

《心理学报》审稿意见与作者回应

题目：跨文化表情识别中的群内优势效应：自发性和呈现方式对气愤和厌恶表情识别的影响
作者：方霞 葛猷勋

第一轮

审稿人 1 意见：

该研究探究了群内优势效应是否受到气愤和厌恶两种表情的表情自发性(表演/自发)和表情呈现方式(静态/动态)的调节,研究问题具有一定的开创性。研究兼具理论和应用价值,有助于丰富情绪方言理论和促进理解真实场景中的跨文化交流。

意见 1: 综述和假设的联系需要加强。本文的前言是一个偏开放性的综述,提出了看似对立的多种假设。这对作者行文的逻辑性提出了比较高的要求。目前,虽然大体框架较为清晰,但是仍有提升空间。具体来说,1.2 和 1.3 的论述分别提出了两种看似竞争的假设,即表演表情识别的群内优势效应大于/小于自发表情;动态表情识别的群内优势效应大于/无差别于静态表情。目前的论述有些不够充分且对立的假设之间区分不够清晰。比如,作者在 1.3 提出“这些动态信息可能是通用的,可以普遍促进所有文化感知者的情绪识别表现”,就这一论述而言,我好像没有找到与之联系紧密的文献综述。另外,建议作者在修改稿中,将对应的综述和假设合并,让读者对假设背后的逻辑一目了然。比如,论述完动态信息的通用性/低利用率等,就可以随之提出动态表情识别的群内优势效应无差别于静态表情的假设。

回应: 感谢专家对行文提出的意见。结合专家意见,我们在修改稿中将综述和假设合并,以便读者更好地理解假设背后的逻辑。修改后的具体内容如下:

有关表情的自发性是否调节表情识别的群内优势效应(见正文 P3-4):

虽然群内优势效应在许多研究中都得以验证,但这些研究主要集中在表演表情上,较少关注自发表情(见 Elfenbein & Ambady, 2002 的元分析)。近年来,有研究发现,在表达相同情绪时,个体的表演表情和自发表情的动作单元有显著区别(Namba, Kagamihara, et al., 2017; Namba, Makihara, et al., 2017)。Fang 等人(2022)则发现,相比于自发表情,表演表情激活了更多的动作单元。这些结果共同表明,表演表情和自发表情在面部肌肉运动模式上可能存在差异。

那么,不同自发性的表情在其识别的群内优势效应上会有差异吗?以往研究发现,表演表情可能比自发表情更易受文化的影响(有关综述,请参阅 Valente et al., 2018)。无论是早期观察研究还是标准化面部肌肉运动测量研究,都发现先天性盲人产生的自发表情与视力正常的个体相似,但表演表情却与正常个体有显著差异。这表明,相比于自发表情,表演表情可能更容易受到社会学习的影响。由此,我们预测表演表情可能比自发表情包含更多与文化相关的情绪方言,从而导致表演表情识别的群内优势效应大于自发表情。该推论获得了 Matsumoto 等人(2009)研究的初步支持。在这项研究中,美国和日本被试对美国和日本运动员在 2004 年雅典奥运会获奖后的快乐或悲伤表情进行识别,结果并未发现群内优势效应。这表明群内优势效应可能不存在于自发产生的表情中。然而, Kang 和 Lau(2013)却有不同发现。他们让欧裔和亚裔美国被试识别欧裔和亚裔美国模特的标准表情(按照 FACS 手册标准制定的表情原型)和自发表情。其中标准表情来自 Japanese and Caucasian Facial Expressions of Emotion (JACFEE; Matsumoto & Ekman, 1988)的表情图片,而自发表情则是通过引导被试

讲述个人经历而获得的动态表情视频片段。结果显示，亚裔美国人在识别标准表情和自发表情时均出现了群内优势效应。值得注意的是，该研究考察的标准表情以静态方式呈现，自发表情以动态方式呈现，且两种表情的情绪类别并不完全相同(模仿表情：悲伤、惊讶、恐惧、气愤、厌恶和愉快；自发表情：悲伤、惊讶、沮丧、气愤和愉快)。因此，表情的自发性对表情识别的群内优势效应是否具有调节作用还有待商榷。

然而，另一方面，相较于自发表情，表演表情具有更高的强度和原型性(Krumhuber et al., 2021; Tcherkassof et al., 2007)。这些特性可能会导致表演表情比自发表情具有更高的识别率。跟该推论一致，以往研究结果确实表明，不论是人类还是机器，都能更好地识别表演表情(Fang, Sauter, et al., 2022; Krumhuber et al., 2021)。因此，我们预测这种高识别特性可能缩小了感知者对不同文化个体的表演表情识别率的差异，从而导致表演表情识别的群内优势效应小于自发表情。

有关表情的呈现方式是否调节表情识别的群内优势效应(见正文 P4):

除了考察表情识别的群内优势效应是否会受到表情自发性的调节外，本研究的第二个目标是探究表情的呈现方式是否影响表情识别的群内优势效应。相比静态表情，动态表情含有独特的动态信息，如表情运动的方向、质量和速度(Bould & Morris, 2008; Cunningham & Wallraven, 2009; Jack et al., 2014)。这些动态信息会启动更高级的认知过程，支持社交和情绪推断，从而促进面部表情的识别(Blais et al., 2017; Sato et al., 2008)。

