

# 基于视听双通道的音乐情绪冲突效应及优势通道研究\*

周欣<sup>1,2</sup>, 郑莹灿<sup>3</sup>, 肖子伦<sup>1</sup>, 郑茂平<sup>4</sup>

(1. 西南大学心理学部, 重庆 400715; 2. 川北医学院精神卫生学院, 南充 637100;

3. 陆军军医大学医学心理系, 军人发展心理学教研室, 重庆 400000; 4. 西南大学音乐学院, 重庆 400715)

**摘要:**为探讨基于视听双通道的音乐情绪冲突效应、冲突情境下的优势加工通道和音乐经验对结果的影响,本研究采用音乐表演视频为材料,比较音乐组和非音乐组被试在一致型和不一致型视听双通道下的情绪评定速度、准确性及强度。结果发现:(1)一致型条件下的情绪评定更准确且更强烈;(2)不一致型条件下,被试更多以听觉通道的情绪线索为依据进行情绪类型评定;(3)非音乐组被试比音乐组被试更依赖视觉通道的情绪线索。结果表明:通道间情绪信息的不一致阻碍了音乐情绪加工;听觉通道是音乐情绪冲突情境下的优势加工通道;音乐经验降低了情绪冲突效应对音乐组被试的干扰。

**关键词:**音乐情绪;视听双通道;情绪冲突效应;听觉优势

**中图分类号:**B842.5

**文献标识码:**A

**文章编号:**1003-5184(2022)01-0036-10

## 1 前言

在传统观念中,音乐是声音的艺术。随着时代的发展,在音乐的基础上演化出了多种艺术形式,如歌剧、音乐剧等。音乐不再只是听觉艺术,还成为了一种视听结合的表演艺术。随着生活水平的提高和网络技术的兴起,人们不再局限于听音乐,观看音乐表演这一欣赏方式在日常生活中越来越常见和多元化,包括线下观看现场的音乐表演,以及线上观看电视台或各类网络平台的音乐表演等。音乐表演中的表情、动作和背景画面等视觉信息与音乐的速度、调式和旋律轮廓等声音特征一样,包含了丰富的音乐情绪信息,影响着听赏者对音乐情绪的识别和体验(Ammirante & Thompson, 2012; Davidson, 2012; Wöllner, 2012; Livingstone, Thompson, Wanderley, & Palmer, 2015; Livingstone & Russo, 2018; Lange, Funderich, & Grimm, 2022)。这种通过整合视、听等多通道的情绪信息,对自己和他人的情绪状态进行判断的现象被称为情绪信息多通道整合(Multisensory Integration, MSI),这一现象在日常生活中也常见,是情绪研究的重要课题(张亮,孙向红,张侃, 2009)。

当来自不同感官通道的情绪信息协调一致,即表达同一情绪时,可提高情绪加工效率,使个体的情绪反应更快速、更准确和更强烈,即多通道整合优

势,也称情绪一致性效应(emotional congruent effect)(Chen, Pan, Wang, Yang, Liu, You, & Yuan, 2016; Klasen, Kreifelts, Chen, Seubert, & Mathiak, 2014)。但是,当通道间情绪信息不一致,即表达不同情绪时,会干扰个体的情绪加工,即情绪冲突效应(emotional conflict effect),且这一干扰在告知被试忽视某一通道时仍发生(Mondloch, 2012; Müller, Habel, Derntl, Schneider, & Zilles, 2011; Zinchenko, Kanske, Obermeier, Schröger, & Kotz, 2015; Misselhorn, Friesse, & Engel, 2019)。个体在生活中常可能面对情绪冲突情境,如面带笑容却语调低沉,或愁眉苦脸却语调轻快,等。需整合各通道的情绪信息,觉察情绪冲突,进行最终的情绪判断,是有难度的加工过程,也是重要的社会技能。因此,情绪冲突效应及冲突情境下的优势通道,即个体以哪个通道的情绪信息为主导进行情绪加工,是情绪信息多通道整合的重要课题。

音乐表演中的音乐情绪信息多通道整合也存在情绪一致性效应和情绪冲突效应两种结果。优秀的音乐表演者可以在演奏或演唱音乐作品的同时,通过与音乐作品情绪一致的面部表情、身体姿势和动作等视觉情绪信息,帮助自己更投入音乐情绪状态的同时,也更有效地向听赏者传达相应的情绪。有研究者发现,视听双通道的音乐情绪加工效率高于

\* 基金项目:中央高校基本科研业务费科研项目(SWU1709123),重庆市社会科学规划基金项目(2019YBYS149)。

通讯作者:郑茂平, E-mail: zmpshxy@swu.edu.cn。

单听觉通道,证实了视听音乐表演优于仅聆听音乐的情绪传达效果(Davidson, 2012; Mitchell & MacDonald, 2014; Platz & Kopiez, 2012)。但也有研究发现,增加视觉信息并不能显著提升情绪加工效率及情绪传达效果,甚至由于视觉信息过多、夸张或与听觉情绪不一致而干扰情绪加工,这些现象在实际的音乐表演中也存在(Krahé, Hahn, & Whitney, 2015; Shoda & Adachi, 2016; Vuoskoski, Gatti, Spence, & Clarke, 2016; 周欣, 郑莹灿, 郑茂平, 2020)。若视觉信息对情绪传达效果无显著影响,是否说明音乐表演中不存在情绪冲突效应?若有显著影响,听赏者在冲突情境下主要依靠音乐表演中的声音特征识别和体验情绪,还是更受表演者的面部表情等视觉线索影响,即优势通道如何?随着音乐表演中表情、动作和背景画面等视觉信息使用的增多,对以上问题进行研究并解答如何合理使用视觉情绪线索达到音乐表演的最佳情绪传达效果,是音乐表演领域的重要课题。

