

## 《心理学报》审稿意见与作者回应

题目：歌词对音乐情绪加工的影响：行为与 ERP 的研究

作者：张伟霞 王莞琪 周临舒 蒋存梅

---

### 第一轮

非常感谢两位审稿专家的宝贵意见和建议！根据审稿专家 1 的意见，我们将本研究问题定位在歌词对音乐情绪加工的影响效应上。为了确保稿件的质量，在行为研究的基础上，我们增补了一个控制更为严格的 ERP 实验，且删除原先略有瑕疵的实验 1。这是因为，在原先的实验 1 中，我们尚未控制音色对音乐情绪加工的影响。我们希望，本次修改将使本稿件的质量得到提升。以下是我们对两名专家意见逐条进行的修改说明。

#### 审稿人 1 意见：

研究者关注的问题是聆听音乐过程中常见发生的现象，确实具有研究的必要性。并且研究不同音乐媒介（声乐和器乐）所产生的心理差异及相关心理机制，具有一定的理论意义，尤其是对于解释音乐的进化起源以及比较音乐和语言的进化起源。但总体来说，作者在该文的研究设计和具体行文部分都存在不少问题，需要做出重新梳理和较大程度修改。

**意见 1：**作者对文章核心概念的解释有误。作者在引言第二段写到的“声乐与器乐的本质差异在于是否存在歌词。”声乐与器乐的根本区别在于音色，即通过不同的发声方式创造出不同的音色，而不是在于歌词。声乐，是指用人声演唱的音乐形式。器乐是相对于声乐而言，完全使用乐器演奏而不用人声或者人声处于附属地位的音乐形式。换言之，歌词只是声乐的附属和伴随特征，而非本质特征，即充分非必要条件。所以，作者在文章中有诸多类似不严谨的表述，请进行更正。

**回应：**非常感谢审稿专家的耐心审读和宝贵意见。我们同意审稿专家的意见。结合审稿专家 1 的意见 2 和意见 3，我们对本研究探究的问题进行了微调。具体来说，本文不再关注声乐与器乐加工差异的问题，而是聚焦于歌词对音乐情绪加工的影响。为了排除原先脑电实验（实验 1）中音色的潜在影响，我们增补了一个 ERP 实验（替换了之前的实验 1）。补充的 ERP

实验请详见修改稿第 7~10 页。此外，我们也对行文中不严谨的表述进行了更正。

**意见 2:** 研究议题的提出到研究结果推导存在逻辑问题。文章的结论部分写到“本研究的主要结论是：虽然听者对声乐和器乐情绪信息都能进行加工，但是对声乐情绪信息的加工在时间进程上滞后于器乐。这种差异缘于声乐中歌词的参与。”研究者力求于证明两种事物（器乐与声乐）的某种非本质区别（歌词）引发了某种效应，而最后却把这种效应认为是导致两种事物（器乐和声乐）产生听觉加工差异区别的原因（歌词的情绪信息效应导致了声乐与器乐在认知上的差异），其结论的推导在逻辑上是存在问题的。综合意见 1 和意见 2，建议作者可以研究目前的问题，但需要考虑换一下核心术语（器乐和声乐）以及考虑更改一下整个立论的写作思路，例如：变为“探讨歌词的情绪信息效应在音乐媒介表达中的作用”等类似主题的写作思路。

**回应:** 在本次修改中，我们更改了本文整个立论的写作思路，使研究结果能够更准确地说明本研究的问题。您的建议对我们有很大的帮助，再次表示感谢！（可参见意见 1 的回应。）

**意见 3:** 文章的“问题提出”部分不够清晰，“研究意义”不够明确。作者只是泛泛而谈“本研究旨在探讨听者对声乐与器乐情绪信息的加工是否存在差异(实验 1)”。但研究这个问题究竟有什么意义，究竟解释了什么理论问题呢？解决了什么理论争议呢？引言部分仅仅用了 500 个字左右进行相关论述，并且论述的逻辑结构较为松散，语言表述方式不太符合科技论文写作的规范。建议查阅更多的文献资料（例如：音乐进化心理学方面的文献），进行重新梳理和细化，补充理论依据和实验论据。

