

《心理学报》审稿意见与作者回应

题目：《新颖词语义韵的发生机制：“双枣树”效应的证据》

作者：吴诗玉 李赞

第一轮

审稿人 1:

意见 1: 本研究通过双枣树效应探究新颖词学习过程中的语义韵产生机制，具有较高的理论价值，写作清晰流畅。在文章组织上有提升空间，另外在方法部分有几处需要澄清。以下是主要问题：

回应: 感谢专家的积极评价和宝贵的意见！

意见 2: 前言部分对外延义和语义韵的界定或许需要更清晰一些，或者两者能否清晰区分也是一个重要的理论问题。外延义是语言学成分，而语义韵是副语言成分，通过双枣树操纵对两者产生了不同的影响，是否也说明了两者的相对独立性呢？

回应: 感谢专家的意见，我们完全同意专家所说的“说明了两者的相对独立性”这一观点。根据专家的建议，我们查阅了大量相关文献，对“引言”的“1.1 语义韵”这一部分内容进行了仔细修改，既仔细修改了措辞，也补充了大段内容。对外延义和语义韵的界定，特别补充了以下内容：

“语义韵一词最早由语料库语言学家 Sinclair 创造，指的是看似中性的单词在频繁发生的搭配中所携带的消极或积极的情感内涵(Sinclair, 1987)。之后，语义韵这一话题吸引了广泛的关注，从概念上看先后经历了“语义传染说”“内涵义说”和“功能说”三种观点的演变(陆军, 2023)。“传染说”认为语义韵是“由搭配所赋予的一种非常一致的语义氛围”(Louw 1993, p. 157)，这个观点进一步拓展了之前 Firth (1957)提出的“词汇结伴说”，强调跨越单词界限的意义传递。“内涵义说”则认为语义韵是“超越单个词边界的内涵色彩的传播”(Partington, 1998, p. 68)，这一观点将传统的难以准确描述的内涵义具体化为有形的、可观察的搭配型式。而“功能说”则将语义韵视为“位于语义-语用连续统上语用一侧的意义”，传递了说话者态度和语用功能(Sinclair, 2004)。非常重要的是，在这里 Sinclair 将语义韵纳入扩展意义单位模型，由此，语义韵特征不再是之前零星、单一或特殊的语言现象，而是语言中普遍和大量存在的多词单位的必要属性。

尽管语义韵这一概念经历了上述演化、发展，但是这些概念无不强调语义韵相较于外延义的独立性，是一个单词从语境中“吸收的”并独立于其外延义的情感内涵。语义韵在话语交际中具有重要意义，这种“超越单个词边界的内涵色彩的传播”传达了说话者/作者在日常交流中的态度和判断(Nevisi et al., 2018)。举中文的“老”(lǎo)字为例。这个字在大部分时间都跟积极词搭配使用，比如老者、老人、老兄、老吴，等等，也因此在中国的语境里都“吸收了”一种积极的语义韵，尽管偶尔也跟消极词搭配使用(如“老东西”)。在中文的话语交际里对一个上了年纪的人说“老人家，让我来帮您拿行李吧”，或者说“老人家，您坐吧”，是非常自然的，但是，在英语国家里，同样的表达(如 Old man, let me help you with the luggage)却会严重冒犯话语接收者，甚至引起他的愤怒，这是因为在英语的话语交际里 old 这个词具

有消极语义韵，有一种“老而无用”的情感内涵。所以，英语世界里一般用 *senior citizens* 来称呼老年群体。

学术界尤其是语料库语言学研究者基于语言现象和事实，对语义韵这一话题开展了较为广泛的研究(甄凤超, 2023)。然而，尽管在概念内涵上学者们都强调语义韵相较于外延义的独立性，但是关于语义韵的发生机制以及其心理现实性却极其缺乏基于实验的证据，很少看到研究者通过实验的方法对这些问题进行探索。”

另外，在文章的讨论部分，我们也增加了内容：

……我们在前文对语义韵这一概念的历史演化、发展进行了较为详细地的介绍，这些概念都从理论的视角强调了语义韵相较于外延义的独立性，本研究通过汉语新颖词学习实验进一步证明了语义韵的心理现实性以及其发生机制中的“双枣树效应”……

意见 3：前言部分提出的两个研究问题分别对应的理论问题是什么？

回应：在文献综述的基础上，我们提出了以下两个研究问题：

- (1) 语境情感是否可通过阅读接触，从语境迁移到单词？
- (2) 语境的情感是否影响单词习得的效果？

第一个问题对应文献综述中介绍的“具身理论”(embodied account)(Barsalou et al., 2008)，而第二个研究问题对应文献综述 1.2 部分介绍的情感与语言学习的交互关系的理论论述(见 Barsalou, 2009; Kanazawa, 2016; 郭晓栋等, 2023)，以及自动警觉假说(automatic vigilance hypothesis)(Sneffjella et al., 2020)。

