

《心理学报》审稿意见与作者回应

题目：项目间语义可整合性对联结再认的影响——来自 ERPs 研究证据

作者：律原、梁九清、郭春彦

第一轮

审稿人 1 意见：

意见 1：以往研究表明了在联结再认记忆实验的编码阶段，项目间“整合性”的不同会影响记忆的提取过程。本研究与以往研究的不同在哪里，即使通过操纵词对使其分为成语和非成语，但终究还是归于项目间的不同“整合性”？

回应：感谢审稿专家的问题。我们的研究与前人相关研究的不同主要表现在以下两点：（1）实验所用的刺激材料：我们选择了非常具有我国文化特色的四字成语来形成高语义整合度的条件，使用这种材料可以使我们的被试更容易将两个图片形成一个整体。据我们查阅相关文献，还没有与我们使用材料相类似的研究。更为重要的是通过使用这样的材料我们有可能深入探讨提高项目间语义可整合性对项目本身提取和项目间关系提取的不同影响，在我们的研究中成语条件的“旧”反应是与传统新旧效应中的“旧”反应相对应的，即相同的项目、相同的关系（测验阶段同时出现的项目在学习阶段出现过，而且都可以构成成语）；而成语条件中的“重组”反应则是我们研究中的特色，与以往研究中的“旧”、“新”反应都不同，成语条件中的“重组”反应对应的关系为“旧”项目加新语义关系（学习阶段与测验阶段呈现的单个项目是相同的，但学习阶段与测验阶段两个项目间的关系发生了改变，学习阶段呈现的两个项目可以构成成语而测验阶段同时呈现的两个项目则不能构成成语），而成语条件“新”反应对应的两个图片本身虽然能构成成语但是呈现的项目本身却是学习阶段没有见过的新项目，与传统“新/旧”效应的“新”反应对应，非成语条件的“旧”反应与非成语条件的“重组”反应则同属于旧项目加旧关系，即项目本身在学习阶段和测验阶段都呈现过，而且不论是“旧”反应还是“重组”反应同时出现的项目都不能用成语来联系，表现出相同的语义整合性，非成语条件下的“新”反应对应的项目没有在学习阶段呈现过而且同时出现的项目也不能构成成语。通过比较成语条件、非成语条件“旧/重组”效应和成语条件与非成语条件“新/旧”效应差异，我们就有可能发现联结记忆中项目语义可整合性的改变对项目提取和项目间关系提取的不同影响。

（2）我们不仅分别探讨了成语条件和非成语条件下新/旧效应和旧/重组效应的 ERPs 差异，还比较了“旧”反应、“重组”反应和“新”反应在成语条件与非成语条件下的 ERPs 差异，并通过地形图标准化的方法分析了成语条件“新/旧”效应与非成语条件“新/旧”效应地形图分布的差异和成语条件“旧/重组”效应与非成语条件“旧/重组”效应地形图分布的差异。通过上述多角度比较分析，我们发现提高项目间语义可整合性确实提高了联结记忆中基于熟悉性和回忆的提取。但是项目间语义可整合性的改变，只对 400~800ms 时间段顶区的“旧/重组”效应发生影响；而对 200~400ms 成语条件和非成语条件的“旧/重组”效应与“新/旧”效应以及成语条件 400~800ms 的“新/旧”效应没有影响，这些结果表明：测验阶段 200~400ms 的脑神经活动与项目提取相关而 400~800ms 的脑神经活动与项目间关系的提取有关。原文中我们对研究目的和结论的表述确实存在不清楚的地方，在修改稿中我们在“1.引言”的最后一段中对研究目的进行了重新阐述，在“4.结论”处也进行了相应的修改（已经用蓝色标出）。请专家再次审阅，谢谢。

意见 2：成语类实验材料的选取如何保证被试在短时间内能够将其整合为成语，是否将选取

的可整合为成语的词对进行前期的测试，再根据测试的结果选取最终的实验材料？

回应：感谢审稿专家的问题。这个问题也是我们在准备实验材料的过程中特别注意的问题。我们采取了2项措施保证了被试能在短时间内将成语条件的图片整合为成语。（1）在准备实验材料阶段我们准备了240对图片对，其中120对图片对可以构成成语，120对不能构成成语。我们随机抽取了不参加正式实验的40名在校大学生对这240对图片进行了图片对的评定。他们首先要评定图片对能否联想成成语。对于可以联想成成语的图片对要写出这些图片对所对应的成语，并评定每对图片与成语的贴切程度（贴切程度用5点进行评价，点数越高说明图片与成语的符合程度越一致），正式实验中我们所使用的60对图片的贴切程度的评定都在4点以上。（2）在正式实验中，成语条件和非成语条件是分开进行的，在成语条件的block中，学习阶段被试需要使用出声报告的方式报告与图片对相对应的成语，这样就能确切知道被试是否将图片对与相应的成语进行了整合。此部分我们已经添加在修改稿中“2.2 实验材料”部分中，并用蓝色字体标出，请专家再次审阅。

