

# 《心理学报》审稿意见与作者回应

题目：学习材料组块方式对相似词长时记忆的影响

作者：张蕾 鲁成柔 林军凤 梅磊磊

## 第一轮

### 审稿人 1 意见：

该研究理论意义明确，具有较高的理论价值和一定的实践应用意义，实验设计较为严谨，操作定义清晰明确，数据分析方法得当，符合心理学科研标准。然而，在实验设计中依旧存在一些疏漏，数据分析的结果到该文所下的结论之间还存在一些非必然性和过度陈述的问题。对此，作者还需要解决以下问题：

**意见 1：**在实验 1 和 2 中，研究者将 112 个英文非词分配至两种学习条件下，即相似词组块和不相似词组块学习条件各 56 个英文非词（14 组×4 个词）。相似词组块条件下词组内 4 个词具有共同的词根（4 个字母），而不相似词组条件下词组内的共同字母保持在 0.32 个以下。然而，作者称本研究的记忆对象（两组学习材料）均为相似词，但不相似词组内的词之间的相同字母却不超过 0.32 个，这是否为相似词？请作者在文章“实验材料”部分明确指出，不相似词组块条件下的相似词是何种定义。

**回应：**非常抱歉，此处文字引起了误解。根据审稿意见，已在正文“2.2.2 实验材料”（论文第 4 页）和“2.2.3 实验程序”（论文第 4-5 页）两部分中重新对两种组块条件下相似词的相似程度进行了详细说明。

首先，两种学习条件下（相似词组块和不相似词组块）的学习刺激为同质的相似词材料。每种组块条件包含 56 个英文非词，包含 14 个相似词组，每个词组由 4 个相似的英文非词组成。每个相似词组中的四个非词间至少前 4 个字母或后 4 个字母相同。两组实验材料中，每个相似词组内的四个英文非词之间的共同字母平均数量分别为 4.25 个和 4.26 个。为了进一步排除实验材料的差异，我们对两种学习条件的材料进行了被试间平衡。

其次，为了考察材料组块方式效应，我们将不相似词组块条件下的 14 组相似词材料的组块方式打乱，形成 14 个不相似词组，每个不相似词组的 4 个不相似词连续呈现，每个不相似词组内的四个英文非词之间的共同字母的平均数量小于 0.5 个。为了避免误解，我们已将相关描述修改为“学习任务包含两种学习条件：相似词组块条件和不相似词组块条件。在

两种学习条件下，被试都需要学习 56 个英文非词，包含 14 组相似词。在相似词组块条件下，学习材料分成 14 个相似词组，每个相似词组的 4 个相似词连续呈现，每个相似词组内的四个英文非词之间的共同字母的平均数量大于 4；而在不相似词组块条件下，将由 14 个相似词组组成的学习材料分成 14 个不相似词组，每个不相似词组的 4 个不相似词连续呈现，每个不相似词组内的四个英文非词之间的共同字母的平均数量小于 0.5 个。”

**意见 2：**实验 1 中，为了防止被试对所学非词产生额外的精细复述，要求被试在学习任务完成后进行 5 分 50 秒的数字判断任务，完成任务后经过 54 分 10 秒后进入再认测试。那么完成数字判断任务到开始再认测试的过程中，被试的情况如何？

**回应：**非常感谢审稿专家的宝贵意见。在完成数字判断任务至再认测试之间的这段时间内，被试被安排在一间休息室中等待再认记忆任务测查。在休息期间，被试可以看杂志或玩手机，但不能回忆学习材料。我们已在“2.2.3 实验程序”部分（论文第 5 页）对被试的相关情况进行了说明。

**意见 3：**实验 2 中，为了探究不相似词组块对相似词记忆效果的促进是基于单个词汇还是基于不同词之间的词根，研究者利用错误记忆范式，发现对不相似词组块下个体对关键诱饵的虚惊率显著更高。那么，被试区分旧词和诱饵的感受性在两种条件下是否存在差异？请进一步分析后报告。

**回应：**非常感谢审稿专家提出的建议。根据您的意见，我们对实验 2 的数据进行重新整理和分析。在该分析中，我们计算了两种条件下被试区分旧词和诱饵词的辨别力指数，并进行了统计分析。结果发现，相似词组块和不相似词组块条件的辨别力指数分别为 0.86 和 0.89，两者没有差异 ( $t(21) = -0.32, p = 0.75$ )。我们已将上述结果添加到实验 2 的实验结果部分（论文第 10 页）。

**意见 4：**实验 2 中，作者对被试的旧词击中率和对诱饵的虚惊率进行了分析，但这在实质上属于信号检测论中操作者的判断标准的变化。在被试感受性没有发生变化的情况下，被试的击中和虚惊一定是同高同低变化，因此这部分的相关分析没有太大意义。建议作者改为直接对被试判断标准的计算和统计分析。

**回应：**非常感谢审稿专家提出的宝贵意见。正如审稿专家所言，在被试区分旧词和诱饵词的感受性没有发现变化的情况下，被试的旧词击中率和诱饵词虚惊率会同向变化。因此，

根据修改建议，我们删除了原文中关于旧词击中率和诱饵词虚惊率的相关分析。

同时，我们增加了对被试的判断标准的分析。结果发现，在判断标准上，相似词组块条件的判断标准 ( $\beta = 1.29$ ) 显著高于不相似组块条件的判断标准 ( $\beta = 0.90$ ) ( $t(21) = 2.96$ ,  $p = 0.049$ ,  $d = 0.45$ , 95%CI[0.04,0.22])。该结果说明，不相似词组块条件下诱饵词的错误记忆率的上升是因为被试在不相似词组块条件下区分旧词和诱饵词的判断标准较为宽松，更容易将诱饵词判断为旧词。我们已经将上述结果添加到实验 2 的结果部分（论文第 10 页）。

**意见 5:** 相比于实验 1 和 2 中的英文非词，实验 3 中所采用的韩字明显在空间频率和结构上更加复杂，如何确定被试在实验过程中的确可以学到，而不会发生地板效应？

**回应:** 非常感谢审稿专家提出的宝贵意见。正如审稿专家所言，相比于英文非词，韩字在空间频率和结构上更加复杂，因而记忆难度可能更大。为了防止出现地板效应，实验 3 选取的实验材料数量（学习材料 64 个）少于前三个实验中英文非词的数量（学习材料 112 个），以降低被试的记忆负担。实验结果也发现，被试在再认测试中的记忆正确率为 0.67，记忆难度处于中等水平，且记忆正确率显著高于概率水平（0.50）( $t(25) = 10.17$ ,  $p < 0.001$ ,  $d = 1.99$ )。该结果说明实验 3 没有出现地板效应。我们已在实验 3 的实验材料部分（论文第 12-13 页）进行了详细说明。

**意见 6:** 对单个韩字的学习能否称为对词的学习？

**回应:** 非常感谢审稿专家提出的宝贵意见。不同的语言关于词的定义存在差异。在字母语言中，关于词的定义比较清楚，词与词之间由空格隔开；而在汉语中，词的定义相对比较模糊，汉语中的词语可以由一个或多个汉字组成（陈嘉映，2007；李敏，2015）。实验 3 选取单个韩字作为词汇记忆的材料主要基于两方面考虑。一方面，实验被试的母语均为汉语，汉语中存在很多单字词（陈嘉映，2007；李敏，2015），因而对这些被试而言，单个韩字也可以视作词；另一方面，韩字的记忆难度较大，使用单字能够很大程度上降低陌生材料的记忆难度。我们已在实验 3 的实验材料部分（论文第 12 页）进行了说明。

**参考文献:**

陈嘉映. (2007). 约定用法和"词"的定义(续). *外语学刊*(6), 13-18.