然而,以往研究极少直接采用音乐表演视频为视听材料,而是通过成对呈现听觉和视觉材料创造视听双通道情境。面部表情图片是常用的视觉材料,笑声、哭声和尖叫声等人声常作为听觉材料。音乐由于情绪信息丰富、准确且高度一致也常作为听觉材料(Lundqvist, Carlsson, Hilmersson, & Juslin, 2009; 李冬冬, 程真波, 戴瑞娜, 汪芬, 黄宇霞, 2012)。以往研究多采用跨通道匹配范式(cross-modal matching),通过操纵各通道材料的情绪类型获得不同匹配条件的视听双通道。两通道的材料表达同一情绪为一致型匹配(congruent),表达不同情绪则为不一致型或冲突型(incongruent)匹配。通过计算公式“ $CE = I(\text{incongruent}) - C(\text{congruent})$ ”来考察情绪冲突效应(Egner & Hirsch, 2005; Chapados & Levitin, 2008; Eerola, Peltola, & Vuoskoski, 2015; Krahé et al., 2015; Platz & Kopiez, 2012)。

以往研究关于优势通道的判定方法则有所不同。有研究不采用视听双通道,直接比较单听觉和单视觉通道,情绪加工效率更高的通道即定为优势通道(Paulmann & Pell, 2011)。有研究在双通道情境下直接指定某一通道为目标通道,或通过指导语引导被试仅注意某一通道,另一通道则为非目标通道。若目标通道的情绪加工效率较其单通道情境有下降,即说明目标通道的情绪加工受到了非目标通道的影响,下降程度代表非目标通道的影响力大小,

影响力更大的非目标通道即为优势通道(Krahé et al., 2015; Petrini, Mcleer, & Pollick, 2010)。以往研究通过上述方法对优势通道做出了不同回答。视觉优势取向认为视觉通道占优势地位,强调面部表情和身体姿势的主导作用(Jessen & Kotz, 2011; Kret, Roelofs, Stekelenburg, & de Gelder, 2013; Strelnikov, Foxton, Marx, & Barone, 2015)。听觉优势取向则表示听觉优于视觉,视觉优势只存在于听觉信息可得性较差的情况(Piwek, Pollick, & Petrini, 2015; Rigoulot & Pell, 2012, 2014)。

音乐情绪冲突效应及优势通道的研究不仅数量少,且结果间也存在听觉优势和视觉优势的争议。有研究者强调视觉信息对音乐情绪加工的主导作用(Thompson, Russo, & Livingstone, 2010; Vuoskoski, Thompson, Clarke, & Spence, 2014)。并将视觉线索细化和分类,总结出音乐家动作的意义及其对音乐情绪的作用(Davidson, 2012)。但有研究者表示视觉信息虽可能影响听觉情绪信息的加工,但其影响并不凌驾于听觉之上(de Gelder & Vroomen, 2000; Krahé et al., 2015; Petrini et al., 2010; Rigoulot & Pell, 2012, 2014; Piwek et al., 2015; Stock, Peretz, Grèzes, & Gelder, 2009)。且上述研究大多仅验证两通道间的相互影响,较少直接证明优势通道,也未明确探讨这种相互影响是否降低音乐情绪加工效率,即是否存在冲突效应。

视觉通道的材料中是仅呈现面孔还是呈现整个身体,是静态图片还是动态画面;听觉材料是采用音乐还是语音,是器乐演奏还是声乐演唱,是否存在言语信息,及其时长,均可能导致结果的差异;视觉和听觉的时间融合性、空间整体性和生态性水平的高低也会影响冲突效应是否出现及其强度如何(Zinchenko, Kanske, Obermeier, Schröger, Villringer, & Kotz, 2018)。

在心理学研究中,实验的生态性非常重要,采用接近日常生活的多通道动态刺激进行视听通道情绪冲突的研究是一个重要的趋势(Zinchenko, Kotz, Schröger, & Kanske, 2020)。音乐表演视频中,来自视觉和听觉两通道的情绪信息共存于一个整体中,是一类生态性较高的视听材料。不仅适用与探索日常生活中的情绪冲突效应,也可用于研究音乐表演中的多通道整合相关问题。除实验材料外,实验方法也有生态性水平的差异。在实验中向被试指定目标通道与日常生活中被试面临的情绪加工情境不

同,而不指定目标通道的优势通道判定方法则更接近日常情境,生态性较高。

根据上述分析,本研究拟采用音乐表演视频为材料,借鉴跨通道匹配范式,进行音乐情绪冲突效应及其优势通道的研究,回答音乐表演中是否存在音乐情绪效应及听赏者会以何种通道为主导进行情绪识别和体验这两个存在争议的问题。

在实验材料方面,本研究选择现代心理学中被认为的最单纯(黄希庭,2007),且在音乐的情绪表达中较典型的开心、悲伤、愤怒和恐惧四种基本情绪类型。以表达这四种情绪类型的音乐表演原始视频为一致型视听双通道材料。将来自不同音乐表演,且分别表达不同情绪类型的声音和画面重新进行剪辑组合,获得12种不同情绪组合方式的新视频,作为不一致型视听双通道材料。这一复杂多样的情绪组合方式较符合日常生活中复杂的情绪冲突情境。在实验任务方面,本研究不指定或通过指导语引导目标通道,而是任被试自由选择评定依据。这一评定方式也更接近真实的日常生活情境和音乐听赏情境。

本实验记录情绪评定的速度、正确率和强度作为情绪加工效率的指标。其中,由于不一致型视听双通道的视觉和听觉材料分别来自不同的音乐表演片段且表达不同的情绪类型,被试对音乐情绪类型的评定可获得两种正确率。即以声音所表达的情绪类型为正确答案的正确率,及以画面所表达的情绪类型为正确答案的正确率。

根据计算公式“ $CE = I(\text{incongruent}) - C(\text{congruent})$ ”比较两通道的情绪加工结果,考察是否存在音乐情绪冲突效应。在音乐情绪冲突情境下的优势加工通道的判定方面,本研究比较不一致型视听双通道的两种正确率,若以听觉通道材料的情绪类型为答案的正确率更高,则说明被试依据声音的情绪信息进行情绪评定时更准确,即音乐情绪冲突情境下的优势通道为听觉通道;反之,若以视觉通道材料的情绪类型为答案的正确率更高,则说明视觉通道为优势通道。

此外,由于音乐训练会影响个体的音乐情绪加工和多感官信息整合交互能力,最终对结果产生影响(侯建成,董奇,2010;Law & Zentner,2012;Kantor, Bigand, & Emmanuel,2013;周欣等,2020)。曾有研究发现,音乐家被试的听觉情绪加工更不易受视觉信息的干扰(Weijkamp & Sadakata,2017)。本研究招募具有不同音乐训练经验的被试,分为音乐