尽管如此，我们仍不清楚动态信息的促进作用是针对特定文化感知者群体，还是普遍适用于所有文化感知者群体。换言之，动态表情中的动态信息是否涵盖了文化特有的情绪方言。如果该动态信息包含情绪方言，那么这些额外的情绪方言将进一步增强表情识别的群内优势效应。因此，我们预期动态表情识别的群内优势效应将大于静态表情识别的群内优势效应。这一猜想在 Elfenbein 和 Ambady(2002)的元分析中初步得到支持。他们发现，使用动态表情作为实验材料的研究(共 43 项研究)所揭示的群内优势效应大于使用静态表情作为实验材料的研究(共 79 项研究)。值得注意的是，该元分析发现的表情呈现方式对群内优势效应的调节仅呈边缘显著。此外，该元分析包含的研究除了在表情呈现方式上存在差异外，还在其他诸多方面存在差异，例如表情的自发性、文化和情绪类别的选择。由于缺乏对其他方面差异的控制，我们无法确定仅仅表情的呈现方式是否会调节表情识别的群内优势效应。

然而，如果动态表情中的动态信息不包含文化特有的情绪方言，而是普遍促进所有文化感知者的情绪识别表现，那么动态表情和静态表情识别的群内优势效应将不会有差异。因此，我们预期表情的呈现方式不会调节表情识别的群内优势效应。

此外，对于“这些动态信息可能是通用的，可以普遍促进所有文化感知者的情绪识别表现”这一表述，确实只是作者的推断，并没有相关文献支持。过去的大部分研究支持动态信息能够促进表情识别(有关综述，请参阅 Krumhuber et al., 2023)，但尚未进行跨文化研究来同时比较静态和动态表情的识别。因此，我们不清楚动态表情中的动态信息是否涵盖了文化特有的情绪方言。为了加强推理的逻辑连贯性，我们将之前引言中有关动态信息无法促进表情识别的少部分研究移至讨论部分。修改后的具体内容如下(见正文 P14-15):

本研究发现，表演表情识别的群内优势效应不受表情呈现方式的调节。然而，与预期不一致的是，该现象并不是因为动态信息普遍促进了所有文化感知者群体的表情识别。相反，在绝大多数情况下(表情种类 × 表达者文化)，感知者对静态表情和动态表情的识别率并无显著差异(见补充材料 2)。该结果说明，本研究中的动态信息对表演表情识别没有促进作用。尽管以往大量研究揭示动态表情的识别率优于静态表情，但也有其它研究并未发现两者的识别率存在差异(如，Fiorentini & Viviani, 2011; Wehrle et al., 2000)。这些不一致的结果可能与面部表情材料及感知情境有关。当面部表情信息有限时，如表情真实性受损(Käsryri et al., 2008; Wehrle et al., 2000)或表情强度较弱(Bould & Morris, 2008; Yitzhak et al., 2018)，动态信

息可以弥补静态表情的信息不足,从而有助于提高表情识别。然而,当面部表情信息充足时,如表情清晰或者强烈,静态表情的识别可能已经达到较高水平,额外的动态信息并不会提高表情识别准确率(Bould & Morris, 2008; Gold et al., 2013)。此外,当面部表情材料呈现时间短且被试应答时间有限时,静态表情比动态表情得到更加充分的加工,反而导致个体对静态表情的识别率大于动态表情(Jiang et al., 2014)。在本研究中,静态表情和动态表情的比较建立在模特表演特定情绪的表情上,而表演表情往往包含了相对强烈的情绪信息(Kayyal & Russell, 2013)。这些强烈的情绪信息导致静态表情和动态表情的识别率相当,进一步导致了表情识别的群内优势效应不受呈现方式调节。

意见 2: 需要突出气愤和厌恶表情是本研究主要的研究对象,并加强与之相关的文献综述。作者在 1.4 指出关注气愤和厌恶这两种表情的原因是由于表情的常见性和材料便利性。除此之外,可能更需要强调的是研究气愤和厌恶背后的理论价值。建议作者在相关文献的论述中增加前人研究了哪些表情的细节信息,论述本研究是否在一定程度上弥补了前人研究的空白。最后,建议在标题字数允许的情况下,将气愤和厌恶加入标题,防止标题信息的过度泛化。

回应: 感谢专家的意见。首先,结合专家的意见,我们对研究关注气愤和厌恶这两种表情的原因进行了更详细的解释,具体内容如下(见正文 P5):

本研究关注气愤和厌恶两种情绪的面部表情。主要考虑如下三点。首先,气愤和厌恶被广泛认为是基本情绪(Ekman, 1992; Izard, 2007),因此预计在东西方文化中都普遍存在且较为常见。其次,尽管气愤和厌恶是不同的情绪,但它们在概念上存在关联(Giner-Sorolla et al., 2018),并且在面部形态上具有相似性(Cordaro et al., 2018; Fang, Sauter, et al., 2022)。这些相似性导致个体容易混淆气愤和厌恶这两种表情(Fang et al., 2018, 2019; Jack et al., 2009; Pochedly et al., 2012)。采用这两种具有混淆度的面部表情可以更好地诱发表情识别的群内优势效应。相反,当采用两种完全不同的表情,如气愤和快乐时,个体对无论来自哪个文化的气愤和快乐表情的辨别都比较好,因此表情识别的群内优势效应不太可能显现。第三,现有表情数据库主要刻画静态表演表情,对动态或自发表情的刻画较少(有关综述,请参阅 Dawel et al., 2022; Krumhuber et al., 2017),同时包含两种表情自发性或者两种表情呈现方式的跨文化表情数据库更少。据作者所知,目前仅有一项跨文化情绪表达的研究同时包含了上述四种条件的表情,为中国和荷兰模特的气愤和厌恶的静态/动态的表演和自发表情(Fang, Sauter, et al., 2022)。

其次,非常感谢专家关于在相关文献的论述中增加前人研究表情的细节信息的建议,我们在文中提及前人研究时补充了其研究表情的说明,具体内容如下(见正文 P3):