李敏. (2015). 关于现代汉语词的定义新探. *海外英语*(11), 208-210.

**意见 7:** 作者认为实验 3 的结果可以说明“不相似词组块不能促进视觉相似词的记忆”，还

是说明“不相似词组块不能促进只具备视觉相似性的词的记忆”？字形相似语音不相似和字形相似但完全没有语音信息这两种情况是否一样？请作者慎重考虑后再下结论。

**回应：**非常感谢审稿专家提出的宝贵意见。我们完全赞同审稿专家关于字形相似语音不相似和字形相似没有语音之间可能存在差异的论述。由于本研究的主要目的是探讨材料组块方式对陌生词汇的影响，而陌生词汇中几乎不太可能找到视觉相似而语音不同的词汇，所以本研究未包含字形相似而语音不相似的词汇。针对审稿专家的意见，我们进行了如下两方面的修改：

首先，我们将实验 3 的结论修改为“不相似词组块不能促进只具备视觉相似性的相似词的记忆”，并在讨论中对相关文字进行相应修改（论文第 14 页）。同时，我们对研究结论的用词进行了弱化，将其调整为“不相似词组块方式对相似词记忆的促进作用可能依赖于学习材料的语音相似性”。

其次，我们在总讨论部分增加了对此问题的讨论（论文第 15 页）。具体来说，由于本研究中的陌生韩字材料不具有语音，所以无法排除视觉相似性在不相似词组块效应中单独起作用的可能性，即无法排除不相似词组块促进视觉相似而语音不相似词记忆的可能性。未来研究需要加入字形相似而语音不相似的陌生词汇作为学习材料，对上述可能性进行进一步探讨。

**意见 8：**请作者仔细检查报告的数据。例如“击中率： $t(21) = -3.50, p = 0.002, d = 0.65, 95\%CI[-0.12, -0.03]$ ；辨别力指数： $t(21) = -3.29, p = 0.003, d = 0.70, 95\%CI[-0.45, -0.10]$ ”。在同样自由度的  $t$  分布中进行假设检验时，为何  $t$  值大的效应反而效应量小，这令人费解。

**回应：**十分抱歉，由于数据计算失误给您造成了困扰。我们已经对全文中的数据进行仔细检查和重新计算，并在论文中进行了修改。例如，在“4.3 实验结果”部分，第一段文字表述修改为“击中率： $t(21) = -3.50, p = 0.002, d = 0.75, 95\%CI[-0.12, -0.03]$ ；辨别力指数： $t(21) = -3.29, p = 0.003, d = 0.70, 95\%CI[-0.48, -0.10]$ ”。

.....  
**审稿人 2 意见：**

本研究采用学-测范式，以英文非词和韩字为材料，操纵学习组块的方式，探讨学习组块方式在不同时间间隔下对相似词的影响及其机制。该研究实验间逻辑清晰，具有一定的应用价值和创新性，但是在行文中存在一些问题。

## 一、重要内容表述不充分

**意见 1:** 摘要内容不完整且不够准确。在摘要中只介绍了本实验研究内容和结果，没有概括研究背景和研究目的以及研究意义。且摘要中提到“通过四个实验考察了学习材料组块方式对相似词记忆的影响，”但只有前两个实验是考察影响，后两个实验考察的影响机制，建议改为“通过四个实验考察了学习材料组块方式对相似词记忆的影响及机制。”

**回应:** 非常感谢审稿专家的宝贵意见。根据审稿意见，我们在摘要中增加了研究背景、目的和意义，并在“通过四个实验考察了学习材料组块方式对相似词记忆的影响”后添加“及机制”（论文第 1 页）。修改后的摘要表述“已有关于材料相似性影响短时记忆的研究提示，不相似材料组块相比于相似材料组块可能促进记忆。为验证该假设，该研究采用学习—测查范式，通过四个实验考察了学习材料组块方式对相似词记忆的影响及机制。结果发现：1) 与相似词组块相比，不相似词组块促进了相似词记忆；2) 不相似词组块的促进效应是通过增强相似词表共同词根的记忆而实现的；3) 不相似词组块的促进效应可能依赖于语音相似性。该结果说明不相似词组块可能是促进相似词汇记忆的有效途径之一。”

**意见 2:** 文章想要探讨不同组块方式对相似词记忆的影响，即相似组块和不相似组块对相似词记忆的影响，但是没有介绍这两种组块方式加工的特点及差异。建议详细介绍和比较，为结果中得到的不相似组块优势提供依据。

**回应:** 非常感谢审稿专家的意见。在“1 前言”部分的第二段中，我们增加了两种组块方式加工的特点及差异。具体的文字表述为“已有研究发现组块策略可以影响词汇的学习与记忆(Gilbert, Boucher, & Jemel, 2014)。组块策略是指在记忆过程中按照相似性或其它原则对学习材料进行组块编码。就学习材料的相似性而言，组块策略可以分为相似材料组块和不相似材料组块。相似材料组块指将视觉或语音相似的材料进行组块编码，而不相似材料组块指将视觉或语音不相似的材料进行组块编码。与不相似材料组块相比，相似材料组块中相似的学习材料相继进行记忆编码，相关脑区的活动会表现出抑制效应，即对相似材料的反应强度降低(Glezer, Jiang, & Riesenhuber, 2009; Glezer, Kim, Rule, Jiang, & Riesenhuber, 2015)。换句话说，与不相似材料组块相比，学习者在相似材料组块中投入的认知资源会减少，因而可能表现出较低的记忆成绩。”

**参考文献:**

- Gilbert, A. C., Boucher, V. J., & Jemel, B. (2014). Perceptual chunking and its effect on memory in speech processing: ERP and behavioral evidence. *Front Psychol*, 5, 220. doi:10.3389/fpsyg.2014.00220
- Glezer, L. S., Jiang, X., & Riesenhuber, M. (2009). Evidence for highly selective neuronal tuning to whole words in the "visual word form area". *Neuron*, 62(2), 199.
- Glezer, L. S., Jiang, X., & Riesenhuber, M. (2009). Evidence for highly selective neuronal tuning to whole words in the "visual word form area". *Neuron*, 62(2), 199-204. doi:10.1016/j.neuron.2009.03.017
- Glezer, L. S., Kim, J., Rule, J., Jiang, X., & Riesenhuber, M. (2015). Adding words to the brain's visual dictionary: novel word learning selectively sharpens orthographic representations in the VWFA. *J Neurosci*, 35(12), 4965-4972. doi:10.1523/JNEUROSCI.4031-14.2015