组和非音乐组,探索音乐经验对音乐情绪冲突效应及其优势通道的影响。

综上,本研究通过较高生态性的刺激材料和实验方法,更真实地探讨情绪冲突情境下的音乐情绪加工机制及其可能的影响因素,为存在争议的研究课题提供新的证据,做出补充和扩展。同时,为音乐表演、音乐审美和音乐教育等环节中视觉情绪线索的合理使用提供心理学依据。

## 2 方法

### 2.1 被试

104名大学生,根据其音乐专业训练背景分为两组:音乐组42人(女生36人),年龄 $21.5 \pm 2.27$ ,均为音乐专业学生,且接受超过五年的声乐或器乐培训;非音乐组62人(女生36人),年龄 $20.44 \pm 1.56$ ,非音乐专业且未接受过声乐、器乐或舞蹈的专业训练。听力正常,视力或矫正视力正常。

### 2.2 实验材料

分一致型(congruous)和不一致型(incongruous)两种。一致型材料来自笔者前期所编制的音乐影像材料库,包括歌剧、交响乐、民乐等多种艺术形式,表达开心、悲伤、愤怒和恐惧四种情绪。音乐曲目来源于以往研究使用过的材料、相关研究结果及音乐专业人士的评定,并排除大众较熟悉的曲目。选择具有高度专业水平的音乐表演视频,剪辑为30秒左右的影像片段(Hunter, Schellenberg, & Schimmack, 2010),保证音乐上的完整感,能完整和清楚地观察到表演者的面部表情和身体姿势,且代表的情绪与音乐曲目的情绪基本一致。通过软件分离声音和画面,获得相应的单听觉和单视觉通道材料。95名大学生被试对三种通道形式的音乐片段进行了情绪类型和强度评定,筛选出强度较高且情绪类型选择集中率较高的64条片段,并根据被试的选择结果确定该材料的情绪类型,作为计算正确率的标准答案。冲突型材料17条,其声音和画面分别来自一致型材料的不同片段且代表不同的情绪类型。为了避免被试对冲突的察觉不是由于通道间的情绪冲突,而是由于表演者动作与音乐节奏的差别太大影响研究结果真实性,在合成视频时尽量保证动作和声音的时间一致性。材料格式为avi,视频帧速率15帧每秒,音频比特率64kbps,音质采样频率44Hz,双声道立体声。

### 2.3 实验设计

音乐情绪冲突效应方面,采用2(通道匹配类型:

一致、不一致) × 2(音乐经验:音乐组、非音乐组)两因素混合实验设计。通道匹配类型为被试内变量,音乐经验为被试间变量。因变量为情绪评定反应时(RT)、正确率(ACC)和强度(Intense),分别反映音乐情绪加工的速度、准确性及情绪反应强度。情绪评定越快、越准确、强度越高说明情绪加工效率越高。

音乐情绪冲突情境下的优势加工通道方面,采用 2(答案通道类型) × 2(音乐经验)两因素混合实验设计。答案通道类型为被试内变量,音乐经验为被试间变量。因变量为不一致型视听双通道的两种正确率,即以视觉通道材料的情绪类型为正确答案的正确率(ACC - A)和以听觉通道材料的情绪类型为正确答案的正确率(ACC - V)。更高的一种正确率说明以该通道的情绪线索为依据的音乐情绪加工更准确,即为音乐情绪冲突情境下的优势加工通道。

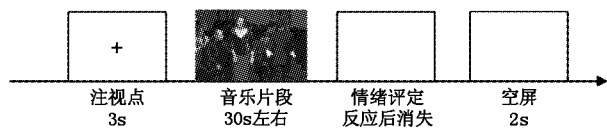


图 1 实验流程图

实验在标准化行为实验室进行,设备的多媒体和音响效果良好,被试可先将音量调到合适大小。采用 E - prime 2.0 编程与执行。由于不一致型材料中的视觉和听觉信息均来自一致通道的材料,视觉或听觉内容再现时间间隔过近或引起被试的注意干扰实验。为了延长信息再现的间隔时长,不同匹配类型的材料分组呈现,一类通道的实验完成后休息 5 分钟再进行下一个实验。匹配类型呈现顺序随机,通道内材料随机呈现。每类通道的实验先进行练习(2 个试次)再进入正式实验(15 个试次)。每个试次由注视点、材料呈现、情绪评定和空屏组成(图 1)。情绪评定包括四个问题:这段音乐表演想表达什么情绪?这种情绪的强度如何?这段表演让您真切的体验到了什么情绪?这种情绪的强度如何?情绪类型有愤怒、开心、悲伤和恐惧 4 个选项,强度评定为 7 点评分。前两个问题指向音乐情绪识别,后两个指向音乐情绪体验(Schubert, 2007; Krahé et al., 2015)。由于顺序不会显著影响结果(Schubert, 2007),且避免改变顺序造成误判,4 个评定项目顺序不变。记录评定正确率(以材料初评结果为正确答案)和强度,及第一个评定项目的反应时。

3 结果与分析

剔除反应时低于 100ms 及高于 4 个标准差以外

的数据后,对反应时和强度进行 2(通道匹配类型) × 2(音乐经验)的重复测量方差分析。不一致型视听双通道的正确率包括以视觉通道材料的情绪类型为正确答案的正确率(ACC - A)和以听觉通道材料的情绪类型为正确答案的正确率(ACC - V)。对两组正确率结果进行 2(答案通道类型) × 2(音乐经验)的重复测量方差分析,并以更高的一种正确率为不一致型视听双通道的情绪评定正确率,进行 2(通道匹配类型) × 2(音乐经验)的重复测量方差分析。

3.1 反应时

反应时通道主效应不显著,  $M: 3.04 < 3.23, F(1, 102) = 2.437, p > 0.05$ 。组间差异不显著,  $F(1, 102) = 0.126, p > 0.05$ 。通道与组间交互作用不显著,  $F(1, 102) = 0.001, p > 0.05$ 。

