

《心理学报》审稿意见与作者回应

题目：部分线索诱发遗忘效应的持久性：项目价值的作用

作者：刘湍丽 张雅静 周菘 邢敏 白学军

第一轮

审稿人 1 意见：

该研究将部分线索范式与价值导向记忆任务相结合，并进一步操纵高、低关联编码条件和测验时程，考察以不同价值的项目作为部分线索对不同价值测验项目再认的影响及其持久性。结果具体表现为，在低关联编码条件下，高价值线索使得测验项目再认受损持续到了最终测验，低价值线索仅使高价值测验项目在即时测验中受损。在高关联编码条件下，仅高价值线索使得低价值测验项目的再认受损从即时测验持续到了最终测验。上述结果表明，项目价值影响部分线索效应，项目价值的作用及其持久性受编码条件的调节，影响着部分线索诱发遗忘效应的出现时间和持续时间。该研究样本量充足，层次分明，具有一定的理论意义，但在方法和理论阐述上存在一些不足。建议修改后发表。

意见 1：引言：对部分概念的定义阐述不足，理论假设阐述不够详尽。P4. 在阐述实验 1 的创新性时，对于项目的价值性这一变量的引入在部分线索效应的理论上有何贡献缺乏描述。创新不仅仅是前人没有研究过，还应具有一定的理论意义和价值。P5. 实验 2 的逻辑必要性阐述不足。在实验 1 的基础上操纵高、低编码关联，并考察持久性的问题，对双机制假设有何补充，需要具体来谈。P5. “优先输出顺序”具体指的是什么？需要结合文献来谈。

回应：非常感谢专家的意见。使我们对上述问题进行重新思考。

实验 1 中的项目价值这一变量对于部分线索效应理论贡献主要体现在：

在记忆活动中，人们会选择性地关注重要的信息。但以往关于部分线索效应的研究中，所有学习材料对于被试来说均是同等重要的，并且部分线索的选择也通常是随机的（通常是学习词表中奇数项/偶数项，以确保在原始集合中均匀分布），因此所有部分线索项目对于被试来说也具有等价性。迄今为止，尚未有研究探索当操纵编码阶段学习项目的重要性，并以不同价值的项目作为部分线索，会如何影响对其他不同价值项目的提取。而价值导向记忆的相关研究发现，对某些项目赋予高分值可以使它们在长时记忆中被优先保持和提取（Hitch et al., 2018; Stefanidi et al., 2018），进而调节记忆表现。这意味着赋予项目不同的价值可能会通过影响加工深度和提取优先级而改变部分线索效应的大小甚至方向。因此本研究中实验 1 将首次探讨以不同价值的项目作为部分线索或者测验项目时对部分线索效应的影响。

在引言部分第五段（P4.）补充了项目价值这一变量对于部分线索效应的理论贡献。

实验 2 的逻辑必要性主要体现在：

双机制假说认为,高、低关联编码的操纵通过学-测次数来实现(Lehmer & Bauml, 2018a, 2018b),而价值导向记忆的相关研究表明,项目价值高低直接影响被试的识记策略,进而影响被试对项目之间关联的建立(Knowlton & Castel, 2022)。据此可以推断,对于不同价值的项目来说,通过相同的学习次数所能达到的项目关联编码程度是不同的,相较于低价值项目,高价值项目更容易建立项目间的关联。同时,被赋予高价值的项目由于在编码过程中得到更多注意加工(Allen, 2019),表征强度更高,不容易随着时间的推移(无论是通过衰减还是干扰)而受损,因此相较于低价值的项目,高价值项目在长延迟上的记忆表现更好。因此,项目价值可能既影响关联编码水平也影响部分线索效应的持久性。也即,当项目价值不同,部分线索效应也应有所不同。同样的,当编码要求不同,相同的项目价值也可能产生不同的加工结果,并进而影响部分线索效应的大小及部分线索作用的持久性。这是双机制假说和前人研究中所没有涉及到的。因此,有必要在实验2中进一步探索在不同编码条件下,不同价值的项目作为线索或者目标(测验项目)时对再认成绩的影响及其影响的持久性,并对双机制假说进行检验。

在引言部分第十段(P5.)补充了实验2的实验逻辑。

“优先输出顺序”是指当对学习材料进行多轮次的学习-测试(即高水平的项目间关联条件下),由于被试反复学习和回忆这些项目,他们的回忆顺序在轮次间变得更加一致(Bäuml & Aslan, 2006; Lehmer & Bäuml, 2018a),当要求被试回忆时,则被试优先以该种顺序(而不是其他顺序)对识记项目进行回忆,即为“优先输出顺序”。在引言部分倒数第二段(P6.)补充了对优先输出顺序的解释。

主要参考文献:

- Allen, R. J. (2019). Prioritizing targets and minimizing distraction within limited capacity working memory: Commentary on “Working memory and attention: a conceptual analysis and review” by Klaus Oberauer. *Journal of Cognition*, 2(1), Article 75.
- Hitch, G. J., Hu, Y., Allen, R. J., & Baddeley, A. D. (2018). Competition for the focus of attention in visual working memory: Perceptual recency versus executive control. *Annals of the New York Academy of Sciences*, 1424(1), 64–75.
- Knowlton, B. J., & Castel, A. D. (2022). Memory and reward-based learning: A value-directed remembering perspective. *Annual review of psychology*, 73, 25-52.
- Lehmer, E. M., & Bäuml, K. T. (2018a). The many faces of part-list cuing-evidence for the interplay between detrimental and beneficial mechanisms. *Frontiers in Psychology*, 9, Article 701.
- Stefanidi, A., Ellis, D. M., & Brewer, G. A. (2018). Free recall dynamics in value-directed remembering. *Journal of Memory and Language*, 100, 18–31.