在这项研究中,美国和日本被试对美国和日本运动员在 2004 年雅典奥运会获奖后的快乐或悲伤表情进行识别,结果并未发现群内优势效应。

然而,Kang 和 Lau(2013)却有不同发现。他们让欧裔和亚裔美国被试识别欧裔和亚裔美国模特的标准表情(按照 FACS 手册标准制定的表情原型)和自发表情。其中标准表情来自 Japanese and Caucasian Facial Expressions of Emotion (JACFEE; Matsumoto & Ekman, 1988)的表情图片,而自发表情则是通过引导被试讲述个人经历而获得的动态表情视频片段。结果显示,亚裔美国人在识别标准表情和自发表情时均出现了群内优势效应。值得注意的是,该研究考察的标准表情以静态方式呈现,自发表情以动态方式呈现,且两种表情的情绪类别并不完全相同(模仿表情:悲伤、惊讶、恐惧、气愤、厌恶和愉快;自发表情:悲伤、惊讶、沮丧、气愤和愉快)。

有关“论述本研究是否在一定程度上弥补了前人研究的空白”,我们相信通过对意见 1 和意见 2 的回应和修改,已经提升了本研究的开创性和重要性。此外,我们在“当前研究”

的开头增加了对过往研究的简短小结，以进一步突出本研究的意义。具体内容如下(见正文 P4-5):

综上所述，关于跨文化表情识别的群内优势效应是否受到表情自发性调节的问题，现有研究虽然进行了初步探讨，但存在一些限制。这些研究要么没有包括表演表情识别任务(因此无法直接比较自发和表演表情识别的群内优势效应; Matsumoto, Ollide, & Willingham, 2009)，要么所包含的自发表情和表演表情在情绪类别和呈现方式等方面存在差异(Kang & Lau, 2013)。此外，现有研究的结果也不一致。至于跨文化表情识别的群内优势效应是否受到表情呈现方式的调节，目前尚未开展相关研究。

最后，按照专家的建议，我们将原标题修改为“跨文化表情识别中的群内优势效应：自发性和呈现方式对气愤和厌恶表情识别的影响”。

意见 3: 需要就表情材料的信效度问题进一步补充说明。根据作者在 3.1.2 的表述，表情选取的标准来源于被试的主观判断（即，选出最能表达相应情绪的那一帧画面），这似乎也是唯一的标准。此外，补充材料里提供的识别准确率似乎也不是很高，普遍低于 0.50。请作者进一步补充有关表情材料本身准确性的说明或证据。比如，是否还有第三方对表情做过评估？本研究的识别准确率在前人的相关研究中处于一个怎样的位置，是否偏低？最后，如果有可能的话，是否可以提供一些示例材料作为参考。

回应: 感谢专家对表情材料提出的意见。关于静态表情选取方法的说明，我们在文中添加如下内容(见正文 P6):

被试完成两种表情后，观看自己的视频，并选择最能表达相应情绪的那一帧画面作为静态表演表情材料(用于实验 2; 参见图 1 示例)。这种方法确保了所选择的静态表情准确地表达了被试意图表达的情绪，避免了引入实验者偏差(类似操作见 Cordaro et al., 2018; Elfenbein et al., 2007)。

其次，感谢专家提出对表情识别准确率的质疑。以下是我们对此的解释：第一，与以往大多数研究使用标准表情(按照 FACS 手册标准制定的表情原型)作为实验材料不同，本研究中的表情要么是被试自由表演的(并非模仿标准表情)，要么是自发产生的。此外，气愤和厌恶表情在概念和形态上具有相似性，这导致个体对这两种表情的识别容易产生混淆(Fang et al., 2018, 2019; Jack et al., 2009; Pochedly et al., 2012)。与以往研究考察表演表情识别的结果相比，其中一项重要研究显示，被试对气愤和厌恶表演表情的识别准确率在 26% 至 54% 之间(Elfenbein et al., 2007)。该结果与本研究中发现的识别准确率 33% 至 62% 相当。考虑到自发表情比表演表情更难识别(Krumhuber et al., 2021)，因此自发表情识别率更低。第二，为了考察本研究中的表情材料是否能被准确识别，我们对实验 1 和实验 2 中不同感知者文化、表达者文化、表情自发性和表情呈现方式的气愤和厌恶表情识别正确率进行单样本 t 检验。结果显示，在绝大多数情况下，气愤和厌恶表情识别正确率均高于随机水平(实验 1 为 1/5 个情绪维度 = 0.2，实验 2 为 1/4 个情绪维度 = 0.25)。这表明，尽管气愤和厌恶表情的识别率较低，但被试仍能以高于随机水平的概率识别这些表情。相关数据已补充至附录中(见附表 1 和附表 3)。

最后，根据专家建议，我们在正文中提供静态表演表情的示例(见图 1)，动态表演和自发表情示例已上传至以下链接：https://osf.io/8f5cd/?view_only=197bfe65e1be4723bf3ba0efb6294ba4，并已在正文中进行了说明(见正文 P6)。由于没有获得荷兰模特的授权，本研究无法公开发布荷兰模特的表情材料。

意见 4: 请补充情绪测量的依据，即考虑了 5 种基本情绪（即，气愤、厌恶、恐惧、悲伤和愉快情绪），为何没有包含其他基本情绪（如，惊讶和轻蔑）的测量。

回应：感谢专家的意见。我们在正文中加入了测量 5 种基本情绪的脚注。具体内容如下(见正文 P24)：

本研究的情绪评价维度涵盖了大部分研究公认的基本情绪(Ortony, 2022)。考虑到以往研究对于惊讶(Bain, 1859/2006; Oatley & Johnson-Laird, 1987)和轻蔑(Matsumoto & Ekman, 2004; Widen et al., 2011)是否为基本情绪存在争议，因此本研究选择不包括这两种情绪评价维度。

