

《心理学报》审稿意见与作者回应

题目：物体颜色对情景记忆的影响

作者：周文洁；邓丽群；丁锦红

第一轮

审稿人 1 意见：

该研究通过 2 个 ERP 实验探讨了输入的颜色和记忆中的颜色知识是否一致对情景记忆的影响，设计简单明了。但仍然存在许多问题需要进一步澄清和改进，以下为具体审稿意见。

意见 1：文章题目为“物体颜色知识对情景记忆的影响”，情景记忆是一个较为广泛的概念，是对特定时间和空间场景下发生的事件的记忆。在文章前言部分并未对情景记忆的定义进行解释和阐述，也没有界定本研究中是研究对什么的情景记忆，建议修改题目和前言，具体说明是对何种情景记忆的影响。同时，按照作者定义，“颜色知识则是该物体的典型颜色信息（如草莓通常是红色的）；它从长期生活实践中获得，以语义或视觉形式存储在长时记忆中。”而研究前言花了大量篇幅来介绍颜色，而不是颜色知识，实验中的操作也是操纵“视觉颜色输入和物体颜色知识之间的一致性”，而不是操纵颜色知识，因此建议对题目进行修改，使之更为准确反映研究的内容。

回应：本文主要探讨物体表面颜色及其所激活的颜色知识对情景记忆中联结记忆编码和提取的影响。本来的思路是：适当的物体表面（视觉刺激）颜色可以激活记忆系统中的颜色知识；颜色知识在记忆编码和提取中发挥重要作用。原文中表述不够准确，应该是“物体表面颜色”而不是“颜色知识”。根据专家建议，我们将题目改成“物体颜色对情景记忆的影响”。本文所探讨的情景记忆主要是项目内联结记忆，即联结记忆。它是指项目与项目本身知觉特征捆绑形成的联结记忆，如单词与单词颜色、物体与物体位置、物体的颜色与形状捆绑成的联结记忆（Mayes, Montaldi, & Migo, 2007）。上述内容已在引言中进行补充和修改。

意见 2: 文章摘要部分,“实验结果表明,颜色在知觉和语义水平上均对项目编码产生影响;同时,颜色信息对物体的知觉表征影响更大,它对物体图片提取中的熟悉性和回想均产生影响;而只影响物体名称提取的熟悉性。此外,颜色影响物体的语义表征,支持激活-扩散模型理论。”建议说明具体的影响(如增大、减小、降低等),而不是直接表述“影响”两字。

回应: 描述不确切部分已在摘要中进行修改。

意见 3: 文章的引言部分有一些问题,首先,作者首先对以往的研究进行了总结介绍,然后直接提出本文的假设,并未对本研究的研究动机、研究意义以及创新性作出清晰的阐述,为什么要研究物体颜色知识对情境记忆的影响,这个研究的价值在哪,以及本研究与前人研究相比,有什么新的推进和发现,这些是文章前言的重要部分。建议作者对前言进行调整,阐述清楚本文的研究意义和创新性。再如,文章引言完全没有介绍与本实验相关的 ERP 研究,即使之前没有相关的 ERP 研究,也要讲清楚本研究想探讨的 ERP 成分,为什么实验中研究的心理过程与这个成分有关以及为什么要使用 ERP 技术来探讨这个问题,目前的文章中仿佛是没有任何假设,通过 ERP 技术发现了什么成分显著就报告什么成分。

回应: 根据专家的意见,我们对引言部分进行了较大幅度的修改。首先,介绍了颜色作为物体的基本属性,在物体识别、记忆表征与提取中的重要作用。其次,情景记忆的再认提取包含熟悉性和回想两个基本成分及其与不同水平加工之间的关系;熟悉性和回想的 ERP 特性。再次,目前研究存在的问题(研究动机),即人们对物体颜色信息的不同加工水平(知觉与概念)与情景记忆提取成分(熟悉性与回想)之间的关系还缺乏全面认识。一方面,颜色在记忆提取中的作用尚存在不同看法;另一方面,现有研究采用的是人工物体,而人工物体与颜色的联结是临时的,与自然物体的颜色在记忆中存储方式有很多差异。它不能反映颜色全面作用。最后,根据以往研究提出了假设,即颜色信息在知觉加工水平上可以促进记忆编码,而在概念水平上则表现出抑制作用;再认提取过程中,颜色信息在知觉水平上对熟悉性和回想均产生较大影响,而在概念水平上,回想受影响更大,熟悉性受到影响较小。同时,也对 ERP 结果进行了预期。上述内容已在引言中进行修改。