**回应:** 基于审稿专家的意见和建议，我们重新查阅、梳理了文献，并调整了本文前言部分的逻辑结构。在本次修改稿中，我们从音乐和语言进化的角度引出话题，以语言与音乐对比研究为背景，结合相关文献，我们提出了本研究的问题，即“歌词（具有明确语义的语言信息）是否影响音乐的情绪加工”。相关内容详见修改稿的前言部分（第 1~3 页）。

**意见 4:** 实验材料控制问题。研究中控制了器乐版本和声乐版本在音乐表达(比如力度、速度以及分句等)上的一致性。请作者进一步描述具体是如何控制速度、力度和分句处理等问题的一致性，其操作定义是什么，例如：力度是指平均响度吗？或是指每一小节的重音处理方式？

**回应:** 在本研究的实验材料控制中，力度不是指平均响度，而是指表演者对音乐音响强弱变

化的处理方式。在本次修改稿中,我们已经对所有相关术语的操作定义进行了明确的解释(详见修改稿第3页“2.1.2 刺激与程序”第一段)。

**意见 5:** 实验程序和数据采集问题。研究程序部分,“音乐播放完毕立即呈现情绪面孔 1000 ms。然后,出现反应界面,要求被试判断音乐与面孔所表达的情绪是否一致,一致按 F 不一致按 J”。作者没有交代记录脑电数据的 mark 打在什么地方。作者是否考虑过被试立即做出行为反应,是否会对脑电数据造成干扰?作者是如何避免这种干扰?

**回应:** 本研究的脑电 mark 打在目标刺激(面孔图片)出现的时间点上,这一信息已经在数据分析部分进行了说明(详见修改稿第7~8页“3.1.2 刺激与程序”部分)。与行为实验不同,在脑电实验中,目标刺激呈现持续 1000 ms,之后我们要求被试进行情绪一致性判断。由于我们仅分析目标刺激呈现 1000 ms 以内的 ERP,因此,这种延时的任务在一定程度上排除了行为反应对脑电数据的影响。此外,由于本研究关注的是情绪启动效应,即情绪不一致条件与一致条件的差异,如果行为反应对脑电数据具有潜在影响,这种影响在不一致条件与一致条件之间应该已经被平衡了,而且在计算 N400 或 LPC 效应时(即不一致条件减去一致条件时),这种干扰已被减去(如果有干扰的话),不会对本研究结果产生影响。

**意见 6:** 行为实验的分析和推导存在重要错误。“实验二中,在反应时方面,2(声乐类型:带有歌词声乐,无歌词声乐) $\times$ 2(一致性:一致,不一致)重复测量方差分析结果:一致性主效应显著( $F(1, 31) = 14.38, p = .001, \eta^2 = .32$ ),情绪一致条件下的反应时( $M = 1051.83$  ms,  $SD = 389.35$ )显著低于不一致条件( $M = 1166.45$  ms,  $SD = 503.22$ );声乐类型主效应显著( $F(1, 31) = 4.42, p = .04, \eta^2 = .13$ ),带有歌词声乐条件下的反应时( $M = 1136.59$  ms,  $SD = 486.83$ )显著长于无歌词条件( $M = 1081.69$  ms,  $SD = 415.88$ )。这说明听者对带有歌词声乐情绪信息的加工比无歌词需要更多的时间”。带有歌词声乐条件下的反应时显著长于无歌词条件。这个效应只能说明加工带有歌词的声学信息速度更慢而已,但并非一定是由于作者所希望探讨的“歌词中情绪信息”所导致的,因为作者所采用的范式下,只有通过不一致和一致的差异才能体现出对情绪语义信息加工的操控,所以,此分析方法和实验结果难以支撑作者得出的结论。建议作者参考以往文献,考虑更换统计方法,例如:可以考虑比较无歌词条件和歌词条件下的情绪信息启动差异量(不一致减去一致)的区别。