意见 4：前言中语义韵的例子均为副词或形容词，而研究中均采用了名词，而且是实物名词，这有没有可能引入无关的抽象性变量？如果有这种可能性，该如何影响结果解读？

回应：感谢专家宝贵的意见！我们在做实验时确实没有考虑到词性差别的问题。但就语义韵这个问题来看，根据学者们提供的定义，它并没有词性的差别。我们在论文中举的是副词或形容词的例子，但“语义韵”一词的创造者语料库语言学家 Sinclair 最早给的例子正好就是名词(短语)，Sinclair 提出：语义韵指的是看似中性的单词在频繁发生的搭配中所携带的消极或积极的情感内涵。他举名词(短语)true feeling 为例，这个词经常跟 scared, conceal 以及 reluctant 等词搭配(scared to show his true feelings; reluctant to show his true feelings)而使其获得一种消极语义韵。

我们同意专家的观点，要关注实验材料测试的都是名词这个问题，但是我们认为这篇论文的实验材料选用的均为名词是实验抽样的结果。毕竟，我们无法在一次实验里考察太多的问题或者控制太多的因素，但是，我们认为本实验的研究结论也适用于其他所有词性，即“在重复阅读带有情感负荷的阅读材料后，语境情感会迁移到词形上，从而使得这个单词获得某种语义韵”。不过，根据审稿人的意见，我们在研究展望中(4.讨论部分的结尾处)指出未来研究可对其他词性进行考察：

“……另外，从实验材料看，本研究只考察了名词，未来研究可以加入对更多词性的考察……”

意见 5: 是否有平衡不同条件的字频? 如果某些条件的汉字字形或发音过于生僻, 可能会对学习记忆产生实验目的无关的影响, 尤其是词汇测试包含了正字法任务。

回应: 我们尽可能地平衡了各种实验条件, 包括审稿专家提出的字频、字形等。就像在“2.2.1 阅读材料”所介绍的:

“通过平衡每一个新词出现的语境情感(3 种情感中的一种)和要表达的外延意义(9 个外延意中的一个), 一共生成了 9 套材料, 在这 9 套材料里确保每一个新词出现在了每一种语境情感里, 并表达了每一种外延意义, 但在同一套材料里, 每一个新词只能出现在同一种情感语境, 表达同一种外延意义。”

也就是说, 每一个词表达了每一个外延义, 出现在了每一种语境情感和语境变异性里, 因此不会出现审稿专家所提出的问题。

意见 6: 词汇评分包含了 4 个任务, 其中只有单词情感效价评分针对语义韵效应, 另外三个则针对外延义学习, 这一层关系能否在问题提出和实验设计中更清楚地展现?

回应: 谢谢专家的意见。我们在实验设计中的“2.2.3 词汇测试”部分对此进行了介绍:

“一共设计了四种词汇测试任务: 单词情感效价评分、单词正字法选择任务(orthographic choice test)、单词定义生成任务和单词定义匹配任务。单词情感效价评分任务主要用于回答第一个研究问题, 即在经过阅读接触之后, 语境的情感是否迁移到单词, 而单词正字法选择任务、单词定义生成任务和单词定义匹配任务则用于回答第二个研究问题, 即语境的情感是否会影响单词习得的效果。”

希望这些文字介绍可以解决专家的疑问。但是, 我们没有在问题提出部分进行更详细的说明, 主要是考虑到在这个地方补充测试任务的介绍可能会显得唐突。

意见 7: 词汇情感评分为 0~9 点评分, 那么图 2 的-2~1 是如何得来的?

回应: 对的, 词汇情感评分为 1~9 点评分。但是, 就像在“3.2 研究问题 2”中的第 2 段所介绍的, “在把这个变量进入模型前先进行标准化处理, 以使模型更容易拟合并更方便解释结果”, 图 2 是基于拟合的模型所生成的图形, 之所以这里的数字是-2 ~ 1 正是对被试的情感评分作标准化处理(即转变成 z 分数)的结果。

意见 8: 单词定义生成任务发现了和正字法结果相反, 应如何解释?