表 1 两组被试的反应时结果

通道类型	<i>M</i> ( <i>SD</i> )	
	音乐组( <i>n</i> = 42)	非音乐组( <i>n</i> = 64)
AV	3.08(0.11)	3.00(0.11)
AVin	3.27(0.12)	3.19(0.15)

3.2 强度

表 2 两组被试的强度结果

项目	<i>M</i> ( <i>SD</i> )	
	音乐组( <i>n</i> = 42)	非音乐组( <i>n</i> = 36)
AV	情绪识别 4.69(0.76)	4.79(0.97)
	情绪体验 4.16(1.08)	4.01(1.17)
AVin	情绪识别 4.54(0.83)	4.51(0.95)
	情绪体验 4.03(1.24)	3.72(1.14)

情绪识别强度通道主效应显著,一致型视听双通道显著高于不一致型视听双通道,  $M: 4.76 > 4.53, F(1, 102) = 15.478, p < 0.001, \eta^2 = 0.133$ 。两组被试均通道差异显著,  $M: 4.69 > 4.54, t(41) = 2.063, p < 0.05, Cohen's d = 0.644; M: 4.83 > 4.52, t(60) = 3.748, p < 0.001, Cohen's d = 0.968$ 。组间差异不显著,  $F(1, 102) = 0.110, p > 0.05$ 。通道与组间交互作用不显著,  $F(1, 102) = 2.096, p > 0.05$ 。

情绪体验强度通道主效应显著,一致型视听双通道显著高于不一致型视听双通道,  $M: 4.11 > 3.88, F(1, 102) = 15.822, p < 0.001, \eta^2 = 0.135$ 。组间差异不显著,  $F(1, 102) = 0.823, p > 0.05$ 。通道与组间交互作用不显著,  $F(1, 102) = 3.061, p > 0.05$ 。

3.3 正确率

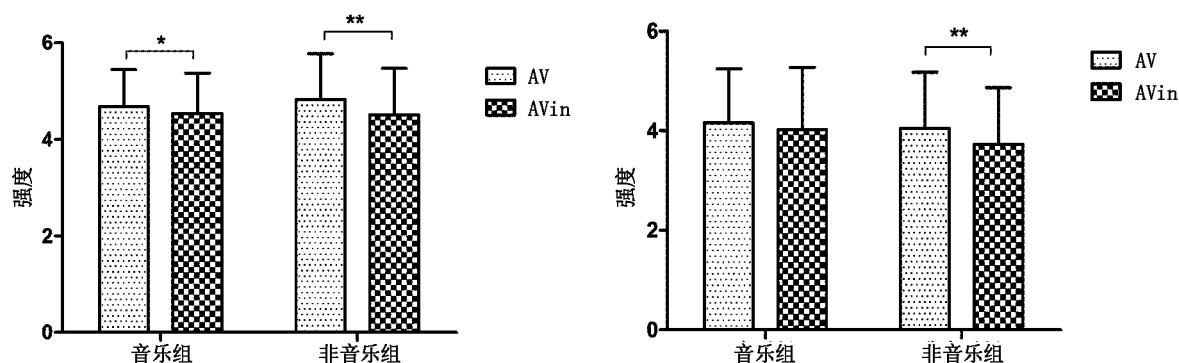


图2 两组被试在不同条件下的情绪强度(左:音乐情绪识别;右:音乐情绪体验)

表3 两组被试的正确率结果

项目		<i>M</i> ( <i>SD</i> )	
		音乐组( <i>n</i> = 42)	非音乐组( <i>n</i> = 36)
AV	情绪识别	0.74(0.13)	0.75(0.14)
	情绪体验	0.72(0.13)	0.73(0.16)
AVin	情绪识别 ACC - A	0.70(0.20)	0.64(0.19)
	情绪识别 ACC - V	0.10(0.10)	0.13(0.10)
	情绪体验 ACC - A	0.72(0.19)	0.63(0.18)
	情绪体验 ACC - V	0.10(0.10)	0.15(0.11)

不一致型视听双通道下情绪识别正确率的答案通道主效应显著,以听觉通道情绪类型为答案的正确率(ACC - A)显著高于以视觉通道情绪类型为答案的正确率(ACC - V), $M:0.67 > 0.12$ ,  $F(1, 102) = 420.219$ ,  $p < 0.001$ ,  $\eta^2 = 0.809$ 。组间差异不显著,  $F(1, 102) = 0.591$ ,  $p > 0.05$ 。答案通道类型与组间交互作用不显著,  $F(1, 102) = 2.812$ ,  $p > 0.05$ 。

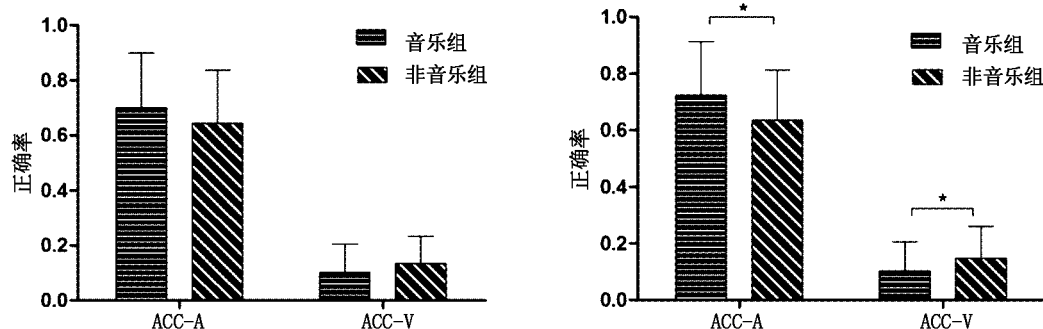


图3 两组被试在不一致条件下的情绪评定正确率(左:音乐情绪识别;右:音乐情绪体验)

不一致型视听双通道下情绪体验正确率的答案通道主效应显著,ACC - A 显著高 ACC - V,  $M:0.68 > 0.13$ ,  $F(1, 102) = 421.969$ ,  $p < 0.001$ ,  $\eta^2 = 0.810$ 。组间差异不显著,  $F(1, 102) = 2.630$ ,  $p > 0.05$ 。答案通道类型与组间交互作用显著,  $F(1, 102) =$

