

《心理学报》审稿意见与作者回应

题目：不同熟练度双语者的颜色范畴知觉效应：来自行为和 ERP 的证据

作者：李杰 何虎 吴柏周 侯友 曹亢 阿如罕

第一轮

感谢审稿专家 1 百忙之中对稿件的审阅，同时感谢审稿专家的宝贵意见和建议。以下是针对审稿意见做出的逐项修改说明，为方便审阅，文中对修改的内容使用红色字体。

审稿人 1 意见：

意见 1：作者采用行为和 ERP 实验探究语言经验对颜色知觉的影响。研究结果发现，汉语熟练度不同的蒙-汉双语者的颜色范畴知觉效应存在差异。作者由此认为，语言经验影响颜色知觉，Spair-Whorf 假设成立。该研究具有一定的理论意义。

回应：感谢专家的肯定。

意见 2：但存在以下几个问题需要进一步的思考与修改。关于语言经验对颜色知觉的影响，过去已有许多单语者和双语者的研究，并且这些研究都认为语言经验会影响颜色知觉。从作者所提出的研究问题上看，该研究更多的在于利用蒙-汉双语者的颜色知觉表现去验证前人的研究结果，而没有更深入，更有意义的研究问题。因此，研究意义也相对有限。

回应：非常感谢专家的意见。确实，有关语言对认知的影响，国内外都有研究，除了颜色知觉，还有时间知觉(Boroditsky, 2001)、运动和空间知觉(Athanasopoulos & Bylund, 2013; Bylund & Athanasopoulos, 2013)、面孔知觉(Fugate, 2013)、分类方式(Athanasopoulos & Kasai, 2007; 张启睿,和秀梅,张积家,2007)和形状知觉(Gilbert, Regier, Kay, & Ivry, 2008)。但是这些研究都是以单语者为研究对象，进行跨语言比较，但各单语者的跨语言比较，会因个体差异等复杂因素的产生干扰。也有少量双语者的研究，讨论语言经验对知觉的塑造，除了颜色知觉，还有图片分类（张积家&张凤玲，2010），运动和空间知觉(Athanasopoulos & Bylund, 2013; Bylund & Athanasopoulos, 2013)、分类方式(Athanasopoulos & Kasai, 2007)。但是以往的双语者研究而在反应指标方面多数研究采用李克特量表进行颜色相似性判断，从脑电角度考察双语者第二语言习得与范畴知觉影响的研究鲜有报道，双语者第二语言对母语颜色范畴干扰的

脑机制尚不清晰，相关的研究结论仍需进一步探索与验证。其次，已有研究提示二语水平可能改变双语者的认知加工能力，但关于颜色知觉的研究的变量主要涉及到与二语水平有关的间接指标，如二语使用频率或在二语国家居住时间等，并未将二语熟练度水平作为直接考量因素。最后，已报道研究中涉及第二语言均为英语，而以汉语为二语的研究报道较少，通过对汉语的研究即可与英语研究进行比对，也可以从广泛的跨语言角度对语言关联性假说进行验证。综合专家的意见，在问题提出部分增加一个参考文献，见前言第五段标红的部分。再次感谢专家的批评。

意见 3：作者在前言倒数第二段最后一句话提到“而以往双语研究中对二语概念范畴对一语概念范畴的影响仅局限于理论层面的推测。”作者对这个问题的阐述比较模糊。建议作者对这一点做更详细的说明。

回应：非常感谢专家的批评。结合前面的文献，特别是 Athanasopoulos, P.等对日语-英语双语者和希腊语-英语双语者的实验研究来看(Athanasopoulos,et al., 2011; Athanasopoulos, et al.,2010)，第二语言的概念范畴对第一语的概念范畴的影响有实验证据，所以修改此部分的表述，见红字，再次感谢专家的批评。