意见2: 方法: 实验流程的阐述存在一定的不足。P6. 对于项目价值评价,建议加入情绪效价和唤醒度的评价。P11. 实验2中干扰阶段,文字标注是“三位数连续减3的运算”,“三位数连续减7的任务”。图中标明的是“连加3数学运算”,“连加7数学运算”。这里存在矛盾。

回应: 感谢专家的建议,根据专家建议我们对实验材料的情绪效价和唤醒度进行了补充评定,评定结果如下:

请30名不参加正式实验的大学生对所选双字词的情绪效价(1代表非常不愉悦,7代表非常愉悦)、唤醒度(1代表非常平静,7代表非常激动)进行7级评定。

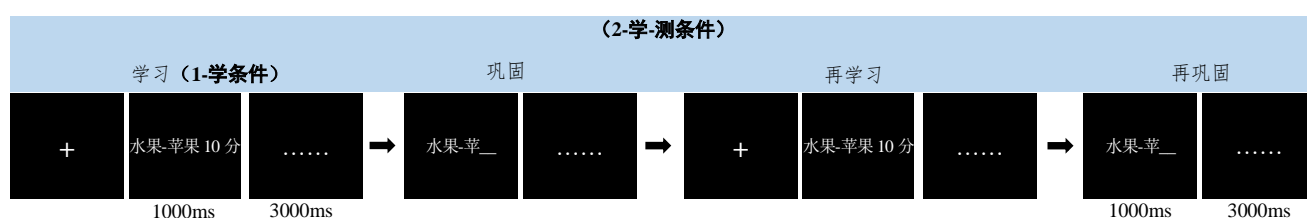
对各词表的情绪效价 ($M_1 \pm SD_1 = 4.06 \pm 0.67$, $M_2 \pm SD_2 = 4.13 \pm 0.56$, $M_3 \pm SD_3 = 4.11 \pm 0.58$)、学习项目情绪效价 ($M_1 \pm SD_1 = 4.07 \pm 0.73$, $M_2 \pm SD_2 = 4.20 \pm 0.55$, $M_3 \pm SD_3 = 4.11 \pm 0.63$)、测验项目情绪效价 ($M_1 \pm SD_1 = 4.11 \pm 0.76$, $M_2 \pm SD_2 = 4.17 \pm 0.58$, $M_3 \pm SD_3 = 4.11 \pm 0.71$) 进行差异检验, 结果表明均无显著差异, $F_s < 1$, $p_s > 0.05$ 。

对各词表的唤醒度 ($M_1 \pm SD_1 = 3.94 \pm 0.52$, $M_2 \pm SD_2 = 3.98 \pm 0.49$, $M_3 \pm SD_3 = 4.00 \pm 0.56$)、学习项目唤醒度 ($M_1 \pm SD_1 = 3.96 \pm 0.52$, $M_2 \pm SD_2 = 4.03 \pm 0.49$, $M_3 \pm SD_3 = 4.02 \pm 0.60$)、测验项目唤醒度 ($M_1 \pm SD_1 = 4.02 \pm 0.54$, $M_2 \pm SD_2 = 3.97 \pm 0.51$, $M_3 \pm SD_3 = 4.05 \pm 0.66$) 进行差异检验, 结果表明均无显著差异, $F_s < 1$, $p_s > 0.05$ 。

在 2.3 实验材料部分 (P7.) 补充了情绪效价和唤醒度的评价结果。

感谢专家的细致阅读! 实验 2 中干扰阶段任务应为“三位数连续减 3 的运算”和“三位数连续减 7 的运算”, 已在图 2 中修改相关描述。另外, 为了更清晰的展示实验流程, 我们对实验 2 的流程图进行了修改, 具体如下:

(a) 学习阶段



(b) 干扰、测验阶段

干扰 1 + 部分线索阶段/干扰

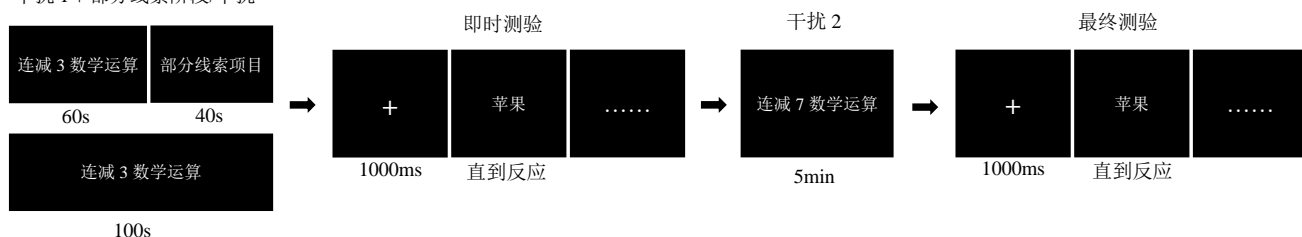


图 3 实验 2 学习与测验阶段实验流程图。(a) 1-学和 2-学-测两种编码条件下学习阶段的流程图; (b) 有线索和线索条件下干扰 1、即时测验、干扰 2 和最终测验的流程图。

意见 3: 结果: 实验结果, 语言表述有时不太准确, 建议一些相对重要的结果以柱状图的形式进行形象化的表达。P8. 低价值线索 ($p = 0.082$) 未能达到显著, 不能以显著的话语来描述。P9. 低价值线索 ($p = 0.064$) 未能达到显著, 不能以显著的话语来描述。

回应: 感谢专家指出了我们对于有的 0.05 边缘显著结果采用了类似显著的描述方式。现对全文中描述不当之处修改如下: P9. 无部分线索条件下的再认正确率高于高价值线索 ($p = 0.001$) 和低价值线索 ($p = 0.082$) 条件下的。修改为: 无部分线索条件下的再认正确率显著高于高价值线索条件下的 ($p = 0.001$), 无部分线索条件下的再认正确率边缘显著高于低价值线索条件下的 ($p = 0.082$)。P9. 高价值线索 ($p = 0.004$) 和低价值线索 ($p = 0.064$)

条件下被试对目标词的反应时显著长于无线索条件。修改为：高价值线索 ($p = 0.004$) 条件下被试对目标词的反应时显著长于无线索条件，低价值线索 ($p = 0.064$) 条件下被试对目标词的反应时边缘显著长于无线索条件。P14. 低价值线索条件下的再认正确率高于高价值线索条件 ($p = 0.061$)。修改为：低价值线索条件下的再认正确率边缘显著高于高价值线索条件 ($p = 0.061$)。测验项目价值主效应显著, $F(1, 28) = 3.82, p = 0.061, \eta_p^2 = 0.12$, 高价值测验项目再认正确率显著高于低价值测验项目。修改为：测验项目价值主效应边缘显著, $F(1, 28) = 3.82, p = 0.061, \eta_p^2 = 0.12$, 高价值测验项目再认正确率边缘显著高于低价值测验项目。

