



跨通道学习及其认知神经机制*

孙洵伟^{1,2)} 孙莹^{1,2)} 付秋芳^{1,2)**}

(1) 中国科学院心理研究所, 脑与认知科学国家重点实验室, 北京 100101; (2) 中国科学院大学心理学系, 北京 100049)

摘要 跨通道学习是指涉及从多个通道获取信息, 对这些多通道的信息进行整合并加以利用的学习。多通道信息整合是跨通道学习的重要基础。尽管跨通道学习条件更接近人类学习的真实环境, 但是目前多数研究依然采用单通道刺激, 跨通道学习相关研究的结果还显得有些凌乱、不够系统。为了较好地概括跨通道学习的特点和机制, 本文首先介绍了多通道信息整合的产生与影响因素, 以及初级皮层具有通道非特异性的实验和理论研究, 之后, 系统梳理了跨通道学习的意识性、表征类型和迁移效应的相关研究, 整理了采用神经元记录、ERP 和 fMRI 等技术探讨跨通道学习神经机制的研究进展。最后, 我们对目前跨通道学习的研究成果进行了总结, 并对这些研究成果的潜在应用以及这一领域未来的研究方向进行了展望。

关键词 跨通道学习, 多通道整合, 神经机制, 认知机制

中图分类号 B845, R338

DOI: 10.16476/j.pibb.2018.0321

跨通道学习是指涉及从多个通道获取信息, 对多通道的信息进行整合并加以利用的学习。在心理学研究中, “通道(modality)”指感觉通道或刺激通道, 如视觉、听觉、触觉和嗅觉等^[1-2]。多个通道往往能比单个通道提供更丰富的信息, 让人类更快速、准确地获得信息并做出反应。例如, 研究表明, 相比于单通道条件, 个体在跨通道条件下往往能更好地识别^[3]、探测^[4]和分类^[5], 有更低的探测阈限和更短的反应时^[6]。并且, 这种跨通道优势在单通道信息冗余的情况下依然存在^[7-8]。

不同感觉通道提供了不同维度的信息, 人类在处理这些信息时, 并不是将其特征简单地相加, 而是整合产生统一的体验。这种不同通道的信息相互影响, 形成新的统一表征的过程, 就被称作多通道信息整合(multisensory information integration)^[9-10]。多通道信息整合强调多通道信息在知觉层面的整合, 而跨通道学习需要对这些信息进行整合并加以利用。因而, 多通道信息整合是跨通道学习的基础, 二者紧密相关但又有一定的区别。此外, 多通道学习和跨通道学习的信息都来自多个通道, 但是, 多通道学习强调学习内容包含多个通道的信息, 不同通道的信息可以发生整合也可以相互独立, 而跨通道学习则强调需结合不同通道的信息才

能完成的学习, 不同通道的信息必须发生整合。加工不同感觉通道信息的能力固然是人所共有的, 但即使是处理同样的多通道信息, 个体之间也存在行为和脑活动的差异^[11], 广泛地表现在言语知觉^[12]、时间数字知觉^[13]、空间知觉^[14-15]等过程中。不过, 对同一个体而言, 其捆绑多感觉通道信息的倾向性具有跨时间的稳定性^[16]。

相比于单通道学习的研究, 跨通道学习的研究更接近人类学习的真实环境。但目前, 多数研究依然采用单通道刺激^[17]。虽然近些年跨通道的研究逐年增长, 但相关研究结果还显得有些凌乱、不够系统。为了较好地概括和阐释跨通道学习的特点和机制, 本文将以视听通道为主, 综述多通道信息整合、跨通道学习的认知加工特点与神经机制, 并对这一领域的研究问题进行总结和展望。