$6.112$ ,  $p < 0.05$ ,  $\eta^2 = 0.058$ 。音乐组的 ACC - A 显著高于非音乐组,  $t(99) = 2.425$ ,  $p < 0.05$ ,  $Cohen's d = 0.487$ ;而 ACC - V 显著低于非音乐组,  $t(99) = -1.997$ ,  $p < 0.05$ ,  $Cohen's d = 0.401$ 。

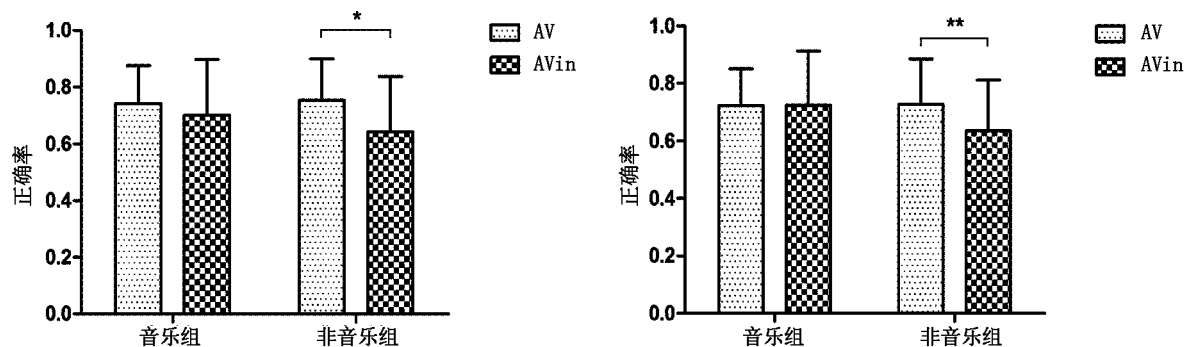


图4 两组被试不同条件下的情绪评定正确率(左:音乐情绪识别;右:音乐情绪体验)

以 ACC - A 为不一致型视听双通道的正确率进行 2(通道匹配类型) × 2(音乐经验)的重复测量方差分析。情绪识别正确率通道主效应显著,一致型视听双通道显著高于不一致型视听双通道,  $M: 0.75 > 0.67, F(1, 102) = 11.346, p < 0.001, \eta^2 = 0.103$ 。组间差异不显著,  $F(1, 102) = 0.769, p > 0.05$ 。通道与组间交互作用不显著,  $F(1, 102) = 2.389, p > 0.05$ 。

情绪体验正确率通道主效应显著,一致型视听双通道显著高于不一致型视听双通道,  $M: 0.73 > 0.68, F(1, 102) = 4.600, p < 0.05, \eta^2 = 0.044$ 。组间差异不显著,  $F(1, 102) = 2.733, p > 0.05$ 。通道与组间交互作用显著,  $F(1, 102) = 4.804, \eta^2 = 0.046, p < 0.05$ 。一致条件下音乐组低于非音乐组,不一致条件下音乐组高于非音乐组。

#### 4 讨论

##### 4.1 基于视听双通道的音乐情绪冲突效应

与以往研究一致,不一致型视听双通道的情绪评定正确率和强度均显著低于一致型视听双通道,反映了情绪加工的效率和音乐表演的情绪传达效果有所降低,说明通道间情绪信息不一致干扰了视听双通道的音乐情绪加工,证实了音乐表演中存在音乐情绪冲突效应。

视听双通道情境下来自听觉和视觉通道的情绪信息均会得到被试的自动化加工并相互影响(Chen et al., 2016; 李萍, 张明明, 李帅霞, 张火垠, 罗文波, 2019)。在不一致型视听双通道条件下,当两通道的情绪信息一致,可相互促进辅助个体更有效率地进行音乐情绪加工;但当两通道的情绪信息不一致时,不仅无法相互促进辅助音乐情绪多通道整合加工,还会造成加工资源的竞争,相互干扰导致音乐情绪加工效率降低(Etkin, Egner, Peraza, Kandel, & Hirsch, 2006; Haas, Omura, Constable, & Canli, 2006)。根据线索一致性模型,视听不一致的音乐表演所表征的情绪符号信息与听赏者对情绪信息的表征一致性降低,使音乐与听赏者间的情绪沟通难度增加,产生音乐情绪冲突效应(Juslin & Sloboda, 2010)。

这提示,音乐表演中的视觉情绪线索若与音乐的声音所表达的情绪不一致,会干扰听赏者对音乐的情绪识别和体验,使音乐情绪传达的准确性和强

度减弱。为避免音乐情绪冲突效应,表演者需考虑面部和肢体动作所包含的情绪信息是否与音乐试图表达的情绪一致,在音乐教育中也可训练表演者进行有目的的视觉情绪表达。

但本实验在反应时指标上并未显示出情绪冲突效应,即情绪冲突效应对音乐情绪加工的速度未产生显著干扰。这可能与本实验的实验程序及实验材料的时长有关。本实验需在被试听赏完 30 秒左右的音乐表演视频后再进行音乐情绪评定。以往研究发现,对情绪冲突的觉察和解决加工也被证实可在较短时间内完成(Paulmann & Pell, 2010; Paulmann, Jessen, & Kotz, 2009; Doi & Shinohara, 2013; Liu, Rigoulot, & Pell, 2015; 王苹等, 2015; 李萍等, 2019)。在本实验的音乐情绪冲突情境中,个体可以在音乐呈现阶段完成对音乐情绪冲突的察觉和解决,不会造成情绪评定速度的音乐情绪冲突效应。

##### 4.2 基于视听双通道的音乐情绪听觉优势

在不一致型视听双通道下,以听觉通道情绪类型为答案的正确率显著高于以视觉通道情绪类型答案的正确率。即以听觉通道情绪信息为依据时,情绪加工更准确。说明当音乐表演中的听觉和视觉通道的情绪信息不一致时,个体更多依据听觉通道的情绪线索进行音乐情绪识别和体验,即音乐情绪冲突情境下的优势通道为听觉通道。这一结果支持了听觉优势取向的观点(de Gelder & Vroomen, 2000; Stock et al., 2009; Petrini et al., 2010; Rigoulot & Pell, 2012, 2014; Krahé et al., 2015; Piwek et al., 2015)。

