

## 《心理学报》审稿意见与作者回应

题目：视觉长时记忆激活度对促进视觉短时记忆的影响

作者：鲍旭辉 姬鸣 黄杰 何立国 游旭群

### 第一轮

#### 审稿人 1 意见：

**意见 1：**作者没有在文中明确说明如何计算“击中率”和“虚惊率”。作者需要定义目标 target 和噪音 noise，以及什么反应是“击中”，什么反应是“虚惊”。

**回应：**作者重新检查了原始数据、整理过程及计算过程，结果发现，由于数据较多，且本研究的数据在处理时直接复制了另外几个多因素实验数据处理的表头和格式，造成直接将不同条件下的正确率和错误率当成了击中率和虚报率，且作者在稿件投出外审的过程中已发现此问题，在此作者再次向各位审稿专家提出的宝贵意见表示感谢，并对数据处理者的草率和粗心以及对信号检测论方法的理解不够透彻表示深深的歉意。修改稿中，作者重新从原始数据开始检查，将“变化”当作信号，“不变”当噪音，从而求得两种条件下对应的击中率和虚报率，进而计算出正确的  $d'$  和  $K$ 。

**意见 2：**作者在计算各条件下的  $d'$  或  $K$  时首先剔除了各条件下 3 个标准差以外的击中率和虚惊率，请问为什么要这样做？为什么不是剔除各条件下 3 个标准差以外的  $d'$  或  $K$  呢？

**回应：**经过了上一步的修改，作者对数据重新进行了处理，剔除了 3 个标准差以外的  $d'$  和  $K$  值，详见 2.2。

**意见 3：**作者计算了各条件下的  $d'$  和  $K$ ，但我很好奇作者如何计算出两个检测条件（即“未变”和“变化”）下的  $d'$  和  $K$ 。若要计算某个条件下的  $d'$  和  $K$ ，则首先需要计算该条件下的击中率和虚惊率，那么，例如“未变”条件下的击中率和虚惊率是如何计算出来的呢？由于击中和虚惊涉及到目标和噪音两个方面，所以我认为不可能只在“未变”或“变化”条件下就能计算出  $d'$  和  $K$ ，无论是把“未变”定义为目标还是把“变化”定义为目标。

**回应：**详见对问题 1 的说明。

**意见 4：**由于“未变”和“变化”两个条件下的  $d'$  和  $K$  不可能计算出来，但是作者却在每个实验中均给出了它们的  $d'$  和  $K$ ，所以本文三个实验的  $d'$  和  $K$  值结果非常值得怀疑，从而导致我不能确定本研究结果是否能支持作者的结论，即只有高度激活的长时记忆表征才能促进短时记忆。

**回应：**详见对问题 1 的修改说明。同时，经过修改后，本研究在统计结果上发生了一定的量表，但并未影响结论的定性。

#### 审稿人 2 意见：

**意见 1：**在研究立意，也就是研究的意义上，若是想考察训练（或者说本文所称的 VLTM 的表征激活程度）是否能影响 VSTM，则本文同 Zimmer 等（2012）以及其它高强度训练发现对 VSTM 存在影响的研究而言，研究结果新的地方在哪里，提供的新的信息在哪里？或者说研究的价值在哪里？若仅仅是为了探究 Jiang 实验室的部分系列实验（通过特殊的操作他们实验室也有研究发现了类似的效应）没有发现 VLTM 对 VSTM 的原因，则很奇怪为何不

采用与 jiang 等相类似的实验材料以增加可比较性？

回应：首先，虽然一些研究者猜想决定 VLTm 是否能够促进 VSTM 的是 VLTm 的激活度，但是，很少有研究通过系统地控制 VLTm 的激活度来验证这一可能性；其次，发现 VLTm 对 VSTM 的促进作用的研究主要是以熟悉/陌生面孔为材料的，而这些材料往往具有较强的语义和情境背景，因此，通过系统地控制 VLTm 的激活度，并采用随机的几何图形能更为有效地验证上述可能性。最后，在发现 VLTm 对 VSTM 具有促进作用的研究中，绝大多数侧重于解释 VLTm 对促进 VSTM 的编码的作用，本研究进一步探索 VLTm 对 VSTM 保持的意义。修改稿在引言最后一段已对此进行了适当修改，详见引言最后一段。