意见 4: P5“以往关于颜色促进效应的争议原因在于,一方面是由不同实验操作导致;另一方面,可能是忽视了物体颜色诊断性因素。”语句不通,建议修改。

回应: 已在原文中进行删除。

意见 5: 实验设计部分应当把最基本的自变量和因变量阐述清楚，脑电实验也应该把每种水平重复多少个试次明确呈现出来。

回应: 实验为被试内重复测量设计，编码阶段自变量为颜色一致性（颜色一致、颜色不一致），提取阶段为 2（颜色一致性：颜色一致、颜色不一致） \times 2（项目新旧：新项目、旧项目）被试内设计。已在原文中进行修改。

意见 6: 实验结果部分脑电数据直接按时间窗口命名报告，不合规范。同时地形图呈现不完全，如实验一学习阶段应当分别呈现学习阶段 380~450ms、510~600ms 下颜色一致和颜色不一致图片产生的地形图分布，而目前只呈现了两个图，不知是何种条件下的，再认阶段地形图以及实验二地形图呈现都不完全。

回应: 编码阶段的脑电结果较少，没有分开。提取阶段改成按熟悉性（FN400）和回想（LPC）两个成分报告。已在原文中更改。

实验目的主要是为了比较颜色一致和颜色不一致条件下的差异。因此，两个实验的编码阶段只选择了颜色一致项目和颜色不一致项目存在差异时窗的差异波地形图。再认阶段只分别选择了颜色一致和颜色不一致条件下相应时窗的新旧效应（旧-新）的地形图。

意见 7: 在实验 1 的学习阶段，相比于颜色不一致图片，被试对颜色一致图片的颜色判断正确率更高；而在实验二中，颜色不一致名称的颜色判断正确率更高。作者给出的解释是，在实验二中，“每个人心理表征中某个物体的典型颜色可能存在差异，因此，在面对颜色一致名称时，判断出现因典型颜色知识偏差而引发的正确率下降”。如果按照这个逻辑推断，那实验一也应当是颜色不一致图片的判断正确率更高，当图片的经典颜色与个人心理表征的典型颜色存在差异，应当也会造成干扰导致准确率下降。那为什么图片和名称会造成这种不一致的结果，应当在讨论中阐述清楚。

回应: 实验 1 与实验 2 实验材料不同，引发的认知过程也不同。实验 1 中，直接呈现物体图片，颜色一致条件下，物体图形直接激活视觉颜色知识，形成物体表征，从而更快的反应和更高的正确率。颜色不一致条件下，视觉颜色知识与表面颜色冲突，降低正确率并增加反应时。实验 2 中，呈现的是不同颜色一致性的物体名称，被试需要通过名称激活长时记忆中的语义颜色知识，进而形成物体的心理表征，不同于实验 1 更直接的心理表征形式，实验 2 被试需要通过名称形成颜色表征，随后进行判断。在这个过程中引入了个体的典型颜色知识偏差。另一方面可能是因为实验 1 的材料包含了颜色、明暗等细节信息，而实验 2 中采用物