* 中国国家自然科学基金和德国研究基金(NSFC 61621136008/DFG TRR-169)和国家自然科学基金(61632004)资助项目。

** 通讯联系人。

Tel: 010-64845395, E-mail: fuqf@psych.ac.cn

收稿日期: 2019-04-15, 接受日期: 2019-05-06

1 多通道信息整合

多通道信息整合有诸多经典效应,例如著名的“麦格克效应(McGurk effect)”表明,当向视觉通道呈现“ba”的口型并向听觉通道呈现“ga”的声音时,听觉会受到视觉信息的影响,人们实际上听到的是第三种声音“da”^[18]。“闪光错觉(sound-induced flash illusion)”的研究则发现,当单次的视觉刺激(黑色圆圈)闪烁伴随多次听觉信号(“哔哔”声)时,人们可能会错误地将单次闪烁知觉为多次闪烁^[19]。“橡胶手错觉(rubber hand illusion)”则是一种视触错觉,它是指将被试的手隐藏在视野外,同时敲击位于被试视野内的一只橡胶手时,会引发橡胶手属于自己的错觉现象^[20]。

那么,多通道的信息进行整合有哪些条件呢?研究发现,空间、时间和相对有效性是影响多通道信息整合的关键因素^[21-22]。其一,空间位置原则,指当不同通道的信息来自于同一事件或者位置相近的多个事件时,多通道信息整合产生易化效应^[22-23],反之则引起抑制效应^[24]。其二,时间同步性效应^[25],指当不同通道信息呈现的时间趋于一致或相近时,有助于人们对于视觉^[26]、触觉^[27]等信息的探测,但当时间间隔超过一定限度(一般在300 ms左右)时,则很难发生多通道信息的整合效应^[28]。其三,多通道信息整合存在反比效应^[21, 29],是指对于来自不同感觉通道的信息,当单通道信息强度较弱时,多通道信息整合的效果更强,相反,当单通道信息强度较强时,多通道信息整合的效果变差或者消失^[30]。更有趣的是,不同通道信息的基本特征还具有跨通道的对应性(cross-modal correspondence)^[31]。例如,人们往往认为高亮度的、面积较小的图形与高频率的声音属于一类,而低亮度的、面积较大的图形与低频率的声音属于一类^[32]。此外,个体认知风格^[33]、先验知识^[34]和感官预期^[35]等也是影响多通道信息整合的因素。

传统观点认为,不同通道的信息首先经过单通道的初级皮层处理,随后在多通道的高级脑区进行加工^[36]。这种观点源于早期研究的结果。例如,某一通道初级皮层的损伤只会导致相应通道的功能受损^[37],某一通道的刺激也只会激活相应通道的脑区^[38]。但近些年研究则表明,即使是传统观点中被严格界定为通道特异的皮层区域也可能处理多通

道的信息^[36, 39]。一方面,有研究发现,某一通道信息相应的初级皮层的激活,会受到同时呈现的其他通道信息的影响^[40-41]。例如,对闪光错觉脑成像的研究发现,V1在加工视觉刺激时也会因同时呈现的听觉信息而表现出活动增强^[42],类似的,早期听觉皮层的激活会因同时呈现的触觉刺激而增强^[43-44]。另一方面,初级感觉皮层在没有相应通道的刺激呈现时,也会被包含有相应通道信息的其他通道刺激所激活。例如,视皮层会被包含空间和形状信息的触觉刺激激活^[45-46],类似的,听觉皮层会被包含声音信息的无声视觉刺激(如无声的乐器演奏视频)所激活^[47-48]。此外,采用脑磁图(magnetoencephalography, MEG)和功能性磁共振成像(functional magnetic resonance imaging, fMRI)技术研究发现,简单的视觉和听觉刺激都能够在几十毫秒内同时激活视、听皮层。但相比于优先加工的信息,非优先加工的信息诱发激活的时间要晚^[49]。Liang等^[50]采用多变量模式分析(multivariable analysis)方法进一步表明,初级感觉皮层在本质上都是多通道的,但不同感觉通道的刺激在某一初级皮层上所激活的空间模式并不相同。

