

《心理学报》审稿意见与作者回应

题目：情境真实性对悲伤移情调节的 ERP 证据

作者：闫志英 卢家楣

第一轮

审稿人 1 意见：

意见 1：本研究问题新颖，实验设计合理。尝试讨论真实情境与电影场景这一因素对悲伤移情的影响。

回应：无

意见 2：移情是源自理解他人的情绪状态或情境而产生的与其相一致的情绪反应(Eisenberg & Strayer。所谓移情，首先要他人的情绪作为源头，其次自身产生的情绪要与他人情绪源同性质，最后，移情者要有明确的自我-他人区分：知道自己的情绪是由于他人的情绪产生的。三个因素，缺一不可。作者在刺激材料的使用中，提出了典型性移情和非典型性移情的说法，来解释刺激材料的分类，但是作者没有给出这一分类的出处。仅仅给出了“Eisenberg 和 Strayer (1987)认为移情是种替代性的情绪反应，它可以由代表他人情绪状态的明显的知觉线索所引发，也可以是在间接情绪线索(如他人所处的情境)的基础上推理的结果。”事实上，这个理论对于解释刺激材料的问题尚不够充分。比如，看到地震中的尸体，产生的悲伤，源头是谁？是已经死去的人么？他们已经没有情绪了，如果他们是移情源头的话，那应该是零情绪。是看到尸体的人么？那不就是移情者自身么？自己对自己的移情是自身的悲伤而非移情悲伤。所谓通过情绪线索产生悲伤，是给出一个他人所处情景的描述，之后由想象产生与他人相同的情绪。比如，我告诉你，小张今天考试没有通过。你通过移情产生了与小张同样的负性情绪。因此，建议作者严格区分两种类型的图片，分别进行处理和讨论。这样既能得出自身悲伤相关的结论也能得出有力的移情相关结论。如果将两种条件进行比较，将很有意义。

回应：这是一个非常关键的问题。关于此问题，我们是基于悲伤的内涵进行考虑的：“悲伤是因自己喜欢、热爱的对象遗失,或期望的东西幻灭而引起的一种伤心、难过的情绪体验”(林崇德, 杨治良, 黄希庭, 2004, p.36)。在本研究中，当被试观看无他人明显情绪线索的地震情境图片时，自己没有任何喜欢、热爱的对象遗失，也没有期望东西的幻灭，因此被试此时的悲伤不是自己的，它来源于他人所处的情境，即移情者通过自己设身处地的推理或想象进而产生了与罹难者当时在此情境中所可能产生的情绪相一致的情绪反应。对于这一解释，我们还可从霍夫曼的理论中找到依据：霍夫曼认为移情的唤醒方式有五种，其中包含条件作用和直接联想。“人们会提这样的问题，当条件作用和直

接联想是由情境，而不是受害者的感受引起的时候，它们是否还是移情唤醒过程。我认为，只要观察者关注受害者，而且在观察者身上引起的感受适合于受害者的情境，而不是观察者的情境，它们就是移情唤醒过程” (马丁·L. 霍夫曼著，杨韶刚，万明译，2002, p.55)。

意见 3: 另外一个建议是，如果作者能做一做溯源分析，例如偶极子或电流密度图，看看结果，并结合 fMRI 在共情上发现的一些脑区结论进行讨论会更丰富有力。

回应: 这是一个非常好的建议，也是本研究进一步的研究方向。接下来，我们将在现有研究的基础上，进行溯源分析和 fMRI 研究，并结合其他的移情研究进行综合分析讨论，以丰富、深化移情研究。

意见 4: 摘要应该补充主要的 ERP 结果

回应: 已对 ERP 结果进行了相应补充。

意见 5: 基线是怎样确定的？从图 3 看，有些记录点的基线校正不够到位。

回应: 刺激前 100ms 用以基线校正。至于有些记录点的基线校正不够到位是由于用 PPT 做图时，用了一个坐标轴，可能有些图形的大小没有调整得跟坐标轴完全吻合，从而造成某些记录点的基线校正不够到位。

审稿人 2 意见:

意见 1: 首先本文书写粗糙，语言表述过于口语化，冗长繁琐，需要作很大程度上的修改，以使文章言简意赅。

回应: 本文已在文字上作了全面修改，克服口语化和冗长繁琐表述，力求言简意赅。

意见 2: 前言部分，文献回顾思路不清晰，问题提出不严谨。需要作者重新整理思路。

回应: 本文针对前言部分进行了重新梳理和修改，增强了前言部分论述的内在逻辑性、问题提出的严谨性，并结合以往研究存在问题提出了本研究的设想。

意见 3: 本研究是用 ERP 方法，但在文献回顾中相关的 ERP 研究过少，大篇幅讲的是磁共振研究。所以前言部分没有紧扣研究主题。建议作者参考 Gross 的认知重评类文献如“The temporal dynamics of emotion regulation: An EEG study of distraction and reappraisal” *Biological Psychology* 87 (2011) 84-92 等。

回应: 首先，对您推荐的文献“The temporal dynamics of emotion regulation: An EEG study of distraction

and reappraisal” *Biological Psychology* 87 (2011) 84-92”进行了阅读与参考；其次，前言部分 ERP 研究过少的问题。虽然移情方面有一些 ERP 研究，但涉及与本研究直接相关的情境与移情关系的 ERP 研究的确较少，但我们也尽量进行了搜索及挖掘，同时减少了 fMRI 研究所占的份量，以使前言部分紧扣研究主题。

意见 4：方法部分，实验材料的选择和评定讲述不清楚，程序图重新画。

回应：程序图已重新画过。实验材料的选择和评定部分重新进行了梳理，并作了相应更改，以求表达更为准确。

意见 5：作者研究悲伤移情，作者仅选择地震图片作为情绪诱发材料不具有普遍性，实验结论的外部效度欠缺。作者应交代清楚为什么选取地震类的图片。

回应：我们之所以选择“地震”作为研究移情情境的原因有两个：其一，地震是一种灾难性情境，具有较强的冲击力，同时具有强烈的社会现实意义。其二，地震为生活中真实发生的、切实可见的现象，特别是近年来通过媒体得知我国和世界各地地震信息不少，影响很大。与以往疼痛移情研究中使用的纯实验室材料（如疼痛手、疼痛脚图片）相比，具有较好的生态化效度。经专家鉴定，可以作为诱发悲伤移情的有效材料。并且，这也是本研究在实验材料选择上的一个特色。

意见 6：在行为结果中，图片类型和情境真实性的交互作用显著，作者的简单效应分析对文章无意义，简单效应分析应该是地震图片在电影和真实情境两种水平上的差异，以及生活图片在电影和真实情境两种水平上的差异。

回应：针对交互作用，简单效应分析已进行了调整，分析了地震（情境）图片和生活（情境）图片在电影和真实情境两种水平上的差异。

意见 7：脑电结果的分析应该对每个成分进行分析，而不是毫无依据得从时间窗口上寻找差异。包括后面的行为结果与波幅做相关也是如此。（时间窗口的选取包含的成分不纯粹，得到的研究成果没有意义。）

回应：本研究中对脑电结果的分析是在参考相关研究的基础上进行的，主要参考了 Fan 和 Han(2008) 疼痛移情研究中的测量方法（平均波幅连续测量法）、区间的划分和电极的选取，此外 Han 等人利用上述测量方法发表了疼痛移情的系列文章，具体如下：

Fan, Y., & Han, S. (2008). Temporal dynamic of neural mechanisms involved in empathy for pain: an event-related brain potential study. *Neuropsychologia*, 46, 160-173.

Li, W., & Han, S. (2010). Perspective taking modulates event-related potentials to perceived pain. *Neuroscience Letters*, 469, 328-332.

Han, S., Fan, Y., Mao, L. (2008). Gender difference in empathy for pain: An electrophysiological investigation. *Brain Research*, 1196, 85-93.

意见 8: 建议作者在平均图上标记出分析的时间窗口，便于读者理解。评价图上没有给出矢状线上的点（Cz、Pz、POz）。

回应: 评价图表示的是主观报告与神经活动之间的相关，即计算每个时间段地震情境图片 ERPs 的平均波幅，并分别计算其与被试观看地震情境图片时主观感受到的悲伤程度(即离线评价 FPS-R 得分)的相关，故不是指某一个具体的电极。

意见 9: 图四，相关图的横坐标和纵坐标分别代表什么，“FPS—R”代表什么作者也没说明。

回应: 图四（现为图 3）的横坐标和纵坐标分别代表：被试对于地震图片主观感受到的悲伤程度——FPS-R 得分和地震图片诱发的 ERP 平均波幅，即主观报告与神经活动之间的相关。“FPS-R”的含义已在图的标题中进行了说明。

意见 10: 讨论部分应紧扣实验结果。作者第二段提到“P1, N1, P1、N1”，在实验结果中，作者并没有对这几个成分进行分析。总体上，讨论不够充分深入。

回应: 已对讨论部分重新进行了修改。参考 Fan 和 Han(2008)的“Temporal dynamic of neural mechanisms involved in empathy for pain: an event-related brain potential study”，本研究同样主要考察的是时程特点，故并未过多强调成分，主要在描述总平均波形图时对成分进行了介绍。

审稿人 3 意见:

本文通过灾难场景刺激以及 erp 技术以诱发对他人悲伤情景的移情。如作者所言，文章选题具有较高的生态意义，且选题扩大了以往共情研究的对象范围，因此具备一定创新意义。

但文章对悲伤移情的操作性定义不清楚。对于如何将悲伤移情与单纯的情绪传(emotion contagion)区分开来，以及两者的连续和区别到底是什么未能交代，从而让读者觉得可能该实验诱发的是情绪传染（也即刺激材料并非诱发共情而是简单的诱发了直接的情绪（或悲伤）反应。即便是该材料诱发的悲伤，本文也缺乏对于该材料是否确实是诱发的悲伤也缺乏标准化。作者采用的评定程序无法确定悲伤情绪（因为作者仅列举了悲伤，高兴与中性三个类别；读者完全有理由认为图片也会诱发恐惧和厌恶）。

其次，本文对 erp 的分析选取了从额部到顶部 21 个电极点，但数据分析显得随意挑选，并未将电极作为一个单独的因素以分析成分的电压分布与因素的交互作用。此外，为何以 20ms 为单位进行连续测量没有清楚的交代（显得很随意，缺乏科学的严谨性）。此外，作者测量了如此多的时间

窗口，却仅仅在 160-180ms 以及 300-350ms 两个区间发现与行为测量的相关。为什么作者预期是在这两个窗口而不是其他窗口得到相关。为什么不是 400-500ms 或者 500-600ms？为什么不是 140-200ms 或 200-280ms？显然，作者的相关是为了求相关而得到的相关，并没有理论假设与预期。这种方式类似与大海捞鱼捞到一个算一个。因此，方法上的缺陷从而使得结论不可靠。这与前面所说的文章问题界定的问题（为什么这个材料诱发的悲伤而不是厌恶与恐惧？为什么这个材料诱发的是悲伤共情而不是情绪传染？）相重合而使得总体结论不可靠。从而使得这项工作更像是一个具体研究的预实验数据而不是达到发表水平之工作的思路，逻辑与严谨程度。

上述意见涉及以下三个问题：

意见 1：但文章对悲伤移情的操作性定义不清楚。对于如何将悲伤移情与单纯的情绪传(emotion contagion)区分开来，以及两者的连续和区别到底是什么未能交代，从而让读者觉得可能该实验诱发的是情绪传染（也即刺激材料并非诱发共情而是简单的诱发了直接的情绪（或悲伤）反应。

回应：一些研究者认为移情需要具有在自己-他人之间进行区分的最低能力(Feshbach, 1978; Kohut, 1978); 而另一些研究者认为这种自己-他人之间的区分并不重要(Hoffman, 1982)。两种观点之间的争论最终决定着情绪感染是移情还是其先导。Decety 和 Lamm(2006)认为移情意味着在自我、他人没有混淆的前提下，自我与他人之间体验到感受的相似性。De Vignemont 和 Singer(2006)对移情的概念进行了明确的限定，认为移情必须具备四个条件：(1) 处于一种情绪状态；(2) 与被移情者的情绪状态相一致；(3) 观察或想象他人情绪时被引发；(4) 知道他人的情绪状态是我们情绪状态的来源。根据第 4 点，移情不等于情绪感染。在上述基础上，本研究采用 Eisenberg 和 Mille(1987)对移情的界定：移情是源自理解他人的情绪状态或情境而产生的与其相一致的情绪反应。从情绪反应、一致、理解这三个关键词可以看出这一定义同样具备 De Vignemont 和 Singer 所提出的四个条件。

具体至本研究中，首先本研究中的地震图片有两种，其中一种是诱发非典型性移情的地震图片，这种图片中并没有明显的表现他人悲伤的情绪线索，故此时不存在“感染”，而需要移情者理解他人的情绪状态或情境，即移情者能够区分自己—他人；其次，在实验程序中，其中 5%的 trail 后要求被试完成一个评价任务：评价刚呈现过的图片的内容——他人的情绪或情境是悲伤的还是不悲伤的，因此移情者能够明确区分自己与他人的情绪状态。