根据专家提出的宝贵意见,我们将实验 1 和实验 2 的结果以柱状图的形式进行形象化表达, 以方便读者阅读。

意见 4: 讨论与结论: 讨论部分没有突出该研究的理论贡献, 对于理论机制的阐述有些不足, 望进一步加以改进。P15. “可能自动化的启动了对线索项目的反向抑制” 需要阐述何为反向抑制, 这种反向抑制是否有文献研究依据。P16. “以上结果部分验证了双机制假说, 也为双机制假说做了必要补充” 这里的解释略微有些笼统, “必要补充” 具体指什么, 这应当是本文创新点和理论价值的集中体现, 但是文中缺乏相应的理论建构。

回应: 感谢专家提出的建设性意见! 帮助我们进一步思考本研究的理论贡献。

关于“反向抑制”的解释如下:

根据部分线索效应的提取抑制假说, 部分线索的呈现会导致被试首先对线索项目进行内隐提取, 这一内隐提取与提取诱发遗忘现象中选择性提取所起的作用一致 (Anderson et al., 1994; Aslan et al., 2007; Bäuml & Aslan, 2006)。在提取诱发遗忘的研究中, 研究者采用竞争强度这一概念用以表征当提取练习不同强度 Rp+项目时, Rp-项目对 Rp+项目形成的干扰(刘旭, 岳鹏飞, 白学军, 2019; Schilling, Storm, & Anderson, 2014; Storm & Bui, 2016)。而提取诱发遗忘效应则是指由于对 Rp+项目的提取练习所导致的测验阶段对 Rp-项目的提取困难, 着重关注的是 Rp+对 Rp-造成的抑制。对于部分线索效应来说, 在部分线索呈现阶段, 当呈现部分线索时, 非线索项目将对线索项目的内隐提取产生竞争, 同时, 为克服非线索项目的干扰以完成对线索项目的内隐提取, 启动抑制机制对非线索项目的记忆表征进行抑制, 这种抑制使得项目本身表征强度降低, 导致在随后的记忆测试中, 非线索项目的回忆受到损害。提取抑制假说所说的“抑制”是指对线索项目的内隐提取所导致的对非线索项目的抑制, 为了区别于此种“抑制”, 经慎重思考及讨论, 本文最终使用“反向竞争”而不再使用“反向抑制”来表达当呈现线索时, 非线索项目对线索项目内隐提取的干扰。在**总讨论部分第三段 (P16.)**, 我们增加了对“抑制”与“反向竞争”的区分解释。

我们也对本文的创新点和理论价值进行了重新思考和梳理:

首先，在以往部分线索效应的研究中，测验方式大多为自由回忆测验。采用自由回忆测验时，有部分线索条件下的惯常做法是：首先呈现部分线索项目请被试阅读，然后要求被试根据线索项目回忆之前学习的剩余项目。此时，部分线索的呈现与回忆过程在时间上有很大的重叠，这导致线索项目与回忆过程之间不可避免的交互影响，使得无法得到纯粹的“部分线索”效应。而本研究则采用了再认测验，一方面可以有效的将线索呈现与项目回忆过程从时间轴上分开（先呈现部分线索，后进行再认测验，进行测验时部分线索不再呈现），避免了线索与回忆的交互影响，更重要的是，相较于自由回忆，再认测验对线索的变化更敏感（Todres & Watkins, 1981; Oswald, Serra, & Krishna, 2006），可以更敏感的测量高价值线索、低价值线索与无线索条件下的差异（事实上，在实施再认测验之前，我们在相同的实验设计中也采用了自由回忆测验进行施测，结果发现再认测验能更敏感的反映高、低价值线索和高、低价值测验项目间的差异）。这是本研究实现理论创新的方法基础。

其次，本研究的理论价值主要体现在：（1）价值导向记忆的相关研究发现，对有价值信息的记忆选择性主要体现在选择性编码和巩固这些信息或在测验时优先提取这些信息，即高价值不仅使得在编码时对这些项目进行更深层次的语义加工（Cohen et al., 2017），也使得在提取时这些项目更易获得优先提取（Stefanidi et al., 2018）。这表明在记忆活动中，人们会选择性地关注重要的信息，但选择性编码的重要性目前还没有得到足够的重视（Knowlton & Castel, 2022）。因此，本研究首次考察了当使用不同价值的项目为线索或目标时对部分线索效应的影响。同时，由于被赋予高价值的项目在编码过程中得到更多注意加工（Allen, 2019），表征强度更高，不容易随着时间的推移（无论是通过衰减还是干扰）而受损。因此相较于低价值的项目，高价值项目在长延迟上可及性水平更高。因此，本研究进一步考察了项目价值影响部分线索效应的持久性。（2）双机制假说认为，在不同的编码情况下，部分线索通过不同机制发挥破坏作用：在低关联编码条件下主要是基于提取抑制持久损害目标项目的回忆；而在高关联编码条件下则是基于提取策略的暂时改变对目标项目短暂的回忆损伤（Bäuml & Aslan, 2006）。实验 2 中，低关联编码条件下低价值线索项目对高价值测验项目的再认损害仅在即时测验中存在，到最终测验消失，此时部分线索的干扰效应是短暂的，低价值线索对低价值测验项目的再认无显著损害效应，这意味着并非在低关联编码条件下就一定表现为损害效应。在实验 2 的高关联编码条件下，无论是即时测验还是最终测验中，高价值测验项目的再认正确率均不受部分线索的显著影响，仅高价值线索使得低价值测验项目的再认受损从即时测验持续到了最终测验。**这表明，部分线索对记忆提取的作用，可能并非如双机制假说所认为的，在低关联编码条件下就一定表现为持久的回忆损伤，在高关联编码条件下就一定表现为暂时的回忆受损，项目价值也影响着部分线索作用的强度及持续时间，在从项目关联编码这一角度界定部分线索对记忆提取的作用时，还需要考虑项目价值的作用。**在讨论部分倒数第二、第三段（P17.），我们增加了本研究理论价值的阐释，在结论部分（P18.），增加了对双机制假说补充的具体说明。

主要参考文献:

- Allen, R. J. (2019). Prioritizing targets and minimizing distraction within limited capacity working memory: Commentary on “Working memory and attention; a conceptual analysis and review” by Klaus Oberauer. *Journal of Cognition*, 2(1), Article 75.
- Cohen, M. S., Rissman, J., Hovhannisyan, M., Castel, A. D., & Knowlton, B. J. (2017). Free recall test experience potentiates strategy-driven effects of value on memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(10), 1581-1601.
- Knowlton, B. J., & Castel, A. D. (2022). Memory and reward-based learning: A value-directed remembering perspective. *Annual review of psychology*, 73, 25-52.
- Liu, X., Yue, P. F., & Bai, X. J. (2019). The correlated costs and benefits problem in retrieval-induced forgetting: Evidence from effects of response inhibition and item competitive intensity. *Journal of Psychological Science*, 42(5), 1039-1046.
- [刘旭, 岳鹏飞, 白学军. (2019). 提取诱发遗忘中的相关代价与效益问题: 反应抑制能力与项目竞争强度的影响. *心理科学*, 42(5), 1039-1046.]
- Oswald, K. M., Serra, M., & Krishna, A. (2006). Part-list cuing in speeded recognition and free recall. *Memory and Cognition*, 34(3), 518-526.
- Schilling, C. J., Storm, B. C., & Anderson, M. C. (2014). Examining the costs and benefits of inhibition in memory retrieval. *Cognition*, 133(2), 358-370.
- Stefanidi, A., Ellis, D. M., & Brewer, G. A. (2018). Free recall dynamics in value-directed remembering. *Journal of Memory and Language*, 100, 18-31.

.....
审稿人 2 意见:

《部分线索诱发遗忘效应的持久性: 项目价值的作用》研究了部分线索价值、项目价值以及编码程度对项目记忆的影响。发现部分线索对项目记忆起干扰作用, 且高价值线索的干扰作用比低价值的大, 其干扰效果持续也更久。项目价值高则记忆效果好。

但稿件仍有以下问题有待商榷。

意见 1: 实验 2 发现“低价值线索仅使高价值测验项目在即时测验中受损”。作者认为“相较于低价值目标项目, 高价值目标项目的再认正确率更高, 反应时更短”, 因此, 怎么解释实验 2 中发现的“低价值线索仅使高价值测验项目在即时测验中受损”, 而“低价值部分线索对低价值测验项目无显著损害”。

回应: 非常感谢审稿专家的宝贵意见! 使我们有机会对这一问题进行深入思考。

考察部分线索效应是通过比较相同价值测验项目在不同线索条件下的成绩差异来实现的, 即高、低价值线索对于测验项目损害与否是通过与无线索条件下相应的高、低价值测验项目再认成绩的比较得出的, 而价值效应则是比较同一线索条件下不同价值测验项目之间的成绩差异。研究假设中所述“相较于低价值目标项目, 高价值目标项目的再认正确率更高, 反应时更短”与实验 2 中发现的“低价值线索仅使高价值测验项目在即时测验中受损, 而低价值部分线索对低价值测验项目无显著损害”并不矛盾, 具体解释如下:

已有价值导向记忆的相关研究表明, 赋予项目不同的价值将改变项目的表征强度(Elliott et al., 2020; Knowlton & Castel, 2022; Villasenor et al., 2021), 高价值项目表征强度更高, 而低价值项目表征强度较低。根据提取抑制假说, 部分线索的呈现会导致被试首先对线索项目进行内隐提取, 这一内隐提取与提取诱发遗忘现象中选择性提取所起的作用一致(Anderson et al., 1994; Aslan et al., 2007; Bäuml & Aslan, 2006)。因此, 当呈现低价值线索时, 高强度

的非线索项目将对低价值线索的内隐提取产生强干扰；而呈现高价值部分线索时，低强度非线索项目则对低价值线索的内隐提取产生弱干扰，甚至无干扰（Storm & Bui, 2016）。本研究将线索呈现与项目再认过程从时间上分离开来，因此在低价值线索条件下，在线索呈现阶段被强干扰过的低价值线索几乎无法对测验项目产生干扰作用，而在线索呈现阶段高强度的非线索项目（测验项目）对线索项目的强干扰导致了非线索项目本身表征强度的受损。同时，为克服非线索项目的干扰以完成对线索项目的内隐提取，启动抑制机制对非线索项目的记忆表征进行抑制，这种抑制使得项目本身表征强度降低，导致在随后的记忆测试中，非线索项目的回忆受到损害。

研究假设中所说的“相较于低价值目标项目，高价值目标项目的再认正确率更高，反应时更短”的结论，是通过比较同一线索条件下高、低价值测验项目之间再认成绩的差异得来的。由于无线索条件下高、低价值项目之间再认成绩的显著差异，为高、低价值线索的作用提供了不同的基线。因此，“低价值线索使高价值测验项目在即时测验中受损，且对低价值测验项目无显著损害”的结论是指低价值线索条件下的高价值测验项目的再认成绩显著低于无线索条件下的高价值测验项目，而低价值线索条件下的低价值测验项目的再认成绩与无线索条件下的低价值测验项目无显著差异。

在**讨论部分第三段（P16.）**，也增加了对上述结果的解释。

主要参考文献：

- Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(5), 1063-1087.
- Aslan, A., & B üml, K. H. (2007). Part-list cuing with and without item-specific probes: the role of encoding. *Psychonomic Bulletin & Review*, 14(3), 489-494.
- B üml, K.-H., & Aslan, A. (2006). Part-list cuing can be transient and lasting: the role of encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(1), 33-43.
- Elliott, B. L., McClure, S. M., & Brewer, G. A. (2020). Individual differences in value-directed remembering. *Cognition*, 201, Article 104275.
- Knowlton, B. J., & Castel, A. D. (2022). Memory and Reward-Based Learning: A Value-Directed Remembering Perspective. *Annual Review of Psychology*, 73(1), 25-52.
- Storm, B. C., & Bui, D. C. (2016). Retrieval-practice task affects relationship between working memory capacity and retrieval-induced forgetting. *Memory*, 24(10), 1407-1418.
- Villasenor, J. J., Sklenar, A. M., Frankenstein, A. N., Levy, P. U., McCurdy, M. P., & Leshikar, E. D. (2021, Aug). Value-directed memory effects on item and context memory. *Memory and Cognition*, 49(6), 1082-1100.