根据“通道精确”假设,各感官通道对感知活动的影响程度与其线索的精确性成正比(文小辉等, 2009)。本实验结果说明,虽然音乐表演中的视觉和听觉通道均包含了丰富的情绪信息,但听觉通道的音乐情绪信息更“精确”。因此,当两通道的信息发生冲突时,听赏者会以更“精确”的听觉通道为主导进行音乐情绪加工。我们的前期研究也发现,单听觉通道呈现的音乐表演比单视觉通道呈现的音乐表演情绪评定更准确和更强烈,即音乐表演中听觉通道的情绪信息更“精确”。这与传统观念中音乐是声音的艺术并将其作为一类常用听觉情绪诱发材料是一致的。这进一步提示,在音乐表演中,若不能保证视觉情绪信息与音乐所表达的情绪一致,应更

专注于声音的情绪表达。

#### 4.3 音乐经验对音乐情绪冲突效应及优势通道的影响

比较两组被试的结果可发现,音乐经验对情绪冲突效应及优势通道产生了影响。一方面,情绪冲突效应对两组被试的情绪加工干扰程度不同。从图2和图4可看出,情绪强度和以听觉通道情绪类型为答案的正确率均显示为非音乐组的通道差异更大。情绪体验强度、情绪识别正确率和情绪体验正确率三项指标均显示为,音乐组通道差异不显著,而非音乐组通道差异显著。且非音乐组不一致双通道的情绪体验正确率显著低于音乐组。说明非音乐组被试的音乐情绪加工受情绪冲突效应的干扰更大,而音乐组被试受到的干扰较小。以往也有研究发现,音乐组被试的听觉加工更不易受视觉信息的干扰(Weijkamp & Sadakata, 2017)。另一方面,两组被试在冲突情境下对视觉和听觉的依赖程度不同。从图3可看出,虽然两组被试均是以听觉通道情绪类型为答案的正确率显著更高,但音乐组被试的两类正确率差异更大。且情绪体验正确率以听觉通道情绪类型为答案时音乐组显著高于非音乐组,以视觉通道情绪类型为答案时音乐组显著低于非音乐组。说明,虽然两组被试在音乐情绪冲突情境下的优势通道均为听觉通道,但非音乐被试对视觉情绪信息的依赖程度比音乐组高,而音乐组被试对听觉情绪信息的依赖程度比非音乐组高。以往研究也曾提出,音乐家被试更少出现情绪冲突效应是因为其对听觉信息的依赖程度更高(Proverbio, Massetti, Rizzi, & Zani, 2016)。

“统计优化模型”和“通道精确”假设提出,当所呈现通道的情绪信息模糊时,个体会通过整合多感官信息获得更多线索来更准确地识别和体验情绪;反之,若某一通道的情绪信息足够精确,个体不需要其他通道信息的协助即可获得准确的情绪识别和体验(文小辉等, 2009)。非音乐组被试可通过音乐的速度,将快速的音乐评定为开心,慢速的音乐评定为悲伤(蔡岳建,潘孝富,庄钟春晓, 2007)。但本研究还需区分开心、恐惧、愤怒和悲伤,对于缺乏音乐经验的非音乐组被试来说,任务较难。

因此,对非音乐组被试来说,听觉情绪信息不够“精确”,需借助所接收的视觉情绪信息完成情绪评

定。若视觉与听觉情绪信息不一致,则会对音乐情绪加工产生干扰,导致情绪评定的准确性和情绪反应强度显著下降。而音乐组被试有丰富的音乐经验,对音乐情绪的评定除了依据速度,还可依靠节奏型、调性、和声和流派等信息。对音乐组被试来说,听觉通道的音乐情绪信息高度“精确”,足以单独依靠其完成情绪评定。在冲突情境下,音乐组被试会专注于音乐表演的声音线索,不受视觉信息的干扰,情绪加工效率不会显著下降。丰富的音乐经验使听赏者在面对不一致的情绪信息时专注于“精确”程度更更的听觉通道信息,保持准确有效的音乐情绪加工。

这提示,接受音乐训练可使听赏者更少受到不一致视觉情绪信息的影响,降低情绪冲突效应对音乐情绪加工的干扰,能更准确地感受音乐所传达的情绪。同时也提醒表演者,向没有音乐经验的听赏者提供的视觉情绪信息必须与听觉情绪信息协调一致,才能保证音乐情绪传达效果。

#### 4.4 研究不足和未来展望

首先,本研究未将情绪类型作为自变量进行考察,而情绪本身的识别难度是有差异的,可能影响实验结果。也有研究曾提出,优势通道可能因情绪类型而有所不同(Paulmann & Pell, 2011)。未来可通过更细致地操纵情绪类型的组合形式来探索情绪类型对结果的影响。其次,本研究未采集生理指标,未来可采用心率、皮肤电等生理指标辅助自我报告结果,获得更准确和客观的结果。最后,本研究仅从外显行为方面对音乐情绪冲突效应及优势通道进行了描述和探讨,未来可结合 EEG、ERP 和 fMRI 脑成像等技术探讨其内在神经机制。

### 5 结论

研究通过比较不同音乐经验的两组被试在一致型和不一致型两种视听双通道情境下的情绪评定速度、准确性和强度,探讨了基于视听双通道的音乐情绪冲突效应、音乐情绪冲突情境下的优势加工通道及音乐经验对结果的影响。结果发现,通道间情绪信息不一致阻碍了视听双通道下的音乐情绪加工;当通道间情绪信息不一致时,个体以听觉通道的情绪线索为主导进行音乐情绪加工;丰富的音乐经验降低了情绪冲突效应对音乐组被试的干扰;音乐表演中各通道情绪信息需协调一致才能保证最佳的情