意见 2：即便是实验材料诱发了悲伤，本文也缺乏对于该材料是否确实是诱发的悲伤也缺乏标准化。作者采用的评定程序无法确定悲伤情绪（因为作者仅列举了悲伤，高兴与中性三个类别；读者完全有理由认为图片也会诱发恐惧和厌恶）。

回应：这是非常好的问题。针对实验材料诱发的悲伤缺乏标准化这一问题，我们对生活图片诱发的悲伤程度进行了补测。每张移情图片均对应认同率和悲伤程度两个指标，其中悲伤程度为所有参评者对该移情图片感受到的悲伤程度评分的平均数。60 张地震图片诱发悲伤程度的平均数为 6.50，标

准差为 0.91；60 张普通生活图片诱发悲伤程度的平均数为 0.22，标准差为 0.44。

对于评定程序无法确定悲伤这一问题，我们的操作是这样的：首先，在实验开始前，我们通过指导语告诉被试，接下来你将看到多少组、多少张地震情境和生活情境两种图片，那么在“地震”这一大的背景下，尤其是多张地震图片的呈现，被试产生的悲伤情绪肯定是为主的；其次，从实验前和实验中对刺激材料的两次评定可以看出地震情境图片的确诱发了悲伤情绪，且诱发的悲伤程度非常明显 [实验前, 60 张地震情境图片和 60 张生活情境图片诱发悲伤程度(0-10 等级)的平均数分别为 6.50 和 0.22; 实验中, 我们又测定了被试对地震和生活两种情境图片主观感受到的悲伤程度, 其平均数分别为 6.62 和 0.37]。这表明在本研究中确实诱发了悲伤情绪；再次，至于图片是否诱发了其他情绪（即是否只诱发了悲伤情绪）：“移情是源自理解他人的情绪状态或情境而产生的与其相一致的情绪反应”，根据这一内涵，我们在前言中明确了判断移情发生的标准——当个体知觉或想象他人在某种情境中的情绪而引发与他人或其情境相应的情绪时移情就发生了。这一标准反映了移情的核心成分——移情者与被移情者之间一致的情绪反应。至于移情者与被移情者之间的情绪反应在多大程度上等同或者说是否要完全等同，关于此问题一直有所争议，故本研究在此不作探讨。

意见 3: 本文对 erp 的分析选取了从额部到顶部 21 个电极点，但数据分析显得随意挑选，并未将电极作为一个单独的因素以分析成分的电压分布与因素的交互作用。此外，为何以 20ms 为单位进行连续测量没有清楚的交代（显得很随意，缺乏科学的严谨性）。此外，作者测量了如此多的时间窗口，却仅仅在 160-180ms 以及 300-350ms 两个区间发现与行为测量的相关。为什么作者预期是在这两个窗口而不是其他窗口得到相关。为什么不是 400-500ms 或者 500-600ms？为什么不是 140-200ms 或 200-280ms？显然，作者的相关是为了求相关而得到的相关，并没有理论假设与预期。这种方式类似与大海捞鱼捞到一个算一个。因此，方法上的缺陷从而使得结论不可靠。这与前面所说的文章问题界定的问题（为什么这个材料诱发的悲伤而不是厌恶与恐惧？（见问题 2）为什么这个材料诱发的是悲伤共情而不是情绪传染？（见问题 1）相重合而使得总体结论不可靠。从而使得这项工作更像是一个具体研究的预实验数据而不是达到发表水平之工作的思路，逻辑与严谨程度。

回应: 本研究中对脑电结果的分析是在参考相关研究的基础上进行的，主要参考了 Fan 和 Han(2008) 疼痛移情研究中的测量（平均波幅连续测量法）、区间的划分和电极的选取方法，此外 Han 等人利用上述方法发表了疼痛移情的系列文章，具体如下：

Fan, Y., & Han, S. (2008). Temporal dynamic of neural mechanisms involved in empathy for pain: an event-related brain potential study. *Neuropsychologia*, 46, 160–173.

Li, W., & Han, S. (2010). Perspective taking modulates event-related potentials to perceived pain. *Neuroscience Letters*, 469, 328-332.

Han, S., Fan, Y., Mao, L. (2008). Gender difference in empathy for pain: An electrophysiological

第二轮

审稿人 1 意见：

意见 1：针对审稿人三的主要问题，作者的解释不尽满意。即使自己没有喜欢的对象遗失，期望的东西幻灭，也不代表悲伤情绪必然来自移情。罹难者在罹难前中会产生最大情绪是恐惧而非悲伤。难以解释观看者的悲伤是与罹难者共情。看到罹难者尸体，与看到地震后的断壁残垣一样，是负性刺激导致的自身诱发的悲伤情绪。近年来对共情的实验操作定义一直在具体化，统一化。没有严格的，统一的，最新的，被广泛接受的实验室操作定义，如何在这一松懈的理论基础上进行数据分析和结果讨论？作者应该加以考虑。另一方面，如同审稿人二提出的问题：在材料评定中，作者仅仅要求被试评定图片的悲伤程度，而事实上地震图片包含了复杂的情绪，悲伤，恐惧，惊讶。作者的材料评定过于单一和强迫选择，因而在实验中诱发的情绪纯度有待商榷。

回应：关于问题的第一个层面，其实我们内部也进行过讨论，甚至争论。最初的考虑为只要移情者的情绪反应与被移情者的情绪反应或与其情境相一致即可，具体而言：对于典型性移情而言，我们能够明显看到有所失去的人在哭泣，如抱着死去的孩子在痛哭；而在“情境”这一情况下，则隐藏了哭泣的那个人，观看者看到的是尸体或废墟下的一段手臂，此时观看者的悲伤依然是与失去罹难者的亲人产生了共情，悲伤的源头是失去死者的亲人。也正因如此，我们在悲伤他们悲伤的同时，往往也会伸出援助之手，而援助的对象也是有所失去但仍然活着的人。

后来进行讨论时，发现这种解释比较难推广到其他移情，而移情应具有一定的普遍性，故经过讨论与一定的调查，我们感觉与罹难者共情这一解释更具有普遍性，并且同样具有一定的依据：通过与地震灾区受灾个体的接触，我们发现当地震来临时，由于情感本身的复杂性，情绪的产生根据时间序列而有所不同。的确，他们第一时间的情绪反应往往首先是恐惧，但恐惧之后产生的主要是悲伤：难过自己的境遇，伤心自己家人、财产的失去或自己鲜活生命的可能幻灭。故在此，我们也想特地请教一下您，对于此问题，您有何看法或宝贵的意见或建议？研究本身也是思考和学习的过程，能够有这样的机会得到您的指导，实乃是一件幸福的事情。

其次，霍夫曼认为移情的唤醒方式有五种，其中包含条件作用和直接联想。“人们会提这样的问题，当条件作用和直接联想是由情境，而不是受害者的感受引起的时候，它们是否还是移情唤醒过程。我认为，只要观察者关注受害者，而且在观察者身上引起的感受适合于受害者的情境，而不是观察者的情境，它们就是移情唤醒过程”（马丁·L. 霍夫曼著，杨韶刚，万明译，2002, p.55）。在本研究中，我们认为观察者产生的感受更适合受害者的情境，而不是自己的情境，因此同样符合霍夫曼的观点。

对于问题的第二个层面，即图片是否仅诱发了悲伤情绪，专家的疑问是完全合理的，我们的考虑是：其一，由于情感本身的复杂性，客观事物在个体身上引起的反应往往不是单一的，就国外疼痛移情研究中常用的疼痛手图片（如剪刀剪手、针刺入手等）而言，它所诱发的情绪可能也不仅是疼痛，也有可能产生害怕、恐惧等其他情绪反应。本研究中通过指导语有意识的引导被试在可能并不单一的情绪反应中产生某种主要情绪，是可行的，也是允许的。具体操作如下：实验开始前，通过指导语告诉被试，接下来你将看到多少组、多少张地震情境和生活情境两种图片，那么在“地震”这一大的背景下，尤其是多张地震情境图片的呈现，被试主要产生悲伤情绪是合乎情理的。这也同2008年汶川地震时，当我们看到电视上母亲抱着死去的孩子痛哭，甚至看到死亡儿童的书包时，我们的眼泪都会夺眶而出的现实经验相一致。相反，如果没有任何指导语，并且只有一两张图片呈现给被试，此时被试产生的情绪反应是复杂的，有可能产生惊奇（发生了什么事情）、恐惧、厌恶等多种情绪反应；其二，至于图片是否诱发了其他情绪，我们在前言中明确了判断移情发生的标准——当个体知觉或想象他人在某种情境中的情绪而引发与他人或其情境相应的情绪时移情就发生了。这一标准反映了移情的核心成分——移情者与被移情者之间一致的情绪反应。至于移情者与被移情者之间的情绪反应在多大程度上等同或者说是否要完全等同，关于此问题一直有所争议，故本研究不作探讨。

意见 2：关于溯源分析的问题，作者实际上回避了这一问题。请根据审稿人要求进行溯源分析，如果结果不够理想，请在回复审稿人信件中附上结果。如果结果合理，请补充在文章中，并根据审稿人建议结合 fMRI 在共情上的文献进行讨论。希望作者认真加以考虑，特别是第二个问题，不是将来怎么样，而是现在就要实际尝试一下。

回应：结合您的宝贵建议，我们克服现实存在的困难对 ERP 数据补充进行了溯源分析（ERP 数据采集自上海，目前已毕业返回昆明），具体结果如下：

（一）本研究的溯源结果

由于真实情境和电影情境中地震情境图片和生活情境图片均存在显著性差异，故采用 BESA 软件分别对两种情境中地震情境图片和生活情境图片诱发的 ERP 差异波作偶极子溯源分析。选定 350-400ms、450-500ms 两个时间窗口，进行主成分分析（PCA）。

350-400ms 时间窗：对于真实情境而言，经 PCA 分析，结果显示有一个成分可以解释总变异的 97.1%，主要定位于后扣带回（posterior cingulate gyrus）（三维坐标以 Talairach 为参考：X=11.8, Y=-20.1, Z=41.1），残差为 7.135%（详见图 1 中的两个图）；对于电影情境而言，经 PCA 分析，结果显示有一个成分可以解释总变异的 95.8%，主要定位于额内侧回（medial frontal gyrus）（三维坐标以 Talairach 为参考：X=13.5, Y=-7.8, Z=49.1），残差为 8.676%（详见图 2 中的两个图）。

450-550ms 时间窗：对于真实情境和电影情境，经 PCA 分析，结果均显示残差过大，分别为 75.167% 和 69.925%。

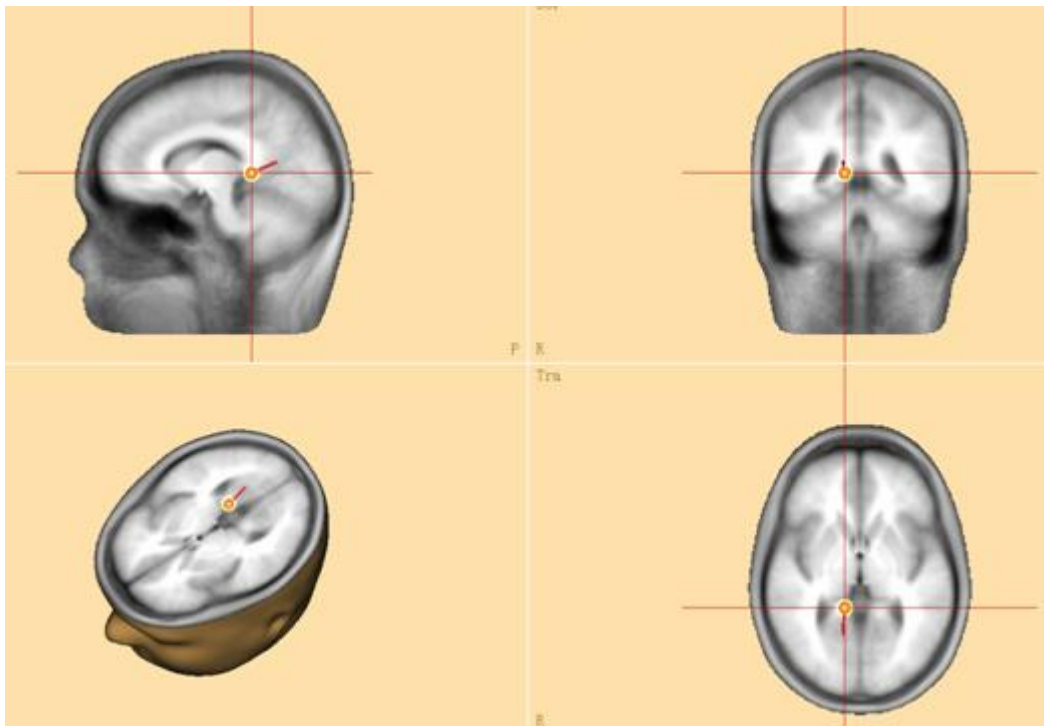
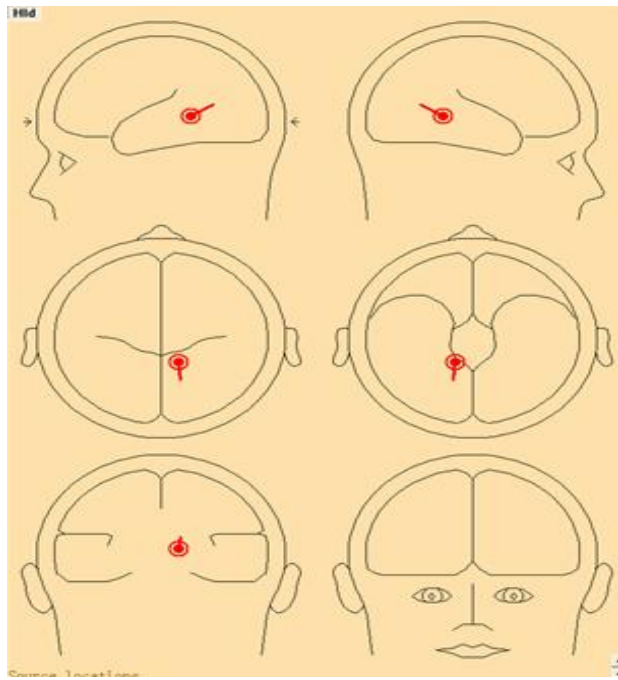