意见 2：行文上部分逻辑不顺畅，更像文献的堆砌而非上下有逻辑的阐述。以第一、二段为例：第一段中，第 1，2 两句讲的是 VSTM 的容量问题，结果第三句就突然转到 VLTm 和 VSTM 的分离了，显得很突兀。建议在第段后顺着第二句“容量并不固定”从 VLTm 在其中可能的作用来引入 VLTm。这样可能逻辑上更顺畅。类似，第一段最后一句突然从 VLTm 和 VSTM 的分离转到“近年来。。。 ”也是较为突兀，望修改。第二段最后一句同整段的逻辑联系在哪里？阐述这个目的在那里？望修改。值得一提的是，我们这里仅指出了第一、二段存在的问题，这并不代表该问题仅仅存在于第一段，请参照这一问题对全文进行梳理。

回应：首先，第一段中虽然提到了 VSTM 容量的问题，但实质上并没有对本文起到作用，因此修改稿删除了这些无关内容，并以本文研究问题提出的理论背景为出发点，使第一段与全文及下文的联系更加紧密。同时，对其它衔接不连贯的地方也进行了修改，使得逻辑更加严密。

意见 3：实验设计上，实际上本文就是一个 3 水平被试间设计的单一实验，将之拆为三个独立的实验，无论是从逻辑上还是从读者理解的容易性上来讲都显得有些生硬，事实上很少可以在文献中看到将基线条件（实验 1）作为一个单独的实验来呈现。建议，这里整合为一个 3 水平被试间实验，无疑对读者清楚的理解本文的关键概念很有帮助。

回应：本文的三个实验的确可以合并为一个  $3(\text{ISI}) \times 3(\text{VLTm 激活度})$  的混合实验设计，其中 VLTm 强度为被试间因素，且合并后使得本文显得更加紧凑，故修改稿在第二部分对此问题进行了相应的处理，原文中“实验 1、2 和 3”分别以“组 1、2 和 3”代替，在进行统计分析与讨论时也进行了相应的调整。

意见 4：实验所采用的范式是局部变化检测范式，该范式存在两个较为棘手的问题，1) 无法排除被试利用 (rational stratage) 推理策略来完成任务，即只对比两幅图（而非若干个 WM 单元）的差异，来完成任务；2) 所测量的成绩可能反应了反应决策水平的差异而非 VSTM 容量上的差异，即 test probe 的项目越多越有可能相互干扰从而影响决策成绩。这两个问题导致了本文在揭示结果上的困难，即训练带来的作业成绩的提升究竟是反应了 VLTm 对 VSTM 的影响还是反应了 VLTm 对上述两类因素的影响？要解决这一问题可能需要更进一步的控制实验，具体可以参见 Luck 等 (1997) 的文章对这两种可能的控制实验。

回应：首先，本研究所采用的变化-检测范式参考和借鉴了相关研究，如：Jackson, M. C., & Raymond, J. E. (2008). Familiarity enhances visual working memory for faces. *Journal of Experimental Psychology: Human Perception and Performance*, 34(3), 556-568.

其次，在实验设计时我们也意识到这个问题，因此在指导语和实验过程中反复强调被试一定要对学习项目中的 6 个项目加以记忆，然后将之与检测刺激中的项目进行逐一比较，否则就达不到实验要求的成绩。假若被试根据推理策略来完成实验任务，例如，将学习刺激和检测刺激看成是两幅图片，那么被试也必需对两幅图片中的 6 个项目进行绑定，从而构成一幅整体的图片，但是，由于变化的图形和“是/否”反应均是随机的，被试事先并不知道什么样的检测刺激会出现，在形成图片的过程中被试必须要对所有的项目进行记忆。同时，检测项目持续的时间足够长（4s 为或按键后消失），被试完全可以在对检测刺激的知觉条件下来提取学习刺激中的项目而进行比较。

另外，若被试真的通过一定的推理策略来完成实验，则随着 ISI 的延长，被试的使用策略的空间更大，其成绩也会随之提高，而我们的研究数据表明，随着 ISI 的延长，记忆成绩反而下降，这是 VSTM 本身的明显特征。因此，我们认为，本实验反映的是 VLTM 对 VSTM 的影响。

意见 5：在实验材料制作上，颜色和形状特征分别只有 8 种，但在“各颜色和形状不重复”前提下，在重复学习实验中（实验 2，3），实际上存在一种可能，若被试意识到在部分的测试种只需要注意剩下的特征（如只要有任意剩余的两种形状特征或颜色特征出现都是 change 条件）即可很好的帮助反应，即判定为与之前记忆不同。因而研究的结果可能并不是反应了学习到的 VLTM 的激活促进了 VSTM，而是反应了该反应策略的学习。这部分可能需对被试进行问卷回访或者增加新的控制实验来解决。