回应: 关于这条意见, 我估计可能是审稿专家的笔误所致, 因为在论文里, 单词定义生成任务和正字法选择任务的结果是一致的, 正如原文中所述:

关于正字法:

“总体上, 随着语境情感效价的提高, 即在越积极的情感语境里, 被试在单词正字法选择任务的准确率也越高, 但是, 变化语境下的准确率始终要高于重复语境, 只有在语境情感效价达到很高值的时候, 两种语境下的准确率才比较接近……”

关于单词定义生成任务:

“在变化语境的条件下，随着语境情感效价的提高，被试在单词定义生成任务的准确率也显著获得提高，也就是说如果被试认为这个词在情感上越积极，他们对这个词的词义记忆就越好。但是，在重复语境下，单词定义生成任务的准确率却几乎没有变化，不受语境情感效价的影响……”

与这两个任务的结果不完全一致的是单词定义匹配任务，正如原文中所介绍的：

“模型拟合的结果显示，语境的变异性没有主效应($\chi^2(1)=2.38, p=0.12$)，但是语境的情感效价有显著的主效应($\chi^2(1)=78.54, p<0.001$)，二者不存在显著的交互效应($\chi^2(1)=0.21, p=0.65$)，说明语境效价的影响不取决于语境变异性的不同水平，即不管是在哪种语境的变异性之下，语境效价的影响都是一样的。”

可见，三个任务所获得的结果总体上是一致的，那就是语境情感效价越高，被试的表现就越好，不存在相反的结果。但是，单词定义匹配任务与另外两个任务的结果确实不完全一致，那就是语境变异性的影响，语境变异性在单词定义匹配任务里没有显著影响。我们认为这种不完全一致的结果可能源于这个任务本身的性质，单词定义匹配任务要求“被试将给定的 18 个汉语词和给定的 18 个定义进行连线匹配”，相比正字法匹配任务以及单词定义生成任务，我们估计单词定义匹配任务所要求的认知负荷可能更小，也因此受到语境变异性的影响也就更小。

意见 9：语境变换的情绪效价是否也是 trial-by-trial 变化的？如果是，那么重复语境下的语义韵效应也可以解释为情感重复暴露效应吧？

回应：感谢专家的意见。我们认同专家的观点，即“语境变换的情绪效价是 trial-by-trial 变化的，那么重复语境下的语义韵效应也可以解释为情感重复暴露效应”。这也就是我们在论文里论述的“双枣树”效应，即“重复啰嗦的接触有情感效价的文本”导致语境情感的迁移，迁移到单词身上”，就像我们在论文里介绍的，“这一结果与之前研究所获得的人类更愿意体验自己熟悉的情感的发现相一致 (Ford & Tamir, 2014; Van Den Bosch et al., 2013)。比如，有研究发现当个人听熟悉的音乐时，自我报告的愉悦感与皮肤电活动(EDA)(情感唤醒的客观指标)之间存在很强的正相关性……”。根据先前的研究，这里的熟悉的音乐即已经反复听过的音乐。

.....
审稿人 2：

研究主要考察了两个问题，1 是语境的情感是否可迁移到单词，2 是语境的情感是否影响单词习得的效果。结果显示，在重复接触相同的阅读材料下，语境的情感可以迁移到单词。对于单词习得，在变化阅读材料的条件下，语境的情感对新颖词词形及词义的学习具有显著的预测作用。总体而言，研究问题具有一定的理论价值，研究结果对于词汇学习也具有一定启发。以下几方面的问题有待提升：

回应：感谢审稿专家的积极评价和宝贵的建设性意见！

意见 1：摘要和题目 从题目来看，“双枣树”效应是本研究的关键，但是在摘要里未见相关内容，建议在摘要里对这一效应做必要说明。此外，摘要只体现了研究方法和结果，并未体现研究价值。

回应：根据建议，我们已经对摘要进行了修改。

意见 2：引言 (1)情绪与情感不同。本研究关注的更像是情绪，而不是情感。(2)作者引用具身认知理论解释语境情绪可能会迁移到单词。除此以外，这种迁移还可能是一种经典条件反射。既然本研究的重点是情绪，建议也从情绪出发，解释词汇情绪学习的机制。

回应：非常感谢专家的意见，对这个问题，我们非常仔细地进行了思考，也查阅了很多资料。我们同意专家的判断，也注意到在心理学研究领域，确实有学者使用“情绪”一词，比如郭晶晶教授等的以下文献：

郭晶晶,王敏帆,方丹.(2019). 语境情绪韵律对新异词汇学习过程的影响. *心理科学* 42(4),876-883.

但是，在外语界，当这个概念跟语义韵相关联并一块讨论的时候，学者们无一例外地都使用了“情感”一词，如：

李晶洁,胡奕阳 & 陶然. (2022). 基于情感倾向分析的语义韵强度算法探析. *外国语(上海外国语大学学报)* (05), 65-74.

李晓红 & 卫乃兴. (2012). 双语视角下词语内涵与语义韵探究. *现代外语* (01), 30-38.

张毓 & 卫乃兴. (2021). 学术论文局部语义韵探索. *外语与外语教学* (06),1-11.