图 1 350-400ms 真实地震情境图片—真实生活情境图片差异波的溯源分析结果

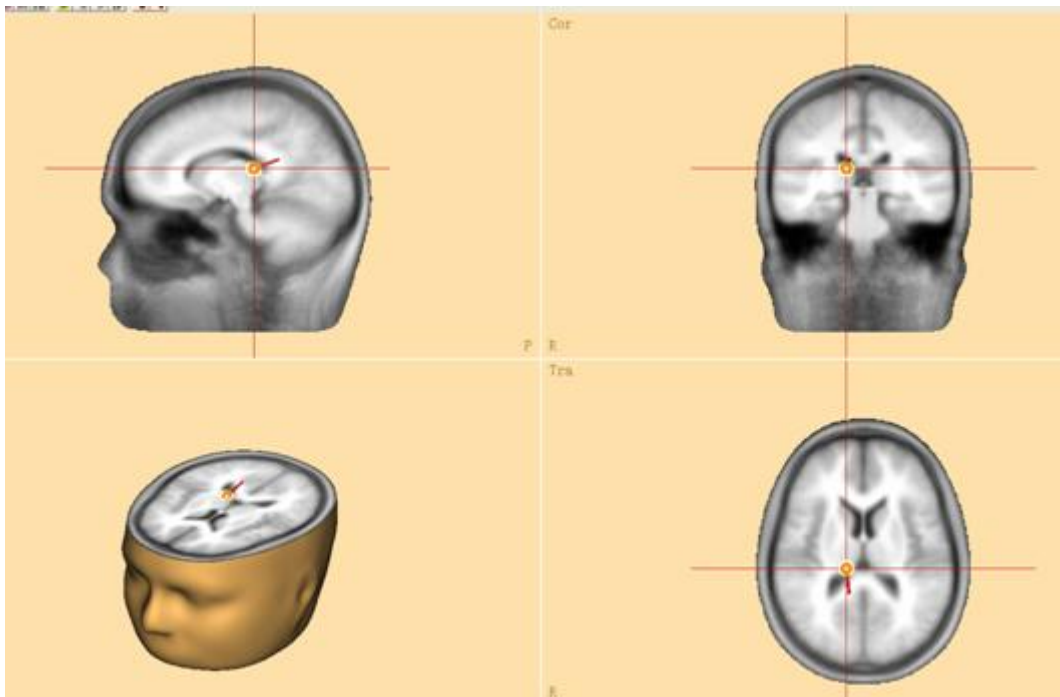
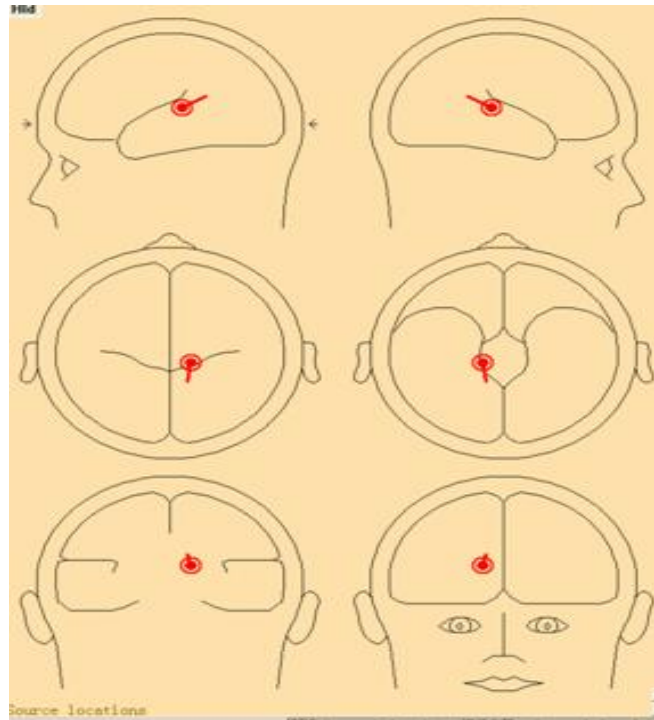


图 2 350-400ms 电影地震情境图片—电影生活情境图片差异波的溯源分析结果

(二) 本研究结果与相关 fMRI 研究结果的对比

本研究的结果显示 350-400ms, 真实情境中地震情境图片和生活情境图片 ERPs 的差异主要定位于扣带回, 偏向后扣带回; 而电影情境中地震情境图片和生活情境图片 ERPs 的差异则主要定位于

额叶(frontal lobe), 偏向于额内侧回。

移情研究的结果表明脑岛—扣带区域在情感分享中起着重要作用。因此, 本研究的溯源结果意味着真实情境中(同电影情境相比)表现出更为强烈的情感色彩; 其次, 认知评价在本质上是一种情绪调节方式, 即使在本研究中并非将其直接用于调节情绪本身。有研究表明背内侧前额叶皮层(dorsomedial prefrontal cortex)似乎与认知和执行过程相关, 并且前额叶皮层(prefrontal cortex)不断被发现与情绪调节有关(Ochsner& Gross, 2005); 眶额皮层(orbitofrontalcortex)在调节移情反应中具有重要作用也被一些 fMRI 研究所证实(Lamm, Batson, &Decety, 2007; Singer et al., 2006)。所以, 电影情境中两种情境图片 ERPs 的差异定位于额叶, 偏向于额内侧回(虽然这一区域并未能与上述研究中的区域完全一致, 可能与溯源带来的误差有关), 这一结果意味着在 350-400ms 时间区间, 在电影情境中, 个体表现出有意识的认知和控制——这是电影情境、不是完全真实的, 从而呈现出与真实情境中有所不同的移情反应。

另两项探讨同样在基于诱导引起的认知评价的基础上形成的情境与移情关系的 fMRI 研究发现: 同知觉正常的疼痛情境相比, 知觉看似疼痛实为非疼痛情境(对麻醉手进行活体组织检查)激活了代表自我—他人区分的颞顶联合区(temporo-parietaljunction)和前额叶皮层; 知觉仅对观察者认为是疼痛的情境(对被移情者为非疼痛情境)激活的区域包括自我—他人区分的背内侧前额叶皮层和认知控制的右侧下额叶皮层(right inferior frontal cortex) (Lamm et al., 2009; Lamm et al., 2007)。本研究溯源的结果同这两项 fMRI 研究的结果不完全相同, 不同之处在于: ①后者激活的区域包括自我—他人区分的部分; ②前者、后者激活额叶的具体位置有所不同。对于第一点不同, 或许并不难理解, 他们研究中设置的情境中均有一类为对于被试为疼痛情境、但对于被移情者并非为疼痛情境, 故移情时需要考虑自己—他人之间的不同, 从而代表自己—他人区分的脑区的激活非常明显; 对于第二点不同, 或许与本研究溯源本身所具有的误差有关, 从而导致了具体位置的偏差。

另一 fMRI 研究也探讨了基于诱导引起的认知评价对疼痛移情神经活动的影响: 告知被试因病人患有某种神经疾病, 需要接受一种会带来疼痛的治疗, 两组被试中有一组的治疗是“有效的”, 而另一组的治疗是“无效的”。结果发现, 同有效组相比, 知觉无效组病人的疼痛时前扣带回下部(perigenual ACC)、腹内侧眶额皮层(ventromedial orbitofrontalcortex)的激活更为强烈(Lamm, Batson, &Decety, 2007)。这同本研究的结果也不尽相同: 相似的是扣带区域的激活, 这同本研究真实情境中两种情境图片 ERP 差异波的定位也在扣带区域相类似; 不同的是腹内侧眶额皮层的激活。两项研究结果的差异或许也与本研究溯源本身所带来的误差有关。

(三) 结论

对基于诱导引起的认知评价而形成的不同情境中的移情反应进行溯源分析, 其结果与移情的 fMRI 研究结果有相似之处, 但不完全相同。

审稿人 2 意见：

意见 1：按照题目一般会将其理解为本研究应该是“情境真实性”和“移情”为两个不同的变量，而事实上，文章是将前者作为后者的一个限定条件而已，容易产生歧义。

回应：根据审稿专家的意见，为避免歧义，如将本文题目改为“情境真实性对悲伤移情调节的 ERP 证据”，不知当否？

意见 2：副标题为“基于诱导……”，那么，可以推断基于诱导对应的另一条件为“基于非诱导”或是与基于其它条件的比较，但从行文中不能很清晰地看出这一逻辑。

回应：考虑到审稿专家的意见，本文删除副标题，不知当否？

意见 3：从摘要看，“电影、真实地震情境图片诱发的 ERP 平均波幅分别在 160-180ms、300-350ms 与主观悲伤程度相关显著，佐证了两种情境的分离”这一描述并没有涉及到移情，有点不能突出主题。

回应：摘要中已进行了相应的补充，以更好地突出“移情”这一主题。

意见 4：在前言中提到“...对疼痛移情的调节发生在不同的时间——分别发生在早期和晚期阶段...”，但是从上下文中并不能很清晰地看出早晚阶段各自的时间范围是多少。

回应：感谢专家的细致，已将早期和晚期阶段的时间范围进行了说明。

意见 5：前言中提到移情的情境有：一、基于诱导引起的认知评价所形成的情境，二、基于客体引起的认知评价所形成的情境。但文章在本实验材料分类时则并没有很好地说明本研究中那些材料是基于诱导的以及那些材料是基于客体的。若按照文章提到的“本研究采用事件相关电位(ERP)技术，探讨生活中常见的、基于诱导引起的认知评价所形成的真实性程度不同的两种情境——真实情境和电影情境对移情的调节发生的时间进程。”这一句话来看，文章的材料均是基于诱导引起的认知评价所形成的情境，而没有基于客体引起的认知评价所形成的情境。是这样吗？若是这样，实际上本研究是如何控制诱导这一条件的呢？此问题的介绍不够清晰。

回应：正如专家所言，本文的“情境”是基于诱导引起的认知评价所形成的情境，不是基于客体引起的认知评价所形成的情境；根据专家意见，已在文中 2.2.4 对“诱导”这一条件的控制进一步进行了明确。

其次，至于您提出的两类情境所对应的实验材料分别是什么，可能是由于我们在前言中因篇幅限制没有进行详细的阐述，只列出了研究者和年份。两类情境所对应的实验材料具体是：基于诱导引起的认知评价所形成的情境，主要是指通过不同的指导语从而诱导被试产生不同的认知评价，进而形成不同的情境，如 Lamm 等（2009）设置了正常疼痛(用针管注射麻醉药到手中)和看似疼痛实为非疼痛(告诉被试图片中是对麻醉了的手进行活体组织检查，故并不是疼痛的)两种情境；基于客

体引起的认知评价所形成的情境，主要是指照相图片（即正常的图片）情境(疼痛手和非疼痛手两种情境)和卡通图片情境(由照相图片情境经技术处理而成)之间的不同(Fan & Han, 2008; Gu & Han, 2007)，这两类情境之间差异主要是由于材料本身特点的不同引起了被试不同的认知评价。

意见 6:“刺激材料类型的确立”部分提到“...采用反映普通生活情境的图片”，照此描述，一定是将“普通生活情境”与“特殊生活情境”或与其它一类生活情境相对应的条件。但这样确定条件的方式是否有一定的依据？是自己界定“普通生活情境的图片”这一条件的还是采用于前人那个文献呢？这个问题需要说明。