意见 2：本研究为多因素设计，实验 2 更是 2*3*2*2 的四因素设计，因此结果比较复杂，目前是用表格来展示反应时和正确率的结果。不知道有没有更好的办法。

回应：非常感谢专家的意见。另一位审稿专家也提出了同样的问题，结合两位专家的意见，我们将实验 1 和实验 2 的结果以柱状图的形式进行形象化表达，以便读者阅读。具体如下：

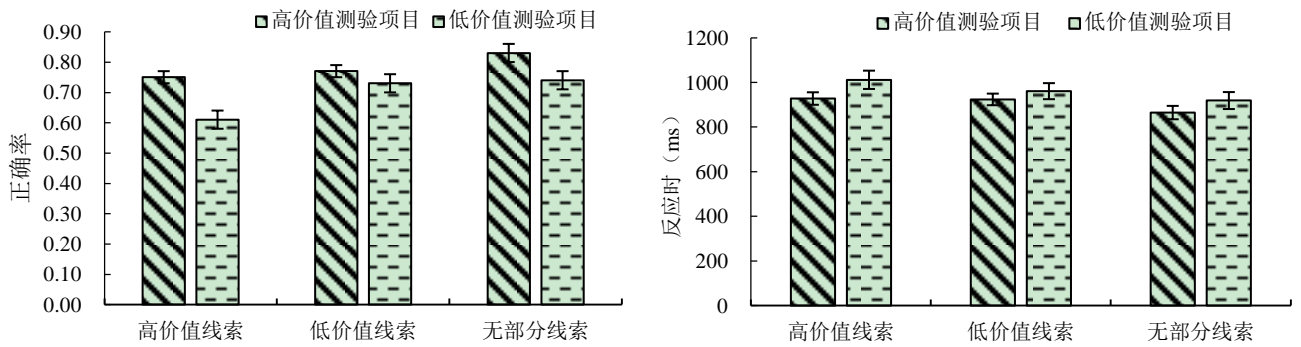


图2 实验1中各线索条件下不同价值测验项目的正确率（左图）和反应时（右图）

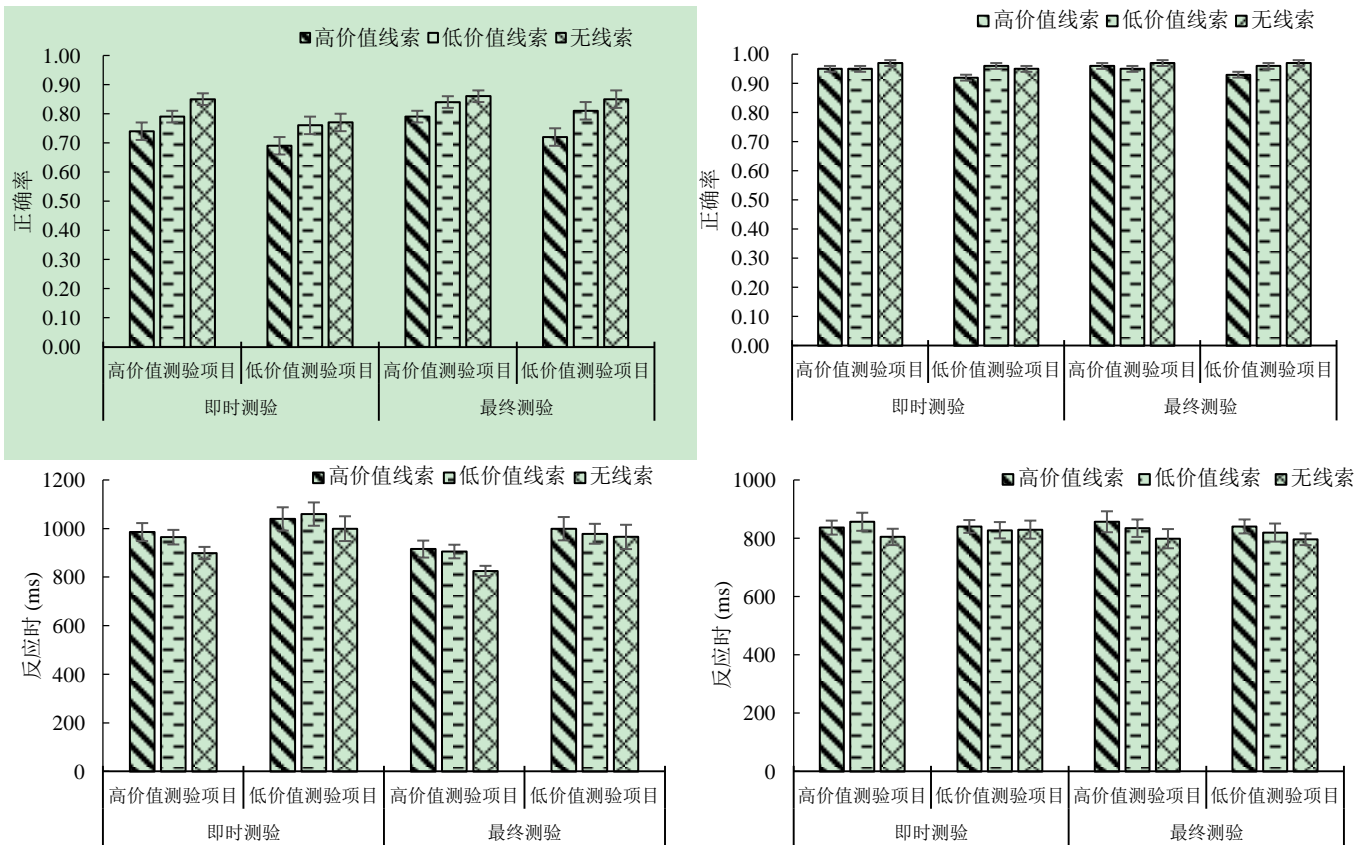


图4 实验2中各线索条件下不同价值测验项目的正确率（左上，1-学条件；右上，2-学-测条件）和反应时（左下，1-学条件；右下，2-学-测条件）

意见 3: “实验程序采用 E-Prime 2.0 软件进行编程，使用联想笔记本电脑呈现刺激。正式实验开始之前，要求被试先进行两组练习：”，请报告笔记本具体型号；请报告练习试次数量，以及练习中使用的词是否在正式实验中出现。