陆军 & 李聪慧. (2020). 英汉口语会话中的语义韵研究. *当代外语研究* (02), 96-110.

经过仔细思考，我们也希望能够继续使用情感，而不修改为情绪，主要有以下几个原因：

(1) 参考与本研究最为关联的国际文献，即：

Snefjella, B., Lana, N., & Kuperman, V. (2020). How emotion is learned: Semantic learning of novel words in emotional contexts. *Journal of Memory and Language*, 115, 104171.

他们在讨论语义韵的学习时，交叉使用了两个英语单词，分别是 affect 和 emotion:

“Either via a definition given through language or via situational contexts in which a word occurs, a person learns to associate sin with negative affect, vacation with positive and cosine with neutral affect...”

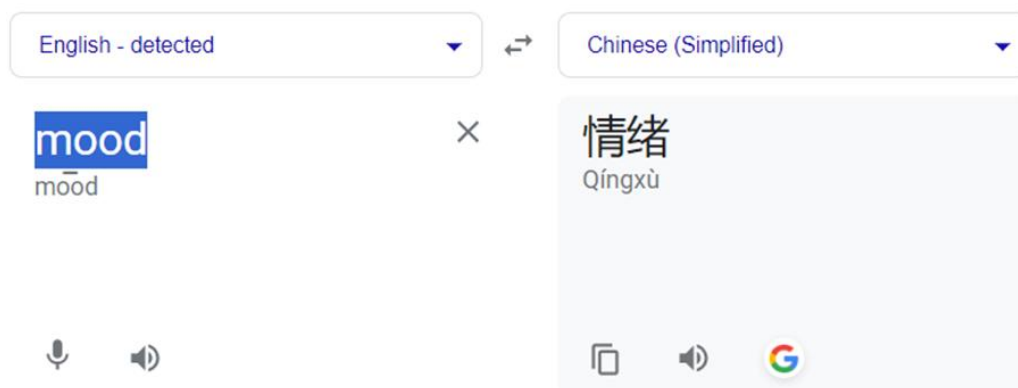
“Clearly, the nature of the linguistic contexts in which novel words are embedded is crucial for whether emotionality of contexts propagates to those words through semantic prosody...”

如果说，emotion 这个词翻译成中文有歧义的话(既可翻译成情感，也可翻译成情绪)，affect 这个词的翻译则非常清楚，都翻译成“情感”。从我们自己的理解来说，本研究探讨的更接近于 affect，故希望继续使用“情感”一词。

(2) 正如我们在文中所论述的，“大量的研究发现，单词会从它们出现的自然语言使用中吸收一些情感极性，而且这种情感极性会最终稳定下来成为这个单词词义的组成部分，称作语义韵产生的情感内涵”，也就是说从语境吸收而最终习得的语义韵是一种稳定持续的语义内涵，从这个角度看，它更接近于“情感”，而不是“情绪”，因为普遍认为“情

感”是持久的，而“情绪”则较为短暂。

- (3) 我们为此也咨询了美国 Alabama 大学的 Dilin Liu 教授，他是语料库语言学专家，精通中英双语，也发表过多篇探讨语义韵的研究论文。Liu 建议使用“情感”一词，他认为情绪在英语中对应的是 mood 这个词，他还使用谷歌翻译向我们展示：



(4) 我们在论文中使用的“语境情感”这个表达接近于现代自然语言处理进行文本挖掘所使用的“情感分析”中的情感，即探讨文本中的这种情感是否会迁移到新颖词，从而习得语义韵，更准确没有歧义的英语单词应该为 *sentiment*，情感分析即为 *sentiment analysis (SA)*。

不过，如果审稿专家认为“情绪”确实更为适合，我们愿意之后再做出修改。

另外，关于经典条件反射。我们查阅乔姆斯基(Chomsky)对行为主义理论的观点，认为本研究讨论的语境情感成功地迁移到单词从而使得本来中性的单词获得某种语义韵的过程可能涉及复杂的认知心过程，就像文献综述里提到的“具身理论”以及自动警觉理论所讨论的情感认知和学习等。这个过程可能并不完全是经典条件反射所讨论的刺激和反应的过程。不过，我们并不是相关领域的专家，理解不周之处，恳请专家进一步指导。

意见 3: 第五页有关语境本身变异性的内容建议单独列为一个小部分。

回应: 已根据建议把这些内容单独列为一个小部分，请参见原文所作修改。

意见 4: 针对两个研究问题的研究假设不清晰。

回应: 我们提出的两个研究问题是：

- (1) 语境情感是否可通过阅读接触，从语境迁移到单词？
- (2) 语境的情感是否影响单词习得的效果？

针对这两个研究问题的研究假设是：

- (1) 语境情感是否可通过阅读接触从语境迁移到单词取决于语境的变异性，根据语境变异假说，只有在变化语境下(多篇)语境情感才可发生迁移。
- (2) 语境情感显著影响单词习得的效果。