回应:此问题是表达的问题，表达不够确切。主要是想挑两类情境图片，一类是反映人类活动的可以诱发悲伤移情的地震情境，一类是反映人类活动的普通生活情境，“普通生活”是与“地震”这一特殊情境相对应。但正如专家所言，这样表达容易引起歧义，故用“生活图片”更为准确，已在文中进行了相应修改。

意见 7:“基于此，本研究把移情划分为两种类型：典型性移情和非典型性移情。”这样划分的依据是什么？有前人文献依据吗？如果没有，为何采用这样的划分？

回应:典型性移情和非典型性移情划分的理论依据主要有以下两个：其一，Eisenberg 等人的理论观点。Eisenberg 和 Strayer(1987)认为移情是一种替代性的情绪反应，它可以由代表他人情绪状态的明显的知觉线索所引发，也可以由在间接情绪线索(如他人所处的情境)基础上的推理结果所引发。其二，卢家楣(2000)将情绪发生的心理机制概括为以境生情和以情生情二类。虽然移情与情绪并不等同，但情绪发生的心理机制对移情具有借鉴作用。

具体而言，典型性移情由他人明显的情绪状态引发，表现为移情者由于直接观察或知觉他人的情绪状态，进而产生了与其相一致的情绪反应，即“以情生情”；非典型性移情由他人所处的情境引发，表现为移情者根据他人所处的情境，通过自己设身处地的推理或想象，进而产生了与他在该情境中所可能产生的情绪相一致的情绪反应，即“以境生情”。

意见 8:材料筛选时，首先让 2 名研究生，2 名副教授对图片进行初步评定，如果按照“移情图片的筛选”这一部分的说明，其实仅是对于图片是否是“地震情境和生活情境”的区分，而没有涉及到移情的层次。但文章直接把这部分工作界定为筛选“移情图片”，不太妥当。

回应:感谢专家的认真与精益求精，已在文中进行相应的修改。

意见 9:在设计 14-15 页中下标列举如下问题时，存在严重不足。如“图片中人物的情绪(或其情境)是()，请在所选答案后的括号内划对号：a 高兴的() b 中性的() c 悲伤的()”。题目仅给出了三个选项，事实上，这种答案有强迫选择的倾向，因为对于负性情绪图片，有可能其情绪特

性是痛苦、厌恶、害怕、惊讶等等的评定，但由于负性情绪选项中只有一项悲伤，被试无法做出其它情绪判断。换言之，如果在给出多种负性情绪选项，而大多数被试都认为相应图片是悲伤的，而不是痛苦、厌恶、害怕、惊讶等其它情绪特性的前提下，才能说明本文的材料筛选是比较合理的。本文没有进行这样的操作，这样的话，将有可能是其它负性情绪色彩的材料一并归为悲伤情绪特点，是非常不妥的。

回应：专家的疑问是完全合理的，我们的考虑是：其一，由于情感本身的复杂性，客观事物在个体身上引起的反应往往不是单一的，就国外疼痛移情研究中常用的疼痛手图片（如剪刀剪手，针管深深刺入手中等）而言，它所诱发的情绪可能也不仅仅是疼痛，也有可能产生害怕、恐惧等其他情绪反应。本研究通过指导语有意识的引导被试在可能并不单一的情绪反应中产生某种主要情绪，是可行的，也是允许的。具体操作如下：实验开始前，通过指导语告诉被试，接下来你将看到多少组、多少张地震情境和生活情境两种图片，那么在“地震”这一大的背景下，尤其是多张地震情境图片的呈现，被试主要产生悲伤情绪是合乎情理的。这也同 2008 年汶川地震时，当我们看到电视上母亲抱着死去的孩子痛哭，甚至看到死亡儿童的书包时，我们的眼泪都会夺眶而出的现实经验相一致；相反，如果没有任何指导语，并且只有一两张图片呈现给被试，此时被试产生的情绪反应是复杂的，有可能产生惊奇（发生了什么事情）、恐惧、厌恶等多种情绪反应。

其二，至于图片是否诱发了其他情绪，我们在前言中明确了判断移情发生的标准——当个体知觉或想象他人在某种情境中的情绪而引发与他人或其情境相应的情绪时移情就发生了。这一标准反映了移情的核心成分——移情者与被移情者之间一致的情绪反应。至于移情者与被移情者之间的情绪反应在多大程度上等同或者说是否要完全等同，关于此问题一直有所争议，故本研究不作探讨。

意见 10：同理，在下标中的“问题 2”和“问题 3”同样是强迫选择任务，因为没有其它负性情绪的选择，被试只能做出悲伤以及悲伤程度的判断。

回应：同第 9 题。

意见 11：根据“每组包含地震情境图片和生活情境图片各 30 张”这一描述，每一条件的图片为 30 张（这一数量并不算多），如果按照 ERP 的要求，每种条件的 trail 数目需要达到一定的数量，文章没有说明在剔除伪迹后每一条件叠加的试次，因此，很难看出文章的 ERPs 叠加结果是否符合发表标准的需要。

回应：电影、真实情境两组图片每组包含地震情境图片和生活情境图片各 30 张，重复 2 次，即真实、电影情境图片的 block 各出现 3 次，即每种条件的 trail 数为 90 个，具体可见“2.3 实验程序”部分。

意见 12：将图片随机分到电影和真实情境两种条件，通过不同的指导语来引起被试产生不同的认知评价。但实际上，在正式实验中，被试只对其中的“5%的 trail 后要求被试完成一个评价任务，评价

刚呈现过图片的内容是否悲伤。”。由此可以看出，被试并不是对所有图片都进行判断，如果是这样的话，图片的叠加是按照什么结果进行的，按照正确判断还是其它依据？

回应:本研究采用 Decety 等人(2010)在其疼痛移情研究中所采用的实验范式,被试的主要任务为看,其中 5%的 trail 后要求被试完成评价任务的目的是为了保证被试集中注意力(结果也表明所有的被试均能完成正确的评价),因此叠加是按图片的类别进行的。

意见 13:“ERP 数据采集与分析”中提到“因额区、中央区 ERP 数据方差分析结果相似,这里仅报告中央区、顶区、枕颞区电极统计结果”,这样描述的逻辑有点混乱,前额区的数据和波形图在文章中并没有呈现,读者无法看出额区的结果是什么,还有,是否因为结果相似就不需要报告了?这样描述的依据是什么?前面提到分析的电极而在后面又不分析额区的结果了。显得逻辑不清。

回应:已增加额区的两个电极的波形图。因额、中央区结果相似,故仅报告了一个脑区。之所以这么描述,因疼痛移情的研究中有类似省略性的描述:例如,Fan 和 Han(2008)认为前部和后部的电极具有相似的结果,故他们仅报告 C3-4(代表额中区)和 PO7-8(代表枕颞区)的统计结果。

意见 14:已增加额区的两个电极的波形图。因额、中央区结果相似,故仅报告了一个脑区。之所以这么描述,因疼痛移情的研究中有类似省略性的描述:例如,Fan 和 Han(2008)认为前部和后部的电极具有相似的结果,故他们仅报告 C3-4(代表额中区)和 PO7-8(代表枕颞区)的统计结果。

回应:电生理数据部分的前两段主要是对波形图的整体描述,后面的分析则主要是分时段进行。

意见 15:在波形图中,并没有说明为何选择列举的 4 个电极、而未列出其它电极。

回应:电极的选择主要是基于每个区选取两个电极的原则进行的,因篇幅限制未列出选取的所有电极。

意见 16:在结果报告中,应该列出差异波地形图,并与波形图相对照,这样的话更容易让读者明晰结果发生的区域和部位。

回应:已增加结果显著脑区电极的差异波地形图。

意见 17:在 ERP 结果分析中,分析了很多时段,但没有很清晰地说明哪些时段为早期阶段,哪些时段为晚期阶段。另外,分析时段与前言中介绍的时段是否一致?似乎并不完全一致?取这些时段的原因是什么?不同时段所代表的意思是什么?文章需要做出交代。

回应:已在文中补充了早期阶段和晚期阶段的时间范围,同时对早、晚期时间段进行了更为详细的阐述;分析时段的选取是在参考相关研究的基础上进行的,主要参考了 Fan 和 Han(2008)疼痛移情研究中的测量方法(平均波幅连续测量法)、区间的划分及电极的选取,也可见 2.4 部分。

意见 18: 在讨论中大篇幅地介绍本文材料的意义, 这部分更应该是放在前言或者实验材料部分说明的。讨论将行为结果和 ERP 结果基本是作为并列重要的内容来说明, 有点重点不突出。不能让读者看明白, 本研究主要侧重于行为还是侧重于 ERP 研究。

回应: 这样处理是因为, 此次实验的一个创新点是材料创新, 故加以突出。正如 B.Sporrow(2001) 在 *Science* 杂志上发文指出的那样, 实验材料创新是实验创新重大四大方面(测验、指导语、材料、被试)之一。基于此, 在讨论中用了较大的篇幅对实验材料的可行性及其与可能相互混淆的概念之间的差异进行了深入的分析, 如放在前言或实验材料部分, 可能不能充分说明这两个问题。

意见 19: 另外, 我同意前面专家提出的意见。对于选择地震图片的意图、图片情绪性的确定、对悲伤移情的界定、分析时段的选择等均是需给与更详细说明的。注意: 需给出文献依据, 如果没有文献依据, 请说明作者自己的理由。

回应: 我们之所以选择“地震”作为研究移情情境的原因有两个: 其一, 地震是一种灾难性情境, 具有较强的冲击力, 同时具有强烈的社会现实意义。其二, 地震为生活中真实发生的、切实可见的现象, 特别是近年来通过媒体得知我国和世界各地地震信息不少, 影响很大。与以往疼痛移情研究中使用的纯实验室材料(如疼痛手或疼痛脚图片)相比, 具有较好的生态化效度。经专家鉴定, 可以作为诱发悲伤移情的有效材料。并且, 这也是本研究在实验材料选择上的一个特色。

其次, 正文在前言的最后一段进一步明晰了移情的界定, 悲伤移情只是移情的一种, 同样符合移情的定义。

再次, 本研究中对脑电结果的分析是在参考相关研究的基础上进行的, 主要参考了 Fan 和 Han(2008)疼痛移情研究中的测量方法(平均波幅连续测量法)、区间的划分和电极的选取, 此外 Han 等人利用上述测量方法发表了疼痛移情的系列文章, 具体如下:

Fan, Y., & Han, S. (2008). Temporal dynamic of neural mechanisms involved in empathy for pain: an event-related brain potential study. *Neuropsychologia*, 46, 160-173.

Li, W., & Han, S. (2010). Perspective taking modulates event-related potentials to perceived pain. *Neuroscience Letters*, 469, 328-332.

Han, S., Fan, Y., Mao, L. (2008). Gender difference in empathy for pain: An electrophysiological investigation. *Brain Research*, 1196, 85-93.

第三轮

审稿人 1 意见:

意见 1: 对审稿意见的回答较满意，同意发表。

回应: 再次感谢您对本文提出的宝贵建议，使笔者受益匪浅，同时感谢您这么久以来的指导与帮助。

审稿人 2 意见:

作者针对审稿人的意见进行了认真的修改，大部分修改是有效的和较为满意的，但是希望对以下几个问题进行深入的思考和修改:

意见 1: 本篇的核心概念是 Empathy，翻译有“共情”和“移情”，而译为“共情”更为常见，建议作者统一。

回应: 首先感谢专家的严谨。我们在正文中进行了相应的说明。对于这一问题，我们也曾进行过思考并查阅了中文的相关文献：移情常被作为精神分析学说的用语，而共情也常被用于临床心理学中咨询师或治疗师身上，但在一般的研究中又未对两者进行严格的区分（本研究亦是如此）。故在此，特地向您请教，两种当中哪个译法更为恰当？如您认为“共情”的译法更为妥当，我们表示接受。

意见 2: 在表示结果时小于号、等于号之类的数学符号的前后要用空格分开。

回应: 已将小于号或等于号的结果格式进行了相应调整。

意见 3: 对交互作用进行事后比较的具体统计方法应该标出来。

回应: 已在正文中进行了补充。

审稿人 3 意见:

意见 1: 在回答专家的问题时，指出“我们的考虑是：其一，由于情感本身的复杂性，客观事物在个体身上引起的反应往往不是单一的，就国外疼痛移情研究中常用的疼痛手图片（如剪刀剪到手，针刺入手等）而言，它所诱发的情绪可能也不仅是疼痛，也有可能产生害怕、恐惧等其他情绪反应。本研究中通过指导语有意识的引导被试在可能并不单一的情绪反应中产生某种主要情绪，是可行的，也是允许的。”，我则认为这是不可以的，因为其反应不是被试的真实情绪体验。

