

---

## 《心理学报》审稿意见与作者回应

题目：无关长时记忆表征能否引导视觉注意选择？

作者：胡岑楼;张豹;黄赛

---

### 第一轮

#### 审稿人 1 意见：

本研究检验了无关长时记忆表征对注意是否具有捕捉效应，该问题具有较大的理论价值。但是对于该研究有以下几点意见：

**回复：** 非常感谢审稿专家的肯定！

**意见 1.** 实验二中将刺激重复两次以上作为长时记忆，这一点是存疑的，尽管最近一些研究提出多次重复会导致工作记忆向长时记忆迁移，但是几次才能完全实现并没有确定结论，所以工作记忆表征重复两次以上就不再捕捉未必是长时记忆的缘故。以重复两次以上就不再捕捉来说明这和工作记忆不同，有可能是长时记忆，所以长时记忆是不能捕捉的，这个论证的逻辑就有一些循环论证的意味。所以我建议最好采用一个明确没有工作记忆干扰的新试验来验证本假设。这个建议也是因为对于阴性结果，我们一般会比较谨慎一些。

**回复：** 感谢审稿专家中肯的意见！诚如专家所说，虽然有研究认为多次重复会导致工作记忆向长时记忆迁，但重复的次数尚没有定论，而且重复次数会受到被试个体差异、所记忆的项目的性质等诸多因素的影响。具体到本研究，当工作记忆项目重复两次之后（例如第 3 次重复）的记忆表征状态可能并非完全属于长时记忆表征，很有可能在部分重复序列中记忆表征仍然是工作记忆表征，但在部分重复序列中已转化为长时记忆表征，因此，在本论文原稿中直接将重复两次以上的记忆表征等同于长时记忆表征确有不妥。但从实验二的结果来看，重复 1 次与重复 2 次时都出现了注意引导效应，但重复 3 次或以上时，注意引导效应消失了，这说明重复 3 次时至少部分重复序列的记忆表征已经由工作记忆表征转化为长时记忆表征，因而导致记忆表征对注意的引导效应减弱甚至消失，此结果应该还是能在一定程度上支持无关长时记忆表征不能捕获视觉注意的结论。修改稿中已对相应的表述进行了修改（p18 蓝色标记部分）。

当然，本研究所证明的问题是基于阴性结果，确实需要非常谨慎，因此，作者在设计实验时希望通过如下两种操纵方法来加强实验结果的可靠性。一方面，设置了工作记忆匹配条件，可以与长时记忆匹配条件进行对比，因为无关工作记忆表征对视觉注意的引导效应得到了大量研究的证实，是一种比较稳健的效应。如果在完全相同实验情境下出现了基于工作记忆表征的注意引导效应，但没有出现基于长时记忆表征的注意引导效应，这种实验结果的分离一定程度能说明阴性结果并非实验误差所致。这也是实验中设置工作记忆匹配条件的重要原因。另一方面，实验 1 和实验 2 采用两种不同的操纵长时记忆表征的方式，如果两种操纵

---

方式都能得到一致的结果，这在某种程度上能增加阴性结果的可靠性。

审稿专家建议采用一个明确没有工作记忆干扰的新实验来验证研究假设，这一建议作者非常认同。“工作记忆表征对长时记忆表征引导注意可能具有干扰作用”，这确实是作者实验设计所忽略的一个额外变量，因此在修改期间遵照审稿专家的宝贵建议，增加了实验3，即在无工作记忆任务的情境下探讨长时记忆表征的注意引导效应，结果还是支持了原来的实验结果，即单独的长时记忆任务条件下依然没有观察到基于长时记忆表征的注意引导效应。

**意见 2.** 实验二为什么没有记录眼动，这使得两个实验无法对比。

**回复：**感谢专家的意见！实验设计之初作者也曾考虑过这一问题，但后来考虑到以下几方面的原因，实验二没有记录眼动。

(1) 担心实验时间太长被试产生疲劳，尤其是眼睛疲劳，会影响实验结果。实验2的试次比较多，行为实验每种实验条件下40试次，每个被试所需的实验时间约75分钟。根据以往文献及研究经验，注意引导效应如果采用首次注视点百分率作为指标，每种实验条件至少60试次，再加上眼动仪的调节与校正以及每个试次进行漂移校正的时间，实验时间至少需要2小时。