意见 5：被试作答的质量是否经过评估，比如是否有识别正确率过低的被试？

回应：感谢专家的意见。通常情况下，除非有充分理由，表情识别研究不会因为识别率低而剔除被试的数据(如，Elfenbein et al., 2007; Krumhuber et al., 2021)。因为较低的识别率并不一定意味着被试的回答质量不佳，可能是个体在表情识别方面存在差异(Freedman, Pisani, & Purves, 2007)。尽管如此，为了确定本研究中是否存在识别率过低的被试以及这些被试数据对研究结果的影响，我们采用了 Z 分数方法剔除异常值(Bakker & Wicherts, 2014)。在数据分析中，我们先分别在每个被试间条件下，按照常见的数据剔除标准剔除了识别准确率低于总平均值减去 3 个标准差($M-3SD$)的被试，结果仅剔除了实验 1 中的 1 名被试。我们再按照更严格的数据剔除标准剔除了识别准确率低于总平均值减去 2.5 个标准差($M-2.5SD$)的被试，这导致实验 1 剔除了 3 名被试，实验 2 剔除了 4 名被试。在此基础上，我们对正文中的分析进行了重复，发现结果与剔除被试前的分析结果相似，说明即使存在识别率过低的被试也并不会影响整体结果。更详细的信息请参见补充材料 3，并已在正文中以脚注的形式进行了说明(脚注修改内容见正文 P24)。

额外的参考文献：

Freedman, D., Pisani, R., & Purves, R. (2007). *Statistics* (4th ed.). New York, NY: Norton.

Bakker, M., & Wicherts, J. M. (2014). Outlier removal, sum scores, and the inflation of the type I error rate in independent samples t tests: The power of alternatives and recommendations. *Psychological Methods*, 19(3), 409–427. <https://doi.org/10.1037/met0000014>

意见 6：请补充表 1 和表 3 中的 df 以及 Cohen's d 的 95%CI

回应：感谢专家的意见。我们已在修改稿中的表 1 和表 3 中增加了 df 和 Cohen's d 的 95%CI(见正文 P7 和 P11)。

意见 7：讨论部分提到，实验 1 中存在的例外（西方感知者识别自发气愤表情时的群内优势效应大于表演气愤表情）可能是因为荷兰表达者的表演气愤表情常用的面部动作单元数量少于其自发气愤表情。请作者对这一论述补充更多的细节，具体的面部动作单元是哪些，这一论述是来源于前人研究还是本研究？

回应：感谢专家的意见。我们在修改稿中对这一论述补充了更多细节，并注明了这一论述的参考文献，以便读者更好地理解，具体内容如下(见正文 P13)：

需要注意的是，本研究中的表演和自发表情材料均来自 Fang, Sauter 等人(2022)的研究。该研究显示，除了荷兰表达者的表演气愤表情的动作单元数量少于自发气愤表情外，其他情况下表演表情的动作单元数量均多于自发表情。这些额外的肌肉运动说明表演表情包含更多信息，更有可能蕴含文化特有的情绪方言。因此，表演表情的群内优势效应大于自发表情。然而，在实验 1 中存在一个例外：西方感知者识别自发气愤表情时的群内优势效应大于表演气愤表情。这可能是因为荷兰表达者的表演气愤表情常用的动作单元数量少于其自发气愤表情，并且两者常用的动作单元相似度较低(仅共享 5 或 6 个动作单元中的其中两个)。具体而言，荷兰表达者的表演气愤表情常用的动作单元有 5 个，分别为 AU4(双眉紧皱)、AU7(内

眼睑缩紧)、AU23(嘴唇紧缩)、AU54(低头)、AU63(双眼上升),而自发气愤表情常用的动作单元有6个,分别为AU1(眉毛内侧抬升)、AU2(眉毛外侧抬升)、AU4、AU7、AU61(双眼左倾)、AU64(双眼下降)。这些结果表明,荷兰表达者在自发气愤表情和表演气愤表情方面存在较大差异。荷兰表达者的自发气愤表情可能涵盖了更多文化特有的情绪方言,从而导致自发气愤表情具有更强的群内优势效应。

.....

审稿人2意见:

本研究通过两项精心设计的实验,探讨了表情自发性和表情呈现方式(静态与动态)是否影响跨文化表情识别中的群内优势效应。实验1聚焦于表演表情和自发表情的识别,而实验2则是关于静态和动态表情的识别。研究选取了气愤和厌恶两种具体情绪,同时招募了来自中国、加拿大和荷兰的参与者。研究结果揭示,表情自发性对群内优势效应有显著影响,然而表情呈现方式的影响并不明显。

论文优点

1. 尽管群内优势效应在跨文化交流中被广泛认为存在,但大多数先前的研究仅限于静态表演表情。因此,本研究的角度具有重要性。
2. 文献综述细致且全面,清晰地阐述了研究现状和理论基础。
3. 统计分析做得详细且严谨。
4. 研究结果对群内优势效应理论的拓展做出了贡献。

意见1: 尽管本文探讨了表情自发性这一变量对群内优势效应的影响,但仅限于检验表情自发性,因而研究问题相对缺乏广度和创新。

回应: 感谢专家的宝贵意见,这些意见提醒我们在原文中未清晰阐述本研究的意义(如广度和创新性)。为此,我们进行了以下修改,以更好地说明本研究与先前研究的区别以及本研究的重要性。

在正文的第二段落,我们修改如下(见正文 P1-2):