**回应:** 非常感谢审稿专家的建议!我们已采纳了审稿专家的建议,进一步对数据做了以下分析:(1)分析带有歌词和无歌词条件下一致与不一致之间差异。简单效应分析表明,带有歌

词条件下情绪一致与不一致差异显著( $F(1, 31) = 11.58, p = .002, \eta^2 = .27$ ), 无歌词条件下差异也显著( $F(1, 31) = 5.57, p = .03, \eta^2 = .15$ )。(2) 比较带有歌词与无歌词条件下的情绪信息启动差异量。配对样本  $t$  检验结果表明, 带有歌词与无歌词条件下的情绪信息启动量不存在显著差异( $p > .05$ )。该结果进一步验证了, 在行为层面, 带有歌词和无歌词的音乐都能启动听者对情绪面孔的加工, 且二者的启动量没有显著差异。我们已在原文补充了以上分析的结果, 详见修改稿第 6 页第 2 段(图 2 上方)。

**意见 7:** 讨论和结论部分存在过度推导的情况。“本研究的主要结论是: 虽然听者对声乐和器乐情绪信息都能进行加工, 但是对声乐情绪信息的加工在时间进程上滞后于器乐。这种差异缘于声乐中歌词的参与。”其中存在过度推导, 且不说根据作者目前的实验结果, 难以证明歌词中情绪信息的效应。就算能够证明, 排除了歌词外, 音色本身也是会引发音乐感受产生差异的重要原因(已经有大量文献支持), 此外, 还可能存在其他影响因素。所以, 不可以做出“这种差异缘于声乐中歌词的参与”的决断。一般情况下, 只能表述为“这种差异与声乐中歌词的参与有关”。

**回应:** 的确, 原先实验 1 的声乐与器乐还存在音色的差异。为了确保稿件的质量, 在行为研究的基础上, 我们增补了一个控制更严格的 ERP 实验, 且删除原先略有瑕疵的实验 1。同时, 我们对论文中不严谨的表述进行了修正。您的建议对我们帮助很大。再次向您表示感谢!

.....

**审稿人 2 意见:**

基于心理学和哲学研究对声乐和器乐在情绪加工作用上的争论, 作者通过脑电相关指标、首次对声乐与器乐情绪加工机制进行了研究。该研究具有新颖性; 全文表达清晰、思路顺畅、实验材料设计严格、结论较充分。由于审稿人水平有限, 在此提出该问题:

**回应:** 非常感谢审稿专家的意见!

**意见 1:** N400 与语义加工有关, 一个音乐片断中某个音高的变化也能产生 N400 (改变乐段表达的语义性信息)。那么, 与情绪一致条件相比, 情绪不一致条件在情感启动范式中诱发更大的 N400, 这是否也与情绪的语义性加工有关, 还是由于其它因素所致? 在本研究中, 诱发的 N400 是因何种原因所致?

**回应：**的确，N400 与语义加工有关。在本研究中，我们使用的是情感启动范式。根据已有研究文献，该范式关注的脑电成分主要是 N400（也包括 LPC），即启动与目标刺激情绪不一致条件比一致条件诱发更大的 N400 波幅。该脑电成分反映出情绪意义的加工，如您所见，这种情绪意义也可以视作情绪的语义。在本研究中，音乐作为启动刺激，面孔作为目标刺激。每一条音乐分别与两个情绪面孔配对；类似地，每个面孔也分别与两条音乐配对，由此形成一致与不一致条件。同时，我们也通过前测（前测 3），确保这两个条件的差异仅在情绪一致与不一致方面。另外，在音乐研究中，据已有研究文献，N400 仅在两种情形下出现：一种是在启动范式中，另一种是音乐记忆任务中。由于本研究的音乐刺激都是被试不熟悉的，因此，N400 的确只能反映情感启动范式中的情绪意义加工。

**意见 2：**同样，对 LPC 的解释，语义不一致诱发 LPC 反映了整合加工，那么本研究中，除了可能反映音乐情绪与面孔情绪的整合加工，还有其它因素所致？

**回应：**的确，通过对已有文献的梳理，我们认为 LPC 除了反映音乐情绪与面孔情绪的整合加工之外，也有可能反映为注意资源的分配。具体来说，以往使用情感启动范式的研究也发现情绪不一致比一致诱发了更大的 LPC，研究者认为，不一致条件下更大的 LPC 是由更多的注意卷入所致，即情绪冲突的信息可能被分配更多的注意资源(Herring, Taylor, White, & Crites, 2011; Hinojosa, Carretié Méndez-Bértolo, Míguez, & Pozo, 2009; Werheid, Alpay, Jentsch, & Sommer, 2005; Zhang, Li, Gold, & Jiang, 2010)。在本修改稿中，我们已经补充了以上内容，详见修改稿第 11 页第 2 段。