初级感觉皮层的“多通道”特性,可能源于皮层神经元的先天特性,也可能源于后天经验的影响。很多研究发现,大脑的初级感觉皮层能够基于后天经验,进行结构与功能的重组^[51-52],从而具有“跨通道可塑性”(cross-modal plasticity)。例如,采用功能性近红外技术(functional near-infrared spectroscopy, fNIRS)的研究发现,4个月左右婴儿的颞叶和枕叶存在功能性耦合,且对于同时出现的视听刺激,颞、枕叶间功能的耦合度可以预测随后的听觉刺激能否激活枕叶,说明同时出现的多感觉经验可能对皮层有塑造作用^[53]。与后天的皮层结构与功能重组不同,“超通道(supramodal)”的观点认为^[54-55],有些视觉皮层区域的发展独立于视觉经验,它们处理特定的信息内容,与特定刺激的通过何种感觉方式传递无关。例如,对明眼人的研究发现,视觉皮层的外侧枕区(lateral occipital area)参与对视觉刺激的物体形状和特征的加工,但对先天性盲人的研究发现,当被试触摸物体(如瓶子)时,该脑区也被激活^[56]。因此,在研究初级皮层时,或许应更多地关注给定的知觉信息或任务,而非仅仅考虑刺激是通过什么通道传入大脑的^[55, 57-58]。

2 跨通道学习的认知机制

2.1 跨通道学习与意识

跨通道学习的一个重要研究是, 意识对于跨通道的加工与学习是否是必要的? 这其中包括两个问题: a. 整合不同通道的信息时, 是否需要意识的参与? 是否只有阈上刺激的多通道信息才能进行整合? b. 在学习跨通道刺激之间的关系或者结构时, 是否需要意识的参与?

关于第一个问题, 早在笛卡尔 (Descartes) 和威廉·詹姆斯 (William James) 等人的论述中, 就有意识与信息整合二者紧密结合的观点^[59-60]. 近年来, Dehaene 等^[61] 的全局神经工作空间理论 (global neural workspace theory) 认为, 远程连接与反馈连接使得不同脑区各类信息的整合成为了可能, 而无意识加工是“概括压缩的”, 不涉及不同脑区间信息的交换. Tononi 等^[62-63] 提出的整合信息理论 (integrated information theory) 认为, 意识就是一个系统所整合的信息. 有实验研究发现, 在清醒状态下用经颅磁刺激 (transcranial magnetic stimulation, TMS) 对右侧前运动皮层进行刺激, 可以观察到脑电波从刺激位点向周围扩散, 而在非快速眼动睡眠状态进行相同操作, 却不会出现上述现象^[64]. 尽管如此, 近年来许多研究表明, 对视觉阈下刺激的加工会受到听觉^[65-66]、触觉^[67] 和本体感觉^[68] 等通道阈上刺激的影响. 但由于上述研究中其整合过程依然包含了阈上刺激, 因而仍不能完全排除意识在多通道信息整合中的作用. Faivre 等^[10] 采用反应启动 (response priming) 范式^[69] 进一步发现, 对阈上视听刺激的判断受阈下视听刺激关系的影响, 说明视听刺激均为阈下水平时也能进行整合. 但由于在实验前被试必须有意识地熟悉实验中的刺激后才能得到上述结果, 因此该实验也未能完全排除意识经验在多通道信息整合中的作用^[70-71]. 另外, 睡眠相关的研究发现, 即使在“完全无意识”的深度睡眠状态下, 被试依然能够获得气味与声音之间的关系. 但是问题在于, 当下没有方法能严格证明睡眠状态下刺激对被试而言是完全无意识的^[72]. 目前, 随着无意识信息整合研究的发展, 研究者开始着手探讨这一过程是发生在感觉水平还是语义水平上, 以及其时间窗口长度等更具体问题^[73].

关于第二个问题, 有研究表明, 在学习跨通道刺激之间的关系或者结构时, 意识不是必需的. 无

意识地获得环境中复杂知识的过程被称作内隐学习^[74]. 传统上多数内隐学习的研究都只采用单通道 (主要是视觉通道) 的刺激. 例如, Ashby 等^[75] 提出的类别学习中言语和内隐系统的竞争模型 (COVIS) 就是基于视觉刺激的研究来提出和进行验证的. 该理论认为, 基于信息整合的类别学习是一种内隐学习. Maddox 等^[76] 在基于信息整合的类别学习中, 采用不同密度的光栅与不同频率的声音刺激. 结果发现, 被试确实能够采用内隐学习的策略来习得跨通道的类别知识^[76-77]. 内隐序列学习的研究也发现, 个体在视、听条件下的学习成绩没有差异, 并且能够学到多通道的序列规则^[78-79]. 有研究发现, 6个月大的婴儿能够通过被动观看的方式, 内隐地建立跨通道的联结, 对出现频率不同的联结表现出脑活动的差异^[80]. 这表明人类在婴儿阶段就已经能够内隐地习得多通道的知识.