以往有关表情识别的群内优势效应研究主要聚焦于表演表情(posed facial expressions)和静态表情(static facial expressions; Elfenbein & Ambady, 2002)。表演表情是指个体有意识地摆出某种表情以传达特定的情绪(Elfenbein & Ambady, 2002),而静态表情则是指在特定时刻捕捉的面部表情。然而,在现实生活中,表情往往是自发的,即个体在交互过程中自然流露的(Matsumoto, Ollide, Schug, et al., 2009),也是动态的,即表情会随着时间的推移发生变化(Krumhuber et al., 2023; 张琪 等, 2015)。近年来,越来越多的研究揭示个体对表演表情、自发表情(spontaneous facial expressions)、静态表情和动态表情(dynamic facial expressions)的表达和知觉可能存在差异(Kang & Lau, 2013; Krumhuber et al., 2021; Matsumoto, Ollide, & Willingham, 2009; Valente et al., 2018),因此有必要考察群内优势效应是否会受到表情的自发性和呈现方式的调节。为此,本文首先回顾群内优势效应及其相关的情绪方言理论,然后探讨表演表情和自发表情识别以及静态表情和动态表情识别可能存在的差异,最后通过两项实验室研究探究表情的自发性和呈现方式是否调节气愤和厌恶表情识别的群内优势效应。本研究首次考察了表情识别的群内优势效应是否受到表情的两个重要特性——自发性和呈现方式的调节,对于拓展表情识别中的群内优势效应研究,并加深我们对跨文化表情识别的理解具有重要意义。

在当前研究部分,我们修改如下(见正文 P4-5):

综上,关于跨文化表情识别的群内优势效应是否受到表情自发性调节的问题,现有研究

虽然进行了初步探讨,但存在一些限制。这些研究要么没有包括表演表情识别任务(因此无法直接比较自发和表演表情识别的群内优势效应; Matsumoto, Ollide, & Willingham, 2009),要么所包含的自发表情和表演表情在情绪类别和呈现方式等方面存在差异(Kang & Lau, 2013)。此外,现有研究的结果也不一致。至于跨文化表情识别的群内优势效应是否受到表情呈现方式的调节,目前尚未开展相关研究。本研究旨在通过两项实验考察跨文化表情识别中的群内优势效应是否受到表情自发性(表演表情和自发表情;实验 1)和呈现方式(静态表情和动态表情;实验 2)的调节。具体而言,实验 1 招募来自东西方文化的被试,以识别来自东西方文化表达者的表演动态表情和自发动态表情;实验 2 招募来自东西方文化的被试,以识别来自东西方表达者的静态表演表情和动态表演表情。

意见 2:“与静态表情相比,动态表情中额外的动态信息可能并不包含文化方言”,这一假设似乎没有得到充分的支持和解释。

回应:感谢专家的意见。对于“与静态表情相比,动态表情中额外的动态信息可能并不包含文化方言”这一表述,确实只是作者的推断,并没有相关文献支持。过去的大部分研究支持动态信息能够促进表情识别(有关综述,请参阅 Krumhuber et al., 2023),但尚未进行跨文化研究来同时比较静态和动态表情的识别。因此,我们不清楚动态表情中的动态信息是否涵盖了文化特有的情绪方言。专家的意见提醒我们应该更加清晰地推导本研究假设,因此我们做了如下修改(见正文 P4):

除了考察表情识别的群内优势效应是否会受到表情自发性的调节外,本研究的第二个目标是探究表情的呈现方式是否影响表情识别的群内优势效应。相比静态表情,动态表情含有独特的动态信息,如表情运动的方向、质量和速度(Bould & Morris, 2008; Cunningham & Wallraven, 2009; Jack et al., 2014)。这些动态信息会启动更高级的认知过程,支持社交和情绪推断,从而促进面部表情的识别(Blais et al., 2017; Sato et al., 2008)。

尽管如此,我们仍不清楚动态信息的促进作用是针对特定文化感知者群体,还是普遍适用于所有文化感知者群体。换言之,动态表情中的动态信息是否涵盖了文化特有的情绪方言。如果该动态信息包含情绪方言,那么这些额外的情绪方言将进一步增强表情识别的群内优势效应。因此,我们预期动态表情识别的群内优势效应将大于静态表情识别的群内优势效应。这一猜想在 Effenbein 和 Ambady(2002)的元分析中初步得到支持。他们发现,使用动态表情作为实验材料的研究(共 43 项研究)所揭示的群内优势效应大于使用静态表情作为实验材料的研究(共 79 项研究)。值得注意的是,该元分析发现的表情呈现方式对群内优势效应的调节仅呈边缘显著。此外,该元分析包含的研究除了在表情呈现方式上存在差异外,还在其他诸多方面存在差异,例如表情的自发性、文化和情绪类别的选择。由于缺乏对其他方面差异的控制,我们无法确定仅仅表情的呈现方式是否会调节表情识别的群内优势效应。

然而,如果动态表情中的动态信息不包含文化特有的情绪方言,而是普遍促进所有文化感知者的情绪识别表现,那么动态表情和静态表情识别的群内优势效应将不会有差异。因此,我们预期表情的呈现方式不会调节表情识别的群内优势效应。

意见 3:实验材料的来源在方法部分没有详细阐述。例如,“动态表演表情材料则在被试选定的那一帧画面的基础上,截取前后共 2 秒的视频片段获得。”建议给出“前后共 2 秒”的理由,并提供示例图片。如果可以的话,希望能够看一下数据库的内容,以便更好的评估其有效性。由于自发表情的识别正确率非常低,所以更需要看一下材料的特点。

回应:感谢专家的意见。之前我们只进行了简要介绍,因为考虑到本研究的实验材料已在 Fang, Sauter 等人(2022)的研究中有详细介绍。然而,专家的意见提醒我们,我们仍然需要对实验材料进行必要的阐述。因此,我们进行了以下修改(见正文 P6)。