回应: 非常感谢专家的建议和提醒。根据专家意见，补充了笔记本电脑的具体型号和练

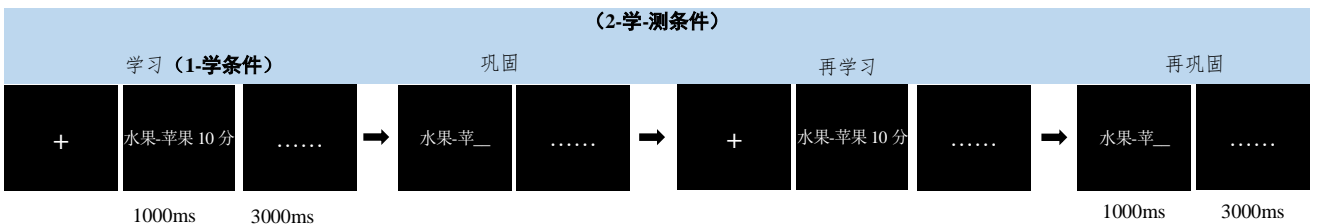
习试次，具体如下：

使用联想笔记本电脑（ThinkBook14 英特尔酷睿版，屏幕 14 英寸，分辨率 1920×1080，刷新率 60Hz）呈现刺激。为便于被试熟悉实验流程，在正式实验开始之前，要求被试先进行两组练习：一组为有部分线索条件，一组为无部分线索条件。每组学习阶段 8 个试次，测验阶段 8 个试次，练习所用实验材料在正式实验中不出现。

意见 4：实验 2 流程图，图 2 上下两部分中间的箭头指向的位置是“中间测验”，会让人有误解为上方流程结束后接下来就是中间测验。图中的“中间测验”，在正文里是即时测验。感觉统一为好。流程图感觉可以画得更好一些。

回应：感谢专家的细致阅读。“中间测验”应为“即时测验”，已在图中更正，同时也检查全文，避免说法不统一的现象。图 2 上下两部分中间的箭头指向的位置确实容易让人误解为上方流程结束后接下来即为即时测验，为了避免读者误解，同时更清晰的展示实验流程，我们对实验 2 的流程图进行了修改，具体如下：

(a) 学习阶段



(b) 干扰、测验阶段

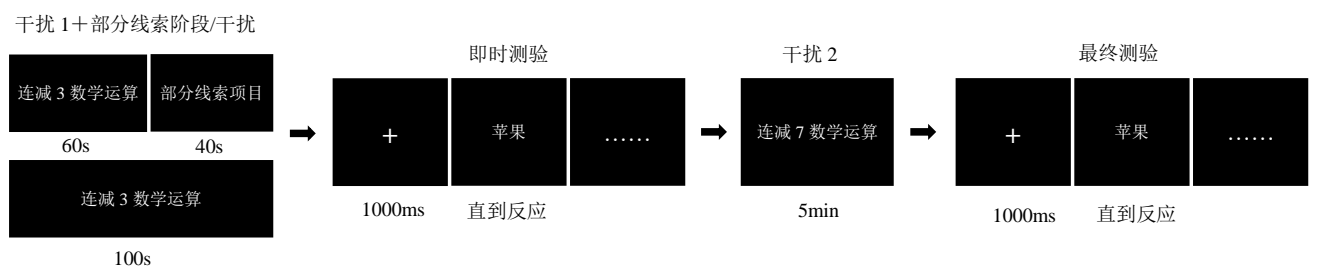


图 3 实验 2 学习与测验阶段实验流程图。(a) 1-学和 2-学-测两种编码条件下学习阶段的流程图；(b) 有线索和线索条件下干扰 1、即时测验、干扰 2 和最终测验的流程图。

意见 5：关于统计：1) 有的.05 边缘显著结果采用了类似显著的描述方式，比如测验项目价值主效应显著， $F(1, 28) = 3.82, p = 0.061, \eta^2 p = 0.12$ 。2) 请核查所有方差分析是否满足球形假设。不满足请做校正。

回应：感谢专家的指正，现对全文中描述不当之处修改如下：

P9. 无部分线索条件下的再认正确率高于高价值线索 ($p = 0.001$) 和低价值线索 ($p = 0.082$) 条件下的。修改为：无部分线索条件下的再认正确率显著或边缘显著高于高价值线

索条件下的 ($p = 0.001$) 和低价值线索条件下的 ($p = 0.082$)。P9. 高价值线索 ($p = 0.004$) 和低价值线索 ($p = 0.064$) 条件下被试对目标词的反应时显著长于无线索条件。修改为: 高价值线索 ($p = 0.004$) 和低价值线索 ($p = 0.064$) 条件下被试对目标词的反应时显著或边缘显著长于无线索条件。P14. 低价值线索条件下的再认正确率高于高价值线索条件 ($p = 0.061$)。修改为: 低价值线索条件下的再认正确率边缘显著高于高价值线索条件 ($p = 0.061$)。测验项目价值主效应显著, $F(1, 28) = 3.82$, $p = 0.061$, $\eta_p^2 = 0.12$, 高价值测验项目再认正确率显著高于低价值测验项目。修改为: 测验项目价值主效应边缘显著, $F(1, 28) = 3.82$, $p = 0.061$, $\eta_p^2 = 0.12$, 高价值测验项目再认正确率边缘显著高于低价值测验项目。

感谢专家的提醒, 已对本研究中的统计结果进行核查, 对于不满足“球形检验”(Mauchly's Test of Sphericity) 假设的结果, 采用 Greenhouse-Geisser 校正 p 值, 多重比较时采用 Bonferroni 校正。

意见 6: 请核实参考文献格式: 比如白学军 等., 2014; 刘湍丽 等., 2019)

回应: 非常感谢专家细致的阅读。我们查阅了《心理学报》杂志官网和最新 APA 出版手册公布的参考文献著录要求, 如果有 3 个或更多个著者, 正文中对于文献的引用只写第一著者的姓(名), 后面用“等”或“et al.”, 具体示例: 赵一 等, 2008; Wolchik et al., 2007。本文对于中文参考文献的引用确实有误, 等字后应加上逗号“,”, 而本文全部误写为点加逗号“.,”。已仔细阅读参考文献著录要求, 并对文中参考文献格式表述不规范的地方进行改正。

第二轮

审稿人 1 意见:

总体而言, 作者对文章的修改较为充分, 对审稿人提出的问题也做了相对妥善的回答, 但是文中仍然有几处小问题需要进一步修改。

意见 1: 结果部分, 在做简单效应分析时, 是否进行了多重比较校正(如 Bonferroni 校正), 以避免假阳性结果的出现?