2.2 跨通道学习的表征类型

表征是信息在大脑中的呈现或存储方式可以是视觉的, 如一只小狗的形象; 也可以是听觉的, 如狗吠的声音; 或者是其他感觉形式的. 在跨通道学习中, 学习者产生的表征是具体的、通道特异性的 (modality-specific)? 还是抽象的、通道一般性的 (modality-general)? 还是两者的结合? 有关这一问题一直存在争议. 有研究支持通道一般性的观点^[81-82]. 例如, 同时给被试呈现人造字符串和发音者的面部视频流, 当二者不对应时会出现经典的“麦格克效应”^[83], 说明统计学习中不同通道是能够相互作用的. 但也有结果支持通道特异性的观点. 例如, 当研究者单独或同时给被试呈现音频流与视频流, 发现被试能够分别提取出其中的统计模式, 且学习成绩没有差异^[84], 这说明即使同时呈现多通道的信息, 它们也不会相互干扰. 再如, 有研究采用人工语法范式, 同时向学习者呈现视听刺激, 但在测试时只对一个通道进行测试, 发现学习者在学习规则时保留了学习通道的信息. 这说明可能存在不同通道平行加工的学习机制^[85-86]. 这一研究结果的产生, 可能是由于学习时间较短, 无法形成抽象的一般性表征而造成的. 但有研究采用了同样的研究范式, 发现在充分学习的条件下, 较长的学习时间只会增加刺激特异性的知识, 而不会形成抽象的表征知识^[87].

实际上, 包括实验刺激、外部环境甚至个体在内的多种因素, 都可能影响跨通道学习所形成的表征类型. 有研究发现, 被试是否能够提取跨通道信

息的统计模式取决于两个通道的信息流的同步性^[88],当视听三联体刺激的边界未对齐时,学习成绩会显著下降.如果学习中的表征完全是通道特异的,那么就不会出现这种学习差异.还有研究表明,有足够的时间进行巩固(consolidation)也是形成抽象表征的重要因素.在经过一段时间的休息(如睡眠)后,学到的知识才能在一定程度上泛化至其他通道^[89-90].这可能是由于睡眠促进了系统层面的记忆的重新组织^[91-93].

2.3 跨通道学习中的迁移效应

个体在学习时往往会表现出迁移效应,即在某一任务的学习后,获得的知识(或其中很大一部分)不与任何特定的知觉特征相联系,由此在新的、未经训练的任务上也表现出了学习效应^[94].在跨通道学习中,如果两个通道的任务有着类似的认知加工过程,或产生了抽象的、通道一般性的表征,那么也应该能表现出通道间的迁移效应.Bratzke等^[95]采用时长辨别任务范式,将被试分为有训练组与无训练组,分别进行前后测.结果发现,相比于无训练组,训练组有显著的成绩提升,且即使采用听觉刺激训练,被试在采用视觉刺激的后测任务中成绩也显著提升,出现了从听觉到视觉的迁移效应.此外,研究结果表明,如果训练与后测的时间间隔越长,成绩的提升则越明显^[96-97].迁移除了发生在视听通道间,也会发生在触觉和其他感觉通道间^[97-98].但也有研究发现,在时长分辨任务中,即使经过大量的学习,在经过刀切法(jackknife)进行数据处理后,没有发现从听觉到视觉的迁移效应^[99],或是即使没有采用刀切法,也只表现出听觉通道内不同维度间的迁移^[100].