意见 4：作者在前言中只简单说明行为实验中采用的是视觉搜索范式，ERP 实验中采用的是 Oddball 范式，而没有对这两种范式在颜色知觉研究中的可行性进行分析。建议作者补充说明为何采用这两种范式研究颜色范畴知觉。

回应：非常感谢专家的意见。已经根据 Gilbert 等(2006)和(Thierry, Athanasopoulos, Wiggett, et al.,2009)研究进行了补充说明，见前言最后一段红字部分。再次感谢专家的意见。

意见 5：作者在前言最后一段，主要对 ERP 实验的结果进行了假设，而没有对行为实验的结果进行假设。建议补充对行为实验结果的假设。

回应：非常感谢专家的意见。已补充，见前言最后一段红字部分。再次感谢专家的意见。

意见 6：本研究重点关注不同汉语水平的蒙-汉双语者，但作者没有说明按照什么标准将被试分为高汉语水平组和低汉语水平组，仅从被试的语言情况自评吗？

回应：确实是仅从被试语言情况进行自评。在没有更好的语言测评材料的情况，自评是被大多数研究采用的方法(姜淞秀,李 杰,刘兴宇等,2015;Athanasopoulos,et al., 2011;

Athanasopoulos, et al. ,2010)。这种方法虽然在信度和效度方面存在一定的问题，但是鉴于目前语言水平量表缺乏，自评量表已经是目前争议最少的语言评价量表，再次感谢专家的意见。

意见 7：行为实验中采用“4 种渐进的蓝色 A、B、C 和 D。在蒙古语中，其中 A 和 B 都是 qinker 色，C 和 D 都是 huhe 色，在汉语中四种颜色都是蓝色。”。这些材料是否在实验前经过蒙古语母语者和汉语母语者的评定？

回应：非常感谢专家的意见。这些材料在实验前的预实验中已经过蒙古语母语者和汉语母语者的评定，预实验是让各母语被试对四个色块进行自由命名，结果显示，其中 A 和 B 都是 qinker 色，C 和 D 都是 huhe 色，在汉语中四种颜色都是蓝色。再次感谢专家的意见。

意见 8：作者在 2.2 材料 提到“目标色块仅出现在色块环的 1、2、3、4 号位置”，从作者提供的图上看，不清楚 1,2,3,4 号位置在哪里。此外，为何只出现在 1,2,3,4 号位置？

回应：非常感谢专家的批评。在文中已经改进，将“仅出现在色块环的 1、2、3、4 号位置”改成“仅出现在色块环最左和最右四个位置”。见相应部分的红字。并在附图中标注“1, 2, 3, 4”。选择四个位置的原因是，最早运用类似的材料时，并不限定目标色块出现的位置，但是 Gilbert 等(2006)的原材料目标色块可能出现位置太多，被试认识负荷较大，实验材料准备麻烦，所以 Zhou 等(2010) 仅选取了其中 4 个极端位置，结果表明，这种改进大大的方便了实验，实验数据也非常好，所以本研究也采用 Zhou 等（2010）的改进，表述也有修改，见红字。再次感谢专家的意见。

意见 9：行为实验共 64 个试次，但不清楚这 64 个试次具体分为几种条件，每种条件包含哪些刺激。

回应：非常感谢专家的批评。在文中已经改进，见相应的红字部分。范畴内和范畴间各 32 次，其中 CB 和 BC 为范畴间各识别 16 次，AB 和 CD 为范畴内各识别 16 次。再次感谢专家的批评。

意见 10：作者在 3.2 材料 提到“预实验显示 H1 和 Q1 都是汉语中的蓝色，但分别是蒙古语中的 huhe 和 qinker 色，G1,G2 在汉语和蒙古语中都表示绿色。”但是，预实验是什么？是评定实验材料吗？另外，H1 和 Q1 之间的物理差异，与 G1 和 G2 之间的物理差异相同吗？