意见3：地形图分布分析表明，对于“新/旧效应”，成语条件与非成语条件的头皮分布没有显著差异；对于“旧/重组效应”，成语条件与非成语条件的头皮分布也没有差异。为何做地形图的分析以及这些无差异说明了什么问题？

回应：感谢审稿专家的问题。在我们的研究中成语条件与非成语条件不但是两个不同的整合水平，而且结合两种条件下“旧”、“重组”和“新”3种不同反应，我们认为存在“旧项目+旧关系”（成语条件“旧”、非成语条件“旧”和非成语条件“重组”）、“旧项目+新关系”（成语条件“重组”）、“新项目+旧关系”（成语条件“新”、非成语条件“新”）三种不同的情况。我们想分别了解提高项目间语义可整合性对联结记忆中项目提取和项目间关系提取的不同影响。所以根据McCarthy和Wood在1985年的研究将成语条件和非成语条件地形图差异波进行标准化以后进行了统计分析，结果发现成语条件和非成语条件“新/旧”效应的地形图分布不存在统计差异；这说明本研究中成语条件下“新/旧”效应和非成语条件下的“新/旧”效应在电极分布上是一致的，也就是说两者表现出相同的脑神经认知过程；同理，由于成语条件和非成语条件“旧/重组”效应的地形图分布也不存在统计差异，说明成语条件下“旧/重组”效应和非成语条件下的“旧/重组”效应在电极分布上是一致的，两者也表现出相同的脑神经认知过程。修改稿中我们在“3.2.4 地形图分析”的第二段和第三段（用蓝色标出），增加了关于此部分的说明，请专家再次审阅，谢谢。

意见4：研究结论为项目可整合性影响联结记忆效率，提高项目间的可整合程度可以促进熟悉性和回忆在联结再认中的作用，也可能提高了概念启动在联结再认中影响。熟悉性和概念启动是两种不同的概念，也是一个争论。研究的目的在于解决争论，为何根据实验结果做出讨论后仍然下一个两可式的结论？

回应：正如专家所言，熟悉性和概念启动是两种不同的概念，也是当前研究的热点问题之一。然而，本研究的目的在于通过操纵编码阶段的项目语义可整合性，来探讨其对联结记忆的影响。不是研究熟悉性和概念启动之间的争论。这个问题在文中表达不够清晰，根据您的意见，在前言部分和讨论部分对本研究目的做了进一步说明，请您审阅（见蓝色部分）。

本研究在讨论部分既提到了熟悉性也提到了概念启动，主要是由于以下原因：

（1）根据双加工理论，在情景记忆提取中存在两种加工：熟悉性和回忆。相关ERPs研究认为在刺激呈现后大约300~500ms左右，在额区出现的“新/旧效应”或称为“FN400效应”，指示熟悉性；刺激呈现后大约500~800ms左右，在顶-枕区出现的“新/旧效应”，指示回忆。（见修改稿“1.引言”中的第二自然段）

在本研究中，我们对200~400ms和400~800ms的ERPs数据进行分析，发现测验阶段，

在刺激出现后约 200ms，无论是成语条件还是非成语条件，均出现“新/旧效应”，持续到 800ms，并且成语条件比非成语条件有更广泛地分布；对于成语条件，刺激出现后 400~800ms 出现“旧/重组效应”，而非成语任何时段均没有出现这种“旧/重组效应”；对于“旧”和“重组”两种反应，刺激出现后约 200ms 开始，成语条件引起的 ERPs 比非成语条件更正；对于“新”反应，两者无显著差异。结合已有的支持双加工理论的 ERPs 研究结果，我们使用熟悉性来解释早期的 200~400ms 的“新/旧效应”，而用回忆来解释晚期的 400~800ms 的“新/旧效应”。这样，可以推知，项目间的语义可整合性高的成语条件之所以促进了联结记忆提取效果，是因为提高了熟悉性和回忆加工。这与先前的大部分研究一致，即提高项目的可整合性会提高熟悉性的作用。（见修改稿“4.讨论”中的第二、三、五自然段）