上述对迁移效应研究结果的不一致,可以通过迁移效应的非对称性(asymmetry)进行解释^[95].有研究者认为,视觉在处理空间信息时起主导作用,听觉则主导时间信息的加工^[101-102].由此,某种信息的迁移效应只出现在从主导通道至非主导通道的迁移中,反之则无法迁移,即跨通道学习中迁移效应是单向、非对称的.具体而言,在空间辨别任务中,学习只能从视觉迁移到听觉,而在时间辨别任务中,学习则只能从听觉迁移到视觉^[103].对比研究证实,在时间顺序判断任务中,视听条件下的学习能够迁移到视觉条件,但不能迁移至听觉条件^[104];也有研究发现,听觉任务训练能显著提升被试视觉节拍感知的成绩,但视觉训练则不能^[105].

3 跨通道学习的脑机制

传统观点认为,只有颞上沟(superior temporal sulcus, STS)、上丘(superior colliculus, SC)等少数脑区和脑结构接收多通道的信息,参与跨通道的加工与学习^[36].但近些年来,利用神经元记录、脑电图和脑磁图、功能磁共振成像等手段的研究表明,包括前额皮层(prefrontal cortex)和后顶叶皮层(posterior parietal cortex)在内的许多脑区和皮层下结构也参与跨通道的加工和学习.下面,我们将根据研究所采用的技术手段,梳理有关跨通道学习的脑机制的研究进展.

3.1 采用神经元记录方法进行的相关研究

神经元记录的方法能够直接测量神经元的生理活动,是早期跨通道研究的核心手段.早期对猕猴的神经电生理学研究表明,颞上沟的前部和后部均有一定数量的神经元对多通道刺激进行反应^[106-107].近年来,有研究采用单神经元测量的方法进一步发现^[108],颞上沟有23%的神经元对视觉生物运动刺激的反应会显著受到同时出现的听觉刺激的调节,证明颞上沟有整合多通道信息的功能.

后顶叶皮层在多通道信息的存储与反馈等过程中起重要作用^[109].首先,后顶叶皮层是多通道信息的缓存器与调节器.在传统的层级感觉加工模型中,后顶叶皮层被看作是感觉信息高级联合区域,仅单向接收和整合初级皮层的的感觉表征^[110].但近期的研究发现,当不同通道的感觉信息存储在记忆中时,后顶叶皮层起到感觉信息缓存的重要作用^[111-112].此外,研究还发现,后顶叶皮层也能够通过反馈来调节不同感觉系统,尤其是视觉与触觉系统间的相互作用^[113-115].其次,不同通道的信息往往具有不同的参考系(例如触觉信息以人体为坐标中心,视听信息有时以外部世界为坐标中心).当个体整合不同通道的信息尤其是对这些信息做出反应时,需要将它们重新投射到一个统一的参考系中,后顶叶皮层在其中起十分重要的作用^[6, 116].

上丘的外层结构接收视觉信息,内层结构接收视、听、触信息,这些信息可来自于皮层和皮层下结构,其整合在很大程度上取决于来自联合皮层前外侧沟(anterior ectosylvian sulcus)和薛氏外沟(lateral suprasylvian sulcus)的信息输入^[117].上丘主要对不同通道的感觉信息进行感觉运动转换,以满足个体运动的需求^[4, 118].在上丘的作用下,个体能够将不同感觉通道的空间地图校正对齐,形成

一张统一的、时时更新的外部空间的地图^[119-120]. 研究者认为, 上丘所形成的外部空间与其他脑区所形成的个体和个体周围空间, 整合为完整统一的空间表征, 成为个体身体意识形成的基础^[121].

此外, 许多皮层下结构也具有处理多通道信息的特征. 例如, 杏仁核 (amygdala) 内有大量对视觉、听觉和触觉反应的神经元. 其中, 对视觉反应的神经元多集中在杏仁核前部, 对听觉反应的神经元则多集中在后部^[122]. 再如, 有研究发现大鼠的部分丘脑 (thalami) 的听觉神经元受视觉信息的调控, 当视觉线索与听觉线索一致时其活动增强^[123]. 而屏状核、脑岛和脑干的神经元也能对多通道信息反应^[124-125].