回应: Lazarus 和 Spisman (1964)曾做过一个实验，安排四组被试分别观看同一部反映某原始社会男子成人仪式的电影，内有不少骇人听闻的残酷画面。对 A 组，用讲解来加强画面的残酷性；对 B 组，用讲解来否定画面的残酷性；对 C 组，用超然态度讲解；对 D 组，则让其观看无声电影。结果发现，同样的画面刺激导致了不同强度的紧张情绪：A 组最强烈，D 组较弱，B、C 组最弱。另有研究发现，通过指导语使被试对情绪信息进行认知重评能够降低观看情绪图片带来的 LPP 波幅升高的现象 (Hajcak & Nieuwenhuis, 2006; Moser, Krompinger, Dietz, & Simons, 2009)。

上述研究表明，外界的诱导的确能够有效影响情绪的产生。本研究与上述研究虽不完全一致，但究其本质是相同的，因此同样可通过操纵指导语使被试产生的情绪趋于悲伤。具体操作：实验开始前，通过指导语告诉被试，接下来你将看到多少组、多少张地震情境和生活情境两种图片，那么在“地震”这一大的背景下，尤其是多张地震情境图片的呈现，被试产生悲伤情绪是合乎情理的。这也同 2008 年汶川地震时，当我们看到电视上母亲抱着死去的孩子痛哭，甚至看到死亡儿童的书包时，眼泪都会夺眶而出的现实经验相一致。相反，如果没有任何指导语，并且只有一两张图片呈现给被试，此时被试有产生悲伤、惊奇（发生了什么事）、恐惧、厌恶等多种情绪的可能性。

意见 2:“基于诱导引起的认知评价所形成的情境”与“而没有基于客体引起的认知评价所形成的情境”在正文中的描述还是不够清晰。

回应:对两类情境的特点进一步进行了补充。

意见 3:正文中“本研究把移情划分为两种类型：典型性移情和非典型性移情”。但并没有标出出处，因此不能看出，这是作者自己的观点，还是借鉴前人的观点？如果是自己的观点其合理性在哪里？尽管前面修改部分也说明了一些，但正文中描述不直观。

回应:这种划分是本研究提出的观点，已在正文中进行了补充。其主要依据为：其一，Eisenberg 和 Strayer(1987)认为移情是一种替代性的情绪反应，它可以由代表他人情绪状态的明显的知觉线索所引发，也可以由在间接情绪线索(如他人所处的情境)基础上的推理结果所引发；其二，卢家楣(2000)将情绪发生的心理机制概括为以境生情和以情生情两大类。虽然移情与情绪并不等同，但情绪发生的心理机制对移情具有借鉴作用。在上述依据的基础上，本研究把移情划分为两种类型：典型性移情和非典型性移情。典型性移情即“以情生情”；非典型性移情即“以境生情”。

意见 4:既然进行了溯源分析，为何在修改后的正文中没给出相应结果和描述？另外，为何修改部分提出的溯源时段与结果分析部分的时段不完全一致？

回应:溯源分析是根据另一评审专家意见而进行的尝试，故并未放在文章正文当中。根据溯源的一般原则，溯源时间段为结果显著部分时间段，故在本研究中具体为 350-400ms 和 450-500ms 两个时间窗口。

意见 5:“电影和真实地震情境”哪个更真实？区分真实与非真实的依据是什么？电影情境难道全是虚构的吗？除指导语外，被试是否还有其它参照可以区分这两种情境，如若不能，本设计的人为性很大，并不能反映被试的真实判断。

回应:Gu 和 Han(2007)采用 fMRI 技术探讨了刺激真实性对疼痛移情神经过程的影响，刺激材料为照相图片情境(疼痛手和非疼痛手两种情境)和卡通图片情境(由照相图片情境经相应技术处理而成)。

Fan 和 Han(2008)采用 ERP 技术进一步探讨了刺激真实性（材料同上，照相图片情境和卡通图片情境）对疼痛移情神经过程的影响。他们认为相对于照相图片情境而言，卡通图片情境的真实性程度有所降低；其次，Lamm 等人（2007，2009）探讨情境与疼痛移情关系的研究中对情境的设置则是通过不同的指导语调控被试对情境的不同理解而实现的。

本研究中情境的设置虽与上述研究有所不同，但研究目的相同，旨在探讨情境真实性对移情神经过程的影响。实验时，告诉被试一部分移情图片来自于汶川地震中的真实地震情境和生活中的真实生活情境，另一部分移情图片来自于电影情境。电影情境并非完全虚构，但其真实性程度相对于真实情境而言有所降低。

意见 6: 在回答前面专家的问题时，有很多不妥之处，如修改部分指出“如将本文题目改为“情境真实性对悲伤移情调节的 ERP 证据”，不知当否？”和“考虑到审稿专家的意见，本文删除副标题，不知当否？”但随后的正文并没有接受这些修改。

回应: 感谢专家的认真与细致，上次正文未进行修改为笔者的疏忽，为此向您致以最为诚挚的歉意；此次已在正文中进行了相应修改。

意见 7: 在修改中既然认为“专家的疑问是完全合理的”。但文章在情绪图片的评定中仅仅有悲伤一种负性情绪，这很明显是不合理的，这种迫选任务并不是被试真实意思的反映。只能是被试迫于要求而做出非真实意思的反应。这样的话，结果真实性有待商榷。尽管作者在回答时提出“我们的考虑是：其一，由于情感本身的复杂性，客观事物在个体身上引起的反应往往不是单一的，就国外疼痛移情研究中常用的疼痛手图片（如剪刀剪到手，针管深深刺入手中等）而言，它所诱发的情绪可能也不仅仅是疼痛，也有可能产生害怕、恐惧等其他情绪反应。本研究通过指导语有意识的引导被试在可能并不单一的情绪反应中产生某种主要情绪，是可行的，也是允许的。”但这样操作的可行性依据是什么？如果作者给不出合理的解释，则只能说明文中采用的作法太主观，而不够客观。

回应: 同问题 1。

意见 8: 作者指出“至于移情者与移情者之间的情绪反应在多大程度上等同或者说是否要完全等同，关于此问题一直有所争议，故本研究不作探讨。”这是有意回避问题，既然本文研究移情，首先应该分析被试究竟是否产生了移情才能进行后续的结果分析，如果被试没有产生合理的移情，那么，后面的结果可靠性就值得商榷。存在的另一个问题是，实验过程中，被试没有对所有图片均作出按键反应，这样的话，被试的想法是否与作者的意思吻合呢？这是一个需要说明的问题。

回应: Singer 和 Lamm (2009)认为当观察者知觉或想象他人情绪并能够部分感受到这种情绪时移情就发生了，但这一判断标准并未被应用到具体研究之中。因此，本研究在参考上述标准的基础上，结合移情的内涵，明确了判断移情发生的标准(当个体知觉或想象他人在某种情境中的情绪而引发与

他人或其情境相应的情绪时移情就发生了)并加以具体使用。这一标准反映了移情的核心关键——移情者与被移情者之间情绪反应的一致性。因此,两者反应具有一致性,意味着移情的发生,而一致性程度则为另一个问题,正如 De Vignemont 和 Singer(2006)所言:“移情研究中另一个存在的问题为移情者与被移情者之间情绪反应在多在程度上是等同的,是粗略的一致(例如,仅效价相同)?还是精细的等同(例如,相同的效价,强度和成份)?神经科学已开始探讨这个问题,但仍未取得明确的答案。”

对于在实验过程中被试没有对所有图片均作出按键反应,被试的想法是否与作者的意思相吻合这一问题:脑电记录结束后,让被试对两类情境图片进行离线评价,即将脑电实验中呈现过的图片再一次呈现(真实情境图片和电影情境图片各 1 个 block),要求被试对自己观看图片时感受到的悲伤程度进行评价,因此能够了解被试的感受。

意见 9: 前言中列举了很多 fmri 研究,但讨论部分基本未涉及。那么,本研究与 fmri 结果之间的异同是什么?不能仅仅在前面修改部分说明,应该在后面的讨论部分进行相关说明,否则就会出现这与前面提到的问题提出部分不呼应的问题,也不能说明本文采用 ERP 的独特性发现。

回应: 已在讨论部分进行了相应补充。

意见 10: “电影、真实情境两组图片每组包含地震情境图片和生活情境图片各 30 张,重复 2 次,即真实、电影情境图片的 block 各出现 3 次,即每种条件的 trail 数为 90 个”。如果是这样的话,如何考虑被试对同一刺激的不同反应呢?换言之,如果被试对相同的材料做出了不同的判断,结果如何筛选呢?同样,在实验中,“叠加是按图片的类别进行的”,这样操作的话,是主试自己的划分标准,而不是被试的标准,不能反映被试的真实体验。

回应: 本研究中让被试进行按键的目的不是为了判断,而是为了使被试集中注意力、避免走神,故其中 5% 的 trail 后要求被试完成一个评价任务。研究范式参考了 Decety 等人(2010)的研究。

意见 11: 既然对“The temporal dynamics of emotion regulation: An EEG study of distraction and reappraisal” Biological Psychology 87 (2011) 84–92”进行了阅读与参考”,但修改后的正文文献并未见该文献。

回应: 感谢专家的细致,已在文中进行参考并列入参考文献。

意见 12: 提出“我们之所以选择“地震”作为研究移情情境的原因有两个:其一,地震是一种灾难性情境,具有较强的冲击力,同时具有强烈的社会现实意义。其二,地震为生活中真实发生的、切实可见的现象,特别是近年来通过媒体得知我国和世界各地地震信息不少,影响很大。”这两点原因的

依据是什么？如何能表明地震的社会现实意义较其它情境更强烈？是作者自己的主观认为还是有客观依据？

回应：这两点原因的依据主要来源于现实生活。地震对于人类而言是一种破坏性极强的灾难性情境，国内大地震远期有唐山地震、近期有汶川地震，国外如 2011 年的日本大地震（引发了海啸），无一不造成重大人员伤亡和财产损失，引发全世界关注，其冲击力不可谓不强；其次，地震是全人类的灾难，每当地震发生时，在诱发其他地区和国家人民悲伤的同时，他们都会伸出援手，帮助灾区人民度过难关，这一点从 2008 年中国汶川地震中便可见一斑。故，地震对于人类而言不仅仅是灾难，它同时激发出了人们团结和爱的巨大潜在能量，正如法国《世界报》2008 年 5 月 27 日所发表的文章：地震令中国社会空前团结。地震的社会现实意义也由此可见。

意见 13：1) 简单照搬前人疼痛研究的时窗不够合理，毕竟疼痛与悲伤不是一样的，其参与脑区也不是完全一致的。2) 两类情境图片的分析时窗不一致，降低了结果的可比性。3) 正如前面的专家所言，不同区域所取时窗不统一，随意性很大，确实有大海捞针的感觉，而文章并没有指出为何选不同时窗的原因。4) 就 ERP 结果而言，由于讨论部分仅仅提出 P3 成分，那么，结果部分的这些不同波形差异能被界定为什么成分呢？

回应：平均波幅测量方法虽也存在一定的不足，只关注了认知加工的总体时间进程，忽略了对 ERP 具体成分的分析 and 比较，从而对结果的解释带来一些困难。然而，这种方法也有其优点，比较客观，标准相对统一，便于不同实验室进行比较（赵仑，2010）。300ms 以前，20ms 也是测量中较为常见的时间区间之一。在此基础上，参考了疼痛移情研究的测量时窗（因无其他可供直接参考的依据），从而确定了本研究的测量时窗。

其次，本研究中两类情境图片（地震情境图片和生活情境图片）的分析时窗是一致的，对不同脑区选取的电极测量的时窗也是一致的。

再次，就最后一个层面而言，主要是因为结果在晚期阶段显著，属 P3 范围，故结合其他相关研究对晚期阶段所代表的心理学意义进行了分析，其中不可避免的提到了 P3 成分。

意见 14：不只一个审稿人提出文章做出选取电极不合理，但修后的结果仍未做修改。

回应：因无其他可供直接参考的研究，本研究中电极的选取参考了疼痛移情的相关研究：Fan, Y., & Han, S. (2008). Temporal dynamic of neural mechanisms involved in empathy for pain: an event-related brain potential study. *Neuropsychologia*, 46, 160-173.