**意见 3：**论文的题目是否可修改为“声乐与器乐情绪信息加工的电生理差异”，请作者斟酌。

**回应：**非常感谢审稿人的建议！根据审稿人的建议，我们删除了题目中的“听者加工”。同时，综合审稿专家 1 的意见，本研究题目最终定为“歌词对音乐情绪加工的影响：行为与 ERP 的研究”。

---

## 第二轮

非常感谢审稿专家的意见和建议！按照专家的意见，我们对论文进行了修改。我们希望，本次修改将使论文的质量得到提升。以下是我们针对专家意见所做的修改说明。

### 审稿专家 1 修改意见:

作者在第一稿基础上,对该文进行了较大程度的调整和修改。文章质量已经有了较大提高。但仍然有以下方面问题需要进一步的修改。

**意见 1:** 作者 ERP 数据分析过程中,是如何得到三阶交互作用?并且在简单简单效应分析下,怎么还会出现 F 值,请作者进一步澄清 ERP 研究中的统计是怎么逐步进行的。

**回应:** 非常感谢审稿专家的建议。针对 ERP 数据,我们分别进行了中线和两侧的分析。中线分析为 2(歌词:带有歌词音乐,无歌词音乐)×2(一致性:一致,不一致)×3(脑区:前,中,后)的三因素重复测量方差分析,两侧分析为 2(歌词:带有歌词音乐,无歌词音乐)×2(一致性:一致,不一致)×3(脑区:前,中,后)×2(半球:左,右)的四因素重复测量方差分析。无论是中线分析还是两侧分析,我们仅发现歌词与一致性的两因素交互作用,没有发现三阶交互效应,因此,本研究没有进行简单简单效应分析。

然而,本研究出现了歌词与一致性的两因素交互,我们进一步做了简单效应分析,并在论文中报告了  $F$  值。这是因为,简单效应分析方法通常为  $F$  检验(舒华,1994)。如果两个因素都仅有 2 个水平(不超过 3), $F$  检验与  $t$  检验的结果是相同的。 $t$  检验被认为是  $F$  检验的特例(张厚粲,徐建平,2004)。

在本研究中,ERP 数据统计分析步骤如下:第一,我们进行重复测量方差分析,如果存在歌词或一致性的显著主效应,则报告该变量在不同水平上的差异情况;第二,如果存在与歌词或一致性有关的显著交互作用,则进行简单效应(或简单简单效应)分析。在本研究中,我们只发现了歌词与一致性的两因素交互,由于歌词与一致性都只有两个水平,因此,简单效应分析既可以使用  $F$  检验,也可以使用  $t$  检验。两种检验方法均得到相同的  $p$  值,而且  $t^2 = F$ 。在论文中,我们报告了简单效应分析的  $F$  检验结果。

### 参考文献:

张厚粲,徐建平.(2004). *现代心理与教育统计学*. 北京:北京师范大学出版社.

舒华.(1994). *心理与教育研究中的多因素实验设计*. 北京:北京师范大学出版社.

**意见 2:** 实验二的行为结果要想说明什么问题?它和实验一的设计有什么区别?如果从作者写出的实验设计来说,两个实验的设计都是 2(歌词:带有歌词音乐,无歌词音乐)×2(一致性:一致,不一致)。但为什么两个实验却会得出不同的结果呢?如果两者的实验设计类似,那么在讨论部分,是否应该解释和说明为什么研究同一个问题,结果却不相同,可能的原因是

什么？

**回应：**感谢审稿专家的建议。在本研究中，实验一旨在从行为层面上探讨听者能否加工带有歌词与无歌词音乐的情绪信息，被试任务是判断目标面孔的情绪，因此，该行为结果反映的是音乐对面孔情绪判断的影响。

实验二旨在通过电生理手段进一步考察大脑加工这两类音乐所传达的情绪是否存在差异。直接对目标刺激进行按键反应会在一定程度上污染脑电数据，为了避免这种情况发生，在实验二中，我们要求被试在反应界面判断启动刺激与目标刺激的情绪一致性，因此，该行为结果反映的是听者对启动刺激与目标刺激之间关系的加工，而不是启动刺激对目标刺激的影响效应。这与实验一存在根本的差异，因为后者的行为结果可以体现出启动刺激对目标刺激的影响效应。