(2) 根据以往的研究结果<sup>1</sup>，当视觉搜索靶子比较显著时，行为反应时与首次注视点百分率的结果非常一致，本研究的结果也支持这一结论，即实验一与新增加的实验三的行为反应时与首次注视点百分率指标所获得的结果几乎一致。

(3) 实验二的目的是采用与实验一不同的实验范式来操作记忆表征的存储状态，可以不直接进行对比。

作者为了增加实验二结果的统计检验力，实验二特意收集了28个被试的数据，其中有效被试数为26个。

此外，实验二的主要目的之一是为了检验实验一中占优势地位的工作记忆对其他记忆表征的阻断作用是否是导致无关长时记忆表征没有表现出注意引导效应的可能原因。新增加的实验三中，由于没有工作记忆匹配条件的存在，因此不存在工作记忆表征对长时记忆表征的阻断作用，结果在行为反应时与首次注视点百分率指标上依然都没有发现无关长时记忆表征对注意的引导效应，这在一定程度上证实了实验二的研究结果。

**意见 3.** 实验一的被试数太少了，一般就是有显著差异的研究被试量少也会担心结果的稳定性，何况是结果为没有显著差异，长时记忆不捕捉的阴性结果，更加需要较多的被试才能减轻大家关于统计检验力的担心。

**回复：**根据审稿专家的宝贵建议，论文修改阶段将实验一被试数增加到24个，所得到的结果与原来的结果几乎完全相同，另外新增加的实验三的结果也与实验一的结果非常一致，说明实验一所得到的结果是比较可靠的。

---

<sup>1</sup> 由于匿名的需要，此处暂不列参考文献，希望审稿专家谅解。

---

## 审稿人 2 意见:

该研究通过两个实验对上述问题进行了探讨,研究结论对进一步理解记忆系统与视觉注意之间的交互关系具有较重要的学术价值。存在以下几个问题,建议修改:

**回复:** 非常感谢审稿专家对论文的肯定!

**意见 1:** 根据 Soto 等的双任务研究范式,一般都是先让被试记住一个项目,然后完成一个视觉搜索任务,最后进行一个记忆探测任务。在实验 1 中,如图 1 所示,作者将视觉搜索任务,工作记忆探测任务和长时记忆探测任务分离开来,这样操作的目的是何在?

**回复:** 感谢审稿专家的意见!在以往相关的研究中,也有研究将视觉搜索任务和记忆探测任务分开,例如在 Kiyonaga, Egner 和 Soto(2012)的研究中,其中 20%的试次在记忆刺激呈现之后直接出现记忆探测任务,而其他试次中记忆刺激之后视觉搜索任务(又见 Soto, Humphreys, Heinke, 2006)。更重要的是,如果视觉搜索任务之后固定会出现记忆探测任务,被试可能会在视觉搜索阶段有意识地对记忆刺激进行复述以提高记忆正确率(Woodman & Luck, 2007),而将视觉搜索任务和记忆探测任务分开,被试知道视觉搜索任务之后不会跟随记忆探测任务,因而无需使用上述记忆复述策略(Dowd, Kiyonaga, Beck, & Egner, 2015)。基于以上考虑,实验 1 将视觉搜索任务和记忆探测任务在不同试次中分开呈现<sup>2</sup>。

**意见 2:** 在实验 2 中,从反应时数据来看,记忆项目第一次重复时视觉搜索的反应时要明显长于第二次重复时视觉搜索的反应时,作者如何解释?

**回复:** 实验 2 中记忆项目第一次重复时的视觉搜索反应时明显长于第二次重复时的视觉搜索反应时,作者认为有两种可能原因:

(1) 与工作记忆任务的资源消耗有关。在实验 2 中,当记忆项目第一次重复的时,该记忆项目作为新项目进行记忆,被试需要对该记忆项目进行编码与存储,而当记忆项目第二次重复时,由于线索(“旧”)的提示,被试只需在原有工作记忆表征的基础上对记忆项目进行简单复述,因此被试在完成视觉搜索任务时,在线保持第一次重复的工作记忆项目所需要消耗的资源显然要多于保持第二次重复的工作记忆项目,由此可能导致视觉搜索变慢。