回应：非常感谢专家的意见。预实验是让各母语被试对四个色块进行自由命名，结果显示，H1 和 Q1 都是汉语中的蓝色，但分别是蒙古语中的 huhe 和 qinker 色，G1,G2 在汉语和蒙古

语中都表示绿色。H1 和 Q1 之间的物理差异，与 G1 和 G2 之间的物理差异，实验中测试发现，H1 和 Q1 的物理差异大于 G1 和 G2 之间的物理差异，这对实验带来了不利的影响，不过在光谱中蓝色光区和绿色光区的波长区本来就不一样，所以没有办法做到物理距离相同，Thierry, Athanasopoulos, Wiggett, et al. (2009) 的研究中，蓝色刺激对和绿色刺激的物理距离也不同。再次感谢专家的意见。

意见 11：常见的 Oddball 范式只有标准刺激和偏差刺激，但作者在实验中设置了标准刺激，偏差刺激和目标刺激。建议作者说明使用目标刺激的目的。

回应：经典的 Oddball 范式确实只有标准刺激和偏差刺激。但是在 ERP 实验中，如果按经典的 Oddball 任务进行，因为被试由偏差刺激的按键反应而产生的脑电成分将污染由偏差刺激本身产生的视觉失匹配负波，所以与本研究相关的实验中(Thierry, Athanasopoulos, Wiggett, et al., 2009; Mo, et al., 2011)，均采用了有目标刺激(靶刺激)的 Oddball 任务，在这种任务中，被试见到目标刺激需要按键，既有利于使被试集中注意力，更好的完成实验，更有利于得到更“纯净”的视觉失匹配负波。

意见 12：脑电实验中 block 的顺序是固定的吗？

回应：非常感谢专家的意见。Block 顺序不固定，4 个 block 的顺序按拉丁方呈现，每个 block 有 64 组，每组 10 个色块，在这 10 个色块中伪随机的呈现 7 次标准刺激，2 次偏差刺激（不会连续呈现），1 次目标刺激，64 组色块，在实验中随机呈现。对原文的表达进行了修改，建相应部分红字内容。再次感谢专家的意见。

意见 13：作者分析了目标刺激和标准刺激诱发的 N2，这里的分析是将两组被试的数据综合起来分析的吗？

回应：非常感谢专家的意见。如专家所说，为了讨论各组被试的注意力集中情况，应该两组被试各条件分开分析 N2，分析的结果如附图 2 和附图 3。研究结果发现，各种条件下被试都激发了较强的 N2，这说明各种条件下被试都集中了注意力在完成实验，没有本质上的差异，既然各组条件激发的 N2 没有本质上的差异，所以便集中分析 N2，说明各组被试都集中注意力完成实验。Mo 等(2011)和 Zhong 等 (2014)也是这样分析的。再次感谢专家的意见。

意见 14：作者在 160-260 ms 时间窗内，对枕区波峰值和潜伏时间进行 2（颜色：蓝，绿）×

2（刺激类型：标准，偏差）重复测量方差分析。这里有两个问题。第一个问题，这里的分析是将两组被试的数据综合起来分析的吗？第二个问题，从方差分析的结果上看，无论是蓝色还是绿色，标准刺激和偏差刺激在这个时间窗的波峰值和潜伏时间的差异都不显著。这个结果说明，蓝色和绿色的偏差刺激都没有诱发 vMMN。那么，后续基于“vMMN”的统计分析也就没有意义了。