我们之所以收集实验二的行为数据，目的是想从正确率角度检验被试是否认真完成了实验任务。实验二的行为结果表明，无论在带有歌词，还是无歌词条件下，听者对启动刺激与目标刺激之间情绪关系的判断均呈现出较高的正确率(均在 83% 以上)，而且不存在差异。这一结果说明，本研究的被试都认真完成了两个条件的任务。基于此，我们仅在 3.2.1 部分对此进行说明。

**意见 3：**建议在实验一统计部分，不再需要讨论交互作用效应，因为效应本身就是边缘显著， $P$  值还很大 ( $p = .07$ )，再加上作者所关心的简单效应分析确实也并没有发现两者（一致与不一致）在不同条件下（有无歌词）存在区别（说明可能是另外一个方向上存在微弱差异，在控制一致和不一致条件下，有无歌词之间）。所以建议删除。在上一稿中让作者修改，是因为需要进一步论证作者的观点，但此稿中由于文章的逻辑已经调整，行为学实验只需说明歌词音乐与无歌词音乐一样都能够启动情绪就可以，那么，情绪一致性主效应显著，就已经能够说明。

**回应：**非常感谢审稿专家的指点，我们接受您的建议，已删除实验一中关于交互作用进一步分析的相关内容。

**意见 4：**作者没有写结论，标准科技论文在结尾部分需要写结论，即本文的主要论点和发现。作者应该进行补充。

**回应：**非常感谢审稿专家的建议，我们已在本次修改稿中补充了研究结论部分。

**意见 5:** 写作有误,  $\eta^2$  应该是  $\eta_p^2$ , 请更正。

**回应:** 感谢审稿专家的细心审阅, 我们已更正了行文中错误的表述。

**意见 6:** 作者在实验二中, ERP 的画图非常不规范, 例如, 没有标出所画 ERP 图的电极点是什么, 例如, 图上没有图例。此外, 为了让读者更好地看到 ERP 效应, 一般来说, 不应该仅仅只画一个电极点的图(一种效应下), 尤其是 ERP 效应是本文最关注的效应以及最重要的发现。目前的画法过于简陋, 地形图对效应的表达的也不是很清晰。建议认真参考已发表文章 ERP 图的画法。对该文中的图表做出认真修改。

**回应:** 非常感谢审稿专家的建议。按照已有相关研究文献, 我们选取了 9 个电极点的波形图(每个兴趣区一个)进行展示, 同时, 通过 BESA Plot 软件画出了带有歌词与无歌词分别在 250~450ms 和 500~700ms 时间窗内的脑地形图。此外, 我们也对图中不规范之处进行修改, 详见第 10 页图 3 和图 4。

**意见 7:** 作者在引言部分的表述, 有些地方还是缺乏严谨性。例如: “对于远古人类来说, 由于缺乏语法, 语言和音乐的主要功能在于情绪表达。随着人类社会化程度的提高, 语言和音乐逐渐分化并朝着不同的进化方向发展。语言发展成为具有明确语义的符号交流系统。由于缺乏类似语言的明确语义性, 音乐则成为情绪表现的重要手段。”这种具有论断性但又难以证伪和证实的话语, 建议需要标明文献出处。并且, 建议不要采用过于绝对的表述方式。

**回应:** 非常感谢审稿专家的意见。我们已注明了文献出处, 并修改了表述方式。

**意见 8:** 引言部分, 作者提出证据“歌词对音乐情绪加工的影响也得到了计算机算法研究的证实(Hu, Downie, & Ehmann, 2009; Laurier, Lartillot, Eerola, & Toiviainen, 2009),。。。。。”这句话放在这个地方很突兀。建议稍微扩充一下, 说明是通过什么指标证实了歌词对音乐情绪加工的影响效应。

**回应:** 感谢审稿专家的仔细审阅。我们已扩充了相关表述, 详见前言第三段。

---