但考虑到这是一篇小论文，篇幅有限，没有把上述研究假设放到论文里。如果审稿专家认为有必要，我们同意把它们放到论文里。

关于方法:

意见 5: 方法 (1)被试选取了高中生, 但并未说明为何选择了高中生群体。或许高中生的选择也是本研究的价值之一。

回应: 感谢专家的建议。当我们在开展实验时选择高中生主要有两个考虑: (1)本研究探讨的是母语词汇习得, 考察高中生要比大学生更有意义, 因为从语言习得角度看, 高中生仍然处在语言能力发展期; (2)高中生更为集中, 收集数据更为便利。在读到专家意见后, 我们也查阅了一些文献, 发现文献中有较多研究发现, 情感学习随着年龄变化, 更年轻的学习者对情感更为敏感。这个发现也支持了我们使用高中生作为被试的选择。我们在论文里补充了以下内容:

“之所以选择高中生而不是大学生主要是考虑到本研究考察母语词汇习得, 而高中生仍然处在语言能力发展的重要阶段, 他们词汇能力的发展值得深入关注, 而从情感学习的角度看, 有研究发现情感学习随着年龄而变化, 更年轻的学习者对情感更为敏感 (Fernández-Aguilar, 2018)。”

但是, 考虑到论文的篇幅, 我们没有在讨论部分进一步分析选择高中生被试的实际意义。

意见 6: 语境的变异性是被试间设计, 各有多少被试是重复语境和变异语境呢? 另外, 实验任务涉及到学习, 是否有考虑与个体学习相关的无关变量的控制呢? 例如学习成绩和工作记忆。

回应: 语境的变异性有两个水平, 因此, 一半的被试(98 人)参加了变化语境下的测试, 另一半参加了重复语境下的测试。

感谢专家关于个体差异性因素的控制问题, 非常具有建设性。在开始实验之前, 我们确实考虑过这个问题, 但考虑的主要是被试阅读时间的差异性, 没有考虑学习成绩和工作记忆, 主要是因为我们认为学习成绩和工作记忆可能对二语学习者来说更加重要, 比如学习者的二语水平、词汇量和工作记忆, 但是对本族语者来说, 重要性可能降低, 另外, 也是更重要的原因是, 如果要同时获得到这些变量的值, 需要耗费更多的时间, 在当前实验, 已经比较复杂的情况下增加这些变量会极大增加被试的压力, 故没有进一步考虑这些变量。我们希望在未来的研究里, 能把这些变量的影响考虑进去。在讨论研究的局限性时, 增加了这部分内容。

意见 7: 测试单词只有 9 个, 基于 9 个单词的结果推出的研究结论是否可靠呢?

回应: 感谢专家的意见。之所以只选择 9 个而不是更多的词作为目标词, 一方面是借鉴了先前很多类似的实验 (Pellicer-Sánchez, 2016: 6 个词; Joseph & Nation, 2018: 6 个词; Webb, 2007: 10 个词; Snefjella & Kuperman, 2016: 9 个词), 另一方面也是因为研究发现人类在短期内的学习能力是有限的 (Miller 1956)。同时, 考虑到本研究被试样本量大, 而且每个词都分别表达了 9 个外延义, 我们认为 9 个目标词可以获得稳定可靠的结果。

意见 8: 阅读过程没有时间限制。这可能使得个体间存在很大的差异。是否考虑过这个问题对实验结果可能造成的影响? 统计分析时是否有把阅读时间作为一个随机效应进行控制呢?

回应: 感谢专家的意见。正如论文原文所报告的, 被试完成阅读和所有测试的平均用时为 34.84 分钟 ($SD=5.23$), 从实验所获得的记录以及最终数据看, 阅读用时的个体差异较小。同时, 我们在拟合模型时也确实尝试了把被试的阅读时间作为预测变量进入模型, 但结果发现被试的阅读时间在任何测试的任何模型里都不显著, 不管拟合的是线性关系, 还是曲线关系

都是如此，故我们拟合的最终模型都没有包含阅读时间。在二语词汇习得研究，有学者提出注意假说(noticing hypothesis)，认为在阅读时被试注意力投入越多（一般通过阅读的时间来衡量），词汇学习效果就越好(Schmidt, 1990)，但是在母语研究里，相关研究结果并不一致，很可能母语阅读个体差异较小，尤其是像本研究，每个阅读小篇章只有两句话组成。另外，我们估计这个问题也可能存在语言差异，是否拼音文字与字母语言可能不完全一样。