3.2 采用脑电与脑磁图进行的相关研究

事件相关电位 (event-related potential, ERP) 与脑磁图可以提供人脑信息加工的功能性信息, 有着较高的时间分辨率. ERP 的早期成分可以反映跨通道刺激的加工与整合. 研究发现, 相比于单通道刺激, 视触刺激同时呈现后的 90~150 ms 内, 视觉 P1 波的波幅增强^[126], 在视听刺激呈现后的 130~150 ms 内, 顶枕区出现波幅更大的 P1 波^[127]. 此外, 对言语信息的研究发现, 视听刺激引发的前中央区 N1 波在 114 ms 达到峰值, Fz 点的 P2 波在 204 ms 达到峰值, 相比于单通道信息, 上述两个波形的潜伏期更短^[128]. 跨通道刺激的一致性也会表现在 ERP 成分上. 相比于先后呈现的面孔刺激和语音刺激, 二者同时呈现后 150~250 ms 诱发了更大的左侧听觉 P2 波^[129], 而相比于效价不匹配的面孔图片与声音刺激, 二者匹配条件下在前额、额中部和枕部的 P2 波的波幅显著增大, 说明在情绪判断的早期发生了视听信息整合^[130].

在对跨通道配对刺激的学习中, 研究发现, 相比于已经学习过的刺激, 当呈现新异的视听配对刺激时, N2-P3 复合波的波幅显著更大, 这一成分可能与输入信息和存储信息间的对比有关^[131], 而即使是在单次学习后, 新异的配对刺激也会在 270~310 ms 产生视觉诱发电位 (visual evoked potentials, VEP)^[132]. 此外, 研究发现, 在完成对跨通道刺激的目标判断任务时, Pz 点 P3 波的潜伏期越短, 视听触刺激的识别正确率越高^[133]. ERP 的晚期成分可以反映跨通道学习的结果. 采用视触刺激的配对学习任务发现^[134], 在刺激呈现后的 480~520 ms, 中央顶叶皮层的 P400 波和顶区的 N400 波, 以及刺激呈现后的 630~670 ms 间的晚期

后侧负慢波 (late posterior negative slow wave) 反映了个体对跨通道刺激的学习. 也有研究发现, 颞中回 (middle temporal gyrus) 的 N400 波与跨通道信息的编码与提取相关^[135-136]. 采用视触^[137]和听触^[138]跨通道延迟样例匹配任务的研究则发现, 与跨通道联结学习相关的晚期正成分 LPC-1 (late positive component) 和 LPC-2, 分别发生在第一个刺激呈现后的 330 ms 左右以及 520~600 ms.

大量脑电图 (electroencephalogram, EEG) 与脑磁图 (MEG) 的研究发现, 多通道信息的加工与大脑的神经振荡存在紧密联系^[139-140]. 例如, 对闪光错觉的研究发现, 第二次声音的出现伴随着视觉皮层 gamma 振荡的增强^[141], 对全脑的大范围功能网络研究也发现, 视听语义的加工伴随着全脑 gamma 频段的同步性提升以及 alpha 和 theta 频段的同步性降低^[142], 这些神经活动很可能对跨通道知觉起到促进作用. 但目前仅通过神经振荡这个单一指标来判断是否存在多通道的整合与加工尚不可行^[143-144].

3.3 采用功能性磁共振成像与神经调控方法进行的相关研究

功能性磁共振成像能够提供大脑的血流量和耗氧量等信息, 有较高的空间分辨率, 可以无创地探究人类不同脑区的激活水平以及脑区间的功能连接, 结合 TMS 和经颅直流电刺激 (transcranial direct current stimulation, tDCS) 等神经调控技术, 更可以为某一脑区的功能提供因果性证据.

大量研究表明, 前额皮层在跨通道信息对应性的调节和加工中起到重要作用^[145]. 具体而言, 当跨通道信息不对应时, 前额皮层的激活会比信息对应时显著增强^[146-147]. 在加工相对复杂、有变化的 3D 物体与声音时, 视听刺激呈现时间的非对应性会增强额叶、颞上沟与初级皮层的功能连接^[25]. 有研究发现, 在收集视听信息形成行为决策的过程中, 额下沟 (inferior frontal sulcus) 会根据不同感觉通道的可靠度及其与决策的相关性, 动态地对其与视听皮层的连接进行加权^[148]; 另有研究表明, 视觉和听觉注意的指向性受外侧前额叶 (lateral prefrontal cortex) 的控制^[149].