体名称材料，颜色更为的单一，与个体的自动激活的颜色知识产生偏差。在实验 2 讨论中做了如下修改：

“实验 2 中，物体名称的颜色以视觉方式输入；而物体名称则激活物体不同特征的语义表征为主（Collins & Loftus, 1975）。视觉输入所形成的颜色表征与物体名称所激活的自然颜色知识（语义表征）之间的比较过程需要更多转换（Kelter et al., 1984），因而耗费更多资源。这反映在实验 2 学习阶段的反应时远大于实验 1（平均相差 97ms）上。此外，由于记忆中物体颜色表征具有多重性（Vurro, Ling, & Hurlbert, 2013），即同一物体可能有多种典型颜色，如香蕉的典型颜色既有黄色，也有绿色；但不太可能是紫色。因此，对紫色“香蕉”的否定判断要比对绿（或黄）色“香蕉”的肯定判断更加确定。此外，Lupyan（2015）认为，当输入精度较高时，表面颜色会由下至上更多的影响心理表征；反之，心理表征则更多受颜色知识影响。即实验 1 物体的心理表征更多的受表面颜色影响，而实验 2 物体的心理表征更多的受颜色知识影响。在判断准确性上表现出颜色不一致项目的辨认成绩优于一致项目。”

意见 8：P18 4.2 部分“在本研究的编码和提取阶段，物体颜色对图片识别与提取的促进作用都比对名称的作用更大；ERP 新-旧效应进一步印证了 Guillaume, Baier 和 Etienne（2020）的研究，即从早期编码到后期提取阶段，额区的 ERP 新-旧效应大小反映了项目-场景不一致性。”其中“编码和提取阶段，物体颜色对图片识别与提取的促进作用都比对名称的作用更大”这一发现说明了什么问题呢，前文有提到这图片和名称分属不同的加工，那这是说明了颜色对于视觉加工和语义加工的不同影响吗？造成不同影响的可能原因是什么？而“额区的 ERP 新-旧效应大小反映了项目-场景不一致性”其背后的逻辑又是怎样的呢？这些问题需要进一步澄清。

回应：尽管颜色在知觉加工和概念加工中都是以相同的视觉方式输入，但它在两种加工中产生作用的方式不同。这种方式上的差异不仅影响了编码过程（学习阶段），同时也对随后的提取（测验阶段）产生不同影响。对编码阶段而言，是编码或表征转换（概念加工）与不转换（知觉加工）的差异。而在再认提取过程中，知觉水平上的颜色信息比语义水平加工对熟悉性的影响更大；相反，回想受语义水平影响更明显。Guillaume, Baier 和 Etienne（2020）发现，项目-场景一致性作用在早期编码和后期提取阶段的 ERP 新-旧效应上都有所表现。与此类似，本研究中的项目-颜色一致性作用也均对编码和提取产生了影响。上述内容已在原文中进行补充。

意见 9: 另外，文章还有许多小细节阐述不够清晰，如 P7“另外，再请 40 名大学生对图片物体颜色与现实物体颜色的一致性进行 5 点等级评定，将一致性等级大于 4.0 的定义为颜色一致图片，一致性等级小于 2.0 的定义为颜色不一致图片，随机选取一种条件的图片进入实验材料库，共选取 256 张物体图片作为实验材料。”，5 点等级评定应当解释清楚评定的数字是从几到几，并且说明等级越大是越一致还是越不一致。另外“随机选取一种条件的图片进入实验材料库”的说法令人疑惑，不明白是何意思；P16 600~800ms 处少一个 s；文中参考文献存在标点符号中英混用、符号重复等问题。建议对文中细节进行检查修改。

回应: “随机选取”指的是随机指定物体颜色为一致或不一致。具体做法是：每一种物体均制作了两个颜色版本（颜色一致与颜色不一致），随机从两种颜色版本中选一种物体作为实验材料。原来的表述不清晰之处已在原文中进行修改，其他方面已在原文中修正。

审稿人 2 意见:

在该研究中，作者采用两个脑电实验考察了颜色信息在物体记忆编码与再认过程中的影响。研究者通过操控颜色与客体自然颜色的一致性，分别采用客体刺激图片与文字考察颜色在知觉和语义水平上对加工项目的影响。结果发现，颜色在知觉和语义水平上对于项目的编码和再认均有影响，在脑电信号中对应于 N400 和 LPC 两个成分的差异。该研究的文字和图表的描述清晰，实验设计合理，探讨的问题具有一定的理论意义。但当前的文章中还存在如下一些问题：