回应：非常感谢专家的意见。针对第一个问题，这里的确实应该分别分析两组被试的枕区波峰值和潜伏时间进行 2（颜色：蓝，绿）×2（刺激类型：标准，偏差）重复测量方差分析。已经将文中的表述进行了修改，见红字部分。针对第二个问题，对两组被试重新分析，“低汉语水平的波峰值的刺激类型的主效应显著 $F(1,13) = 10.26, p < 0.001, \eta^2 = 0.42$ ，颜色类型的主效应不显著，二因素交互作用显著 $F(1,13) = 24.68, p < 0.001, \eta^2 = 0.63$ ；高汉语水平的波峰值的刺激类型的主效应显著 $F(1,14) = 89.56, p < 0.001, \eta^2 = 0.85$ ，颜色类型的主效应也显著 $F(1,14) = 60.45, p < 0.001, \eta^2 = 0.80$ ，二因素交互作用不显著。潜伏时间分析结果显示主效应和交互作用都不显著，均在 200ms 左右。因为两组被试的偏差刺激所激发的负波都显著大于标准刺激所激发的负波，所以用枕区偏差刺激的波形减去标准刺激的波形，在 160ms-260ms 时间窗内所得的差异波为视觉 vMMN。”已经在论文对应部分进行修改，见红字部分。再次感谢专家的意见。

.....

感谢审稿专家 2 百忙之中对稿件的审阅，同时感谢审稿专家的宝贵意见和建议。以下是针对审稿意见做出的逐项修改说明，为方便审阅，文中对修改的内容使用蓝色字体。

审稿人 2 意见：

意见 1：本研究采用前人经典行为和 ERP 实验范式，研究了第二语言（蒙汉双语者的汉语）水平对颜色知觉加工（颜色范畴效应）的影响。研究发现，相对于高汉语水平者，低汉语水平者对蒙语不同范畴颜色（但是汉语相同范畴颜色）的行为分辨能力更好，与颜色差异自动检测相关的 ERP 成分 vMMN（160-260ms）的振幅也更大，表现出更强的颜色范畴效应。该研究设计总体合理，在刺激材料上不仅采用了蒙语不同范畴、但是汉语相同范畴的颜色作为实验条件，而且采用了蒙汉语都属于相同范畴的颜色进行对照，这一点是值得肯定的。实验结果总体上较清晰（但是 ERP 信噪比一般），符合研究预期，为语言（二语程度）影响颜色知觉提供了新的可能研究证据。

回应：感谢专家的肯定。

主要存在以下问题和需要改进的地方：

1. 研究方法部分：

意见 2：缺乏对高/低汉语者的操作性定义：通过哪些筛查指标对蒙汉双语者个体进行高或低汉语水平分类？

回应：非常感谢专家的意见。语言水平根据被试语言情况自评等级进行分类。自评是被大多数同类研究采用的方法(姜淞秀,李 杰,刘兴宇等,2015;Athanasopoulos,et al., 2011; Athanasopoulos, et al. ,2010)。这种方法虽然在信度和效度方面存在一定的问题，但是鉴于目前语言水平量表缺乏，自评量表已经是目前争议最少的语言评价量表，再次感谢专家的意见。

意见 3：颜色知觉有较大的个体差异和主观性，本研究中不同颜色（例如 qinker 和 hube 色）的 RGB 值具体是如何选择的？（对于某个特定 RGB 颜色值，是否可能有的蒙语者认为是 qinker 色，而有的认为是 hube 色？）

回应：非常感谢专家的意见。这些材料在实验前的预实验中已经过蒙古语母语者和汉语母语者的评定，预实验是让各母语被试对四个色块进行命名，结果显示，其中 A 和 B 都是 qinker 色，C 和 D 都是 hube 色，在汉语中四种颜色都是蓝色。已在实验部分添加相应内容，见蓝字。再次感谢专家的意见。

意见 4：脑电实验方法：脑电记录时的参考电极是什么？离线参考为什么选用 TP9 和 TP10 的平均？

回应：非常感谢专家的意见。参照文献脑电记录的参考电极是 Referece 点(姜淞秀,李 杰,刘兴宇等 ,2015;Athanasopoulos,et al., 2011; Athanasopoulos, et al. ,2010)。离线分析时 Athanasopoulos 选取的是平均参考电极，但是根据赵仑所著的《ERPs 实验教程》所言平均参考电极基于的头颅模型是基于西方人种理想头颅模型，并且还会产生本不存在的“幽灵电极”，这导致误差较大，TP9 和 TP10 是安放在双耳乳突的电极，在非听觉任务中，用双耳乳突作为参考电极有信号稳定，信号失真度小，所以目前应用非常非常广泛(姜淞秀,李 杰,刘兴宇等,2015)。再次感谢专家的意见。

2.实验结果部分：

意见 5: 请仔细检查统计结果, 例如, ERP 实验总共 29 名被试, 方差检验的自由度为什么是 $F(1, 55)$? 请通篇仔细检查数据!!!