**意见 3:** 实验 2 想要探讨不相似词组块的记忆优势，且应用了错误记忆范式，但未提及与相似组块相比，不相似组块存在什么记忆优势，以及如何用错误记忆范式探讨不相似组块的记忆优势。建议引用前人文献，仔细说明这一过程。

**回应:** 非常感谢审稿专家提出的重要建议。实验 1 结果发现，与相似词组块相比，不相似词组块能够提高陌生相似词的记忆。模糊痕迹理论(Fuzzy Trace Theory)认为，再认判断主要基于两类记忆表征：一类是在学习中对项目之间共同部分加工而形成的较笼统的记忆表征；另一类是对各个细节特征进行加工而形成的较明确的记忆表征(Reyna & Brainerd, 1995)。并且，这两种记忆表征在错误记忆中具有相反的作用，即前者会提高与学习材料相似的项目的错误记忆，而后者会降低错误记忆(Gallo, 2010; 肖红蕊, 黄一帆, 龚先旻, 王大华, 2015)。依据模糊痕迹理论，不相似词组块对相似词记忆的促进作用既可能是因为增强了相似词共同词根的笼统记忆，也可能是因为提高了单个词汇的细节性信息的记忆。

为了区分上述两种可能性，在实验 1 的基础上，实验 2 使用错误记忆范式，在再认记忆测试阶段加入与学习材料相似的诱饵刺激，对不相似词组块的促进效应进行探讨。如果不相似词组块的记忆优势是因为不相似词组块促进了相似词表共同词根的笼统记忆，那么不相似词组块条件下诱饵刺激的错误记忆率高于相似词组块条件；如果不相似词组块的记忆优势是因为不相似词组块促进了单个词汇独特的细节性信息的记忆，那么不相似词组块条件的错误记忆率更低，或两种条件的错误记忆率没有差异。

在论文前言部分的最后一段（论文第 3 页）和实验 1B 的讨论部分（论文第 8 页），我们对上述实验逻辑进行了阐述。

#### 参考文献:

- Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: an interim synthesis. *Learning & Individual Differences*, 7(1), 1-75.
- Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition*, 38(7), 833-848.
- 肖红蕊, 黄一帆, 龚先旻, & 王大华. (2015). 简化的联合再认范式中情绪对错误记忆影响的年龄差异. *心理学报*, 47(1), 19-28.

#### 意见 4: 缺乏对实验结果的预期。

**回应:** 非常感谢审稿专家的意见。根据审稿意见,我们在前言最后一段中增加了对实验 1 的结果预期(论文第 3 页)。具体的文字表述为“依据材料相似性影响短时记忆的研究发现(Avons, 1999; Logie et al., 2000; Logie et al., 2016; Saito et al., 2008; Smyth et al., 2005),预期不相似词组块相比于相似词组块能够促进相似词的记忆”。

实验 2 和实验 3 的结果相对比较开放,没有非常明确的结果预期。因此,我们在前言部分最后一段中增加了对实验结果的各种可能性的分析(论文第 3 页)。具体来说,实验 2 增加的文字为“模糊痕迹理论(Fuzzy Trace Theory)认为,再认判断主要基于两类记忆表征:一类是在学习中对项目之间共同部分加工而形成的较笼统的记忆表征;另一类是对各个细节特征进行加工而形成的较明确的记忆表征(Reyna & Brainerd, 1995)。依据模糊痕迹理论,不相似词组块对相似词记忆的促进作用既可能是因为增强了相似词共同词根的笼统记忆,也可能是因为提高了单个词汇的细节性信息的记忆。由于共同词根的笼统记忆表征的增强会提高错误记忆,而单个词汇的细节性信息的记忆表征增强会降低错误记忆,所以使用错误记忆范式可以有效区分上述两种可能性(Gallo, 2010; 肖红蕊,黄一帆,龚先旻,王大华, 2015)。如果不相似词组块的记忆优势是因为不相似词组块促进了相似词表共同词根的笼统记忆,那么不相似词组块条件下诱饵刺激的错误记忆率高于相似词组块条件;如果不相似词组块的记忆优势是因为不相似词组块促进了单个词汇独特的细节性信息的记忆,那么不相似词组块条件的错误记忆率更低,或两种条件的错误记忆率没有差异。”实验 3 增加的文字为“如果学习材料组块方式效应依赖于语音相似性,那么相似韩字的学习不会表现出不相似词组块的记忆优势;反之,则会表现出不相似词组块的记忆优势。”

#### 参考文献:

- Avons, S. E. (1999). Effects of Visual Similarity on Serial Report and Item Recognition. *Quarterly Journal of Experimental Psychology*, 52(1), 217-240.
- Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition*, 38(7), 833-848.
- Logie, R. H., Della Sala, S., Wynn, V., & Baddeley, A. D. (2000). Visual similarity effects in immediate verbal serial recall. *Q J Exp Psychol A*, 53(3), 626-646.
- Logie, R. H., Saito, S., Morita, A., Varma, S., & Norris, D. (2016). Recalling visual serial order for verbal sequences. *Mem Cognit*, 44(4), 590-607. doi:10.3758/s13421-015-0580-9
- Saito, S., Logie, R. H., Morita, A., & Law, A. (2008). Visual and phonological similarity effects in verbal immediate serial recall: A test with kanji materials. *Journal of Memory and Language*, 59(1), 1-17. doi:10.1016/j.jml.2008.01.004
- Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: an interim synthesis. *Learning & Individual Differences*, 7(1), 1-75.
- Smyth, M. M., Hay, D. C., Hitch, G. J., & Horton, N. J. (2005). Serial position memory in the visual-spatial domain: reconstructing sequences of unfamiliar faces. *Quarterly Journal of Experimental Psychology*, 58(5), 909-930.
- 肖红蕊, 黄一帆, 龚先旻, & 王大华. (2015). 简化的联合再认范式中情绪对错误记忆影响的年龄差异. *心理学报*, 47(1), 19-28.

**意见 5:** 作者在实验 1B 采用与实验 1A 相同的实验方法, 但将学习和测查之间的时间间隔延长至一周, 以探究学习材料组块方式对相似词记忆的影响是否能长时保持。但是, 为什么实验 2 中仅设置学测间隔为一周, 实验 3 中仅设置学测间隔为一小时。这样看来, 学测间隔会成为影响记忆成绩的潜在因素, 从而干扰组块方式的效应。

**回应:** 非常感谢审稿专家提出的重要意见。由于实验 1 结果发现无论是间隔 1 小时还是 1 周, 不相似词组块对相似词记忆的促进效应都是稳定存在的, 所以实验 2 和实验 3 只设置了一种学习-测试间隔。实验 2 的学习-测查间隔设置为 1 周的主要原因是实验 1 中间隔 1 小时的再认记忆测试的成绩偏高 (0.82), 而间隔 1 周的成绩适中 (0.70)。而过高或过低的记忆成绩都可能减弱两种组块条件之间的差异。实验 3 的学习-测查间隔设置为 1 小时主要是因为韩字的记忆难度较大, 较短的学习-测查间隔可以防止地板效应的发生。并且, 1 小时的学习-测查间隔可以与实验 1A 的结果进行比较。我们已在实验 2 和实验 3 的实验程序部分对学习-测查间隔的选择原因进行了说明 (论文第 9 页和第 13 页)。

尽管如此，正如审稿专家所言，学习-测查间隔的差异仍存在影响实验结果的可能性。因此，在总讨论的倒数第二段，我们将三个实验的学习-测查间隔的差异作为本研究的一个局限进行讨论（论文第 16 页）。具体的文字表述为“值得注意的是，因为实验 1 发现在不同学习-测查间隔条件下（1 小时和 1 周），材料组块方式效应都稳定存在，所以在实验 2 和实验 3 中，只设置了一种学习-测查间隔条件。具体来说，实验 2 的学习-测查间隔为 1 周，实验 3 的学习-测查间隔为 1 小时。虽然实验 1 的结果说明材料组块方式效应不会受到学习-测查间隔的影响，但是三个实验在学习-测查间隔上差异仍存在影响实验结果的可能性。未来研究需要使用完全一致的学习-测查间隔，进一步验证该研究的结果。”