(2) 之所以我们也说可能反映了概念启动，是因为正像您所说，熟悉性和概念启动还存在争议。的确，关于早期 FN400“新/旧效应”反映了熟悉性还是概念启动是目前研究的热点问题之一。大部分的学者支持了前者，认为反映了熟悉性；但是也有学者（如 Paller 和 Voss 等人）认为其反映了概念启动。针对这一争议，我们实验室的其他研究者专门对这一问题进行了研究，采用甲骨文、金文等古文字作为实验材料，通过让被试进行意义评级，将这些材料分为高意义和低意义两类，采用 R/K 范式、信心判断范式等进行，发现高意义的材料存在 FN400“新/旧效应”，而低意义的材料不存在该效应。相关论文在 Brain Research 上发表 (Neural correlates of familiarity and conceptual fluency in a recognition test with ancient pictographic characters, 2013, 48-60)。然而，我们目前的研究并不试图去解决这个争议，也无法解释这个争议。我们主要是探讨项目可整合性对联结记忆的影响。对于本研究结果讨论时借用他人的结果来进行讨论和解释。为了更全面的讨论其背后的机制，我们认为项目语义可整合性只所以提高了联结记忆成绩，也不排除有语义上概念启动的影响。针对专家的这一问题，我们对前言和讨论中的内容进行了修改，请审阅。

意见 5: 摘要需要精炼地从四个要素概括，结论部分只需精炼地报告研究的主要结论。

回应: 感谢审稿专家的建议。在修改稿中根据专家的意见我们重新撰写了摘要，请专家再次审阅。

意见 6: 实验仪器及其参数未报告，统计检验力未报告

回应: 感谢审稿专家的建议。实验仪器及其参数我们在“2.4 ERPs 记录及数据处理”中进行了报告（已经用蓝色标出），在修改稿中，我们报告了所有统计量的统计检验力，所有的检验力已经用红色字体标出。

审稿人 2 意见:

意见 1: 本研究以视觉图片为材料，操作记忆图片间的“语义可整合性”，来探讨整合性对联结再认影响的神经机制。具体讲，作者采用成语图片对和非成语图片对构成两类整合程度不同的刺激材料进行联结再认实验，再认时设置“旧”、“重组”和“新”3 种情况，同时记录 ERP。结果发现记忆刺激的可整合程度对早期项目提取（200~400ms）和晚期关系（400~800ms）提取都有影响。论文书写较规范，参考文献引用全面。

回应: 感谢审稿专家的肯定。

意见 2: 前言过于繁琐，读起来抓不住重点，建议进一步理清脉络。另外，本研究的研究动机、理论意义缺乏阐述，仅仅在前沿最后一段的第一句简单讲“然而，以往研究大多以词或图片为材料，探测熟悉性在不同类型联结记忆中的作用。那么，在相同的维度内，随着整合性程度的提高，是否熟悉性的作用也会更加突出”。研究这个问题的意义在哪？为什么要研

究这个问题？审稿人认为这些方面均需要进一步加强或者阐述。

回应：感谢审稿专家的问题。我们研究中设置的成语条件与非成语条件不但是两个不同的整合水平，而且结合两种条件下“旧”、“重组”和“新”3种不同反应，区分了“旧项目+旧关系”（成语条件“旧”、非成语条件“旧”和非成语条件“重组”）、“旧项目旧+新关系”（成语条件“重组”）、“新项目新+旧关系”（成语条件“新”、非成语条件“新”）三种不同的情况。我们想分别了解提高项目间语义可整合性对项目提取和项目间关系提取的不同影响。

根据专家的意见我们对重新撰写了引言，对与本研究相关的文献重新进行了梳理，并明确了研究目的与研究假设（修改见“1.引言”部分的最后一段，修改的部分已经用蓝色标出）请专家再次审阅。