颞上皮层也是跨通道学习加工的核心脑区, 其中颞上回 (superior temporal gyrus) 负责听触加工^[150], 颞上沟在视听信息加工和学习, 尤其是言语加工中, 起重要作用^[151]. 结合神经调控的方法, 研究发现, 相比于不施加或在其他脑区施加 TMS,

在颞上沟施加单脉冲TMS,能在单通道言语任务不受影响的情况下,显著减少麦格克效应出现的次数^[152],对颞上沟施加tDCS,会降低颞上沟的激活水平,减少麦格克效应出现的可能性^[153].这些研究为颞上沟在多通道言语加工中的重要作用提供了因果性证据.此外,对颞上沟施加兴奋性tDCS刺激,可以增强个体知觉到的闪光错觉^[154],施加抑制性tDCS则出现相反的效果.脑成像与功能连接的研究也表明,闪烁错觉中视听刺激的时间一致性会增强颞上沟的激活水平^[155],并通过反馈通路作用于初级皮层^[156].这也印证了早期的研究结果^[157-158],即感觉特异性的初级皮层不仅通过前馈通路将信息传递至多感觉汇集区域,同时也受到后者的反馈影响.

此外,研究发现上丘和后顶叶皮层也参与跨通道的学习.当被试产生橡胶手错觉时,能够引起身体同侧上丘的激活,同时增加上丘与其他参与身体自我意识形成的脑区之间的功能连接^[159-160].同时,对双侧纹状体皮层缺失者的研究发现,患者能够习得听觉线索与红色视觉刺激之间的联结,但不能习得听觉线索与紫色视觉刺激的联结,由于上丘对于紫色不敏感,说明上丘在视听联结学习中起重要作用^[161].而结合神经调控与fMRI的研究证明,后顶叶皮层也参与不同通道信息加工的注意分配和视听言语加工过程.例如,研究发现后顶叶皮层活动增强时,会减少视听跨通道刺激的探测反应时^[162],也增加了麦格克效应出现的可能性^[163].

4 小结与展望

人类无时无刻不在处理多感觉信息,学习来自不同通道的知识.在过去的20年中,越来越多的研究者探究跨通道学习的现象,揭示其认知加工机制以及脑机制.研究的主要成果有以下几个方面.

首先,传统观点认为,初级皮层只处理与其通道对应的信息,不同通道信息的整合加工只发生在部分高级脑区.但近年的研究发现,单一通道的信息也能够激活其他通道的初级感觉皮层,或在伴随有其他通道信息呈现时初级皮层的激活更强.初级皮层能够加工多通道信息,并不意味着某一通道的初级皮层在跨通道加工时一定处理其他通道的信息,或者说某一通道的刺激一定会激活其他通道的初级皮层.目前的研究只能表明,初级皮层并不像早先认为的那样只处理本通道的信息.初级皮层之间也能够进行交互作用,在多通道加工中起作用.

第二,意识在跨通道信息的加工和学习中可能并不是必需的.有研究发现,在不同通道的刺激均处于阈下水平时,个体依然能够完成信息的整合,而内隐学习的相关研究也表明,在个体掌握跨通道的知识后,这些知识也不一定会完全进入意识层面.然而,现有的结果只能说明在某些跨通道学习条件下,意识的作用很小或很不明显.这一领域的研究将有助于我们了解意识是如何产生的,以及意识究竟有什么作用等问题.

第三,跨通道学习获得的知识究竟是抽象的、通道一般性的?还是具体的、通道特异的?在统计学习等领域的研究中,目前的研究结果仍存在分歧.对此可能的一种解释是,包括刺激、环境甚至个体在内的多种因素都会影响实验结果,因此应当具体情况具体分析.已有研究发现休息巩固和不同通道信息间的同步性,是影响跨通道学习获得何种表征的重要因素.对于学习获得的知识能否进行跨通道的迁移,大量研究表明,知识的迁移能够在视、听、触等通道间进行,但迁移可能是非对称的:即时间信息往往只能由听觉通道迁移到其他通道,而空间信息往往只能由视觉通道迁移到其他通道.