## 二、文章部分内容存在逻辑跳跃

**意见 6:** 在前言中，作者引用三个前人研究对比相似项目和不相似项目的记忆，结果发现不相似材料的记忆优于相似材料。然后就得出“学习材料的组块方式会影响词汇的记忆”这一结论。作者直接从材料类型过度到组块方式，衔接较为突兀，建议作者先阐明相似材料通过相似词组块的方式记忆，不相似材料通过不相似词组块的方式记忆，然后再过度到这一结论。

**回应:** 非常感谢审稿专家的宝贵意见。根据修改意见，我们对材料类型到组块方式的衔接进行了修改，并明确说明了相似材料通过相似材料组块方式记忆，而不相似材料通过不相似材料组块方式记忆（论文第 2 页）。具体的文字表述修改为“在上述研究中，相似项目序列的学习材料以相似材料组块方式呈现，即一组视觉相似的学习材料相继呈现；而不相似项目序列的学习材料以不相似材料组块方式呈现，即一组视觉不相似的学习材料相继呈现。因此，这些研究提示，与相似材料组块相比，不相似材料组块可能促进词汇的记忆。尽管如此，由于前人研究中不相似材料组块的记忆优势都是基于两组不同学习材料（相似材料和不相似材料）的记忆成绩对比而发现的，所以其研究结果既可能反映了不相似材料组块的记忆优势，也可能反映了两组实验材料的记忆难度差异。具体来说，与视觉不相似的学习材料相比，视觉相似的学习材料之间容易相互混淆，因而记忆难度更大(Nairne, 1990)。为了排除实验材料记忆难度的差异，本研究拟使用两组记忆难度匹配的相似词为材料，探讨学习材料组块方式（相似词组块和不相似词组块）对相似词记忆的影响及作用机制。这是本研究拟解决的第一个研究问题。”

**意见 7:** 在 4.3 的结果中通过相关分析发现，学习材料击中率的差值和诱饵词错误记忆率的

差值正相关。然后就得出“不相似词组块方式对相似词记忆效果的提升是以等量诱饵词的错误记忆率的提高为代价”这一推论。然后在 4.4 的讨论中进一步推论“不相似词组块方式对相似词记忆效果的提升并不是通过增强单个学习材料的特异性记忆，而是通过增强相似词组的共同词根的记忆而实现的”。这些推论在逻辑上具有跳跃性。而且通过相关分析推测增强相似词组的共同词根促进相似词的记忆，也无法排除单个学习材料的特异性记忆发挥的作用。作者需要解释清楚错误记忆率与相似词记忆成绩以及共同词根的记忆间的关系。同时在推论的过程中一定要严谨，避免过度推论或夸大结果。

**回应：**非常感谢审稿专家提出的重要意见。根据修改意见，我们进行了如下几个方面的修改：

首先，在“4.3 实验结果”部分，删除了学习材料击中率和诱饵词虚报率的相关分析，同时也删除了关于“不相似词组块方式对相似词记忆效果的提升是以等量诱饵词的错误记忆率的提高为代价”的推论。在实验结果和讨论部分只保留了“不相似词组块方式在促进相似词记忆效果的同时，提高了诱饵词的错误记忆率”的推论。

其次，在前言（论文第 3 页）、实验 1B 的讨论（论文第 8 页）、实验 2 的讨论中（论文第 11 页），对错误记忆率的提升与相似词共同词根的记忆间关系的逻辑进行了充分的阐述。具体来说，模糊痕迹理论(Fuzzy Trace Theory)认为，再认判断主要基于两类记忆表征：一类是在学习中对项目之间共同部分加工而形成的较笼统的记忆表征；另一类是对各个细节特征进行加工而形成的较明确的记忆表征(Reyna & Brainerd, 1995)。并且，这两种记忆表征在错误记忆中具有相反的作用，即前者会提高与学习材料相似的项目的错误记忆，而后者会降低错误记忆(Gallo, 2010; 肖红蕊,黄一帆,龚先旻,王大华, 2015)。依据模糊痕迹理论，不相似词组块对相似词记忆的促进作用既可能是因为增强了相似词共同词根的笼统记忆，也可能是因为提高了单个词汇的细节性信息的记忆。

为了区分上述两种可能性，实验 2 使用错误记忆范式，在再认记忆测试阶段加入与学习材料相似的诱饵刺激，对不相似词组块的促进效应进行探讨。如果不相似词组块的记忆优势是因为不相似词组块促进了相似词表共同词根的笼统记忆，那么不相似词组块条件下诱饵刺激的错误记忆率高于相似词组块条件；如果不相似词组块的记忆优势是因为不相似词组块促进了单个词汇独特的细节性信息的记忆，那么不相似词组块条件的错误记忆率更低，或两种条件的错误记忆率没有差异。实验 2 结果发现不相似词组块条件下诱饵刺激的错误记忆率高于相似词组块条件。该结果支持“不相似词组块对相似词记忆的促进作用是通过增强相似词

共同词根的记忆而实现的”假设。

### 参考文献:

Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition*, 38(7), 833-848.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: an interim synthesis. *Learning & Individual Differences*, 7(1), 1-75.

肖红蕊, 黄一帆, 龚先旻, & 王大华. (2015). 简化的联合再认范式中情绪对错误记忆影响的年龄差异. *心理学报*, 47(1), 19-28.

### 三、讨论不够深入

**意见 8:** 实验 2 的结果说明不相似组块方式对相似词记忆效果的提高是通过增强相似词组的共同词根实现的。但是对于这一结果在总讨论和 4.4 中的讨论内容几乎一模一样，缺乏深入讨论。建议结合相似组块和不相似组块的加工特点，深入探讨不相似组块的记忆优势。

**回应:** 非常感谢审稿专家的宝贵意见。根据修改意见，在“6 总讨论”部分的第三段，依据模糊记忆理论的观点对组块方式效应的产生机制进行了深入讨论（论文第 15-16 页）。同时，增加了关于相似词连续呈现的脑成像研究证据，进一步从脑机制上对不相似词组块的促进效应进行了讨论。本部分讨论的内容修改为“更为重要的是，本研究还探明了不相似词组块方式对相似词记忆的促进作用的机制。本研究发现相比于相似词组块方式，不相似词组块方式对相似词的记忆具有促进作用。如前所述，模糊痕迹理论认为再认记忆判断主要依赖于两类记忆表征：对学习材料共同部分概括、笼统的记忆表征和对学习材料各个细节特征明确、清晰的记忆表征(Reyna & Brainerd, 1995)。这两种记忆表征在错误记忆中具有相反的作用，即前者会提高与学习材料相似的项目的错误记忆，而后者会降低错误记忆(Gallo, 2010; 肖红蕊,黄一帆,龚先旻,王大华, 2015)。根据模糊痕迹理论，本研究发现的不相似词组块的记忆优势既可能是因为不相似词组块方式增强了相似词中共同词根的笼统记忆，也可能是因为不相似词组块方式提高了单个词汇的细节性信息记忆。本研究的实验 2 使用错误记忆范式对上述两种可能性进行了验证。(Roediger & Mcdermott, 1995; Ye et al., 2016)。实验结果发现，不相似词组块方式在促进相似词记忆的同时也提高了相似诱饵刺激的错误记忆。该结果说明，不相似词组块方式对相似词记忆效果的提升并不是通过增强单个学习材料的细节性信息