绪传达效果。

### 参考文献

- 蔡岳建,潘孝富,庄钟春晓.(2007).音乐的速度与调式对大学生情绪影响的实证研究.《心理科学》,30(1),196-198.
- 侯建成,董奇.(2010).音乐认知能力的发展及其大脑可塑性研究.《星海音乐学院学报》,(3),79-84.
- 黄希庭.(2007).《心理学导论(第2版)》(普通高等教育十一五国家级规划教材).人民教育出版社.
- 李冬冬,程真波,戴瑞娜,汪芬,黄宇霞.(2012).情绪音乐库的初步编制与评定.《中国心理卫生杂志》,26(7),552-556.
- 李萍,张明明,李帅霞,张火垠,罗文波.(2019).面孔表情和声音情绪信息整合加工的脑机制.《心理科学进展》,27(7),1205-1214.
- 文小辉,刘强,孙弘进,张庆林,尹秦清,郝明洁,牟海蓉.(2009).多感官线索整合的理论模型.《心理科学进展》,17(4),59-66.
- 张亮,孙向红,张侃.(2009).情绪信息的多通道整合.《心理科学进展》,17(6),1133-1138.
- 周欣,郑莹灿,郑茂平.(2020).基于视听双通道的音乐情绪加工机制及情绪类型和音乐经验的影响.《心理科学》,43(1),39-45.
- Ammirante, P., & Thompson, W. F. (2012). Continuation tapping to triggered melodies: Motor resonance effects of melodic motion. *Experimental Brain Research*, 216(1), 51-60.
- Chapados, C., & Levitin, D. J. (2008). Cross-modal interactions in the experience of musical performances: Physiological correlates. *Cognition*, 108(3), 639-651.
- Chen, X. H., Pan, Z. H., Wang, P., Yang, X. H., Liu, P., You, X. Q., & Yuan, J. J. (2016). The integration of facial and vocal cues during emotional change perception: EEG markers. *Social Cognitive and Affective Neuroscience*, 11(7), 1152-1161.
- Davidson, J. W. (2012). Bodily movement and facial actions in expressive musical performance by solo and duo instrumentalists: Two distinctive case studies. *Psychology of Music*, 40(5), 595-633.
- de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition & Emotion*, 14(14), 289-311.
- Doi, H., & Shinohara, K. (2015). Unconscious presentation of fearful face modulates electrophysiological responses to emotional prosody. *Cerebral Cortex*, 25(3), 817-832.
- Eerola, T., Peltola, H. R., & Vuoskoski, J. K. (2015). Attitudes toward sad music are related to both preferential and contextual strategies. *Psychomusicology Music Mind & Brain*, 25(2), 116-123.
- Egner, T., & Hirsch, J. (2005). The neural correlates and functional integration of cognitive control in a stroop task. *NeuroImage*, 24(2), 539-547.
- Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R., & Hirsch, J. (2006). Resolving emotional conflict: A role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron*, 51(6), 871-882.
- Haas, B. W., Omura, K., Constable, R. T., & Canli, T. (2006). Interference produced by emotional conflict associated with anterior cingulate activation. *Cognitive Affective & Behavioral Neuroscience*, 6(2), 152-156.
- Jessen, S., & Kotz, S. A. (2011). The temporal dynamics of processing emotions from vocal, facial, and bodily expressions. *NeuroImage*, 58(2), 665-674.
- Juslin, P. N., & Sloboda, J. A. (2010). The past, present, and future of music and emotion research. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 933-955). New York: Oxford University Press.
- Kantor-Martynuska, Bigand, J., & Emmanuel. (2013). Individual differences in granularity of the affective responses to music. *Polish Psychological Bulletin*, 44(4), 399-408.
- Karin, P., Frances, C., Carol, S., & Pollick, F. E. (2011). The music of your emotions: Neural substrates involved in detection of emotional correspondence between auditory and visual music actions. *Plos One*, 6(4), e19165.
- Klasen, M., Kreifelts, B., Chen, Y. H., Seubert, J., & Mathiak, K. (2014). Neural processing of emotion in multimodal settings. *Frontiers in Human Neuroscience*, 8(8), 822.
- Krahé, C., Hahn, U., & Whitney, K. (2015). Is seeing (musical) believing? the eye versus the ear in emotional responses to music. *Psychology of Music*, 43(1), 140-148.
- Kret, M. E., Roelofs, K., Stekelenburg, J. J., & de Gelder, B. (2013). Emotional signals from faces, bodies and scenes influence observers' face expressions, fixations and pupil size. *Frontiers in Human Neuroscience*, 7(810), 810.
- Lange, E. B., Fünderrich, J., & Grimm, H. (2022). Multisensory integration of musical emotion perception in singing. *Psychological Research*.
- Law, L. N. C., & Zentner, M. (2012). Assessing musical abilities objectively: Construction and validation of the profile of music perception skills. *Plos One*, 7(12), e52508.