回应: 非常感谢专家的意见。已经修改了, 见蓝字部分。再次感谢专家的意见。

意见 6: ERP 结果中请提供 vMMN 的地形分布 (包括不同条件之间差异波的地形分布)

回应: 非常感谢专家的意见。已经修改了, 见正文的图和附图。再次感谢专家的意见。

意见 7: 讨论部分: 第 3 段, “语言...还影响感知觉这样的初级认知过程, 不仅影响有意识的认知过程, 还影响无意识的内隐认知过程”, 这个推论值得商榷。本研究的 Oddball 范式, 不同范畴颜色虽然与任务无关, 但是并不一定是无意识的。

回应: 非常感谢专家的意见。确实如专家所说, 有是否无意识的方面有争论, 同样的任务 Thierry, et al.(2009)认为是无意识的, Mo, et al.(2011)认为是前注意(preattentive), Gábor, et al.(2014)认为是否是无意识和前注意有争论, 参照各方的意见, 将表述修改, 见蓝字。再次感谢专家的意见。

意见 8: 请仔细检查全文, 纠正语言表达错误, 例如: (1) 2.3 过程, 第一段, “要求被试在 3000ms 判断目标色块出现的位置并做出反应”, 是否应为“3000ms 内”? (2) 3.3.1 脑电数据记录, “以为 Reference 点参考电极”? (3) 3.3.3 实验过程, “将实验用的颜色作为图形或方形色块”, “图形”应为“圆形”? 等等