的记忆，而是通过增强相似词组的共同词根的笼统记忆而实现的。与此一致，脑成像研究发现相似陌生词汇的连续呈现会引起脑活动的抑制效应，即对相似词汇的反应强度降低(Glezer et al., 2009; Glezer et al., 2015)。换句话说，与不相似词组块相比，学习者在相似词组块条件下对相似词汇的共同词根投入的加工资源会减少，即对相似词共同词根的记忆会减弱，因而表现出较低的记忆成绩。”

#### 参考文献:

- Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition*, 38(7), 833-848.
- Glezer, L. S., Jiang, X., & Riesenhuber, M. (2009). Evidence for highly selective neuronal tuning to whole words in the "visual word form area". *Neuron*, 62(2), 199-204. doi:10.1016/j.neuron.2009.03.017
- Glezer, L. S., Kim, J., Rule, J., Jiang, X., & Riesenhuber, M. (2015). Adding words to the brain's visual dictionary: novel word learning selectively sharpens orthographic representations in the VWFA. *J Neurosci*, 35(12), 4965-4972. doi:10.1523/JNEUROSCI.4031-14.2015
- Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: an interim synthesis. *Learning & Individual Differences*, 7(1), 1-75.
- Roediger, H. L. & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology Learning Memory & Cognition*, 21(4), 803-814.
- Ye, Z., Zhu, B., Zhuang, L., Lu, Z., Chen, C., & Xue, G. (2016). Neural Global Pattern Similarity Underlies True and False Memories. *J Neurosci*, 36(25), 6792-6802. doi:10.1523/JNEUROSCI.0425-16.2016
- 肖红蕊, 黄一帆, 龚先旻, 王大华. (2015). 简化的联合再认范式中情绪对错误记忆影响的年龄差异. *心理学报*, 47(1), 19-28.

**意见 9:** 实验 3 得到对陌生词的记忆，语音相似性比字形相似性更重要。在总讨论中也只是重复叙述了前面得到的结果，缺乏深入讨论出现这一结果的原因。且对于这一结果，前人研究中是存在争议的，建议作者解释出现分歧的原因，进一步为实验结果提供依据。

**回应:** 非常感谢审稿人的宝贵意见。根据修改意见，在“6 总讨论”部分第二段中，从两个方面对本研究没有发现视觉相似性效应的原因进行了讨论（论文第 15 页）。一方面，视觉相似性的作用可能会受到词汇相似性程度的调节。已有研究发现，当词表内词汇之间的相似性较低时（词表内部分词汇相似），语音相似性和视觉相似性在词汇记忆中都会起作用

(Lin et al., 2015; Saito et al., 2008); 而当词汇之间的相似性较高时(词表内所有词汇都相似), 语音相似性对词汇记忆起主要作用, 视觉相似性的作用不明显(李轩,刘思耘, 2012)。由于本研究选用的词表内所有词汇之间都相似, 即相似性程度较高, 所以没有发现视觉相似性的作用。另一方面, 由于本研究中的陌生韩字材料不具有语音, 所以无法排除视觉相似性在不相似词组块效应中单独起作用的可能性, 即无法排除不相似词组块促进视觉相似而语音不相似词记忆的可能性。未来研究需要加入字形相似而语音不相似的陌生词汇作为学习材料, 对上述可能性进行进一步探讨。

#### 参考文献:

- Lin, Y. C., Chen, H. Y., Lai, Y. C., & Wu, D. H. (2015). Phonological similarity and orthographic similarity affect probed serial recall of Chinese characters. *Mem Cognit*, 43(3), 538-554. doi:10.3758/s13421-014-0495-x
- Saito, S., Logie, R. H., Morita, A., & Law, A. (2008). Visual and phonological similarity effects in verbal immediate serial recall: A test with kanji materials. *Journal of Memory and Language*, 59(1), 1-17. doi:10.1016/j.jml.2008.01.004
- 李轩,刘思耘. (2012). 汉语短时序列回忆中的语音相似性和视觉相似性效应. *心理学报*, 44 (12), 1571-1582.

#### 四、文章中部分表述缺乏依据

**意见 10:** 文章中部分表述缺乏依据。如讨论中提到的“不相似词组块方式的促进作用一方面可能是因为不相似词组块方式增强了单个学习材料的特异性记忆, 进而促进了相似词的记忆; 另一方面也可能是因为不相似词组块方式增强相似词中共词根的记忆, 进而促进了相似词的记忆,” 缺乏前人研究的支持, 建议参考相关文献。

**回应:** 非常感谢审稿专家的宝贵意见。在修改后的论文中, 我们依据模糊痕迹理论的观点提出不相似词组块促进相似词记忆的两种可能机制。具体实验逻辑请参见对您意见 3 和意见 7 的回复。

#### 五、方法部分问题

**意见 11:** 被试量是如何确定的? 被试量偏少, 建议根据前人研究的效应量以及本研究想要达到的 power 值确定被试人数。

**回应：**非常感谢审稿专家提出的重要意见。本研究四个实验的被试量分别为 22、22、22 和 26。被试量的确定有两方面依据。一方面，以往类似研究的被试量为 16-24 人(李轩, 刘思耘, 2012; Lin et al., 2015; Saito et al., 2008), 本实验的被试量与这些研究相当。另一方面, 一般认为, 理想的统计检验力需高于 0.8 (Cohen, 1988)。如下图所示, 通过 G-power 软件 (<http://www.gpower.hhu.de/>) 计算统计检验力 0.8 所对应的被试量为 15 人。因此, 本实验的被试量具有足够的统计检验力。我们已在实验 1A 的被试部分详细说明了上述依据 (论文第 4 页)。

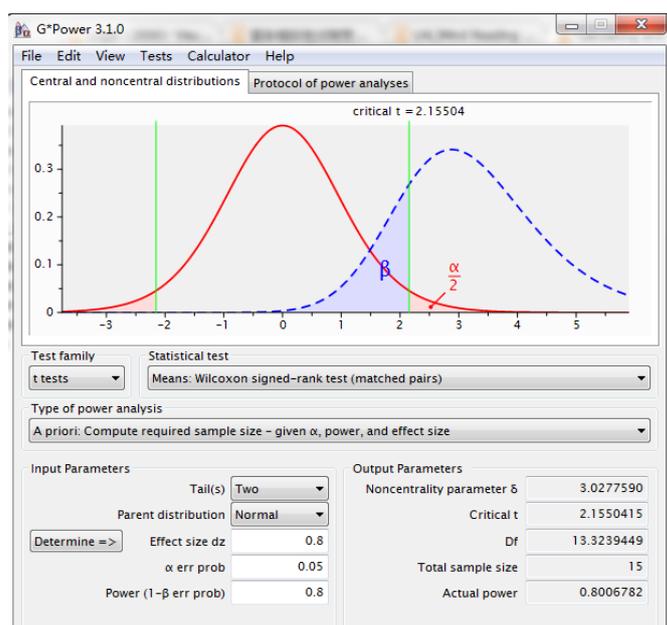


图 1 预期被试样本结果

#### 参考文献：

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. 2nd ed: L. Erlbaum Associates.
- Lin, Y. C., Chen, H. Y., Lai, Y. C., & Wu, D. H. (2015). Phonological similarity and orthographic similarity affect probed serial recall of Chinese characters. *Memory & Cognition*, 43(3), 538-554. doi:10.3758/s13421-014-0495-x.
- Saito, S., Logie, R. H., Morita, A., & Law, A. (2008). Visual and phonological similarity effects in verbal immediate serial recall: A test with kanji materials. *Journal of Memory and Language*, 59(1), 1-17. doi:10.1016/j.jml.2008.01.004
- 李轩, 刘思耘. (2012). 汉语短时序列回忆中的语音相似性和视觉相似性效应. *心理学报*, 44 (12), 1571-1582.