意见 15：“观看者看到的是尸体或废墟下的一段手臂，此时观看者的悲伤依然是与失去罹难者的亲人产生了移情，悲伤的源头是失去死者的亲人”，这句话是没道理的，尽管我们可以推测“失去死者的亲人”可能是悲伤的，但读者没有看到“失去死者的亲人”，并不能了解他们的真实情绪体验。

回应：这是笔者初期阶段的想法，最终的解释为我们会产生与罹难者当时相应的情绪反应(可能在程度上不如当时罹难者强烈)。对于另一专家提出罹难者在罹难前会产生的最大情绪是恐惧而非悲伤，我们认为：根据与地震灾区受灾人的接触，我们发现当地震来临时，由于情感本身的复杂性，情绪的产生根据时间序列而有所不同。的确，他们第一时间的情绪反应往往是恐惧，但恐惧之后产生的主要是悲伤，难过自己的境遇，伤心自己家人、财产的失去或自己鲜活生命的可能幻灭。而最近的马航 370 的失联，我们可从电视上目睹“失联者”家属的情绪反应——随着时间的推移而明显呈现出不同的层次，再次表明情感的复杂性与多维性。

意见 16：讨论部分大篇幅讨论“4.1 采用地震情境图片诱发悲伤移情的可行性”，这在 ERP 研究中是极其少见的，也是不怎么提倡的作法。讨论部分提到“4.2 情境真实性对移情调节的效应性分析”，那么，我想知道的是，究竟产生了什么效应呢？文章却没有给出。

回应：这样处理是因为，此次实验的一个创新点是材料创新，故加以突出。正如 B.Sporrow (2001) 在 *Science* 杂志上发文指出的那样，实验材料创新是实验创新重大四大方面（测验、指导语、材料、被试）之一。基于此，在讨论中用了较大的篇幅对实验材料的可行性及其与可能相互混淆的概念之间的差异进行了深入的分析，如放在前言或实验材料部分，则可能不能充分说明这两个问题。

其次，至于讨论部分“情境真实性对移情调节的效应性分析”中的效应，可能并未准确表达本研究的思想，已更改为“情境真实性对移情的调节”，再次感谢您的严谨与细致。

意见 17：摘要和结论部分都没有很好地回回应：“情境真实性是如何对移情的调节的”这一题目。题目和内容不符合。

回应：这是一个非常好的问题，再次体现出评审专家的敏锐。同涉及其他因素对移情调节的研究相类似(Decety et al., 2010; Fan & Han, 2008; Han et al., 2008; Li & Han, 2010)，本研究主要关注的是“调节的结果”，而未关注“调节的内在机制”（即如何进行调节）。因此，您的问题为我们提供了下一步的研究方向。

第四轮

审稿人 1 意见：

意见 1：题目中出现两个描述“移情”和“共情”，那么到底作者是关注哪个呢？二者是等同的吗？如果对前人的审稿意见表示接受（或不接受），应该说明理由，而不是在文章一开始就放在这里，会引起更多的理解麻烦。纵观全文，基本采用了“移情”这一描述，但也有有的地方采用了“共情”，如流程图中，这容易造成一定的混淆或理解麻烦。

回应：本研究原本用的是移情，上次评审专家提出共情更为常见，我们不知如何取舍，故在题目中

出现了移情和共情两个描述以表示在此两者均为“empathy”，以供专家进行评判。如让我们自己选择，我们偏爱“移情”这一译法，其理由如下：移情强调情绪的移动，并产生共同的情绪体验，即由移而同；共情强调共鸣，即共同的情绪体验，但缺少了“移动”这一动态的色彩。

其次，流程图中用了“共情”是由于替换时图片中的文字无法自动替换造成的，是笔者的疏忽（已在正文中进行了修改），为此向您致以诚挚的歉意并感谢您的认真、细致。

意见 2：文章没有让被试对移情刺激进行判断，也就没有报告被试的反应时，这样的话，本文分析的 ERP 成分是反映了其它心理加工还是反映了移情，这是值得商榷的。如果被试对图片的判断反应更慢，则对于什么阶段才是移情的是需要谨慎的。

回应：本研究采用了 Decety 等人（2010）疼痛移情研究中的实验范式和任务，他们采取了让被试自然观看疼痛图片的任务，并未关注被试的反应时；其次，他们同样在 ERP 实验之后，安排被试进行离线评价。

意见 3：本文对研究的理论意义的描述不够。

回应：已在正文中进行了补充。

意见 4：讨论部分：不知道本文的重心是什么？如果重心在 ERP 研究，有较大篇幅分析实验材料的可行性等，这有点重点不突出。另外，本文对于 350-400ms 和 450-550 ms 等这些不同时段的分析，能说明什么呢？他们是不同的成分？如果是不同的，本文对于他们各自反映了什么样的心理过程的描述上不充分。

回应：就本文的重心而言，首先我们研究的对象是悲伤移情，而 ERP 只是手段，采用地震情境图片是否的确能够诱发悲伤移情的发生，即采用地震情境图片诱发悲伤移情的可行性这一问题便成为本研究的核心，也是 ERP 研究的基础和前提，故有必要对实验材料的可行性进行分析；其次，此次实验的一个创新点是材料创新，故加以突出。正如 B.Sporrow（2001）在 Science 杂志上发文指出的那样，实验材料创新是实验创新四大方面（测验、指导语、材料、被试）之一。基于上述原因，在讨论中用了较大的篇幅对实验材料的可行性进行了较为深入的分析。但是，考虑到专家的意见，这部分占用的篇幅也的确较长，我们将这部分的内容进行了删减，仅保留了核心部分。

其次，350-400ms 和 450-550ms 两个时间段均属于 P3 的范围，对其进行分析说明情境真实性对悲伤移情的调节发生在晚期阶段。

意见 5：“来自社会现实的地震情境图片可以作为诱发悲伤移情的刺激材料，以提高生态化效度”这个结论的依据是什么？如果本文的两类图片不经过指导语的很好调控，被试是很难知道哪个更生态的。因为即便是来自电影的图片也有很多是真实生活的反映。在实验中，其实被试对材料的判断受诱导

所影响，而实验材料的真实情绪性是不是跟由于诱导产生的情绪一致呢？其实不然，如果不一致的时候，可能反而降低了相应的生态效度。如：对于“问题 1：图片中人物的情绪（或其情境）是（ ），请在所选回应：案后的括号内划对号：a 高兴的（ ） b 中性的（ ） c 悲伤的（ ）”这样的判断，很明显，备择选项没有包含了真实的所有情绪条件，被试只能根据选项做出可能不是材料所能导致的情绪而出现偏差，这样的话其实是降低了生态效度的。为何仅仅列出一种负性情绪选项呢？

回应：我们这里所说的提高生态化效度是与国外疼痛移情研究中常用的疼痛手、疼痛脚等图片相比而言的，并非地震情境图片同生活情境图片相比而言的。其主要依据为：地震情境是生活中真实发生的、切实可见的，且较为常见的现象；而疼痛手、疼痛脚等图片在现实生活中则并不常见，多出现在实验室情境之中，具有强烈的实验室研究的特色。正如有研究者所言：“目前，疼痛共情(即移情, empathy)研究多采用脑科学的实验室研究法，……实验材料多采用疼痛表情和疼痛手、脚的图片或录像，与个体实际的生活状态有较大的差距，缺乏情境性” (孟景, 陈有国, 黄希庭, 2010, p.438)。因此，将来自社会现实中的地震情境图片作为诱发悲伤移情的刺激材料，以提高生态化效度，正是本研究在实验材料上的一个改进和创新。

其次，指导语调控不涉及生态化效度问题，而是操纵真实、电影两种情境图片的真实性程度：在要呈现真实情境图片时，通过指导语告知被试，图片来源于汶川地震中的真实地震情境和生活中的真实生活情境；在要呈现电影情境图片时，通过指导语告知被试，图片来源于电影中的地震情境和电影中的生活情境。

至于为何只列出一种负性情绪选项，是因为在“地震”这一大背景下(呈现多张地震情境图片)，个体产生的情绪主要就是悲伤。这同 2008 年汶川地震时，当我们看到电视上的母亲抱着死去的孩子痛哭，甚至看到死亡儿童的书包时，眼泪都会夺眶而出的现实经验是相一致的。并且，在评定实验材料时，通过与被试的接触、交流，我们就发现了被试产生的负性情绪主要就是悲伤这一现象，否则在材料评定时我们便会进行相应修改以做再次评定；正式实验后，通过又一次与被试们的接触与交谈，再次发现了被试们主要出现的是悲伤的情绪体验。这也就证实了我们只列出“悲伤的”一个负性情绪选项是合乎现实的。

意见 6：文章中提到前人采用了“照相图片情境和卡通图片情境”两种刺激条件，从物理特征上和常识的角度，他们确实是两类存在明显差异的刺激。本文则将“地震情境图片”与“生活情境图片”相区分，这并不是能提高生态效度的体现。在本文中，“地震情境图片”与来自电影的图片，其实都还是有相似的，尤其是在物理特征上。还有，即便是电影图片中其实好多也是真实生活的反映，即他们也是真实的生活情境的体现。另外，结果分析和描述也比较主观，主试可能认为两类图片的确能诱导出不同的移情结果，但由于实验过程中被试没有相应的判断，是不是在实验过程中真正出现了期

待的移情，还是有待商榷的。即便实验后被试对刺激有评估，但并不能保证两次产生的情绪是一致的。

回应：本文采用地震情境图片而非以往移情研究中常用的疼痛手、脚等图片是为了提高生态化效度，而非地震情境图片与生活情境图片相区分是提高生态效度的体现。真实情境图片与电影情境图片是说明情境真实性程度的不同，即来源于现实的地震情境图片的真实性程度高于来自电影的地震情境图片，这也不涉及生态化效度的问题，恰是本文操作自变量的两种水平。并且真实、电影两种情境的形成是通过指导语来调控的（真实情境和电影情境中的图片本就完全相同，无非是真实、电影两种情境图片的分配在被试间进行了平衡），而非通过图片物理特征上的差异来调控的（这与前人采用的“照相图片情境和卡通图片情境”两种不同物理特征刺激作为操作自变量的两种水平完全不同，前人是通过实验材料本身的不同来操纵情境真实性程度，而本文则是通过指导语不同来操纵情境真实性程度）。

就实验过程中被试没有相应判断这一问题：其一，个体的行为反应要慢于脑电反应（即当个体做出行为反应时，脑电反应可能已结束），因此行为任务不能成为确认移情发生的关键性指标，而材料的有效性才是关键——即实验材料是否的确能够有效诱发悲伤移情的发生。例如，Fan 和 Han（2008）的疼痛移情研究中，让被试分别执行两种行为任务：疼痛判断任务和计算手的数量，这意味着两种行为任务下均能够产生移情，并非只有判断任务才能确保移情的产生。其二，本研究采用了 Decety 等人（2010）在其疼痛移情研究中的实验范式和任务，被证明是有效的。其三，就此问题，我们在实验前曾咨询过具有丰富 ERP 研究经验的学者，他认为让被试认真地自然观看更为恰当。并且，5%的 trail 后（被试并不知情，他们并不知晓判断任务何时出现）有一个判断任务，进一步提醒被试要集中注意力。

意见 7：从实验设计可以看出，实验过程的材料多次重复，是不是首次与之后重复阶段能诱导出同样的效果呢？如果不能，在实验之后的被试对图片的评估就显得意义更不足了。文章只是主观上认为是一致的，但是事实上是不是一致的呢？由于没有相应的判断结果可以比较。这在一定程度上忽略了多次接触可能对情绪评估的影响。

回应：由于 ERP 研究的特点，实验材料往往会重复出现，这一现象在具体研究中经常出现：如 Fan 等人（2008）、Han 等人（2008）的研究，实验材料包括疼痛照相图片和中性照相图片合计 40 张、疼痛卡通和中性卡通图片合计 40 张（由照相图片经技术处理而成），而这 80 张图片共出现 1280 次，每张图片重复出现 16 次；Li 和 Han（2010）的研究中，疼痛图片和中性图片各 40 张、合计 80 张，每张图片均重复出现 8 次，共 640 次。并且，上述研究同样在实验后让被试对实验材料进行主观评估。由此，也可看出实验材料的重复蕴含着研究者普遍认可同一刺激在诱发被试反应上是具有一致性的。

意见 8：“此外，为了使被试集中注意力，其中 5%的 trail 后要求被试完成一个评价任务，评价刚呈现出图片的内容是否悲伤”。这样的话，实验并不能保证被试对于其他更多的 95%的刺激是否也是集中注意了。如果没有很好地集中了注意，会影响诱导效果。由于没有可供参考的行为值，所以很难得出是否被诱导了。还有，被试对于那些 5%的 trail 的判断结果也没有报告，这是为什么？”

回应：本研究的研究范式和任务参考了 Decety 等人(2010)的研究，此范式和任务被证明是有效的；其次，何时出现评价任务——即“5%的 trail”，被试并不知情，我们告诉被试的是实验过程中随时会出现评价任务，请认真地自然观看。此外，对于 5%trail 后的判断结果已在正文中进行了补充。

意见 9：“FPS-R 离线评价”并没有诱导说明，这与实验过程中进行了的诱导条件不是完全匹配的，因此即便 FPS-R 离线评价结果在两类条件（“地震情境图片”与“生活情境图片”）间有差异，也与实验过程不同的。

回应：我们更感兴趣的是真实性程度有所不同的真实情境与电影情境中被试移情反应的差异，并非简单的地震情境图片与生活情境图片之间的差异。真实情境和电影情境正是通过指导语的诱导而形成的，具体表现为：实验开始前，有一个总指导语，即在被试对实验设置一无所知的情况下，指导语需要对被试进行诱导真实情境图片与电影情境图片之间的差异；在 ERP 实验结束时，被试对真实、电影两种情境图片的设置已非常熟悉。因此，离线评价时（ERP 实验结束后随即进行）不用再次在离线评价指导语中对真实、电影情境图片的设置进行赘述。并且，每个 block（共 2 个）前依然有本 block 来源的说明（说明本 block 内的图片是来源于真实情境还是电影情境），每个 trial 的流程也同正式实验完全相同。因此，离线评价同正式实验过程是相同的，无非在离线评价指导语中省略了真实、电影两种情境图片设置的详细信息，在 ERP 实验刚结束时这种省略是符合现实和逻辑的。