回应: 非常感谢专家的意见。已经修改了, 见蓝字部分。再次感谢专家的意见。

第二轮

审稿人 1 意见: 作者较好地解释了一审意见中所提出的问题, 但有些问题作者只在修改说明中做了解释, 建议作者在文章中也做进一步的修改。

意见 1: 建议作者在 2.1 被试 和 3.1 被试 这两个部分补充说明语言背景调查的依据。

回应: 非常感谢专家的意见。已经补充, 见红字。

意见 2: 建议作者在 2.2 材料 和 3.2 材料 这两个部分补充说明实验材料的评定是如何进行的, 具体结果如何。

回应：非常感谢专家的意见。已经补充，见红字。

意见 3：建议作者在 2.2 材料 补充说明为何“目标色块仅出现在色块环的最左和最右四个位置”。

回应：非常感谢专家的意见。已经补充，见红字。

意见 4：此外，文中有多处地方表述不清晰，建议作者进一步通读全文，修改表述不清晰的地方。

回应：非常感谢专家的意见。已经补充，见红字。

审稿人 2 意见：作者对于第一次审稿意见的回复仍然存在一些问题：

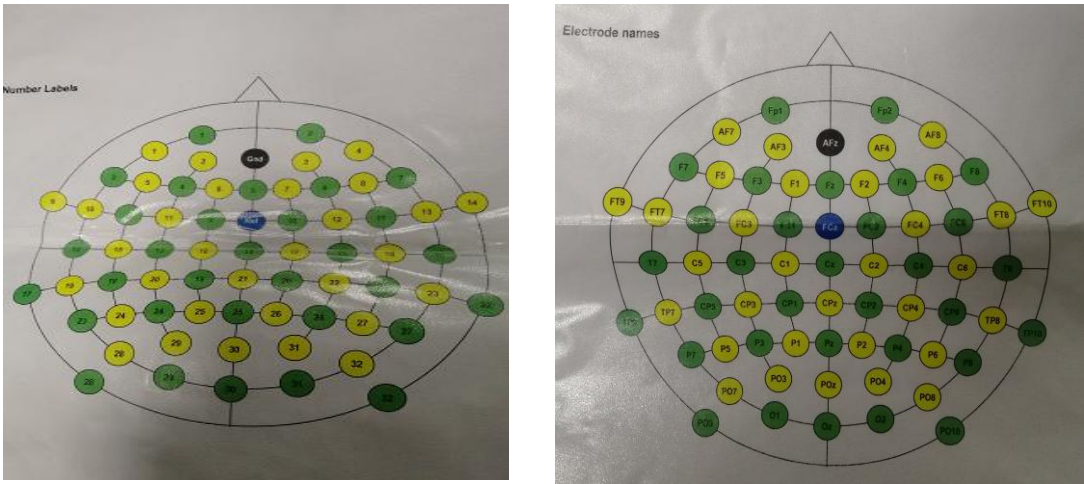
意见 1：高/低汉语水平的操作性定义：研究显示了两组被试在多个筛选指标（包括蒙古语听力/口语/阅读/写作/使用频率，汉语听力/口语/阅读/写作/使用频率，以及开始习得二语的年龄/能用汉语交流年龄）的平均成绩和标准差，但是没有说明是根据哪些指标对被试进行高或低汉语水平分类的（例如，在哪个/些指标达到多少分数以上属于高汉语水平者？）。关于这个问题，作者的回复是“本研究采用的标准是被试自评语言情况，评价内容涉及从日常学习到平素生活使用和语言接触等各方面，尽量涵盖和语言水平评价有关的内容，并可以进行互相参照反映被试自评的可信度。尽管这种方法在信度和效度方面仍存在一定的不足，但是目前争议最少的语言评价量表的方法，同时也是参考和借鉴相应的文献(姜淞秀,李 杰,刘兴宇等,2015; Athanasopoulos,et al., 2011; Athanasopoulos, et al. ,2010) 进行选择的”；“语言水平根据被试语言情况自评等级进行分类。这种方法受方法学本身有限，虽在信度和效度方面存在一定的不足，已是目前争议最少的语言自评量表，也有相应的文献供参考(姜淞秀,李 杰,刘兴宇等,2015;Athanasopoulos,et al., 2011; Athanasopoulos,et al., 2010)，所以研究中借鉴并选择了该方法。”这个回复存在以下问题：（1）没有在正文中体现（2）请提供研究所使用的语言自评量表（可以在附录中给出）（3）文中表 1 是“自评量表”的测量结果统计值吗？如何根据“自评量表”测量结果对各个被试进行高/低汉语水平分类？“汉语熟练度”是本研究的关键操纵因素，需要在文章中详细说明本研究对于这个因素是如何操作性定义的。

回应：非常感谢专家的意见。（1）已在正文讨论部分体现，见蓝字。（2）已在附录部分给出，见 46 页。（3）分类标准为，蒙古语听说读写自评量平均得分减去汉语听说读写的平均得分

小于等于 1 的为高汉语水平组，大于 1 的为低汉语水平组，已在正文中补充，见“视觉搜索实验被试部分”红字。再次感谢专家的意见。

意见 2: 脑电记录时的参考电极究竟是什么？“以 Reference 点为参考电极”(第 16 页第一段) 是什么意思？？

回应：非常感谢专家的意见。本研究脑电数据采集的帽子是 actiCAP 电极帽，它是把 FCz 点标成 reference 点，已将正文中的表述修改，见相应部分的蓝字。如图。