**意见 12：**在实验 1B 中只有 2 名男生，在本实验中，性别是否会对实验结果造成影响？

**回应：**非常感谢审稿专家提出的宝贵意见。正如审稿专家所言，实验 1B 中男性被试偏少，这主要是因为师范大学的男生比例较低，而且参与行为实验的积极性不高。为了确认性别是否会对实验结果造成影响，我们分别计算了实验 1A、实验 2 和实验 3 中男生和女生在两种实验条件的击中率。结果发现，男生和女生的数据模式类似，即实验 1A 和 2 中都表现出不相似词组块的记忆优势，而在实验 3 中男生和女生都没有明显的组块方式效应（见下表）。基于上述结果，我们认为本研究的实验结果不太可能受到性别的影响。

**表 1 实验 1A、2 和 3 中不同性别在两种条件下的平均击中率（标准差）**

	男生		女生	
	相似组块	不相似组块	相似组块	不相似组块
实验 1A	0.76 (0.12)	0.82 (0.08)	0.81 (0.14)	0.88 (0.11)
实验 2	0.61 (0.17)	0.66 (0.18)	0.65 (0.15)	0.74 (0.16)
实验 3	0.87 (0.11)	0.88 (0.11)	0.75 (0.15)	0.77 (0.14)

**意见 13：**为了检验同质性，选取 9 名被试进行记忆，结果没差异。没差异是否是由于被试量太少的原因。建议增加被试数或选取更精确的检验方法。

**回应：**非常感谢审稿专家提出的宝贵建议。为了检验两组学习材料的同质性，我们在主实验之前招募了 9 名被试参与预实验。由于被试量较小，所以在进行基于被试的分析同时增加了基于项目的分析，即将每个学习材料 9 名被试的得分进行平均，然后对两组学习材料的差异进行检验。与基于被试的分析相比，基于项目的分析对于实验材料的变异性更敏感。结果发现，无论是基于被试的分析还是基于项目的分析都发现两组学习材料之间没有差异（基于被试的分析： $F(1,8) = 0.30, p = 0.602$ ；基于项目的分析： $F(1,110) = 0.30, p = 0.585$ ）。该结果说明两组材料之间是同质的。我们已将上述分析结果补充到实验 1 的实验材料部分（论文第 4 页）。为了进一步排除学习材料的差异，我们在实验中对两种组块条件的学习材料进行了被试间平衡。因此，本研究的实验结果应该不会受到学习材料差异的影响。

**意见 14：**在实验程序中介绍“每个刺激呈现 2 遍，两次呈现之间间隔 1~5 个其它刺激”两次呈现之间是间隔 1-5 个刺激还是组块？如果是刺激，那相似组块组的 4 个相似词之间会被其他词断开吗？

**回应：**非常抱歉，该部分文字表述不清。实验程序中介绍的“1-5 个其它刺激”指的是刺激，

相似词组块内的 4 个相似词之间是连续呈现的，不会间隔其它词。为了避免误解，我们已将相关描述修改为“在实验中，每个组块内的四个英文非词连续呈现，每个英文非词呈现 3 s，之后呈现 1~3 s（平均 2 s）的注视点。被试的任务是努力记住呈现的每一个英文非词。每个组块连续呈现两次，但两次呈现时组块内英文非词的呈现顺序不同，同一英文非词两次呈现之间间隔 1~5 个其它英文非词（平均 3 个），两种组块条件的重复间隔匹配。”

## 六、 语言表达问题

**意见 15:** 4.4 的讨论中“我们如果不相似词组块方式的促进效应是因为增强了单个学习材料的特异性记忆，那么其错误记忆率应该与相似词组块方式没有差异，甚至更低”多了一个“我们”。类似问题请自查。

**回应:** 非常感谢审稿人提出的意见，已对正文的所有文字内容进行检查，并且进行了相应的修改。例如在“4.4 讨论”部分，文字表述修改为“如果不相似词组块方式的促进效应是因为增强了相似词共同词根的笼统记忆，那么其错误记忆率应该高于相似词组块方式；如果促进效应是因为增强了单个词汇的独特细节性信息的记忆，那么其错误记忆率应该与相似词组块方式没有差异，甚至更低。”

---

## 第二轮

### 审稿人 1 意见:

**意见 1:** 部分内容叙述重复冗余。如 2.4 讨论部分，第一段的最后一句和第二段的第一句表述的是同一内容，建议删去一处。又如用错误记忆范式研究不相似组块的加工优势的原因及实验假设，作者在引言、4.4 的讨论以及总讨论中反复提及，建议作者一次表述清晰，后面概括性提及既可。

**回应:** 非常感谢审稿专家提出的宝贵意见。根据修改意见，我们进行了如下几点的修改:

首先，删除了“2.4 讨论”部分第二段的第一句表述，“2.4 讨论”部分第二段的具体文字表述修改为“实验 1A 的结果首先重复了前人研究关于不相似词组块的记忆优势的发现 (Avons, 1999; Saito et al., 2008)。更为重要的是，与前人研究通过两组不同学习材料（相似学习材料和不相似学习材料）的比较而发现的不相似词的记忆优势不同，本研究发现，当两组学习材料同为相似词时，不相似词组块方式也能促进相似词的记忆。该结果说明不相似词组块可能是促进相似词汇学习和记忆的一种有效途径。”（论文第 8 页）。

其次，对 4.4 讨论和总讨论中关于错误记忆的实验逻辑文字表述进行了修改，使用概括性语句代替了与前言中重复的内容。具体而言，在“4.4 讨论”部分的第二段，文字表述修改为“实验 2 的结果首先重复了实验 1 的结果。具体而言，与实验 1 的结果一致，实验 2 发现与相似词组块方式相比，不相似词组块方式促进了相似词的记忆效果，并且这种优势可以长期保持。更为重要的是，实验 2 通过使用错误记忆范式，加入与学习材料非常相似的诱饵刺激，对不相似词组块方式对相似词记忆的促进机制进行了探究。结果发现，相比于相似词组块条件，不相似词组块方式同时提升了学习材料的击中率和相似诱饵词的错误记忆率，且不相似词组块方式对诱饵词的错误记忆率提升是由于个体区分旧词和诱饵词的判断标准降低引起的，即在不相似组块条件下更倾向于将诱饵词判断为旧词。如前言所述，依据模糊痕迹理论(Reyna & Brainerd, 1995)，实验 2 的结果直接说明不相似词组块方式对相似词记忆的促进效应是通过增强相似词组的共同词根的笼统记忆而实现的。”(论文第 13 页)。在“6 总讨论”部分的第三段，具体文字表述修改为“更为重要的是，本研究还探明了不相似词组块方式对相似词记忆的促进作用的机制。本研究发现相比于相似词组块方式，不相似词组块方式对相似词的记忆具有促进作用。如前所述，依据模糊痕迹理论(Reyna & Brainerd, 1995)，本研究发现的不相似词组块相对于相似词组块的记忆优势既可能是因为不相似词组块方式增强了相似词中共同词根的笼统记忆，也可能是因为不相似词组块方式提高了单个词汇的细节性信息记忆。”(论文第 17 页)。