意见 3：统计繁琐，而且具体的统计与作者方法中的阐述不一致。作者在方法中提及“分析时,选取 12 个电极 (F3/Fz/F4/C3/Cz/C4/P3/Pz/P4/O1/Oz/O2)，以各时段的平均电压值为因变量,进行 2(条件：成语/非成语)×3(反应类型：旧/重组/新) ×12(电极：F3/Fz/F4/C3/Cz/C4/P3/Pz/P4/O1/Oz/O2)重复测量方差分析”（P7，第二段第二句话）。审稿人这是应该采取的分析思路。然后，在具体的分析中，作者完全抛弃了该思路，完全按照不同的思路进行分析，不仅让审稿人费解，而且会极大的提高犯一类错误的概率。此外，作者在统计分析时对结果的解释往往脱离具体的统计，让审稿人非常难以阅读。如行为结果分析时，作者发现主效应、交互作用均显著。因此，需要对主效应，交互作用做进一步分析，前者如果是 3 个水平需要做事后检验，后者需要做简单效应分析。然而，作者仅仅用“进一步分析表明”来讲具体的分析。该简单分析是如何进行的，针对的内容是什么，这个让审稿人花了很长时间去理清。

回应：感谢专家的意见，我们在统计分析过程中的实际统计顺序与专家指出的方法一致，但是在行文中确实有顺序不当和表述模糊的情况。现已经重新进行了修改，主要在“3.1 行为结果”部分和“3.2 ERPs 结果”部分，修改处已经用蓝色标出，请审稿专家再次审阅。

意见 4：作者研究中一个重要的概念或者变量是“整合性”。但是作者在论文中并未介绍何为整合性，以及如何考察整合性。这对于读者理解该研究是极其重要的，建议补充。

回应：感谢审稿专家的问题，联想记忆中同时出现的项目间的可整合性是指同时呈现的两个项目可以被整合成一个项目的程度。如果同时呈现的两个项目可以通过某种方式整合成一个项目，那么这两个项目在提取时就会更多的表现出与项目记忆提取相类似的特点。我们的研究中，用能否将两个项目用成语联系作为整合性高低的标准。如果两个项目可以被成语联系则认为项目间的可整合性高，而不能用成语联系则可整合性低。具体的考查方式是让在编码阶段报告出图片对所对应的成语。在修改稿中，我们在“1.引言”的第 4 段中添加了关于可整合性的意义的解释，请专家再次审阅。

意见 5：作者对于整合性的操纵，体现在呈现的图片能否构成成语上。审稿人认为，该种操纵在某种程度上可能可反应整合性的高低，但同时存在很多其他的解释，比如加工深度，语义加工的不同等。因此，该种操纵是否有效、能否回答作者的问题，至少需要讨论。此外，作者提到前人研究也有用图片的，那么该项研究相对于这些前人研究，有何区别？在讨论中，建议至少需讨论相关的内容。

回应：感谢审稿专家的问题。根据专家的意见我们在“4.讨论”部分增加了概念启动与熟悉性的相关的讨论，但是我们研究不关注提高项目间关系的可整合性是影响了熟悉性还是概念启动，而是想分别探讨项目间语义关系的改变对项目提取和项目间关系的影响，我们已经对引言和结论分别进行了修改，请专家再次审阅。

意见 6: 统计分析建议报告效果量；行为数据建议以图的形式而非表格的形式呈现，便于读者更加直观的看数据。实验的持续时间，有无休息需要报告。

回应: 感谢审稿专家的问题和建议，本研究的正式实验分为 8 个 block，其中 4 个 block 为成语条件，4 个 block 为非成语条件，每个 block 分为学习阶段和测验两个阶段，学习阶段由 15 个 trial 组成，每个 trial 的呈现时间为 5000ms，ISI 为 900~1100ms，在学习阶段完成后有一个持续时间为 60 秒的倒减三过程，然后进入测验阶段，每个测验阶段由 45 个 trial 组成，每个 trial 呈现时间为 2000ms，ISI 为 900~1300ms，每个 block 的持续时间约为 5 分钟，完成一个 block 被试可以休息 2 分钟。整个正式实验的持续时间约为 1 小时 20 分钟左右（含休息）。这部分已经添加到“2.3 实验设计及程序”的第 1 段中，用蓝色标出，并在“图 1：实验程序图”中进行了相应的标注，请专家再次审阅。实验的行为数据已经使用图的形式进行报告（参见图 2）

第二轮

审稿人 1 意见:

意见 1: 作者根据审稿意见对论文进行了认真修改，建议发表。

回应: 感谢专家的肯定。

审稿人 2 意见:

意见 1: 中文摘要过长，过于详细，学报要求中文摘要 200 字。建议作者精简。英文摘要建议找专业人士润色。

回应: 谢谢专家，我们对中文摘要进行了再次修改，并找相关专业人士对英文摘要进行润色。详见中英文摘要。

意见 2: P9 第一行(Rugg & Yonelinas, 2003)格式错误。

回应: 谢谢专家,我们对文中的内容作了修改,见引言第二自然段绿色字体部分。

意见 3: 图 2 图片质量较为粗糙，建议作者进一步发表前进一步修改。

回应: 谢谢专家，我们对图 2 及其它图进行了再次绘制，提高了作图质量，请审阅。

意见 4: 论文的理论贡献、创新性目前仍不是非常明确，如作者在前沿中提到，“在相同的语义加工层面，随着整合性程度的提高，是否熟悉性的作用也会更加突出？是否同时也增加了回忆的作用？联结再记忆的提取过程中对项目本身特征的提取和对两个项目间关系的提取是否对应着不同的脑神经机制”。个人建议，作者需要解释为什么解释这些问题重要以及并在讨论中讨论回答这些问题对已有理论的贡献在哪。此外，对特征的提取和对项目关系的提取，脑机制应该不一样，作者需解释为何提出该问题；此外，审稿人在正文中似乎未找到对该问题的回答。

回应: 谢谢专家提出的问题。我们认为专家的第 4 个问题可以拆分成 3 个小问题，需要分别进行回答。

问题①: “在相同的语义加工层面，随着整合性程度的提高，是否熟悉性的作用也会更加突

出？是否同时也增加了回忆的作用？

目前，熟悉性和回忆在联结记忆提取中的作用还存在争议。较早的研究认为联结记忆提取中仅有回忆的作用（Donaldson 和 Rugg, 1999； Hockley 和 Consoli, 1999； Yonelinas, 1997），而新近的很多研究发现，如果两个项目是可以整合的，熟悉性也可以有助于联结记忆提取（Yonelinas, 2002； Rhodes 和 Donaldson, 2008； Mollison 和 Curran, 2012）。另外，尽管操作了项目可整合性，不同的实验条件，熟悉性和回忆对联结记忆的影响还是有很大不同。如 Bader 等研究发现整合性的联结记忆只有熟悉性的作用，非整合性的联结记忆只有回忆的作用。

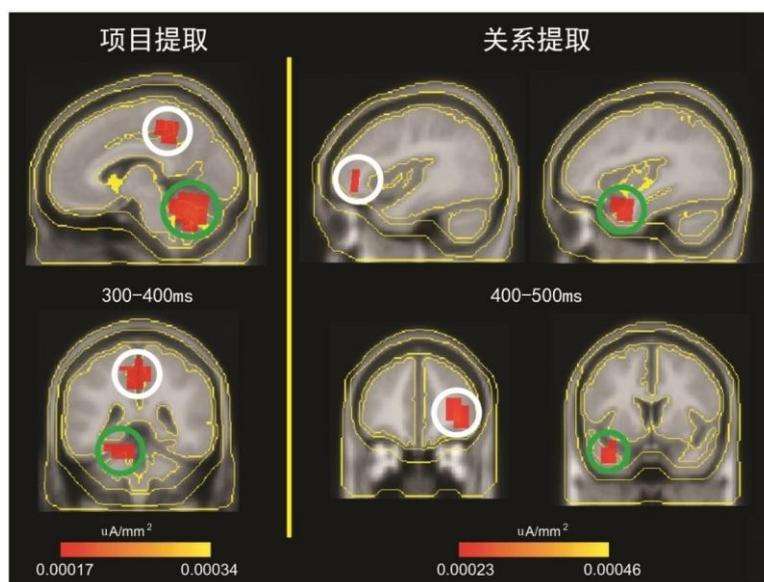
我们的研究表明熟悉性和回忆都对联结记忆都有贡献。当提高项目语义整合性时，会促进回忆的作用，也会促进熟悉性对联结记忆的效率。这些也会反映在 ERPs 早期额区新/旧效应和晚期顶区新/旧效应上。从这个假设出发，我们操作了图片材料，在编码阶段创新性地让被试进行成语判断任务，区分高项目语义整合的成语条件和低项目语义整合的非成语条件。结果：对 200~400ms 和 400~800ms 的 ERPs 数据进行分析，发现测验阶段，在刺激出现后约 200ms，无论是成语条件还是非成语条件，均出现“新/旧效应”，持续到 800ms，并且成语条件比非成语条件有更广泛地分布；对于“旧”和“重组”两种反应，刺激出现后约 200ms 开始，成语条件引起的 ERPs 比非成语条件更正；对于“新”反应，两者无显著差异。结合已有的支持双加工理论的 ERPs 研究结果，我们使用熟悉性来解释早期的 200~400ms 的“新/旧效应”，而用回忆来解释晚期的 400~800ms 的“新/旧效应”。这样，可以推知，项目间的语义可整合性高的成语条件之所以促进了联结记忆提取效果，是因为提高了熟悉性和回忆加工。问题②：作者需解释为何提出对特征的提取和对项目关系的提取，脑机制应该不一样，这一问题？