意见 1: 问题的提出和意义不明确。根据作者总结的前人研究结果可以推测，知觉和语义水平上的颜色知识加工可以对记忆编码产生促进作用。然而，对于知觉和语义水平分别对于记忆再认和回想存在差异是基于什么提出的？作者提到一句，前人研究中发现“再认的熟悉性和回想成分对于知觉和语义加工有特异性”，这个特异性具体是什么？和当前的假设提出有什么样的关系（或异同）？请作者进一步说明。

回应: 根据专家的意见，我们对引言部分进行了较大幅度的修改。首先，介绍了颜色作为物体的基本属性，在物体识别、记忆表征与提取中的重要作用。其次，情景记忆的再认提取包含熟悉性和回想两个基本成分及其与不同水平加工之间的关系；熟悉性和回想的 ERP 特性。再次，目前研究存在的问题，即人们对物体颜色信息的不同加工水平（知觉与概念）与情景记忆提取成分（熟悉性与回想）之间的关系还缺乏全面认识。一方面，颜色在记忆提取中的

作用尚存在不同看法；另一方面，现有研究采用的是人工物体，而人工物体与颜色的联结是临时的，与自然物体的颜色在记忆中存储方式有很多差异。它不能反映颜色全面作用。最后，根据以往研究提出了假设，即颜色信息在知觉和语义水平上都可以促进记忆编码；而再认提取过程中，颜色信息在知觉水平对熟悉性的影响更大；相反，回想受语义水平加工影响更明显。同时，也对 ERP 结果进行了预期。上述内容已在原文中进行修改。

有关“特异性”的内容表述不清，已在文中进行更正。原意指大脑左右半球具有加工的不一致性，大脑右半球偏好加工知觉信息，左半球偏好加工抽象记忆表征。客体感知特征的改变会影响与熟悉性相关的 FN400 成分；进行语义、抽象信息检索时，半球不对称性更多的影响回想，而不是熟悉性（Küper & Zimmer, 2018）。

意见 2：脑电数据分析的方法部分存在信息缺失。请作者说明学习阶段和再认阶段的时间窗是根据什么选择的，为什么 250~400ms, 400~600ms, 600~800ms 分别可以对应到语义激活、熟悉性和回想三个认知过程？另外，这些时间窗的选择是否都是相对提示信号（+）为零点的，也请作者阐明。

回应：针对时窗选择问题，我们参考以往研究，重新进行了划分并进行相应的统计分析。说明如下并已添加到“2.1.5 脑电数据分析方法”中：

“在学习阶段，刺激呈现 300~500ms 期间，完成物体表征（Schendan & Kutas, 2002）；在词汇理解中这正是获得词义的 N400 所在时段（Balass, Nelson, & Perfetti, 2010）。参照 Lu 等（2010）的研究，本研究选择 375~475ms 时段作为 N400（物体识别早期阶段）的时窗。同时，将 500~700ms（P600）作为物体晚期识别的指标（Schendan & Kutas, 2002, 2003）。在提取阶段，则根据 Curran（2000）及 Rugg 和 Curran（2007）等研究分别选取刺激呈现 300~500ms 和 500~800ms 作为熟悉性和回想所对应的时窗。”

时间窗的选择是以刺激开始呈现为零点。

意见 3：行为结果的几个问题：

（1）实验一学习阶段采用颜色比较任务（与物体自然颜色是否一致），颜色不一致的情况下被试的行为表现较一致情况下显著更差，这个结果是否说明，当颜色不一致的刺激，被试不能很好的辨别其与自然颜色的差异，而倾向于报告“一致”？这是否暗示，测试阶段再认-旧项目条件下发现的颜色一致 vs. 不一致差异是由于编码阶段的差异导致的？

(2) 实验二中学习阶段的颜色一致性条件间差异与实验一中的结果相反。作者对实验二的解释是“当颜色一致时存在较大的个体颜色知识偏差，而当颜色不一致时，所用颜色的差异性反而减少了颜色本身的影响”(p16)，以及“实验二以语义加工为主，激活颜色需要耗费更多资源(p17)”。这些解释似乎更倾向于预测两个颜色条件间没有差异，对于如何能够直接导出当前结果并不明确。