意见 3: 补加上的地形图是用的 160-260ms 的 ERP 均值进行绘制的吗？请在文中具体说明。

回应：非常感谢专家的意见。已经在文中说明，见蓝字。

意见 4: 本研究 vMMN 测量使用了峰值分析，但是本研究 vMMN 波形较宽广，信噪比不强，在这种情况下，通常不适合进行波峰分析；如果一定要使用波峰分析，分析时间窗口须包括各个条件的波峰，本研究 vMMN 的分析时间窗口为 160-260ms，但是在一些条件下 vMMN 平均波峰在 260ms 或 150ms 左右（图 2），超出了测量范围，测量结果会存在一定问题。本研究的 ERP 信噪比不强，对主要结果的可靠性造成了一定影响，这一点需要注意并加以讨论。

回应：非常感谢专家的意见。根据专家的指导建议，对本研究的脑电数据进行了重新统计分析。结合各条件下峰值的平均潜伏时间，将分析的时间窗口改为 180ms-260ms，对 180-260ms 时间窗内的 vMMN 波形进行 2 汉语水平（高，低） \times 2 颜色（蓝，绿）平均波幅分析发现，汉语水平和颜色的主效应不显著，汉语水平和颜色的交互作用显著， $F(1,28) = 25.08$, $p <$

0.001, $\eta^2 = 0.46$, 简单效应分析显示, 对于蓝色刺激产生的 vMMN 的平均波幅, 高汉语水平显著小于低汉语水平, $p < 0.05$, $\eta^2 = 0.12$, 同时低汉语水平被试在识别蓝色产生的 vMMN 平均波幅也显著大于绿色, $p < 0.001$, $\eta^2 = 0.46$ 。在 180-260ms 时间窗内, 峰值分析和平均波幅的结果一致, 相互验证, 基本符合实验预期。非常感谢专家诚恳、有效的指导意见, 关于信噪比和刺激材料改进的问题, 已经加以讨论, 见蓝字。

第三轮

审稿人 1 意见:

意见 1: 在 3.3.2 脑电数据分析部分, 关于 vMMN 的介绍“160-260ms 内的波峰便是失匹配负波”, 但是在 vMMN 的分析中, 为什么将时间窗口范围设置为 180-260ms?

回应: 非常感谢专家的意见。遵照专家的意见, 为了文章表达前后一致, 已经修改, 见红字。再次感谢专家的意见。

意见 2: 脑电数据分析中, 对偏差刺激和标准刺激的差异进行了原始波峰分析 (19 页第 2 段), 但是缺乏相应的 ERP 原始波形图 (各组被试各个颜色的标准刺激和偏差刺激诱发的 ERP 原始波), 请补充。

回应: 非常感谢专家的意见。绝大数研究 vMMN 的文献, 在正文中只提供最终的差异波形图和统计检验结果, 遵照专家的指导, 我们将所有的原始波形图在附件中呈现, 见附图。再次感谢专家的意见。

意见 3: 请提高结果图的分辨率和清晰度 (例如, 图 2 的两条 ERP 虚线容易混淆, 文字也不够清晰)

回应: 非常感谢专家的意见。已经改正, 见图。

意见 4: 结果未达到统计显著的也请提供统计数据, 例如第 20 页第一段, “高汉语水平被试蓝色刺激和绿色刺激产生的 vMMN 平均波幅差异不显著”, 请补充 p 值和效应量。

回应: 非常感谢专家的意见。已经改正, 见红字。

意见 5: 高汉语水平被试和低汉语水平被试的蓝色刺激诱发的 vMMN 之间存在显著差异吗?

回应：非常感谢专家的意见。对于蓝色刺激产生的 vMMN 的平均波幅，高汉语水平显著小于低汉语水平， $p < 0.05$ ， $\eta^2 = 0.12$ 。再次感谢专家的意见。

审稿人 2 意见：作者较好地根据二审意见修改了文章。建议作者再对以下几个问题做一定的修改，修改后即可考虑录用。

意见 1：作者在 20 页的第二行写到：“高汉语水平被试蓝色刺激产生的 vMMN 平均波幅显著大于低汉语水平”。但是从结果图和文章其他内容上看，这里的结果应该是“高汉语水平被试蓝色刺激产生的 vMMN 平均波幅显著小于低汉语水平”。请作者核对此处结果。