关于结果:

意见 9: 请具体列出每一个分析的混合效应模型和变量的编码方式。

回应: 已经以附件的形式，把每一个模型放在附件。在拟合最终模型时，使用的是 R 的默认编码方式为即 dummy coding (即 treatment coding)，但是，正如正文所汇报的，文章使用 *mixed()* 函数计算主效应和交互效应，在计算主效应和交互效应时，*mixed()* 函数会自动把自变量的编码方式改为 *contr.sum*，以准确计算变量的主效应和交互效应。

意见 10: 3.2.1 和 3.2.2 有关交互作用的结果没有具体的统计量。

回应: 感谢建议，已经补充相应的统计量。

意见 11: 讨论 研究不足与展望一般是放在讨论的最后，而不是结论。

回应: 谢谢！已根据要求修改。

第二轮

审稿人 1: 感谢作者的详细回复和修改。以下问题还希望能继续改进:

回应: 感谢审稿专家的耐心和肯定，并帮助我们进一步提升稿件质量！

意见 1: 意见 4: 前言中语义韵的例子均为副词或形容词，而研究中均采用了名词，而且是实物名词，这有没有可能引入无关的抽象性变量？如果有这种可能性，该如何影响结果解读？

补充: 这个问题的意思是否有可能结果只适用于实物名词，而非抽象名词，因为当前材料的新词都是以实词的形式出现的。建议增加一定的讨论和文献支持。

回应: 非常感谢审稿专家的耐心解释，之前，我们确实没有理解清楚这一重要建议的含义。经再次仔细查阅了文献，我们发现先前的研究确实发现情感对具体的实物名词和抽象名词的学习会产生不同的影响。比如，根据情感具身说的观点(Affective Embodiment Account, Borghi et al., 2017)，实物名词的学习需要调用更多的运动感觉信息(sensorimotor information)，而抽象名词的学习调用更多的则是情感信息(emotional information)，因此，在情感语境下，抽象词的学习效果更可能要好于具体的实物名词。从这个角度看，本研究的发现可能对具体的实物名词更具有普遍意义，而对抽象的名词是否如此则仍需要更多的研究加以证实。

另外，选择实物名词作为实验的目标词，一个重要的原因是实物名词更容易操控，在材料设计上更便利，另外一个非常重要的原因是，前人研究发现，与抽象词相比，实物词在效价评分任务(valence rating task)中表现出更强的敏感性(Lana, 2021)，因此，比较容易观察到实验操控的效果。

根据审稿人的建议，我们在文章的两个部分补充了内容或讨论。第一个部分是在方法部

分的“2.2.2 目标词”部分，增加了以下内容：

上文已经介绍过，9个新颖词嵌入到135篇阅读材料后，表达的外延义都是日常生活中常见的大语义范畴，从词性看都是实物名词。选择实物名词作为目标词的一个重要原因在于前人的研究发现，与抽象词相比实物词在效价评分任务(valence rating task)会表现出更强的敏感性(Lana, 2021)，比较容易观察到实验操控的效果。

参考文献如下：

[1] Lana, N. (2021). *Learning abstract words: Role of valence in linguistic context (Unpublished master thesis)*. McMaster University.

另一个部分是在论文讨论部分(4.讨论)的末尾：

正如我们在材料设计部分所介绍的，本研究的目标词从词性看都是实物名词。需要引起注意的是，根据先前已有研究的发现，情感对具体的实物名词和抽象名词的学习会产生不同的影响。比如，根据情感具身说的观点(Affective Embodiment Account, Borghi et al., 2017)，实物名词的学习需要调用更多的运动感觉信息(sensorimotor information)，而抽象名词的学习则更多调用的是情感信息(emotional information)，因此，在情感语境下，抽象词的学习效果可能要好于具体的实物名词。从这个角度看，本研究的发现对具体的实物名词可能更具有普遍意义，而对抽象的名词则仍需更多研究加以证实。未来研究可以直接对词的抽象性进行操控，以考察这一变量在不同的情感语境下对词汇习得的影响。

参考文献如下：

[1] Borghi, A. M., Binkofski, F., Castelfranchi, C., Cimatti, F., Scorolli, C., & Tummolini, L. (2017). The challenge of abstract concepts. *Psychological Bulletin*, 143(3), 263-292.