**意见 2：部分内容表述不清晰。** 前言的最后一段中，作者表述“依据模糊痕迹理论，不相似词组块对相似词记忆的促进作用既可能是因为增强了相似词共同词根的笼统记忆，也可能是因为提高了单个词汇的细节性信息的记忆”，不相似词组块对相似词的促进作用是相对于相似词组块而言的，无论是共同词根笼统记忆的提高还是细节性信息记忆的提高也都是相对于相似词组块而言，且最终需要对比不相似词组块和相似词组块的错误记忆虚报率，建议作者在此处及后文的表述中清楚的表达出这种比较关系。

**回应：**非常感谢审稿专家提出的重要意见。正如审稿专家所言，不相似词组块对相似词的促进作用是相对于相似词组块而言的。因此，根据修改建议，我们已对文中相应的表述进行了修改。例如，在“1 前言”最后一段，文字表述修改为“依据模糊痕迹理论，与相似词组块方式相比，不相似词组块对相似词记忆的促进作用既可能是因为增强了相似词共同词根的笼统记忆，也可能是因为提高了单个词汇的细节性信息的记忆。”(论文第 3 页)；在“3.4 讨论”部分第二段，部分文字表述修改为“如前所述，与相似词组块相比，不相似词组块对

相似词记忆的促进作用既可能通过增强单个词汇的特异性细节信息的记忆促进了相似词的记忆,也可能是通过增强相似词中共同词根的记忆而促进了相似词的记忆。”(论文第 10 页);在“6 讨论”部分第三段,部分文字表述修改为“如前所述,依据模糊痕迹理论(Reyna & Brainerd, 1995),本研究发现的不相似词组块相对于相似词组块的记忆优势既可能是因为不相似词组块方式增强了相似词中共同词根的笼统记忆,也可能是因为不相似词组块方式提高了单个词汇的细节性信息记忆。”(论文第 17 页)。

.....

**审稿人 2 意见:**

**意见 1:** 文中存在一些对引用文献的过度陈述问题。例如,作者在前言的第二段中提到:“相似材料组块指将视觉或语音相似的材料进行组块编码,而不相似材料组块指将视觉或语音不相似的材料进行组块编码。与不相似材料组块相比,相似材料组块中相似的学习材料相继进行记忆编码,相关脑区的活动会表现出抑制效应,即对相似材料的反应强度降低(Glezer, Jiang, & Riesenhuber, 2009; Glezer, Kim, Rule, Jiang, & Riesenhuber, 2015)。换句话说,与不相似材料组块相比,学习者在相似材料组块中投入的认知资源会减少,因而可能表现出较低的记忆成绩。”作者在讨论中又重复了上面问题 1 中所提到的这段陈述,两段陈述几乎完全相同。此处的“相关脑区的活动”是否能够与“投入的认知资源”划等号?请作者进一步思考,并认真检查全文,对类似的陈述进行修改。

**回应:** 非常感谢审稿专家提出的宝贵意见。我们同意审稿专家关于脑区活动强度与投入的认知资源不完全对等的意见。为了避免引用文献的过度陈述,我们删除了关于“投入的认知资源”的表述,代之以“脑区的参与程度”。具体来说,在“1 前言”部分第二段,相关文字表述修改为“与不相似材料组块相比,相似材料组块中相似的学习材料相继进行记忆编码,相关脑区的活动会表现出抑制效应,即对相似材料的反应强度降低(Glezer, Jiang, & Riesenhuber, 2009; Glezer, Kim, Rule, Jiang, & Riesenhuber, 2015)。换句话说,与不相似材料组块相比,学习者在相似材料组块条件下记忆编码相关脑区的参与强度较低,因而可能表现出较低的记忆成绩。”(论文第 1 页)。而在“6 总讨论”部分第三段,相关文字表述修改为“与此一致,脑成像研究发现相似陌生词汇的连续呈现会引起脑活动的抑制效应 (Glezer et al., 2009; Glezer et al., 2015)。因此,学习者在相似词组块条件下记忆编码脑区的活动强度低于不相似词组块条件。换句话说,与不相似词组块相比,学习者在相似词组块条件下会降低记忆编码脑区在共同词根编码时的参与程度,因而表现出较低的记忆成绩”(论文第 17 页)。

**意见 2:** 记忆包含了许多不同类型。根据测试任务要求不同,个体对记忆信息的提取方式也可分为不同类型。不同记忆类型和提取方式之间差别很大。作者认为四个实验中的检测是个体对相似词的记忆,这种陈述过于笼统。请作者阐明以下问题:研究中被试获得的记忆属于什么记忆类型?其信息提取是属于什么方式,例如是属于熟悉性还是回想?

**回应:** 非常感谢审稿专家提出的宝贵意见。正如审稿专家所言,记忆包含许多不同的类型。根据保持时间的长短,记忆可以分为感觉记忆、短时记忆和长时记忆。依据学习-测查的时间间隔(1小时和1周),本研究所探讨的记忆类型属于长时记忆。长时记忆又可进一步分为外显记忆和内隐记忆(Squire, 1996; 朱磊等, 2003)。本研究要求个体外显地记忆学习材料,因而学习材料的记忆包含外显记忆。同时,由于每个学习材料重复学习2次,所以材料的学习包含重复启动效应,因而也包含内隐记忆的成分(Turk-Browne et al., 2006)。本研究无法区分学习材料组块方式效应是发生在外显记忆层面还是发生在内隐记忆层面。为了避免结论扩大化,我们认为本研究发现的学习材料组块方式效应主要体现在相似词的长时记忆上。我们在文章标题、摘要和正文中明确说明了本研究主要探讨学习材料组块方式对相似词长时记忆的影响。

在信息提取方式方面,再认记忆的双加工理论(dual-process theory)认为再认记忆包含熟悉性(familiarity)和回想(recollection)两种独立的加工过程(Rugg, 2003; Wixted, 2010; Smith, 2011)。为了考察学习材料组块方式对熟悉性和回想过程的影响,我们增加了对两种组块条件的高、低自信击中率的分析。具体来说,高自信击中率是指对旧词进行6的按键反应的项目数与旧词数量的比值,低自信击中率是指对旧词进行4~5的按键反应的项目数与旧词数量的比值。一般认为,与低自信记忆项目相比,高自信记忆项目的提取包含更多回想过程(Squire et al., 2007)。因此,如果学习材料组块方式主要影响回想过程,那么只有高自信记忆得分会表现出组块方式效应;反之,则高低自信记忆得分都会表现出组块方式效应。实验1a、1b和2的结果都发现,与相似词组块条件相比,不相似词组块条件的高自信击中率更高,而两种组块条件在低自信击中率上没有显著差异(表3)。该结果说明,学习材料组块方式可能主要影响基于回想的再认记忆。我们已在论文的前言、方法、结果和讨论部分增加相关陈述(论文第3-4、6-9、11、15和17页)。尽管如此,由于高自信判断的项目的记忆提取同时包含熟悉性和回想两种过程,所以本研究通过高低自信记忆成绩的比较无法完全排除学习材料组块方式会影响基于熟悉性的再认记忆的可能性。因此,该结果还需进一步使用回忆与再认记忆的比较、记住/知道范式等多种分离熟悉性和回想的实验范式进行验证