回应：本实验只有 8 种实验刺激，因此，要记住每次实验中的 6 个项目，不如只记住本次实验中没有出现的两个项目或特征。但是，学习刺激中的 6 个项目是随机出现的，因此剩余的特征或项目也是随机的，要这样做，被试必须首先记住并排除已经出现的 6 个（这 6 个随机出现）才能寻找出没有出现的那两个，由于学习刺激呈现时间和 ISI 均相对较短，因此被试做出这种推理的空间不大，故他们只能依据对学习刺激和检测刺激的对比来完成实验。另外，我们在实验前也意识到这个问题，故在实验中反复强调被试一定要根据我们的实验要求来进行实验，否则被试也得不到报酬（组 2：20 元，组 3：30 元），而事实上，在实验完成后每名被试均获得了同样的报酬，且事后也详细询问了每名被试是如何来完成实验任务的，回访前告诉被试他们的真实操作过程与实验结果和报酬无关，要求他们对真实的实验操作过程进行回忆和说明，被试均表示是根据指导语来操作的，且实验后也对他们的实验操作进行了详细的询问。鉴于实验数据均为有效的，故在初稿中忽视了说明。

意见 6：除开将实验 1 作为基线外，可能同时也比较实验 2 和 3 的结果，即将实验 2 作为实验 3 的另一个基线条件也有一定的意义，实验 1 同实验 3 相比，虽然在关键的变量上（重复强度）上存在差异，但同时在可能刺激数目上也存在差异（实验 1 种每次随机组合，同一个 item 出现的概率很低），实验 2 无疑同实验 3 相比没有这部分的差异，却保留了关键变量上的差异（VLTM 形成程度）。

回应：修改稿将三个实验组合成了一个  $3(\text{ISI}) \times 3(\text{VLTM 激活度})$  混合实验设计，在结果分析时包含了对原来实验 2 和实验 3 的比较，且结果发现实验 3 的 VSTM 成绩与实验 2 的 VSTM 成绩达到边缘显著水平 ( $p=0.07$ )。

**意见 7:** 文中一直将好的 VLWM 成绩成为高激活的 VLWM, 这样的提法是否合适, 请斟酌。毕竟可能激活更多的是反映在提取阶段。是否考虑为其它的提法?

**回应:** 本文的确是以 VLTM 的好坏作为激活度的直接指标, 这主要因为对相关文献的参考和借鉴, 例如, 以面孔为材料的研究将名人面孔当作高激活水平, 而陌生面孔当作低激活水平。作者也试图以 VLTM 的“强度/质量”来代替“激活”一词, 但部分地方还是无法避免使用“激活”一词, 这样一来, 文中会同时出现“强度”和“激活”两词, 反而不利于文正的整体结构和理解。故修改稿保留了“激活”的概念。

**意见 8:** 文中实验设计部分称“根据 Logie, Brockmole 和 Jaswal(2011)的结论, 本研究所有实验中均设置 1s、1.5s 和 3s 三种间隔时间, 以保证 VSTM 充分形成。。。”, 但实际上 Logie 等的文中发现 ISI 要到 1.5s 后 (visual sensory memory) 的作用才趋于消失, 因而文中这一提法很容易误导读者, 建议修改, 并阐明该设置的相关理由。这里也涉及到结果解释上的一点不确定: “ISI 主效应显著,  $F(2, 46)=5.23, MSE=0.08, p<.01, \eta^2p=.19$ , 说明随着 ISI 的延长, VSTM 逐渐消退”, 因为上述理由, 这样的推理很难说一定成立, 望修改。

**回应:** 作者详细核实了原文中引用的这条文献, 的确发现引用有误, 在修改稿中删除了词条文献及相关内容, 由于要检验 VLTM 在 VSTM 的保持中的作用, 故设置了三种 ISI。

**意见 9:** 在对前人研究结果进行介绍的时候最好能说清楚关键的实验条件, 如“此外, Olson, Jiang 和 Moore (2005) 则发现, 由重复训练所形成的 VLTM 能够促进 VSTM, 他们在实验中要求被试判断检测刺激中的位置是否发生了变化, 结果表明, 当相同的位置重复发生变化时, 对其的检测成绩明显提高”, 处就很是让人疑惑, 基本上没有看过原文的人很难理解这究竟是在说什么。“当相同的位置重复发生变化时”并为适当的描述出 Olson 研究的关键操作。