(3) 图 3，图 7 中的 figure legend 颜色（浅蓝 vs.白色）与柱状图颜色（浅蓝 vs.深蓝）不一致。图 3a 中再认-旧项目的柱状图颜色和其他柱状图颜色不一样（浅蓝 vs.白色）。请作者检查。

回应：

(1) 学习阶段，呈现颜色不一致刺激时，会激活长时记忆中的物体颜色知识，与当前感知冲突，增加认知负荷，干扰编码过程。提取阶段颜色不一致条件下行为结果较差可能同时来源于两方面，一方面来源于编码阶段的差异；另一方面，进行信息检索时，被试刚刚形成的形状-颜色联结会受到长时记忆中固有的较为牢固的形状-颜色联结干扰，阻碍提取。即高颜色诊断性客体的颜色知识同时影响情景记忆的编码和提取阶段。

(2) 实验 1 与实验 2 实验材料不同，引发的认知过程不同。实验 1 中，直接呈现物体图片，颜色一致条件下，物体图形直接激活视觉颜色，形成物体表征，从而更快的反应和更高的正确率。颜色不一致条件下，视觉颜色与表面颜色冲突，降低正确率并增加反应时。实验 2 中，呈现的是不同颜色一致性的物体名称，被试需要通过名称激活长时记忆中的语义颜色知识，进而形成物体的心理表征。不同于实验 1 更直接的心理表征形式，实验 2 中被试需要通过名称形成颜色表征，随后进行判断，在这个过程中引入了个体的典型颜色知识偏差。另一方面可能是因为实验一的材料包含了颜色、明暗等细节信息，而实验 2 中采用物体名称材料，颜色更为的单一，与个体的自动激活的颜色知识产生偏差。在文中做了如下修改：

“实验 2 中，物体名称的颜色以视觉方式输入；而物体名称则激活物体不同特征的语义表征为主(Collins & Loftus, 1975)。视觉输入所形成的颜色表征与物体名称所激活的自然颜色知识(语义表征)之间的比较过程需要更多转换(Kelter et al., 1984)，因而耗费更多资源。这反映在实验 2 学习阶段的反应时远大于实验 1（平均相差 97ms）上。此外，由于记忆中物体颜色表征具有多重性(Vurro, Ling, & Hurlbert, 2013)，即同一物体可能有多种典型颜色，如香蕉的典型颜色既有黄色，也有绿色；但不太可能是紫色。因此，对紫色“香蕉”的否定判断要比对绿（或黄）色“香蕉”的肯定判断更加确定。此外，Lupyan (2015) 认为，当输入精度较高时，表面颜色会由下至上更多的影响心理表征；反之，心理表征则更多受颜色知识影

响。即实验 1 物体的心理表征更多的受表面颜色影响，而实验 2 物体的心理表征更多的受颜色知识影响。在判断准确性上表现出颜色不一致项目的辨认成绩优于一致项目。”

(3) 已修改。

意见 4: 脑电结果的几个问题:

(1) 结果的报告方式有待统一，如 p10: 380~450ms 结果部分，作者没有报告脑区位置的主效应，而是直接比较了“额区和额中区 vs.中央区与顶区”的差异；而在 510~600ms 结果部分，作者只报告了脑区位置主效应，对于“额区和额中区 vs.中央区与顶区”的差异缺少相应的统计结果。在“再认阶段 ERPs 分析”中也存在相似的问题。

(2) 实验二中采用文字刺激，而文字本身和颜色信息的绑定相对于自然客体图片应该更弱。建议作者在引言中，或实验一与二的过渡部分阐明实验二的必要性，以及与实验一的差异。

回应:

(1) 已在文中补充修改。

(2) 已在引言和实验 2 过渡部分中进行修改。

意见 5: 作者在结论中指出(p19)，“颜色在知觉和语义水平上均会对项目编码产生影响，颜色一致有助于视觉系统快速、有效地辨别物体”。但该结论与实验二的结果并不相符（实验二中学习编码阶段，颜色一致性对应的行为比不一致条件更差）。