意见 2：意见 8： 单词定义生成任务发现了和正字法结果相反，应如何解释？

补充：这个问题可能没有表述清楚，确切得说是两个任务下的语境效应相反，即正字法任务中的语境差异发生在低情绪效价，而单词定义生成任务的语境差异发生在高情绪效价。多遍还是多篇也是本研究关注的问题之一，那么对正字法和语义任务产生不同的多遍/多篇效应进行讨论，是必要且有意义的。

回应：感谢审稿专家的细致和耐心，这确实是一条非常重要，值得我们好好思考，之前我们对这条建议也确实有误解。从图 2 和图 3 可以看到，尽管变化语境对单词正字法和单词语义这两种知识的学习都比重复语境更有优势，但是，在不同的语境变异之下，情感效价对单词正字法和单词语义的学习产生了不同的影响，就像审稿专家提出的“正字法任务中的语境差异发生在低情绪效价，而单词定义生成任务的语境差异发生在高情绪效价”。

我们对这种差异进行了仔细思考和分析，在原文的讨论部分增补了一长段来解释这种差异：

“不过比较图 2 和图 3 我们发现，尽管变化语境对单词正字法和单词语义两种知识的学习都比重复语境更有优势，但是，在不同的语境变异之下，情感效价对单词正字法和单词语义的学习产生了不同的影响，看起来，正字法任务中的语境差异主要发生在低情感效价，而单词定义生成任务的语境差异则更明显地体现在高情感效价。我们认为，一方面这种差异说明了词形和语义学习对语境变异性的依赖程度不同。先前的研究发现，单词的词形和语义学习在一定程度上是相互独立的(Sommers & Barcroft, 2013)。根据语境变异性假说，学习者在通过不同语境多次接触同一词汇时，会在记忆系统中留下关于该词义和语境的记忆痕迹。这

些痕迹会随着不同接触而得到强化，最终形成独立于语境的核心语义表征。然而，词形的学习则不一定需要通过多种不同语境(即多篇)的接触来完成，重复接触(即多遍)也是词形学习的有效途径 (Bolger et al., 2008)，心理语言学的重复启动实验为此提供了明确的证据(Wagner et al., 2000)。另一方面，这种差异也说明了正字法选择任务和定义生成任务所测试的核心内容存在本质上的差异。相对于正字法选择任务，定义生成任务涉及更多的词汇提取(retrieval)活动，更容易受到注意力的影响(Hao et al., 2021)。早期研究发现，变化语境之下(即多篇)学习者需要投入更多的认知资源来处理信息，包括注意、搜索、提取和评估(McKeown et al., 1985)。因此，在多篇阅读中，学习者的信息提取能力可能会随着注意力的增强而提高。而根据积极心理学的拓展-构建理论(Fredrickson, 2003)，积极情感可以通过扩展个体的注意、认知和行为资源来促进学术表现。此外，人类对积极信息的偏好性也可能导致他们在扎根或具身认知的基础上更加关注和深层次地加工积极信息(Hinojosa et al., 2016)。因此，在多篇阅读和高情感效价的阅读中，可能会出现积极效应的叠加，从而显著提高核心语义的学习效果。而正字法选择任务则可能并不需要调动大量的注意力资源，主要依赖于短时记忆，相比于多篇，多遍重复的阅读会帮助学习者对词形形成单一、强化的记忆痕迹(Bolger et al., 2008)，并在积极情感下逐渐加强，从而缩小了上述变化语境(即多篇)的优势。”

参考文献如下：

- Bolger, D. J., Balass, M., Landen, E., & Perfetti, C. A. (2008). Context variation and definitions in learning the meanings of words: An instance-based learning approach. *Discourse Processes*, 45(2), 122-159.
- Fredrickson, B. L. (2003). The value of positive emotions. *American Scientist*, 91, 330-335.
- Hao, S., Liang, L., Wang, J., Liu, H., & Chen, B. (2021). The effects of emotional context and exposure frequency on L2 contextual word learning. *International Journal of Bilingualism*, 25(5), 1280-1296.
- Hinojosa, J. A., Rincón-Pérez, I., Romero-Ferreiro, M. V., Martínez-García, N., Villalba-García, C., Montoro, P. R., & Pozo, M. A. (2016). The Madrid Affective Database for Spanish (MADS): Ratings of dominance, familiarity, subjective age of acquisition and sensory experience. *PLoS One*, 11(5), e0155866.
- McKeown, M. G., Beck, I. L., Omanson, R. C., & Pople, M. T. (1985). Some effects of the nature and frequency of vocabulary instruction on the knowledge and use of words. *Reading Research Quarterly*, 20, 522-535.
- Singh, L., Morgan, J. L., & Best, C. T. (2002). Infants' listening preferences: Baby talk or happy talk? *Infancy*, 3(3), 365-394.
- Sommers, M. S., & Barcroft, J. (2013). Effects of referent token variability on L2 vocabulary learning. *Language Learning*, 63(2), 186-210.
- Wagner, A. D., Koutstaal, W., Maril, A., Schacter, D. L., & Buckner, R. L. (2000). Task-specific repetition priming in left inferior prefrontal cortex. *Cerebral Cortex*, 10(12), 1176-1184.
-