(Squire et al., 2007; Smith et al., 2011)。我们已在讨论部分对此局限进行了讨论（论文第 17 页）。

表 3 再记忆测查中两种组块条件的高、低自信击中率（标准差）及 t 检验结果

	相似词组块	不相似词组块	<i>t</i>	<i>p</i>
实验 1A	2020.76 (757.13)	1908.26 (735.72)	1.55	0.136
实验 1B	2151.65 (781.03)	2128.97 (727.22)	0.33	0.747
实验 2	2386.15 (986.30)	2336.87 (925.58)	0.62	0.545
实验 3	2286.98 (803.35)	2441.74 (1200.42)	-0.84	0.409

### 参考文献：

- 朱磊, & 杨治良. (2003). 多种记忆分类之研究. *心理科学*, 26(4), 694-697.
- Rugg, M. D., & Yonelinas, A. P. (2003). Human recognition memory: a cognitive neuroscience perspective. *Trends in Cognitive Sciences*, 7(7), 313-319.
- Smith, C. N., Wixted, J. T., & Squire, L. R. (2011). The hippocampus supports both recollection and familiarity when memories are strong. *Journal of Neuroscience the Official Journal of the Society for Neuroscience*, 31(44), 15693.
- Squire, L. R., Wixted, J. T., & Clark, R. E. (2007). Recognition memory and the medial temporal lobe: a new perspective. *Nature Reviews Neuroscience*, 8, 872-883.
- Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences of the United States of America*, 93(24), 13515.
- Turk-Browne, N. B., Yi, D.-J., & Chun, M. M. (2006). Linking Implicit and Explicit Memory: Common Encoding Factors and Shared Representations. *Neuron*, 49(6), 917-927.
- Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of remember/know judgments. *Psychological Review*, 117(4), 1025-1054.

**意见 3：**作者在进行 **priori power analysis** 来确定样本量时，设定的效应量是多少？这并未在文章中报告。在对审稿专家 2 的意见回答中可以查看到，作者选取的效应量为 **0.8**，那么这个依据是什么？请在文中澄清。

**回应：**非常感谢审稿专家提出的建议。在使用 priori power analysis 来确定样本量时，设定的效应量大小为 0.8（见下图）。一方面，Cohen’s d 的评价标准为“小效应( $\geq 0.2$  且  $< 0.5$ )；中等效应( $\geq 0.5$  且  $< 0.8$ )；大效应( $\geq 0.8$ )” (Cohen, 1988)；另一方面，以往研究在确定被试人数时所采用的效应量为 0.8 (Mcgrath, 2006)。因此，设定的效应量为 0.8。我们已在“实验 1A”部分的“2.2.1 被试”中说明了上述依据（论文第 4 页）。

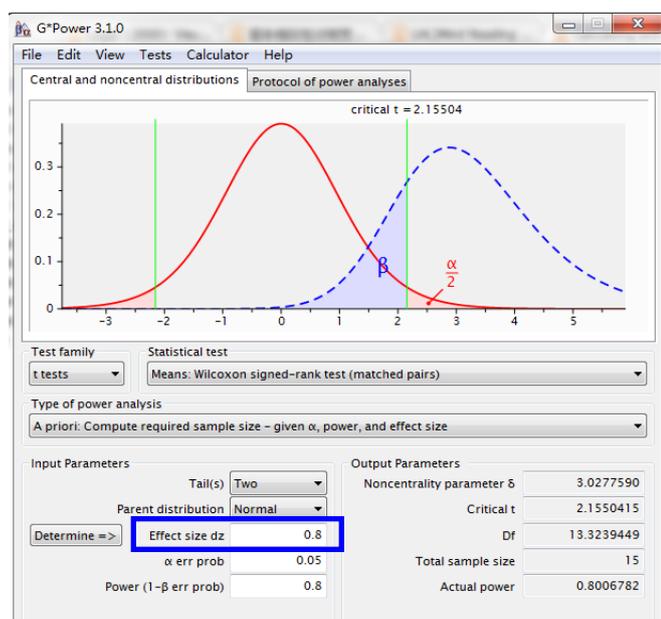


图 1 预期被试样本结果

#### 参考文献：

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. 2nd ed: L. Erlbaum Associates.
- Mcgrath, R. E., & Meyer, G. J. (2006). When effect sizes disagree: the case of r and d. *Psychol Methods*, 11(4), 386-401.

**意见 4：**作者均对四个实验的正确率以及辨别力进行了分析，然而并没有对被试接受测查时的反应时进行分析。那么，作者如何确定“速度-准确性”权衡没有对实验结果产生影响？

**回应：**非常感谢审稿专家提出的宝贵意见。为了确认“速度-准确性”权衡是否会对实验结果造成影响，我们分别计算了实验 1A、实验 1B、实验 2 和实验 3 中两种学习条件的反应时。结果发现，在反应时上，四个实验均没有发现明显的组块方式效应（见表 2，论文第 7 页）。基于上述结果，我们认为本研究的实验结果不太可能受到“速度-准确性”权衡的影响。我们已在 4 个实验的结果部分补充了相关分析（论文第 7、9、12 和 15 页）。

表 2 再记忆测查中两种组块条件的平均反应时（标准差）及 t 检验结果

	相似词组块	不相似词组块	<i>t</i>	<i>p</i>
实验 1A	2020.76 (757.13)	1908.26 (735.72)	1.55	0.136
实验 1B	2151.65 (781.03)	2128.97 (727.22)	0.33	0.747
实验 2	2386.15 (986.30)	2336.87 (925.58)	0.62	0.545
实验 3	2286.98 (803.35)	2441.74(1200.42)	-0.84	0.409

**意见 5:** 文中一些数据呈现的方式还可改善。例如，表 1 中的显著性 *p* 值一列均为 0.00，这种数据呈现没有实际意义，建议此处呈现效应量大小。

**回应:** 非常感谢审稿专家提出的宝贵意见。我们已对文中的表 1 进行了修改，具体如下：首先，删除了显著性 *p* 值这一列，并在 *t* 值右上方用“\*\*\*”表示“ $p < 0.001$ ”；其次，增加各个 *t* 检验的效应量大小，四个实验记忆正确率的效应量大小分别为 3.63、2.10、2.49 和 1.99（论文第 6 页）。

---

### 第三轮

#### 审稿人 2 意见:

基本解答了我的问题，可发表。

#### 编委复审

两位审稿人均为本领域的权威专家，同意他们的意见，建议修改后录用。

#### 主编终审

研究以英文非词和韩字为材料，采用学习-测验范式，通过操纵学习组块的方式，探讨学习组块方式在不同时间间隔下对相似词识记的影响及其内在机制。总体而言，研究具有较好的理论意义和一定实际应用价值，研究方法和过程符合科学研究标准。经过几轮修稿，文章在行文质量和规范上有较大提高，现已符合学报发表论文的相关要求，建议发表。