**回应:** 已针对此问题进行了修改。

**意见 10:** “ $M \pm SD = 0.79 \pm 0.04$ ” 鲜少见这样的写法, 是否考虑简化 “ $M \pm SD$ ”?

**回应:** 已对文中类似的写法进行了更正, 改“ $M \pm SD = 0.79 \pm 0.04$ ”为“ $M = 0.79, SD = 0.04$ ”。

### 审稿人 3 意见:

**意见 1:** 请作者详细解释是如何计算变化与不变条件下的  $d'$  与  $K$  值。审稿人目前无从得知。

**回应:** 作者重新检查了原始数据、整理过程及计算过程, 结果发现, 由于数据较多, 且本研究的数据在处理时直接复制了另外几个多因素实验数据处理的表头和格式, 造成直接将不同条件下的正确率和错误率当成了击中率和虚报率, 且作者在稿件投出外审的过程中已发现此问题, 在此作者再次向各位审稿专家提出的宝贵意见表示感谢, 并对数据处理者的草率和粗心以及对信号检测论方法的理解不够透彻表示深深的歉意。修改稿中, 作者重新从原始数据开始检查, 将“变化”当作信号, “不变”当噪音, 从而求得两种条件下对应的击中率和虚报率, 进而计算出正确的  $d'$  和  $K$ 。

**意见 2:** 作者的前沿第一段存在逻辑跳跃, 请作者重新思考下。此外, “受到客体所承载的信息负荷(Alvarez & Cavanagh, 2004)、客体复杂性(Awh, Barton, & Vogel, 2007)”此处有误, 在 Alvarez 的论文中, 信息负荷和客体复杂度是等价的概念。而 Awh (2007) 则否认了客体复杂度的影响。

**回应:** 作者经过认真思考与重新阅读文献, 删除了无关内容及存在争议的文献, 如“受到客体所承载的信息负荷(Alvarez & Cavanagh, 2004)、客体复杂性(Awh, Barton, & Vogel, 2007)”等; 同时, 原文第 2 段本是要说明本文所采用的实验范式的, 但其与第一段内容不能很好地

衔接，本次修改对这一问题进行了修改，不专门对实验范式进行介绍，而将具体的实验步骤和评论穿插在方法与讨论部分，在保证前言部分的逻辑流畅的同时，又不失去应有的信息。

**意见 3:** 论文的图片质量较差，建议作者重新做过。

**回应:** 已对文中不清晰图片进行了修改。

**意见 4:** VLTM 是指代记忆系统中的一个模块，而作者很多地方却用以指代 VLTM 的表征。审稿人认为较为不妥，建议修改。

**回应:** 已对文中相关表述进行了修改，不再使用“VLTM 表征”一词。

## 第二轮

**审稿人 1 意见:**

**意见 1:** 作者在回答我第一轮审稿时提出的问题时说道：“由于数据较多，且本研究的数据在处理时直接复制了另外几个多因素实验数据处理的表头和格式，造成直接将不同条件下的正确率和错误率当成了击中率和虚报率”。但由于作者在初稿中的统计分析完全是和表格中的数据对应的，所以我觉得这不可能仅仅是表格的错误，而是作者之前根本没有理解信号检测论。

**回应:** 作者再次对信号检测论的理解及数据处理过程中的失误给审稿专家带来的不便表示抱歉，并感谢审稿专家的宝贵意见。

**意见 2:** 作者根据组 3 和组 1 之间的记忆绩效的显著差异认为高度激活的长时记忆对短时记忆保持起到了促进作用。但是正如另一位审稿人之前所说，组 3 和组 1 的记忆刺激是有差异的，组 1 中有  $8 \times 8 = 64$  个图形，而组 3 中只有固定的 8 个刺激。因此，组 3 的记忆任务难度要比组 1 小得多。这样以来，组 3 的记忆绩效比组 1 高也就很好理解了。即组 3 和组 1 之间的差异可能仅仅是记忆任务难度不同所致，而不是长时记忆激活程度不同所致。由于组 3 和组 2 之间只有边缘显著的绩效差异，所以本文的数据目前尚不能支持本文的结论“高度激活的长时记忆促进了短时记忆的保持”。

**回应:** 的确，本文组 2 和 3 中的实验刺激材料只有 8 个，而组 1 中的组合达到了 64 个之多，从数量上看，差距的确较大。但是，本文中，在一次变化-检测实验过程中，被试需要对短暂呈现的学习刺激中的 6 个客体进行记忆（根据现有研究文献及本研究结果，VSTM 容量低于 6），然后在此基础上来判断检测刺激是否与之相同，而非是对实验过程中全部的刺激进行记忆。另一方面，组 2 中同样采用的是 8 个客体，但其成绩却与组 1 没有差异，这也说明刺激材料的数目并不是导致组 1 和组 3 成绩差异的原因。