(2) 与任务转换的难度有关。当由记忆任务转换到视觉搜索任务,需要付出一定的转换代价。由上面的分析可知,记忆项目第一次重复时记忆任务的难度明显要大于第二次重复

---

## <sup>2</sup>参考文献

Dowd, E. W., Kiyonaga, A., Egner, T., & Mitroff, S. R. (2015). Attentional guidance by working memory differs by paradigm: an individual-differences approach. *Attention, Perception & Psychophysics*, 77(3), 704-712.

Kiyonaga, A., Egner, T., & Soto, D. (2012). Cognitive control over working memory biases of selection. *Psychonomic Bulletin and Review*, 19(4), 639-646.

Soto, D., Humphreys, G. W., & Heinke, D. (2006). Dividing the mind: The necessary role of the frontal lobes in separating memory from search. *Neuropsychologia*, 44(8), 1282-1289.

Woodman, G. F., & Luck, S. J. (2007). Do the contents of visual working memory automatically influence attentional selection during visual search? *Journal of Experimental Psychology: Human Perception and Performance*, 33(2), 363-377.

---

时的记忆难度,因而,注意从记忆任务上脱离并转换到视觉搜索任务上所需要的时间也更长,因而导致第一次重复时的反应时要长于第二次重复时的反应时。

**意见 3:** 图 1 中的流程图的文字不清晰。

**回复:** 实验 1 的流程图在修改稿中已进行修改。

**意见 4:** 参考文献引用存在不规范的地方,建议按照 APA 格式认真核对修改。

**回复:** 参考文献已按照 APA 格式进行修改。

---

## 第二轮

**审稿人 1 意见:**

作者比较好的解答了我的问题,我有一些新的修改建议:

**意见 1:** 实验一和二中被标为长时记忆的条件或者重复多次后可能转化为长时记忆都只是一种可接受的假设而不是得到验证的事实,所以作者最好在文中表达出这一意思,而不是看起来像是确定的事实。

**回复:** 感谢审稿专家具有建设性的意见!根据专家意见,修改稿已在讨论最后部分对本研究所操纵的长时记忆表征进行了说明(p17)。

“需要说明的是,本研究中长时记忆表征是在实验室情境下要求被试长时间地存储于长时记忆中的表征,相对于自然情境中人类长时记忆表征的存储来说,这种长时记忆存储只限于被试在完成实验过程中的暂时性存储,未来可采用更具有生态效度的长时记忆范式对此问题予以进一步探讨。”

**意见 2:** 实验 3 虽然更有力地说明了作者想要表达的意思,但是被试数比较少。这里会有两个问题,第一,实验 3 的检验力不足;第二,三个实验的被试数差别太大,给人以作者是在看结果显著性决定被试数的印象,这是应该绝对避免的。

**回复:** 非常感谢审稿专家的意见。首先需要向专家解释的是,由于上次修改过程临近期末,学生需要准备与参加期末考试,招募被试比较困难,最终只收集到 16 名被试的数据。从统计分析的结果来看,16 名被试所表现出来的结果模式较为一致,故修改版本中只提供了 16 名被试的结果。但诚如专家的意见,原文中实验 3 的被试数量要明显少于实验 1 与实验 2,可能会让读者误以为是根据显著性结果来决定被试数。而且需要证明的结果为阴性结果,确实需要增加被试数以提高统计检验力。本轮论文修改阶段将实验 3 的被试数增加至 25 名,

---

所得到的结果与原来的结果一致。长时记忆匹配条件和中性条件下的首次注视点百分率分别为 13% 和 12.03%,差异不显著, G power 软件计算结果显示两者差异的效应量为 0.18, 达到显著所需的被试数估计为 374; 反应时在两种匹配条件下只相差 6 ms(926 ms vs.932 ms), 差异不显著, G power 软件计算结果显示两者差异的效应量为 0.04, 达到显著所需的被试数估计为 6442。因此, 实验 3 的结果表明无关长时记忆并不能像工作记忆表征一样引导视觉选择注意。