此外，不少关于情绪调节的 ERP 研究也经常会出现指导语省略的情况。这类研究往往在实验前给予被试关于所采用的情绪调节方式的详细指导语及训练，而在正式实验中往往出现简略的提示词，如“增加（情绪反应）”或“降低（情绪反应）”，这表明实验过程中重要的是被试的理解及掌握，并非指导语本身形式上是否完整(Hajcak & Nieuwenhuis, 2006; Moser, Kropf, Dietz, & Simons, 2009)。

意见 10：“FPS-R(Face Pain Scale-Revised)”的效度值得商榷，由于这与本文研究的主题并不吻合。

一、本文的材料并不是面孔，二、本文也不是研究疼痛的。其效度不够，因此这个评估量表并不适合于本研究。另外，Face Pain Scale-Revised 是一个量表，应该不只一个调研问题吧。本文采用的评价却是有一个判断问题，为什么呢？

回应：“FPS-R”的确只有一个问题，其实质是一个评估疼痛程度的标尺，常见的有 11 点（0-不疼痛，10-非常疼痛）和 6 点，常用于非面孔（疼痛手或脚作为刺激材料）疼痛移情研究中（Fan & Han, 2008; Han, Fan, & Mao, 2008; Li & Han, 2010）。本研究参照的只是“FPS-R”的形式，而非此量表

本身，并且将单位（原为疼痛）进行了更换（现为悲伤：0-不悲伤，10-非常悲伤）。

意见 11：“正式实验前设有练习部分”。什么样的练习呢？

回应：在实验前，为了让被试熟悉正式实验的流程，提前进行 2 个缩减版（告知被试）block 共 20 个 trial 的练习，不记录 EEG。

意见 12：图片的像素是多少？不同条件的图片像素统一吗？

回应：不同条件下的图片像素都是统一的。对所有的图片进行了统一处理：像素为 433*315,分辨率为 72 像素/英寸，均为 BMP 格式，亮度、对比度保持一致，详见 2.2.2。

意见 13：既然做了定位，为何结果部分不报告呢。

回应：溯源分析是根据另一评审专家意见而进行的尝试，主要目的是尝试，故提出该意见的专家并未要求将此结果放至正文中。

意见 14：“我们进行了刺激材料的评定, 60 张地震情境图片和 60 张生活情境图片诱发悲伤程度(0-10 等级)的平均数分别为 6.50 和 0.22; 在实验中, 我们又测定了被试对地震和生活两种情境图片主观感受到的悲伤程度, 其平均数分别为 6.62 和 0.37”这些结果与“FPS-R 离线评价”的结果是否一致？

回应：不完全一致，但在统计结果上无显著性差异。就像 IAPS 图片系统在具体研究中的应用，如让参与的被试重新评定选取图片的效价和唤醒度，与原图片系统对应的数值也不会完全一致，但是如两次评定结果在统计上不具有显著性差异，就可认为是一致的。

意见 15：“Hoffman, 2001/2002, p.55”是什么意思？为何与后文参考文献的页码不一致？另外，有这样表示参考文献年代形式的吗。

回应：这本书是我国翻译过来的书，具体对应的参考文献为：霍夫曼, M. L. (2002). 移

情与道德发展：关爱和公正的内涵 (杨韶刚, 万明译). 哈尔滨：黑龙江人民出版社.

这个表示方法笔者也不太确定，是参考了《心理学报》参考文献的格式，但不确定是否完全正确。

具体如下：

翻译的书

译者姓名写在书名后面的括号中。

■ 原著的出版时间写在末尾的括号中。

■ 正文引用中给出原著出版时间和翻译时间，用斜杠隔开。

Laplace, P.-S. (1951). *A philosophical essay on probabilities* (F. W. Truscott, & F. L. Emory, Trans.). New York: Dover. (Original work published 1814)

[拉普莱斯, P.-S. (1951). *概率哲学* (张三, 李四译). 北京: 未名出版社.]

正文引用: (Laplace, 1814/1951)

意见 16: 在有些条件, 对于人来讲可能看到其面孔更能判断其情绪。而对于后文列举的图片中的“附录 2: 生活情境图片(样例)”中的第 2 张图片(敲击键盘的), 被试如何感知被感知者的情绪呢? 因为, 这个被感知者可能在悲伤地敲击键盘也可能是正愉快地敲击键盘。换言之, 被试并不能真正移情。

回应: 生活情境图片是作为对照图片使用的, 即中性图片(同疼痛移情研究中的非疼痛图片、IAPS 中的中性图片的作用)。这类图片本身就不诱发(或几乎不诱发)移情的产生。

意见 17: 什么叫“情绪感染”?

回应: 有关情绪感染的概念很多, 也存在争议, 但大致都认为情绪感染是由他人情绪引起的并与他人情绪相匹配的情绪体验, 是一种情绪传递的过程(张奇勇, 2013)。其中, 情绪感染的一个典型例子是当婴儿听到其他婴儿的哭声时, 自己也会哭泣。因此, 有研究者认为移情与情绪感染之间的关键性区别在于自我-他人之间的区分, 即个体有能力能够辨别自己的情绪体验是来源于自己还是他人(de Vignemont & Singer 2006; Decety&Jackson 2004; Decety&Lamm2006)。

意见 18: “本研究对先前涉及情境与移情关系的 fMRI 研究进行了补充, 在揭示二者关系脑区空间特征的基础上, 提供了基于诱导引起的认知评价所形成的更具现实意义的第一类情境对移情的调节发生时程特点的实证依据”这句话可以作为一个段落吗? 放在文章中的目的是什么?

回应: 感谢专家的严谨, 已删除了这句话, 的确有画蛇添足之嫌。

意见 19: 不明白对于不同区域使用不同时段的目的, 另外前面分析了不同区域的不同时段, 为何在后面只针对少量的时段进行讨论?

回应: 对于不同脑区, 分析的时段都是相同的; 但不同的脑区, 结果显著的时段则有所不同, 而我们讨论的时段正是针对结果显著的时段进行的。

意见 20: 前面的波形图呈现了额区 (F3, F4)、中央区 (C3, C4) 和顶枕区 (PO3, PO4), 为何后面的差异波则呈现顶区的结果呢。

回应: 波形图是额区、中央区、顶区和枕颞区各选了 2 个电极, 但只有顶区的结果是显著的(图片情境内容×图片情境真实性在顶区电极的交互作用的显著, 意味着真实情境比电影情境诱发了更为强烈的移情反应), 故差异波仅呈现了顶区的结果, 这也是同 ERP 研究中差异波呈现的原则相符合的。

意见 21:后面分析的时段怎么不是连续的呢?如“同生活情境图片相比,地震情境图片诱发了更大的 ERP 波幅;在 100-140ms ($F(1,13) = 7.00, P < 0.05$)、260-300ms ($F(1,13) = 12.36, P < 0.01$)、300-350ms ($F(1,13) = 4.95, P < 0.05$)”中间的时段缺少,如 141-259ms 的结果。这是为什么呢?

回应:因为呈现的只是结果显著的时间段,故不是连续的:即在上述时间段,同生活情境图片相比,地震情境图片诱发了更大的 ERP 波幅;而在 141-259ms 等其他时间段,结果不显著,故没有必要进行报告。换言之,ERP 研究一般只报告结果显著的时间段,那么其余的、未报告的时间段是结果不显著的。

意见 22:“从图 2 可以看出,在中央区,所有条件下的刺激在 60-120ms 之间诱发了一个负成分(N110);紧随 N110 之后,在 140-200ms 之间是一个正波(P180);P180 之后,在 200-280ms 之间是一个负向的波峰(N240);N240 之后,在 300-340ms 之间是另一个负波(N320),随之在 340-800ms 之间紧接着是一个较大的正偏向(P3)。”这样描述不同成分是否合理?前人对 N1、P1 等成分的描述是否与本文的描述区域一致?

回应:本文的描述参考了疼痛移情研究中的描述方法(Decety, Yang, & Cheng, 2010; Fan & Han, 2008; Han, Fan, & Mao, 2008; Li & Han, 2010),时间区间同上述研究基本一致,有细微差异。事实上,上述疼痛移情研究之间的时间区间也存在细微不同:Fan 和 Han(2008) 的研究发现, N110 的时间区间为 90-130ms, P180 的时间区间为 140-200ms, N240 的时间区间为 200-280ms, 随之为波峰在 340ms 的负波 N340, 晚正偏向 P3 的时间区间为 360-800ms; Li 和 Han(2010) 的研究发现, N110 的时间区间为 80-120ms, 随之的正偏向 P160 的时间区间为 140-180ms (而非 P180), N240 的时间区间为 220-270ms, 随之的负波 N320 的时间区间为 310-350ms (而非 N340), 晚正偏向 P3 的时间区间为 340-740ms。

第五轮

审稿人 1 意见:

意见 1:对于 ERP 分析中,不同时段的数据有的报告了,有的则未报告,作者回回应:前面的审稿意见时指出是因为仅仅报告了显著的结果。以我的看法,即便是不显著的时段也应该报告出来,以便让读者能看到的确实是不显著的。

回应:关于这一问题,我们是这样考虑的:首先,采用平均波幅测量法的 ERP 研究,其结果往往只呈现显著部分的时间段,本研究因此也参考了相关的研究(Fan & Han, 2008; Han, Fan, & Mao, 2008; Li & Han, 2010; 孟迎芳, 2012) 对其中 ERP 结果的报告作同样方式的处理;其次,就本研究的具体情况而言,在 60 ms-300ms 区间(以 20ms 为间隔进行测量),如果全部时间段都要呈现的话可能

要呈现 12 个时间段的结果；在 300 ms-600ms 区间(以 50 ms 为间隔进行测量)则可能要呈现 6 个时间段的结果，若再加上额区、中央区等每个区均要呈现接近二十个时间段的结果，那么本文结果部分的呈现会变得过分冗长而失当，才作如此处理。

意见 2: 对于“本研究采用平均波幅连续测量法: 在 60 ms-300 ms 区间, 以 20 ms 为间隔进行测量; 在 300 ms-600 ms 区间, 以 50 ms 为间隔进行测量”这样的操作是否有前人参考依据还是作者自己的独特操作? 为何选取的间隔不同呢?

回应: 首先, 本研究中时间区间的选取参考了 Fan 和 Han(2008)的疼痛移情研究。其次, 应用较多的间隔时间窗有 20ms、50 ms、100 ms、150 ms, 且在不同的时间范围内时间间隔是可以变化的(赵仑, 2010)。

意见 3: 对于“Hoffman, 2001/2002, p.55”这样的描述肯定是不合适的, 作者应该查阅相关文献给出确切的日期, 你参看了哪本版本就按照哪个版本来处理。而不是(假设将来接受后)留给编辑部来处理。

回应: 根据《心理学报》的要求——正文引用中给出原著出版时间和翻译时间并用斜杠隔开, 我们对该格式再次进行了核对, 引用格式应该没有错误: 前一时间是原著出版时间, 后一时间是译著出版时间。

注: 学报关于译著引用的格式要求如下:

译者姓名写在书名后面的括号中。

原著的出版时间写在末尾的括号中。

正文引用中给出原著出版时间和翻译时间, 用斜杠隔开。

Laplace, P.-S. (1951). *A philosophical essay on probabilities* (F. W. Truscott, & F. L. Emory, Trans.). New York: Dover. (Original work published 1814)

[拉普莱斯, P.-S. (1951). *概率哲学* (张三, 李四译). 北京: 未名出版社.]

正文引用: (Laplace, 1814/1951)

意见 4: 有些文献没有报告页码, 如下面的一个文献:

Lu, J. M. (2000). *The psychology of affective teaching*. Shanghai: Shanghai Education Publishing House.

[卢家楣. (2000). *情感教学心理学*. 上海: 上海教育出版社.]