**意见 3:** 作者为什么在实验中设置三个 ISI？这个问题另一个审稿人之前也提过，作者的回答是“由于要检验 VLTM 在 VSTM 的保持中的作用，故设置了三种 ISI。”我想问的是：设置三种 ISI 到底如何能够帮助作者检测 VLTM 在 VSTM 保持中的作用呢？作者的实验假设是什么呢？从本文的统计分析结果来看，VSTM 绩效随着 ISI 增加而降低，同时 ISI 与 VLTM 激活度没有交互作用。这个结果完全可看作是 VLTM 激活度不影响 VSTM 的证据，因为如果 VLTM 激活度能促进 VSTM 的话，那么组 3 中的 VSTM 绩效应该是随着 ISI 增加而提高或保持不变，而不应该同组 1 一样随着 ISI 增加而降低。

**回应:** VSTM 的一个特点是记忆痕迹随着时间的推移而迅速消退，本研究设置三种 ISI 的目的是为了更大范围地探索 VSTM 随着 ISI 的增加而发生的变化。本研究是系列研究中的一部

分，其作用在后续研究中更加突出。

本文通过设置不同的 ISI 来考察高激活 VLTm 对 VSTM 的保持所起的作用需要借助 3 组实验的结果来看，主要是在不同激活度的 VLTm 背景下，ISI 主效应效应值的变化，在高度激活的 VLTm 中，ISI 的效应值要小于无 VLTm 和中度 VLTm 中的效应值，由此，说明在高度激活的 VLTm 背景下，VSTM 的消退受 ISI 的延迟的影响有所减小。然而，本文通过变化-检测范式所考察的仍然是对学习刺激中的 6 个客体的 VSTM，虽然组 3 具有高度激活的 VLTm 背景，但变化-检测范式中考察的毕竟不是 VLTm，因此，变化检测的成绩随着 ISI 的增加而下降，这是 VSTM 固有的特性，是量而非质的变化。

**意见 4：**有一些小错误。如在论文第 10 页：“(F 或 J，按键在被试间随机平衡，一半按 F 键表示“是”，另一半按 J 表示“否”)”这句话是不符合逻辑的，既然 F 和 J 在被试间平衡，那么就应该是有一半人按 F 表示“是”，另一半人按 J 表示“是”。还是在论文第 10 页，关于“否”反应和“是”反应的定义也是错误的。

**回应：**已对这些小的错误进行了修正。其中，本研究中的实验任务是要求被试判断检测刺激中是否有一个客体发生了变化，“是”反应代表检测刺激与学习刺激不同，即有一个客体发生了变化，而“否”反应表示检测刺激与学习刺激相同，不存在变化。

**审稿人 2 意见：**

**意见 1：**建议合理利用句号和逗号。如论文的第一段的第一句话是发一个非常长的句子。

**回应：**作者已对存在此问题的地方进行了修改。

**意见 2：**数据分析极端，审稿人认为直接进行 3\*3 的混合方差分析即可。

**回应：**本文从整体上而言的确是一个 3\*3 的混合设计方差分析，之所以还对三组实验单独进行分析，原因是考察在不同激活度的 VLTm 背景下自变量的效应值，并通过效应值的变化来判断 VLTm 激活度的作用。

**意见 3：**K 值的计算，由于采用全部报告法，应当采用 Pashler 的方法。

**回应：**首先，作者对前期采用不适当的公式表示道歉，在修改稿中，作者已对推荐文献进行了阅读，并采用了 Pashler 的容量计算公式，内容详见修改稿。

### 第三轮

**审稿人 1 意见：**

**意见 1：**建议合理利用句号和逗号。如论文的第一段的第一句话是发一个非常长的句子。

**回应：**作者已对存在此问题的地方进行了修改。

**意见 2：**数据分析极端，审稿人认为直接进行 3\*3 的混合方差分析即可。

**回应：**本文从整体上而言的确是一个 3\*3 的混合设计方差分析，之所以还对三组实验单独进行分析，原因是考察在不同激活度的 VLTm 背景下自变量的效应值，并通过效应值的变化来判断不同激活度的 VLTm 的作用。

**意见 3：**K 值的计算，由于采用全部报告法，应当采用 Pashler 的方法。

**回应：**首先，作者对前期采用不适当的公式表示道歉，在修改稿中，作者已对推荐文献进行了阅读，并采用了 Pashler 的容量计算公式，内容详见修改稿。