实验材料来源于 Fang, Sauter 等人(2022)的研究。该研究收集了表演表情和自发表情。在表演表情条件中, 被试被要求在摄像机前摆出特定的表情, 以便“他们的朋友能够轻松理解他们的感受”。被试完成两种表情后, 观看自己的视频, 并选择最能表达相应情绪的那一帧画面作为静态表演表情材料(用于实验 2; 参见图 1 示例)。这种方法确保了所选择的静态表情准确地表达了被试意图表达的情绪, 避免了引入实验者偏差(类似操作见 Cordaro et al., 2018; Elfenbein et al., 2007)。动态表情材料则基于被试选择的那一帧画面, 截取前后共 2 秒的视频片段。¹ 选取这个时间段既保证了被试能够清楚看见表情的变化, 又避免了引入其他非目标表情。此外, 已有研究证明, 2 秒的表情变化片段反映了较为自然的表情变化(Fang et al., 2018, 2019; Hoffmann et al., 2010)。总共有 45 名中国模特(22 名女性)和 49 名荷兰模特(36 名女性)进行气愤和厌恶表情表演, 获得了 188 个静态/动态表演表情材料。在自发表情的采集中, 被试被告知将协助开发一个能够理解人类情感的机器人。他们需要回忆过去生活中与气愤或厌恶情绪相关的个人经历, 并向机器人详细讲述这些经历。在讲述过程中, 被试的面部表情被全程记录。研究表明, 这种情绪重现法(re-lived-emotion method)能有效诱发自发表情(如, Siedlecka & Denson, 2019; Tsai & Chentsova-Dutton, 2003)。其余程序与表演表情相同。共有 35 名中国模特(24 名女性)和 35 名荷兰模特(24 名女性)进行了气愤和厌恶故事讲述, 共获得了 140 个动态自发表情刺激。

此外, 根据专家建议, 我们在正文中提供静态表演表情的示例(见图 1), 动态表演和自发表情示例已上传至以下链接: https://osf.io/8f5cd/?view_only=197bfe65e1be4723bf3ba0efb6294ba4, 并已在正文中进行了说明(见正文 P6)。由于没有获得荷兰模特的授权, 本研究无法公开发布荷兰模特的表情材料。

意见 4: 自发表情没有出现群内优势效应的部分解释可以更加深入, 当前的解释显得不够充分。

回应: 感谢专家的意见。具体修改内容如下(见正文 P13-14):

此外, 尽管表情识别的群内优势效应在大部分情况下都存在(表情种类 × 表情自发性 × 感知者文化), 但东方感知者识别自发气愤和厌恶表情时并未出现群内优势效应。该结果可能由两个原因导致。首先, 除情绪方言对跨文化情绪沟通的影响外, 历史异质性是另一个常见的影响因素(historical heterogeneity; Niedenthal et al., 2018)。历史异质性指的是过去 500 年中对一个国家的当前人口做出贡献的来源国的数量(Rychlowska et al., 2015; Wood et al., 2016)。历史异质性越高的社会拥有更多样化的价值观和信仰。为了在如此多元化的社会环境中生存, 个体需要直接而准确地传达他们的情感和意图。因此, 相较于来自历史同质性社会(例如中国和日本)的个体, 来自历史异质性社会(例如美国和加拿大)的个体在使用面部表情表达不同情绪时往往更加清晰(Fang, Rychlowska, et al., 2022; Wood et al., 2016), 因此更容易被其他个体识别。在本研究中, 与表演表情相比, 历史异质性可能对自发表情产生了更大的影响。对于中国感知者来说, 尽管情绪方言使得他们能更好地辨认中国的自发表情(相比于荷兰的自发表情), 但历史异质性使得他们能更好地辨认荷兰的自发表情(相比于中国的自发表情)。这两种作用相互抵消, 从而导致中国感知者识别自发气愤和厌恶表情时并未出现群内优势效应。

第二, 除了历史异质性外, 集体主义/个人主义也会影响情绪的跨文化沟通(Matsumoto et al., 2008)。相较于西方文化, 东方文化的个体更具集体主义倾向, 更注重维护群体的和谐与稳定, 在交流时可能更倾向于弱化对消极情绪的表达(Ekman, 1971; Matsumoto et al., 2008)。Fang, Sauter 等人(2022)的研究在一定程度上支持了这一推论。他们发现, 中国个体在自发表达气愤和厌恶表情时常用的动作单元(气愤表情仅有 2 个, 厌恶表情仅有 3 个)远远少于荷兰个体(气愤表情有 6 个, 厌恶表情有 7 个)。弱化的情绪表达提高了表情识别的难度。因此,

即使对于中国感知者来说,中国表达者的自发消极表情也可能较难识别,导致中国感知者在识别自发表情时并未出现群内优势效应。未来的研究需要更深入地探讨这些可能的解释。

意见 5: 实验任务被设置为评分量表,与部分文献使用的分类判断存在差异。需要进一步阐述评分量表的合理性。(同时可考虑一下厌恶和愤怒这两种表情的易混淆特点)

回应: 感谢专家的意见。我们在正文中补充了对此方法的说明(见正文 P7):

这里采用评分量表的方式是因为感知者可能会在表情中识别出多种情绪(Fang et al., 2018, 2019; Hess et al., 2016)。与迫选法不同,评分量表允许被试表达在感知特定表情时识别出的不同情绪的强度,而不仅仅是选择一种情绪。

意见 6: 本文中使用的是气愤和厌恶这两个表情,但是更常见的是高兴、惊讶、悲伤等等。作者似乎从可获得性的角度来选择这两种表情,建议能够更加充分地阐述理由。

回应: 感谢专家的意见。结合专家的意见,我们对研究关注气愤和厌恶这两种表情的原因进行了更详细的解释,具体内容如下(见正文 P5):