**意见 3:** 实验 1 增加了 14 名被试后眼动捕获率平均水平发生了明显提高, 作者应该说明两组 14 人被试的结果模式是否一样, 最好补充相关信息决定是否真的可以将两组人直接合并检验。

**回复:** 感谢审稿专家的意见! 根据审稿专家的意见, 作者对增加的 10 名被试的数据进行分析, 结果如下:

正确率: 10 名被试的工作记忆任务的正确率、长时记忆任务的正确率以及视觉搜索任务的正确率均在 95% 以上。

对首次注视点百分率的单因素重复测量方差分析发现: 匹配类型的主效应显著,  $F(4, 36)=8.73$ ,  $p<0.001$ ,  $\eta_p^2=0.493$ , 对首次注视点百分率进行进一步的多重比较发现, 工作记忆匹配条件显著大于中性条件 (24.37% vs. 11.56%,  $p<0.05$ ), 但长时记忆项目匹配条件与中性条件没有显著差异 (12.97% vs. 11.56%,  $p=0.564$ )。另外, 长时/工作记忆匹配条件下与工作记忆表征匹配的干扰刺激捕获的首次注视点百分率显著大于中性条件 (22.34% vs. 11.56%,  $p<0.05$ ), 表现出注意引导效应, 长时/工作记忆匹配条件下与长时记忆表征匹配的干扰刺激所捕获首次注视点百分率显著低于中性条件 (7.34% vs. 11.56%,  $p<0.5$ )。

对视觉搜索任务的平均反应时进行单因素重复测量方差分析发现: 匹配类型的主效应显著,  $F(3, 27)=5.18$ ,  $p<0.01$ ,  $\eta_p^2=0.366$ , 进一步的多重比较发现, 工作记忆匹配条件下的平均反应时显著长于中性条件下 (1025ms vs. 975ms,  $p<0.05$ ), 表现出明显的注意引导效应, 而长时记忆项目匹配条件下的平均反应时与中性条件下的平均反应时却没有显著差异 (992ms vs. 975ms,  $p=0.157$ ), 另外, 长时/工作记忆匹配条件下的平均反应时也显著长于中性条件 (1038ms vs. 975ms,  $p<0.05$ )。

由上述结果可以看出, 增加的 10 名被试所表现出的结果模式与原来 14 名被试所表现出的结果模式完全一致。

---

**意见 4:** 建议对三个实验分析其 effect size 及其被试数估计, 这样可以更直观说明即使有效果, 效果也是非常微小的。

**回复:** 修改稿中作者已根据审稿专家的宝贵意见, 采用 G power 软件对三个实验中的阴性结果进行了效应量分析和被试数估计。实验 1 和实验 3 的长时记忆匹配条件以及实验 2 中工作记忆重复次数三次及以上时的记忆匹配条件与中性条件进行比较, 其差异的效应量都比较小 ( $<0.2$ ), 要达到显著时所需要被试量估计数都超过 370。因此, 作者认为三个实验的结果还是比较可靠的。

**意见 5:** 多重比较有无校正。

**回复:** 多重比较时已采用 Bonferroni 校正。

**意见 6:** 有些表达很奇怪, 比如“以缓减有限注意资源加工信息时的负荷”, 还需要好好检查下文字表达。

**回复:** 感谢审稿专家的意见, 在修改稿中已经对全文的表达进行了检查与修改, 例如“以缓减有限注意资源加工信息时的负荷”已修改成“以缓减资源有限的注意加工系统加工信息时的认知负荷”。

---

### 第三轮

**主编意见:** 研究有一定新意, 对审稿人意见也做了较好的回应。建议进一步完善写作, 如总讨论和结论的顺序应分别为 5 和 6, 总讨论的第一行就有落字 (应为长时记忆)。

**回复:** 非常感谢主编的意见, 修改稿中已对总讨论与结论的编号进行了调整, 对正文、英文摘要进行了全面的检查与修改。