回应: 感谢您的认真与细致, 我们再次核对了《心理学报》参考文献著录格式详细要求中的 10 [著者自己著的书张三. (2008). *心理学史*. 北京: 未名出版社.], 并没有要求页码, 倒是论文集集中的论文或主编丛书中的章节要求在书名后给出论文或章节的页码范围。并且, 正文中也并非原文引用, 故也未标示出具体页码。对于此问题不知笔者是否理解有误, 还望专家赐教, 再次对您的指导表示

最为衷心的感谢。

意见 5: 数据分析中选取电极的命名存在错误。即在“Fz, F3-F4, FCz, FC3-FC4 (额区); Cz, C3-C4, CPz, CP3-CP4 (中央区); Pz, P3-P4 (顶区); POz, PO3-PO4, Oz, PO7-PO8 (枕颞区)”这一语句中, 有些区域的命名存在错误, 如 CPz, CP3-CP4 (应该是中央-顶区)。还有, 上面的电极中, 为何有些要加“-”? 貌似这些加“-”的电极为一个电极, 即 F3-F4 等是一个电极。

回应: 关于电极的命名, 从不同的研究目的出发, 往往会有些许的差异, 如有研究者将 F3, FZ, F4, FC3, FCZ, FC4, C3, CZ, C4 全部归为额中区(王沛, 杨亚平, 赵仑, 2010); 而有的则将电极所对应的脑区进行了如下划分: 前额区(FP1、FPz、FP2)、额区(F3、Fz、F4)、中央区(C3、Cz、C4) (孟迎芳, 2012)。本研究为了方便与疼痛移情的研究结果进行比较, 将电极的选取和命名同样参考了 Fan 和 Han(2008)的疼痛移情研究。此外, 对于电极的表示方法已在正文中进行了修改。

意见 6: 对于“在中央区, 所有条件下的刺激在 60-120ms 之间诱发了一个负成分(N110); 紧随 N110 之后, 在 140-200ms 之间是一个正波(P180); P180 之后, 在 200-280ms 之间是一个负向的波峰(N240); N240 之后, 在 300-340ms 之间是另一个负波(N320), 随之 340-800ms 之间紧接着是一个较大的正偏向(P3)。在顶区、枕颞区, 在 80-110ms 之间是一个正波(P1); P1 之后, 110-160ms 之间是另一个负波(N130); N130 之后, 200-600ms 之间是一个很明显的正偏向(P3)”这样的描述思路不统一, 大多数时段关注的是成分, 在 200-280ms 之间则是关注的波峰(N240), 为什么呢?

回应: 感谢专家非常仔细地发现和指出这一问题, 我们已据此在正文中作了修改, 使那段文字统一用正波或负波加以描述。

意见 7: “脑电记录结束后, 进行图片的离线评价: 将脑电实验中呈现过的图片再一次呈现(真实情境图片和电影情境图片各 1 个 block), 要求被试对自己观看图片时感受到的悲伤程度进行评价, 工具采用面孔疼痛量表的修订版(Face Pain Scale-Revised), 简称为 FPS-R (悲伤程度的评估分为 10 个等级: 0—不悲伤, 10—非常悲伤), 它改编自《面孔疼痛量表》(Face Pain Scale)(Bieri, Reeve, Champion, Addicoat, & Ziegler, 1990)”。在上述这段描述中, 很明显文章采用了 FPS-R 对本研究的材料进行评估, 因为实验材料和实验任务都不同, 尽管作者在回答前面专家提出的问题时指出是采纳了前人的问卷形式, 但实际则不是这样的。文章中明明是继续沿用的 FPS-R 这样的表达。

回应: 再次感谢您的严谨与细致, 已在正文中进行了相应的修改。

意见 8: “其三, 判断移情发生标准的明确, 在参考相关研究的基础上, 拟明确判断移情发生的标准, 并将其具体运用于本研究, 以确保本研究为真正意义上的移情本体研究, 克服以往移情研究中存在的判断移情发生的标准缺少或模糊的不足”这一句描述中提到, 前人存在相关的“标准缺少或模糊的

不足”为何能得出这样的描述呢？在该语句前述的描述中，并没有对相关问题有较为清晰的描述，因此在这里提出该描述有些突兀。

回应：感谢您的严谨，已在正文中进行了修改，以使表述更为准确。

意见 9：格式不统一，如“ms”与前面的时间有的有间隔有的无空格。

回应：已在正文中进行了统一。

意见 10：既然前面的时间段是以 20ms 为间隔的？为何在这些结果有的是该间隔的好几倍，有的间隔则仅仅是一个 20m 呢，这样操作的依据是什么？如“100-140”、“160-180”等。

回应：具体的计算方法是每 20ms (300ms 之前)为间隔进行计算的，但如果在两个连续的时间间隔内结果均显著，如 100-120ms 和 120-140ms，那么则将两者进行合并，计算 100-140ms 时间间隔内的平均波幅。此处理方法与其他相关研究(Fan & Han, 2008; Han, Fan, & Mao, 2008;Li & Han, 2010)是一样的。

意见 11：引言中提及的“其二，研究材料的改进，以现实生活中真实发生的、较为常见的，具有强烈社会现实意义的表现地震情境的图片作为诱发悲伤移情的实验材料，突破以往移情研究中常用的疼痛手、脚等完全局限于实验室材料的限制，提高研究材料的生态化效度”与结论部分的“的料期阶段，认知评价的影响进一化分化、具体化，(2) 来自社会现实的地震情境图片可以作为诱发悲伤移情的刺激材料，以提高生态化效度”，这两句描述中提到一方面是提高了研究材料的生态效度？前人研究是“以往移情研究中常用的疼痛手、脚等完全局限于实验室材料的限制”这样描述的根据是什么？是否有客观指标表明以往研究的生态效度更低？

回应：之所以这样描述的依据为：地震情境是生活中真实发生的、切实可见的，且较为常见的现象；而疼痛手、疼痛脚等图片在现实生活中则并不常见，多出现在实验室情境之中，具有强烈的实验室研究的特色。正如有研究者所言：“目前，疼痛共情(即移情，empathy)研究多采用脑科学的实验室研究法，……实验材料多采用疼痛表情和疼痛手、脚的图片或录像，与个体实际的生活状态有较大的差距，缺乏情境性” (孟景, 陈有国, 黄希庭, 2010, p.438)。

意见 12：讨论部分对“160-180ms”的描述并不深入，而在结论部分第 3 点提“160-180ms……”，这样的话，可能存在主次不分的问题。

回应：已对正文中的表述进行了相应修改。

第六轮

审稿人 1 意见：

意见 1：文中提到：“其二，研究材料的改进，以现实生活中真实发生的、较为常见的，具有强烈社会现实意义的表现地震情境的图片作为诱发悲伤移情的实验材料，突破以往移情研究中常用的疼痛手、脚等完全局限于实验室材料的限制，提高研究材料的生态化效度”，这样描述的依据是不充分的。作者提出这样描述的依据是什么，有前人可供参考的文献这样描述了吗？换言之，如果不是四川等地发生大地震，也许很多国人（包括被试在内）都可能觉得地震情境并不会引起更多的熟识感。疼痛手、脚更应该是大家熟悉的才对。即便作者在回答前面专家的问题时已经举证了“目前，疼痛共情(即移情，empathy)研究多采用脑科学的实验室研究法，……实验材料多采用疼痛表情和疼痛手、脚的图片或录像，与个体实际的生活状态有较大的差距，缺乏情境性” (孟景，陈有国，黄希庭，2010, p.438)这样的观点，而事实上，地震情境也与被试个体的实际生活差距很大，并不能由此证明地震情境的生态效度高。

回应：提高生态效度的方法有很多，主要是尽量保持实验设计的情境与自然情境的相似性（徐敏，华正昌，辛丽娜，2010）。本研究之所以将地震图片作为实验材料来提高生态化效度，出于两点考虑：第一，本研究中的实验材料直接取材于地震和日常生活中真实发生的自然情境，而非疼痛移情研究中常用的刻意、人为制作的情境（见下图）。因此，首先从材料的来源来看，前者的生态性要高于后者。第二，地震情境是人类生活中经常发生的现象。事实上通过媒体得知世界各地地震的信息实在不少，就拿我国来说，几乎每年都有关于地震灾害的报道，尤其是随着电视新闻的普及，使地震情景与普通百姓的现实生活已不再遥远，因而以此作为实验材料具有相当的普遍性；至于疼痛移情研究中常用的疼痛手或脚图片，在日常生活中并不多见，我们通过和周围人的接触与了解，都认为如图所示的现象不具有普遍性（偶然在生活中有剪破手指的，但剪断手指实属罕见！）。换言之，如果不是疼痛移情研究采用此类图片作为实验材料，大家或许根本不会意识到此类情境的存在。



意见 2：“评价工具参照了面孔疼痛量表修订版(Face Pain Scale-Revised, FPS-R)的形式(Fan & Han, 2008; Han et al., 2008; Li & Han, 2010)，并将其单位(疼痛)更换为悲伤(0-不悲伤, 10-非常悲伤)”。很明

显，文章在随后的附录中同样采用了“FPS-R”这样的表述，这确实是不合理的，因为本文的材料并不是只有面孔。因此，继续沿用“FPS-R”这样的表述是不够合理的。

回应：再次感谢专家的认真与细致，确为笔者的疏漏，已在文中进行了修改。

意见 3：从图 4 可以看出，FPS-R 得分与电影地震情境图片、真实地震情境图片诱发 ERP 平均波幅的相关差异其实并没有那么大，尽管提到某一相关值明显，而另一相关值尽管不显著，但实际上是 0.48 和 0.49 均是不小的相关值。因此，这些相关值应该是处于边缘显著状态的。因此，若要得出“电影、真实地震情境图片诱发的 ERP 平均波幅分别在 160-180ms 和 300-350ms 与主观感受相关显著，佐证了两种情境的分离”这样的结论尚需谨慎。如果再次检查数据后，0.48 和 0.49 并非边缘显著的相关值，要得出上述描述还是可以的。

回应：相关显著的意思是通过此样本所获得的相关系数是能够代表总体的相关性的。至于相关系数的大小，在显著的情况下，相关系数越大，相关性越大，但在不显著的情况下，相关系数再大也不能用来说明总体上是否相关。对于本研究而言，因样本量不大，相关显著的临界值就会比较大，所以才出现 0.48、0.49 这样不小的相关值却不显著的情形。

本研究中 0.48、0.49 所对应的 P 值的确处于 0.05~0.1 之间，但 P 小于 0.1 又大于 0.05 的情况，用所谓“边缘显著”的说法，主要是为了强调研究者预计“如果加大样本，是很可能达到显著的，但本研究由于条件有限，暂时不进一步加大样本来检验了”。因此，在现有的条件下，仍是可以得出上述结论的。

意见 4：“从图 2 可以看出，在额区、中央区，所有条件下的刺激在 60-120 ms 之间诱发了一个负波(N110); 紧随其后的是 140-200 ms 之间的正波(P180)及峰值在 200-280 ms 之间的负波(N240); N240 之后，在 300-340 ms 之间是另一负波(N320)，最后呈现的是 340-800 ms 之间的非常明显的正波(P3)。在顶区、枕颞区，在 80-110 ms 之间是一个正波(P1); P1 之后，110-160 ms 之间是另一个负波(N130); N130 之后，200-600 ms 之间是一个很明显的正波(P3)”这句话中，将不同区域出现相同的成分进行描述，这个可能需要再检查一下，不同区域（如，在顶区、枕颞区）是否可能记录到同样峰值的 N130 等成分？另外，在这段描述中提到“340-800 ms 之间的非常明显的正波(P3)”和“200-600ms 之间是一个很明显的正波(P3)”，为何同样是提到 P3 但其起始和结束时间点不同？另外，这与前人提到的 P3 是否是一致的？

回应：ERP 研究结果往往是前部的电极呈现出相似的波形，而后部的电极呈现出相似的波形，在以往的研究中并不少见，例如疼痛研究中常将额区和中央区用额-中区加以描述：在额-中区，所有条件下的刺激在 90-130ms 之间诱发了一个负波(N110)，紧随其后的是 140-200ms 之间的正波（P180），……最后是 360-800ms 之间的正波（P3）；在枕颞区，……最后为 200-450 ms 之间的正波（P320）（Fan & Han, 2008; Han, Fan, & Mao, 2008）；所有刺激在额区于 80-120ms 诱发了一个

负波(N110)，紧随其后的是额中区于 140-180ms 之间的正波（P160），……最后是 340-740ms 之间的最大波幅在中央区的正波（P3）（Li & Han,2010）。

其次，对于在顶区、枕颞区是否可能记录到同样峰值的 N130 等成分，因顶区、枕颞区均对应后部的电极，故呈现出相似的波形（成分），但 N130 的峰值并不一定相同（波形图相似，但对应具体数值未必相同）。此外，由于额、中区和顶区、枕颞区所处位置不同，两个 P3 的起始和结束时间点也就自然不同；本研究中 P3 与疼痛移情研究中 P3 的时间范围稍有出入，这并不奇怪：从上文可看出，疼痛移情研究之间对应波的时间区间也存在一定差异。

意见 5： .波形图并不统一，图 2 是彩图，图 3 是黑白的。

回应： 已在正文中统一用黑白图形进行表示。

意见 6： 根据现在《心理学报》的投稿标准，统计检验应该给出相应的效应量。

回应： 已在正文中进行了补充。

第七轮

审稿人 1 意见：

意见 1： 请作者对本文进行修改，修改主要涉及对文章的论述部分进行精炼，整体论文字数控制在 8000 字左右。然后提交，仍然由编委再审。

回应： 已对文章进行修改，正文部分 7600 字左右。因英文摘要和参考文献部分较长，整体字数有所超出（9500 字左右）。