回应: 非常感谢专家的意见。根据上一轮专家的意见, 我们已对本研究的统计结果进行核查, 对于不满足“球形检验”假设的结果, 采用 Greenhouse-Geisser 校正 p 值, 多重比较时采用 Bonferroni 校正。在本轮修改中, 我们在 **2.6 结果与讨论部分第二段 (P9.)** 对统计校正进行了说明。

意见 2: 在结果部分, 报告的反应时数据是否剔除 3 个标准差以外的 trial, 以使结果更稳定?

回应：非常感谢专家提出的问题。文中报告的反应时数据是剔除了3个标准差以外及反应错误的试次，且已在 **2.6 结果与讨论部分第一段(P8.)**和 **3.6 结果与讨论部分第一段(P12.)**进行过说明。

意见 3：在柱状图上建议对数据差异的显著性进行标注，以便读者阅读。

回应：非常感谢专家的宝贵建议！根据专家的建议，我们已在柱状图上对数据差异的显著性进行了标注。因在正确率指标上，大部分情况下线索条件与其他变量的交互作用是显著的，而反应时指标上，则主要表现为显著的主效应，因此正确率和反应时结果图上横坐标变量的设置有所不同，以方便标注显著性结果。具体修改如下：

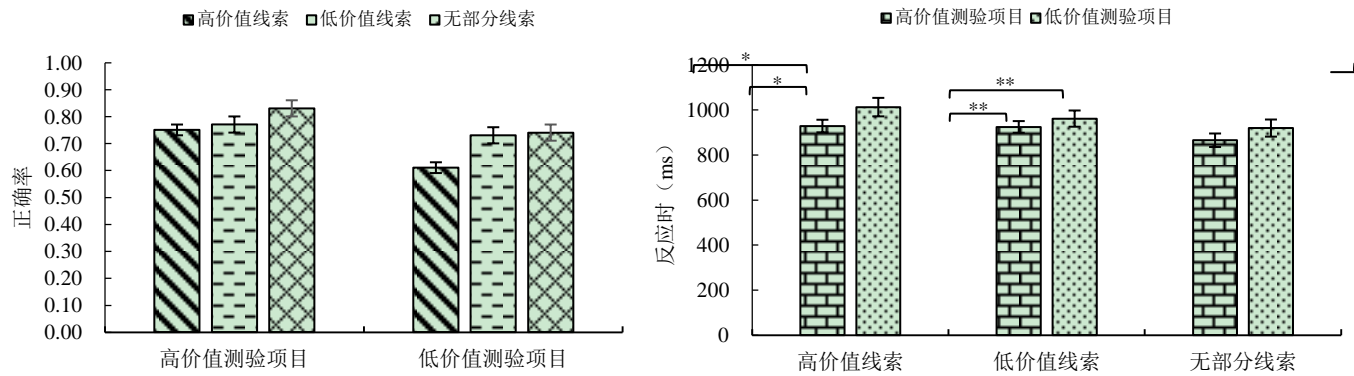


图2 实验1中各线索条件下不同价值测验项目的正确率(左图)和反应时(右图)($^{\dagger}p < 0.07$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$)