- Livingstone, S. R. , & Russo, F. A. (2018). The Ryerson audio – visual database of emotional speech and song (RAVDESS) : A dynamic multimodal set of facial and vocal expressions in North American English. *PLoS ONE*, 13(5), e0196391.
- Livingstone, S. R. , Thompson, W. F. , Wanderley, M. M. , & Palmer, C. (2015). Common cues to emotion in the dynamic facial expressions of speech and song. *The Quarterly Journal of Experimental Psychology*, 68(5), 952 – 970.
- Liu, P. , Rigoulot, S. , & Pell, M. D. (2015). Culture modulates the brain response to human expressions of emotion: Electrophysiological evidence. *Neuropsychologia*, 67, 1 – 13.
- Lundqvist, L. O. , Carlsson, F. , Hilmersson, P. , & Juslin, P. N. (2009). Emotional responses to music: Experience, expression, and physiology. *Psychology of Music*, 37(1), 61 – 90.
- Misselhorn, J. , Frieze, U. , & Engel, A. K. (2019). Frontal and parietal alpha oscillations reflect attentional modulation of cross – modal matching. *Scientific Reports*, 9(1), 5030.
- Mitchell, H. F. , & Macdonald, R. A. R. (2012). Listeners as spectators? audio – visual integration improves music performer identification. *Psychology of Music*, 42(1), 112 – 127.
- Mondloch, C. J. (2012). Sad or fearful? the influence of body posture on adults' and children's perception of facial displays of emotion. *Journal of Experimental Child Psychology*, 111(2), 180 – 196.
- Müller, V. I. , Habel, U. , Derntl, B. , Schneider, F. , & Zilles, K. (2011). Incongruence effects in crossmodal emotion integration. *NeuroImage*, 54, 2257 – 2266.
- Paulmann, S. , & Pell, M. D. (2011). Is there an advantage for recognizing multi – modal emotional stimuli? *Motivation and Emotion*, 35(2), 192 – 201.
- Paulmann, S. , Jessen, S. , & Kotz, S. A. (2009). Investigating the multimodal nature of human communication. *Journal of Psychophysiology*, 23(2), 63 – 76.
- Petrini, K. , Mcaleer, P. , & Pollick, F. (2010). Audiovisual integration of emotional signals from music improvisation does not depend on temporal correspondence. *Brain Research*, 1323(8), 139 – 148.
- Piwek, L. , Pollick, F. , & Petrini, K. (2015). Audiovisual integration of emotional signals from others' social interactions. *Frontiers in Psychology*, 9, 611.
- Platz, F. , & Kopiez, R. (2012). When the eye listens: A meta – analysis of how audio – visual presentation enhances the appreciation of music performance. *Music Perception*, 30, 71 – 83.
- Proverbio, A. M. , Massetti, G. , Rizzi, E. , & Zani, A. (2016). Skilled musicians are not subject to the McGurk effect. *Scientific Reports*, 6, 30423.
- Rigoulot, S. , & Pell, M. D. (2012). Seeing emotion with your ears: Emotional prosody implicitly guides visual attention to faces. *Plos One*, 7(1), e30740.
- Rigoulot, S. , & Pell, M. D. (2014). Emotion in the voice influences the way we scan emotional faces. *Speech Communication*, 65(7), 36 – 49.
- Schubert, E. (2007). Locus of emotion: The effect of task order and age on emotion perceived and emotion felt in response to music. *Journal of Music Therapy*, 44(4), 344 – 368.
- Shoda, H. , & Adachi, M. (2016) Expressivity, affective nuance, and presentation modality in a performer – to – audience communication. *Psychomusicology: Music, Mind, and Brain*, 26(2), 167 – 178.
- Stock, J. V. D. , Peretz, I. , Grèzes, J. , & Gelder, B. D. (2009). Instrumental music influences recognition of emotional body language. *Brain Topography*, 21(3), 216 – 220.
- Strelhnikov, K. , Foxton, J. , Marx, M. , & Barone, P. (2015). Brain prediction of auditory emphasis by facial expressions during audiovisual continuous speech. *Brain Topography*, 28(3), 1 – 12.
- Thompson, W. F. , Russo, F. A. , & Livingstone, S. R. (2010). Facial expressions of singers influence perceived pitch relations. *Psychonomic Bulletin & Review*, 17(3), 317 – 322.
- Vuoskoski, J. K. , Gatti, E. , Spence, C. , & Clarke, E. F. (2016). Do visual cues intensify the emotional responses evoked by musical performance? a psychophysiological investigation. *Psychomusicology Music Mind & Brain*, 26(2), 179 – 188.
- Vuoskoski, J. K. , Thompson, M. R. , Clarke, E. F. , & Spence, C. (2014). Crossmodal interactions in the perception of expressivity in musical performance. *Attention, Perception, & Psychophysics*, 76(2), 591 – 604.
- Weijkamp, J. , & Sadakata, M. (2017). Attention to affective audio – visual information: Comparison between musicians and non – musicians. *Psychology of Music*, 45(2), 204 – 215.
- Wöllner, C. (2012). Is empathy related to the perception of emotional expression in music? a multimodal time – series analysis. *Psychology of Aesthetics Creativity & the Arts*, 6(3), 214 – 223.
- Zinchenko, A. , Kanske, P. , Obermeier, C. , Schröger, E. , & Kotz, S. A. (2015). Emotion and goal – directed behavior: ERP evidence on cognitive and emotional conflict. *Social Cognitive and Affective Neuroscience*, 10(11), 1577 – 1587.

Zinchenko, A. , &Kanske, P. , Obermeier, C. , Schröger, E. , Villringer, A. , & Kotz, S. A. (2018). Modulation of cognitive and emotional control in age – related mild – to – moderate hearing loss. *Front. Neurol*, 9, 783.

Zinchenko, A. , Kotz, S. A. , Schröger, E. , & Kanske, P. (2020). Moving towards dynamics: Emotional modulation of cognitive and emotional control. *International Journal of Psychophysiology*, 147(2020), 193 – 220.

## Emotional Conflict Effect and Dominant Modality Based on Audio – Visual Modality

Zhou Xin<sup>1,2</sup>, Zheng Yincan<sup>3</sup>, Xiao Zilun<sup>1</sup>, Zheng Maoping<sup>4</sup>

(1. Faculty of Psychology, Southwest University, Chongqing 400715;

2. School of Mental Health, North Sichuan Medical College, Nanchong 637100;

3. Developmental Psychology for Armyman, Department of Medical Psychology, Army Medical University, Chongqing 400000;

4. School of Music, Southwest University, Chongqing 400715)

**Abstract:** The aim of the current study was to investigate emotional conflict effect and dominant sensory and the influence of musical training experience based on audio – visual modality. Participants were divided into two groups, one comprising musician who had been musically trained for more than 5 years, and the other non – musicians with no musical expertise. The original musical performance video was used as congruous audio – visual(AV) material. The video which integrated the sound and picture from different segments and represents different emotional types was used as incongruous AV material. Participants were instructed to identify the type and the intensity of emotion recognition, as well as emotion felt. The reaction time(RT), accuracy(ACC), and intensity of every assessment were recorded and analyzed. Results: 1) congruous condition has higher intensity and accuracy in than incongruous condition. 2) There was a higher accuracy when answer with auditory emotion type than with visual emotion type. 3) musician participants had higher accuracy when answer with auditory emotion type, while non – musician participants showed higher accuracy when answer with visual emotion type. Which illustrated that, the incongruity of emotional information between sensory modalities hindered the processing of musical emotion in AV modality. Auditory was the dominant modality of musical emotion processing under the emotional conflict condition. Musical expertise reduced the interference of emotional conflict effect on musician participants.

**Key words:** musical emotions; audio – visual; emotional conflict effect; auditory dominance