我们提出这一问题希望我们的研究能更好的用认知心理学解释现实生活的常见现象。现实生活中，我们接触的事物一般都不是独立存在的，而同时出现的事物都或多或少的存在某种语义联系。两个同时出现的事物如果具有某种明确的语义联系，则它们之间的关系发生变化时我们是很容易觉察这种变化的；与之相反，两个同时出现的事物如果没有明确语义联系，则它们之间的关系发生变化时我们可能很难发现这种变化。举一个例子，平时我们经常将牙膏和牙刷放在一起，那么如果将牙刷和球鞋放在一起我们能立即意识到这种变化，但是如果将牙刷和水杯放在一起可能就不太容易发觉这种变化。上面这个例子说明联结记忆中项目的知觉特征的改变和项目间关系的改变对联结再认的影响可能是不同的。我们的研究希望能在这一方面有所突破。为此我们在修改稿中增加了脑内溯源分析，证明联结记忆的再认过程中项目提取和关系提取在脑内的定位是同的。

问题③：联结再认记忆的提取过程中对项目本身特征的提取和对两个项目间关系的提取是否对应着不同的脑神经机制？

如果仅以地形图的电压值进行分析，尚不能回答有关“对项目本身特征的提取和对两个项目间关系提取”脑神经机制的不同。为进一步说明联结记忆中项目提取和关系提取的脑机制具有不同定位，在论文修改的过程中，依据 ERP 波形特点（见图 3），在 200~400ms 时窗，旧与重组条件的 ERP 重合，因此，选择成语条件下新/旧效应（ERP 差异波）作为项目本身

特征的提取；在 400~800ms 时窗，成语条件旧与重组条件的 ERP 出现分离，两者的 ERP 差异波反映的是项目间关系提取。利用 Curry6.0 软件，进行电流密度分析。使用 LORETA 溯源分析对‘项目提取’和‘关系提取’进行脑区定位。依据 LORETA 溯源分析原理，选择差异波最大 GFP 对应时间窗口（100ms）能较准确的反映出 ERP 定位特点。对成语条件新/旧效应 300~400ms 时间段（代表项目提取）和成语条件旧/重组效应 400~500ms 时间段（代表关系提取）进行了溯源分析。溯源分析的结果表明，在 300~400ms，项目提取主要定位于楔前叶（Precuneus）和齿状回（Dentate Gyrus）；在 400~500ms，关系提取主要定位于额下回（Inferior Frontal Gyrus）和颞上回（Superior Temporal Gyrus）。已有研究表明楔前叶与情景记忆有关，并且与根据空间细节进行再认判断有关（Fletcher et al., 1995；Wallentin M et al. 2007）。而齿状回被认为与构建新的记忆有关，表现出项目记忆提取的特点（Nakashiba, T. et al., 2012）。另外，也有研究表明额下回与言语和语义关系理解密切相关（Winhuisen L et al., 2005）、颞上回和前额皮层都与语言和社会认知有关（Bigler, E. et al., 2007；Adolphs, R., 2003）。综合已有的相关研究结果和 ERP 溯源分析结果，我们认为，联结记忆再认过程中刺激呈现后 200~400ms 的 ERPs 活动主要反映了项目提取，而 400~800ms 的 ERPs 活动主要反映了关系提取。这一结果证明了联结记忆中项目提取和关系提取具有不同的认知神经机制。详见下图和 3.2.6 溯源分析。



针对专家提出的要求，我们对前言的和讨论中的相关内容进行了修改（见绿色字体部分），请专家指正。

第三轮

审稿人 1 意见：

意见 1：较好地修改或回答了审稿意见，建议发表。

回应：感谢专家的肯定。