本研究关注气愤和厌恶两种情绪的面部表情。主要考虑如下三点。首先,气愤和厌恶被广泛认为是基本情绪(Ekman, 1992; Izard, 2007),因此预计在东西方文化中都普遍存在且较为常见。其次,尽管气愤和厌恶是不同的情绪,但它们在概念上存在关联(Giner-Sorolla et al., 2018),并且在面部形态上具有相似性(Cordaro et al., 2018; Fang, Sauter, et al., 2022)。这些相似性导致个体容易混淆气愤和厌恶这两种表情(Fang et al., 2018, 2019; Jack et al., 2009; Pochedly et al., 2012)。采用这两种具有混淆度的面部表情可以更好地诱发表情识别的群内优势效应。相反,当采用两种完全不同的表情,如气愤和快乐时,个体对无论来自哪个文化的气愤和快乐表情的辨别都比较好,因此表情识别的群内优势效应不太可能显现。第三,现有表情数据库主要刻画静态表演表情,对动态或自发表情的刻画较少(有关综述,请参阅 Dawel et al., 2022; Krumhuber et al., 2017),同时包含两种表情自发性或者两种表情呈现方式的跨文化表情数据库更少。据作者所知,目前仅有一项跨文化情绪表达的研究同时包含了上述四种条件的表情,为中国和荷兰模特的气愤和厌恶的静态/动态的表演和自发表情(Fang, Sauter, et al., 2022)。

意见 7: 表情自发性的诱发方法值得进一步探讨,特别是当前使用的情绪回忆法是否能有效地诱发自发表情。这可能看到材料之后有个更直观的了解。

回应: 感谢专家的意见。首先,实验室中常见的情绪诱发技术包括呈现视觉刺激、呈现听觉刺激、回忆自我经历(即情绪重现法)、情境设置和想象场景等(Siedlecka & Denson, 2019)。

Siedlecka 和 Denson(2019)对涉及诱发特定目标情绪的文献进行分析,认为回忆自我经历这种方法能有效诱发气愤和厌恶。我们在文中进行了相应修订(见正文 P6):

研究表明,这种情绪重现法(relived-emotion method)能有效诱发自发表情(如,Siedlecka & Denson, 2019; Tsai & Chentsova-Dutton, 2003)。

其次,为了考察本研究中的表情材料是否能被准确识别,我们对实验 1 和实验 2 中不同感知者文化、表达者文化、表情自发性和表情呈现方式的气愤和厌恶表情识别正确率进行单样本 t 检验。结果显示,在绝大多数情况下,气愤和厌恶表情识别正确率均高于随机水平(实验 1 为 1/5 个情绪维度 = 0.2, 实验 2 为 1/4 个情绪维度 = 0.25)。这表明,尽管气愤和厌恶表情的识别率较低,但被试仍能以高于随机水平的概率识别这些表情。相关数据已补充至附录中(见附表 1 和附表 3)。

希望上述解释能解答专家的疑惑。

意见 8: 摘要部分可以进一步扩展, 以更全面地描述研究的意义。

回应: 感谢专家的意见。我们对摘要的修改如下(见正文 P1):

随着全球化进程的加速, 跨文化交流变得尤为重要。在不同语言文化群体之间的交流中, 非语言方式如面部表情在传递情感和意图方面起着重要作用。以往研究发现, 识别同文化背景个体的面部表情比跨文化背景个体更准确, 这被称为群内优势效应。然而, 以往研究主要关注静态表演表情, 对动态表情和自发表情识别中的群内优势效应知之甚少。为了探究群内优势效应是否受表情自发性(表演和自发)和呈现方式(静态和动态)的调节, 本研究招募了中国和加拿大/荷兰被试, 分别对中国和荷兰模特表演和自发的气愤和厌恶表情(实验 1), 以及静态和动态表演的气愤和厌恶表情(实验 2)进行识别。研究结果显示, 在多数情况下, 表演和自发表情的识别都呈现群内优势效应, 且表演表情的群内优势效应显著高于自发表情; 静态和动态表演表情的识别也出现群内优势效应, 但两者总体上没有显著差异。这些结果表明, 表情识别的群内优势效应会受到表情自发性的调节, 但可能不受表情呈现方式的影响。本研究首次考察了表情识别的群内优势效应在表情自发性 and 呈现方式这两个重要特性的调节下的情况, 对于扩展群内优势效应研究和深化跨文化表情识别理解具有重要意义。

意见 9: 建议在关键词中添加“文化差异”等词汇。

回应: 感谢专家的意见。已在关键词中添加“文化差异”这一词汇, 以凸显研究的跨文化属性。

意见 10: 部分正文语句需要重新组织, 以提高逻辑性。

回应: 感谢专家的意见。我们认真通读了全文, 并对表达不清、逻辑不流畅的语句进行了调整和修改, 并邀请了几位专家对全文进行了仔细的阅读和检查, 以确保行文流畅、逻辑清晰, 易于读者阅读和理解。凡涉及到修改的内容均已在修改稿中用灰色底纹标注。

第二轮

审稿人 1 意见:

在仔细审阅提交的修订稿件后, 我非常高兴地看到作者对之前提出的审稿意见给予了充分的考虑, 并据此进行了有效的改进。这些修改显著提高了文章的质量和学术价值。

我有一点小小的建议希望作者能考虑。在本次修订稿中, 作者详细介绍了按照 FACS 手册标准制定的标准表情原型。我建议能进一步阐明本研究所选取的自发和表演表情材料相对于这些标准表情, 在理论和实践层面上的优势。如果可能的话, 希望能在文章正文中补充相关内容, 以增强论文的深度和说服力。

再次感谢您对改进工作的努力, 我期待看到您的进一步修订稿。

回应: 非常感谢审稿专家对于我们修改稿的肯定以及提出的建议。我们在引言中进行补充说明如下(见正文 P4):