回应: 原来的结论仅从 ERP 结果上考虑，编码阶段颜色一致条件下，被试有更正的波幅，表明相关语义知识和物体名称的激活，即颜色一致促进了颜色语义知识的激活。而忽视了行为结果，存在缺陷。因此，结合行为结果，将结论改为“（1）颜色在知觉和语义水平上对项目编码产生不同影响，颜色一致有助于知觉水平的物体辨别，而阻碍语义水平的辨别物体。”

第二轮

审稿人 1 意见:

作者较好地回复了本人第一轮审稿意见，以下为一些补充意见:

意见 1: 说明提到了有双侧乳突参考电极，但脑电的 offline 参考方式没有交待清楚，平均参

考还是双侧乳突或是其他？

回应：离线分析是以双侧乳突的平均电压值进行重新参考。已添加在“2.1.5 脑电数据分析方法”中。

意见 2：全文结果的文字报告中作者都没有提供具体的数值，只有图，而图上很多地方读者是看不清楚各个条件的具体数字，因此有必要报告这些数值。

回应：具体数值（主要是行为结果）已在原文中进行增补。

意见 3：两个实验中各个脑电成分的时间窗口选取“随意性”很大，例如图 5 中作者将 300~500ms 和 500~800ms 定义为两个不同的脑电成分，500ms 这个转折点的选取看起来很主观。作者提到有前人参考文献，但在这个领域中，是否所有文献都选取这些时间窗来定义这两个脑电成分？作者可能需要承认这点是这个领域 ERP 研究的局限，或者提供数据驱动的方法（例如聚类，主成分分析等）来说明这些时间窗的确对应了不同的脑电成分，反映了不同的心理加工过程。

回应：是的，在脑电研究中，有时时窗划分会存在一定差异。在记忆提取的相关研究中，大多数研究将再认提取分为熟悉性和回想两个基本成分（Brainerd, Reyna, & Kneer, 1995; Hintzman & Curran, 1994; Jacoby, 1991; Mandler, 1980; Yonelinas, 1994）。熟悉性是对信息强度（量）的提取（Evans & Wilding, 2012; Yonelinas, 2002），主要表现为刺激呈现后 300~500ms 左右时间窗额区（FN400, Frontal Negativity）的 ERP 新旧效应，峰值大约出现于刺激呈现后 400ms（Curran, 2000; Rugg & Curran, 2007）。而回想则是对记忆信息性质（质）的提取（Evans, & Wilding, 2012; Yonelinas, 2002），更多指向概念水平加工（Bruett & Leynes, 2015; Yonelinas, 2002），主要表现为刺激呈现后 500~800ms 左右时间窗顶区（LPC, Late Positive Component）的 ERP 新旧效应，峰值大约出现于刺激呈现后 600ms（Curran, 2000; Rugg & Curran, 2007）。虽然个别实验可能因任务不同，时窗划分存在差异，但大多数记忆提取研究均采用 300~500ms、500~800ms 这两个时窗来观察 FN400 和 LPC 成分，这基本上是研究者的共识。在本文第一版文稿中结合以往研究和本实验数据，将回想成分划分为早期、晚期两个部分进行分析。但在后续修改过程中，考虑到要和传统研究相一致，提高可比性，因而仍然沿用了经典的 300~500ms、500~800ms 的时窗划分方法。这种时窗划分的变化对实验结果和结论没有很大改变。

.....

审稿人 2 意见：

作者已基本针对之前的意见进行了修改。

作者在结果部分仅报告统计结果（如有无主效应和交互作用），但对于结果表明的含义没有说明（直到讨论中才提及）。建议作者在每次报告结果后，补充类似“该结果表明 / 提示”之类的解释性语句，增强读者对结果的理解。

回应：已在稿件的“结果”中增加了相应的统计结果的初步解释，以增进对数据结果的理解。