审稿人 2：作者总体上对论文内容做了较好的修改，论文质量有了较大的提升。以下几个问题建议再做进一步考虑：

回应：感谢审稿人的积极评价，也感谢审稿人提出的进一步修改建议！

意见 1：还是建议作者能够根据有关理论和过去的研究发现，在论文中对研究问题提出明确的研究假设。

回应: 感谢审稿专家的建议, 我们做了相应修改, 在研究问题之后, 补充了研究假设, 如下: (需要注意的是, 根据意见 3, 我们对两个研究问题使用了新的语言表述, 见下文)

根据前文所介绍的理论以及已有研究的发现, 我们提出以下两个研究假设:

- (1) 语境情感可以通过阅读接触, 从语境迁移到单词, 使读者习得单词的语义韵, 但是这种迁移同时受语境变异性的影响。
- (2) 单词语义韵的习得效果显著影响词形及词义的习得效果, 但是, 这种影响同时也受语境变异性的调节。

在论文的讨论部分, 我们综合论文的研究发现, 又再次提到了这两个假设:

“上述结果支持了前文提出的两个研究假设……”

意见 2: 虽然统计分析只发现了语境的主效应, 但建议作者能像图 2 和图 3 一样, 在图 4 呈现重复和变化语境下的结果, 这样有助于更好地了解整体的结果。

回应: 感谢审稿人的细致和耐心! 我们已经根据要求重新作了图, 展示了语境的情感效价在两种不同的语境变异性的影响。

意见 3: 从目前的修改稿内容来看, 论文的标题似乎只体现了部分研究内容 (第一个研究问题), 建议作者可以再斟酌一下标题, 看看如何能够更全面地涵盖研究内容, 特别是和第二个研究问题有关的内容。

回应: 感谢审稿人的细致和耐心! 确实, 看起来, 论文中的研究问题 1 和研究问题 2 彼此独立, 分别回答了不同的问题。但是, 我们认为这主要是这两个研究问题的语言表述所导致的结果, 原来的两个问题是:

- (3) 在经过几次阅读接触之后, 语境的情感是否可迁移到单词?
- (4) 语境的情感是否影响单词习得的效果?

看起来, 第一个研究问题回答的是语境情感是否可迁移, 而第二个问题回答的是语境情感的影响, 这也使得论文的标题看起来不够全面, 没有涵盖问题 2。但实际上, 我们认为这两个研究问题的核心都是语义韵的习得问题, 即标题表述的单词语义韵发生的问题。因为第一个问题中的语境的情感是否可迁移回答的是语义韵是否可以习得或如何习得, 而在第二个问题, 我们在回答语境的情感是否影响单词习得的效果时, 在操作上, 正如我们在 3.2 小节开头部分所介绍的, 不是把语境情感当作分类变量, 而是“……把语境的情感表征操控为在研究问题 1 里被试对每个单词所作的情感评价的分数……”, 这里的情感效价的分数实际上就是语义韵的习得效果, 因此, 概括起来, 这两个问题中的第一个问题回答的是语义韵是否可习得的问题, 而第二个问题是语义韵的习得效果是否影响单词词形及语义的习得效果, 核心都是语义韵的习得, 即语义韵发生的问题。这样的话, 现在的论文题目则可以涵盖两个研究问题, 正是因为这个原因, 我们希望能够继续保留现在的题目, 因为我们认为它较好地概括了这个研究论文最核心的内容。但是, 为了更好地体现论文中两个研究问题的关联性, 我们现在把它们改为:

- (1) 语境情感是否可通过阅读接触, 从语境迁移到单词, 使读者有效习得单词的语义韵?

- 如果是，这种迁移是否受语境变异性的影响？
- (2) 单词语义韵的习得效果是否影响词形及词义的习得效果？ 如果是，这种影响是否受语境变异性的调节？

感谢审稿人，希望这个修改可以解决您的关切。

第三轮

审稿人 1 意见：没有进一步意见了

审稿人 2 意见：作者对提出的问题做了较好的回应与修改，建议发表。

编委意见：该文经两位审稿专家三轮详细的审阅，作者对所有问题进行了详细的答复，并对文章进行了相应的修改，文章的质量有显著提升。我同意审稿人的意见，建议接收。

主编意见：本研究以中文母语者为研究对象，对语境情感能否通过阅读接触迁移到单词，并对这种迁移效应是否影响单词习得效果进行了考察。本论文的选题具有较强创新性，研究框架清晰，数据处理过程科学规范。