其中, 表演表情是通过让模特摆出特定表情以便于他们的朋友能够轻松理解他们的感受所得, 而自发表情则是通过让模特回忆并讲述过去生活中与特定情绪相关的个人经历所得。相较于前人使用的标准表情(如, Kang & Lau, 2013), 本研究中采用的表演表情更加符合各自文化中的情绪表达规范, 并且避免了按照统一标准表演的面部表情掩盖文化差异可能导致群内优势效应消失的问题(Elfenbein et al., 2007)。

审稿人 2 意见:

原先提出的大部分问题得到比较好的解答。目前还有一个问题需要进一步讨论:

作者在文中引用的支持“两秒钟”的前人文献并不能提供足够的支撑,例如 Hoffmann 这篇文章并不是针对的自发性的表情,且时间也不是 2 秒。

我看了作者提交的视频材料,材料的质量并不高,例如愤怒的自发表情,似乎并不是一个完整的表情过程。

回应: 非常感谢审稿专家提出的意见。如果完整的表情过程意味着从中性到特定情绪再回到中性,那么本研究中的大部分研究材料都不满足这个标准。因为在表演或自发条件下,模特并没有被要求从中性开始变化到特定情绪再回到中性。特别是在自发条件下,模特是在讲述与特定情绪相关的过去生活经历,他们的表情是连续动态变化的,并不存在所谓的“完整的表情过程”。即使延长时间,大多数情况下也不会有“完整的表情过程”。相反,延长时间可能会混入其他非目标情绪(从目标情绪变为另一种非目标情绪)。为了增加表情材料的生态效度,本研究的表情呈现并不像按照 FACS 标准制作的那样精确。尽管如此,本研究中的表情材料均能被有效识别(见附表 1 和附表 3),而且识别率与先前类似的研究相当(Elfenbein et al., 2007)。

关于表情刺激呈现时间长度的问题,很抱歉之前没有将 Hoffmann 等人(2010)这篇文章的佐证说清楚。根据 Hoffmann 等人(2010)的研究,从中性开始变化到特定情绪再回到中性的表情持续时间大约为 1649ms-3108ms(从中性变为特定情绪的时间+从特定情绪变为中性的时间)。考虑到以下因素:1)确保被试能够清晰地观察到表情的变化,2)避免引入其他非目标表情,3)所有的表情材料片段长度相同,我们最终选取了 2 秒的表情变化片段。

综上,我们在文中进行了如下修改:

选取这个时间段既保证了被试能够清晰地观察到表情的变化,又避免了引入其他非目标表情。此外,约 2 秒的表情变化片段反映了较为自然的表情变化(Fang et al., 2018, 2019 采用了 2000ms 的表情转变时间; Hoffmann et al., 2010 采用了 1649ms-3108ms 的从表情起始到结束的时间)。

第三轮

审稿人 2 意见:

作者的工作非常好,回复非常认真,推荐修后发表。目前还有一两个点比较在意。

作者在回应中指出:选取这个时间段(2 秒钟)既保证了被试能够清晰地观察到表情的变化,又避免了引入其他非目标表情。这就意味着,在回忆的过程中,模特有很多几乎同时出现的表情,如愤怒+恐惧,作者会不会选择有利于结果的表情?另外,对于愤怒和厌恶来说,两秒钟是不是太长了?我希望作者能够在这个问题上进一步讨论:应该固定时长还是选择一个相对完整的表情。

回应: 非常感谢审稿专家对我们工作的肯定。本研究采用的静态表情材料是被试观看自己表演的视频时自主选择的最能表达相应情绪的那一帧画面,而动态表情材料则是所选画面前后共 2 秒的视频片段。实验者并没有干预被试的选择,也无法“选择有利于结果的表情”。此外,本研究对于“情绪识别的群内优势效应是否受表情自发性和呈现方式的调节”这一问题形成了竞争性假设,因此并不存在所谓的“选择有利于结果的表情”。

关于 2 秒的表情片段是否太长的问题: 在表演条件下,被试往往会在他们认为已经摆出特定的情绪面孔后维持一段时间(通常为 3 秒甚至更久),但会伴随着强度的变化。如果选

取更短的时间，动态的表演表情可能没有发生任何变化。

关于选择固定时长还是相对完整表情片段的问题：正如我们第二轮回应中所述，大部分情况下并不存在所谓的“完整的表情过程”，因此我们无法按照此标准选取表情片段。此外，表情刺激呈现的时长不一致也意味着被试对其的感知时长不一致，这可能会作为一个混淆因素影响实验结果。综上考虑，我们选择了统一时长。

为了更加清晰地阐释我们的选择理由，我们在文中进行了相应修改(见正文 P24，以脚注形式呈现，用紫色标注；对应说明内容见正文 P6 的第 22 行)：

在表演条件下，被试往往会在他们认为已经摆出特定的情绪面孔后维持一段时间(通常为 3 秒甚至更久)，但会伴随着强度的变化；而在自发条件下，因为被试讲述与特定情绪相关的过去生活经历，其表情呈连续的动态变化。我们选择了 2 秒的时间段，既能保证被试能够清晰地观察到表情的变化，又避免引入其他非目标表情。此外，约 2 秒的表情变化片段反映了较为自然的表情变化(Fang et al., 2018, 2019 采用了 2000ms 的表情转变时间；Hoffmann et al., 2010 采用了 1649ms-3108ms 的从表情起始到结束的时间)。

再次感谢您的宝贵意见和建议。如果还有任何其他问题或疑虑，请随时提出，我们将非常乐意回答。

编委意见：作者已经很好地回答了审稿人的最后的意见。建议发表。

主编意见：该文作者就多轮审稿的意见进行了详尽而认真的回复，审稿专家均表示满意。整个流程规范严谨，论文经修改达到心理学报的发表标准，同意接受。