回应：非常感谢专家的意见。专家意见正确，已经改正，见蓝字。

意见 2：文中仍有多处地方表述不清，请作者再次通读全文，修改表述不清的地方。

回应：非常感谢专家的意见。已经改正，见蓝字。

编委复审意见

经过几轮的审稿，作者对文章作了重要的修改，目前两审稿人均认为文章小修后可以发表。我本人也详读了论文，总体上同意审稿人的意见。目前对稿件仍有一些建议，尤其表现在以下三方面：

意见 1：摘要部分最后一句，建议改为“本研究支持了 Spair-Whorf 假设”

回应：非常感谢专家的意见。已经改正，见红字。

意见 2：P15 第二段“首先，从脑电角度考察双语者第二语言习得与范畴知觉影响的研究鲜有报道，双语者第二语言对母语颜色范畴干扰的脑机制尚不清晰，已有的研究结论仍需进一步探索与验证”，这一评论需要表达得更清楚，鲜有报道并不是值得研究的重要原因，为何需要脑电研究的理论切入点需要写得更清楚，例如强调技术的实时性对研究加工实时进程的作用。

回应：非常感谢专家的意见。遵照专家的意见，已经修改，见红字。

意见 3：作者的写作不够认真，引用格式，标点符号格式均需要按照学报投稿要求仔细校对。

例如有些引用既加了姓，又加了名，请把名的缩写去掉，许多逗号用了半角，但后未加逗号等等，有些双字姓名之间又有空格，等等，不一一赘述，请作者务必仔细检查。这些问题修改后可发表。

回应：非常感谢专家的意见。由于作者对 APA 引文格式非常不熟练，所以出现了大量的格式错误。已经遵照专家的意见并对照新版的 APA 指导手册修改，包括：将“et.al”改成“et al.”；删除正文引用出现的名；删除双字姓名之间的空格；规范正文同时引用多篇文章的排序方式；改变文中半角和全角混用的符号；规范了“和”与“&”的用法，再次感谢专家的指导。

主编终审意见

经过几轮的审稿，审稿专家均认为该文章经过小修后可以发表。我本人也详读了文章，基本同意审稿专家的意见。但是，该文章在一些细节上仍然存在问题，需要进一步修改后才能发表。

意见 1：虽然经过审稿专家的提醒，作者对参考文献格式进行了修改，但是仍然存在问题。请参考已经在心理学报发表的论文的格式进行修改。

回应：非常感谢专家的意见。由于作者对心理学报引文格式非常不熟练，所以出现了大量的格式错误。已经遵照专家的意见并对照“心理学报‘参考文献著录要求’”对参考文献进行了修改，主要纠正了 1) 参考文献中的期刊名简称问题，2) 进一步规范了“和”“&”“等”“et al”的用法，3) 正文引用中的作者数量问题，4) 修改“-”和“—”的问题。

意见 2：文章中既包含行为实验（实验 1），又包含脑电实验（实验 2），所以标题建议修改成“不同熟练度双语者的颜色范畴知觉效应：来自行为和 ERP 的证据”。

回应：非常感谢专家的意见。遵照专家的意见，已经修改。

意见 3：英文标题和摘要存在语法等问题，如，英文的标题就有问题，建议进一步润色。

回应：非常感谢专家的意见。遵照专家的意见，已经找相关机构润色，并提供编辑证书。

意见 4：数字间表示范围的符号应为“~”，如“150-250ms”应为“150~250ms”，“0.01-70 Hz”应为“0.01~70 Hz”。

回应：非常感谢专家的意见。遵照专家的意见，已经修改。

意见 5：“图 2”与“3.5 讨论”之间排版好像有问题。不知是不是 word 版本导致的。建议作者核查。

回应：非常感谢专家的意见。遵照专家的意见，已经修改。