码。

“提取”解释则认为，中期测试的提取过程影响了 RES 的产生。首先，根据记忆再巩固理论，提取信息破坏了原始记忆的稳定状态，使记忆更易受错误信息影响。其次，提取本身也促进了上下文分离，并减少了先前材料的干扰。最后，提取流畅性假说认为，RES 的产生是由于错误信息在最终测试中更易被提取，因此被试在面对错误选项时反应更快且更有信心。

(2) 就“3.1 记忆强度”至“3.4 迁移适当加工理论”部分的内容而言，我们在阐述所有 PET 相关理论之后，进一步从“编码”和“提取”两个角度对这些理论进行了比较和分析，具体修改内容如下：

相应的，对 PET 的理论解释依然可从“编码”和“提取”两个角度加以概括。从“编码”的角度来说，记忆强度理论认为“强大”的原始记忆可以降低暗示感受性。在进行中期测试时，被试有机会对原始信息进行额外编码，这增强了被试的原始记忆(Yang et al., 2021)，促进了 PET 的产生。此外，中期测试通过促进差异检测降低了被试的暗示感受性，并表现为花费更多时间阅读错误信息。这说明差异检测增强了被试对错误信息的某种批判性的编码。

从“提取”的角度来看，提取努力理论侧重于提取的直接好处——考察提取过程的努力程度，且认为花费的认知努力越多，记忆强度提升越多。而迁移适当加工理论则将“提取”看作一种特殊的技能，使被试学会“如何测试”，从而在后续测试中表现更好。因此，迁移适当加工理论更注重中期测试与最终测试的格式是否一致。

相关参考文献：

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.

(3) 最后，为了方便读者阅读和理解，我们从中期测试的作用阶段和解释角度两方面对 RES 和 PET 的相关理论进行了系统梳理，并在“3.6 统合 RES 和 PET 的总模型”部分以表格形式加以呈现。具体修改如下：

上述分析结果表明，有关 RES 和 PET 现象的相关理论不仅内容丰富，且从不同角度为中期测试对错误信息效应的影响及其作用机制提供了多样化解读。总体来看，这些理论之间的关系可从两个方面加以概括：一是中期测试效用出现在错误信息范式的阶段有所不同；二是从“编码”或“提取”不同角度阐释了中期测试对错误信息的作用机制。为便于理解，我们从“作用阶段”和“解释角度”两个维度对相关理论进行了比较与整合，具体见表 1。

表 1 RES 与 PET 的理论

现象	理论	作用阶段	解释角度
RES	记忆再巩固理论	原始信息	提取
	干扰与分离	中期测试	提取
	注意捕捉假说	叙述信息	编码
	编码重置理论		编码
	编码策略理论		编码
	动机理论		编码
	提取流畅性假说	最终测试	提取
PET	记忆强度	原始信息	编码
	提取努力理论	中期测试	提取
	差异检测理论	叙述信息	编码
	迁移适当加工理论	最终测试	提取

第四轮

编委 1 意见:

同意发表。

编委 2 意见:

针对中期测试对错误信息的影响，本论文系统梳理了与之相关的研究，并针对性的提出了新理论模型，还为未来研究指明了方向。该论文逻辑框架清晰，文字表述流畅。

主编意见:

同意发表。