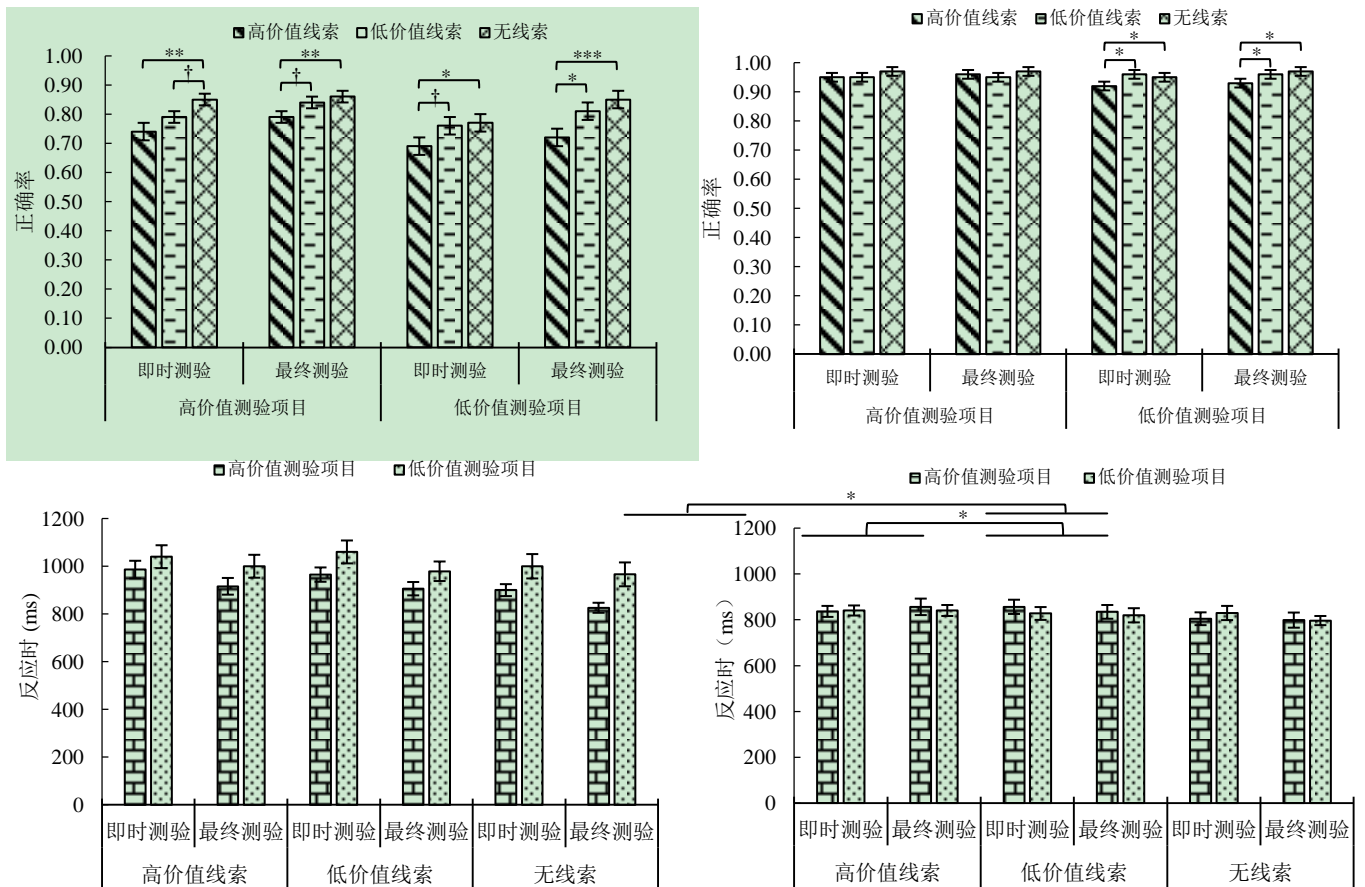


图4 实验2中各线索条件下不同价值测验项目的正确率(左上, 1-学条件; 右上, 2-学-测条件)和反应时(左下, 1-学条件; 右下, 2-学-测条件) ($\dagger p < 0.07$, $* p < 0.05$, $** p < 0.01$, $*** p < 0.001$)

意见 4: 英文摘要需要经过有熟练英文写作经验的作者进一步修改。

回应: 非常感谢专家的宝贵建议! 为使英文摘要表达更为地道, 我们请了麦考瑞大学 (Macquarie University) 心理科学学院的老师对摘要进行了修改润色。

Persistence of Part-list-cuing-induced Forgetting:

The Role of Item Value

The part-list cuing effect refers to the phenomenon that when providing a subset of previously learned items as retrieval cues, people's recall performance for the remaining items is often worse compared to when retrieval cues are absent. Memory research also showed that items with high value are generally better remembered than items with low value. However, it is unclear how the values of items affect the part-list cuing effect and its persistence. Through two experiments, this study investigated the influence of item value on the part-list cuing effect.

Experiment 1 employed a part-list cuing paradigm in a value-directed memory task. During the learning phase, participants were asked to study category exemplars which were assigned different values (1 or 10 points). Participants were then asked to make an old/new judgement after

the presentation of part-list cues. Experiment 2 further manipulated the encoding condition (i.e., 1-study encoding vs. 2-study-test encoding) and the test schedule (i.e., immediate test vs. final test). In the 1-study condition, participants received only one study cycle, but went through two study–test cycles in the 2-study-test condition. The immediate test phase is the same as Experiment 1; the final test involved a final recognition test after a 5min distractor task.

Results from the two experiments collectively showed both the assigned values of cued and test items affected the item recognition performance: cue items with high value resulted in poorer target item recognition performance than those with low value; however, the recognition accuracy was higher for target items with high- than low-value, and the high-value target items were more sensitive to the presentation of part-list cuing. The emergence and persistence of part-list cuing was also modulated by item values. Under the 1-study condition, the high-value cues led to worse target item recognition regardless of the values of the target items, and this detrimental effect was observed in both immediate and final tests. In contrast, the low-value cues only caused poorer recognition of high-value targets in the immediate test. Under the 2-study-test condition, only high-value cues caused recognition impairment of the low-value targets in both immediate and delayed tests. The above results partially validate the two-mechanism account of part-list cuing, and also are a key supplement to this hypothesis: the role of part-list cuing on memory retrieval is not necessarily manifested as a lasting impairment in the low associative coding condition, or a transient impairment in the high associative coding condition, and the item value also influences the strength and persistence of the role of part-list cuing, and it is also necessary to take into account the role of item value when defining the role of part-list cuing on memory retrieval from the perspective of item associative encoding.

Keywords: part-list cuing; item value; encoding conditions; test schedule

第三轮

审稿人 2 意见：修改较为完善，建议发表。

编委意见：

意见 1：文中的结果部分有许多 $p>0.05$ 的统计检验结果都算作显著，算作“边缘显著”也不合适。应当修改更正。

回应：感谢编委专家提出的问题。鉴于本文几处“边缘显著”结果皆有差异显著的趋势（效应量较大）（Cohen, 1988; 胡竹菁, 戴海琦, 2011; 温忠麟 等, 2022），为了提高实验结果的稳定性，我们将实验 2 的被试人数由原来的 59 人增加至 74 人，以确保每种条件下被试达到 30 人以上。更新后的结果请见“3.6 结果与讨论”部分。

另外，我们查阅了近几年心理学领域关于统计结果中“边缘显著”问题的相关文献（e.g., Olsson-Collentine, van Assen, & Hartgerink, 2019; Pritschet, Powell, & Horne, 2016），这些文献认为“ p 值在 0.05 和 0.10 之间”证据价值较低，不建议将这些结果报告为边缘显著，因此本文中仅在 $p < 0.05$ 时才视为差异显著，而不再采用“边缘显著”的说法，并对全文中的结果陈述方式进行了统一。

主要参考文献：

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Olsson-Collentine, A., van Assen, M. A. L. M., & Hartgerink, C. H. J. (2019). The prevalence of marginally significant results in psychology over time. *Psychological science*, 30(4), 576–586.
- Pritschet, L., Powell, D., & Horne, Z. (2016). Marginally significant effects as evidence for hypotheses: Changing attitudes over four decades. *Psychological science*, 27(7), 1036–1042.
- 胡竹菁, 戴海琦. (2011). 方差分析的统计检验力和效果大小的常用方法比较. *心理学探新*, 31(3), 254–259.
- 温忠麟, 谢晋艳, 方杰, 王一帆. (2022). 新世纪 20 年国内假设检验及其关联问题的方法学研究. *心理科学进展*, 30(8), 1667–1681.

意见 2：描述统计的图和表数据重复，只需要用图或表其中一种即可。

回应：感谢专家细致和严谨的审阅。结合您及前两轮外审专家的意见，本文中描述统计的结果保留了形象化表达的柱状图，以方便阅读。

同意修改后发表。

第四轮

主编意见：同意发表。