尽管跨通道学习吸引了越来越多的研究者的注意,但相关研究还不够系统.有些问题值得做进一步深入细致的研究.未来的研究应关注以下几个方面.

首先,对于跨通道的信息在大脑中是如何整合的这一问题,还缺乏全面的、系统的认识.通过神经元记录和脑成像等技术,研究发现颞下沟、后顶叶皮层等脑区以及上丘等一些皮层下结构在跨通道信息加工和学习中起重要作用.但目前对于跨通道学习局部的脑网络以及相关脑区之间的功能连接等脑机制仍了解不多.这是一个复杂庞大的问题,可以结合脑成像、脑电和神经调控等不同技术的优势来进行探究.同时,同一领域内不同的研究往往采用不同的实验范式、实验参数和刺激,再加上受被试群体个体差异的影响,研究间往往没有较好的可比性.因而,有必要适当控制实验条件的一致性.

其次,应进一步探究意识、注意等因素与跨通道学习的关系.对于意识,其在跨通道学习中的具体作用机制和相关脑机制仍不明确.未来的研究应更多聚焦于意识对跨通道学习的影响.而在跨通道学习中不同通道的信息是否存在注意资源的竞争,即它们是否有着独立的注意资源库,这也是亟待解

决的一个重要问题^[164]。

再次, 多通道的信息是如何被编码, 存储于记忆系统的? 这涉及到跨通道信息的表征类型、知识迁移和加工存储机制等一系列问题。在未来研究中, 应进一步加强迁移学习认知机理的研究。通过仅建立较少跨通道联结的婴幼儿为对象进行, 可以为这一问题的结果提供新的证据和启示。

最后, 大脑跨通道加工的脑机制和相关的神经可塑性的关系, 仍然需要更多的研究, 并将研究成果应用于实际。例如, 有研究发现, 跨通道学习功能受损对读写障碍、孤独症等病症有重要的影响^[165-166]; 同时也有证据表明, 训练能够提升跨通道学习的行为表现, 增强相关的脑区活动^[167-168]。如果能进一步研究并将这两类研究结合起来, 有助于特殊人群的认知训练和治疗, 并对人工智能和感觉替代装置的研发起到重要的推动作用。

参 考 文 献

- [1] Calvert G A. Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cerebral Cortex*, 2001, **11**(12): 1110 - 1123
- [2] Shams L, Seitz A R. Benefits of multisensory learning. *Trends in Cognitive Sciences*, 2008, **12**(11): 411-417
- [3] Molholm S, Ritter W, Murray M M, *et al.* Multisensory auditory - visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Cognitive Brain Research*, 2002, **14**(1): 115-128
- [4] Stein B E, Meredith M A. *The Merging of The Senses*. Cambridge: The MIT Press, 1993
- [5] Chen Y-C, Spence C. When hearing the bark helps to identify the dog: Semantically-congruent sounds modulate the identification of masked pictures. *Cognition*, 2010, **114**(3): 389-404
- [6] Bolognini N, Maravita A. Proprioceptive alignment of visual and somatosensory maps in the posterior parietal cortex. *Current Biology*, 2007, **17**(21): 1890-1895
- [7] Girard S, Pelland M, Lepore F, *et al.* Impact of the spatial congruence of redundant targets on within-modal and cross-modal integration. *Experimental Brain Research*, 2013, **224**(2): 275-285
- [8] Laurienti P J, Burdette J H, Maldjian J A, *et al.* Enhanced multisensory integration in older adults. *Neurobiology of Aging*, 2006, **27**(8): 1155-1163
- [9] Brunel L, Carvalho P F, Goldstone R L. It does belong together: cross-modal correspondences influence cross-modal integration during perceptual learning. *Frontiers in Psychology*, 2015, **6**(4): 358
- [10] Faivre N, Mudrik L, Schwartz N, *et al.* Multisensory integration in complete unawareness: evidence from audiovisual congruency priming. *Psychological Science*, 2014, **25**(11): 2006-2016
- [11] Gullick M M, Booth J R. Individual differences in crossmodal brain activity predict arcuate fasciculus connectivity in developing readers. *Journal of Cognitive Neuroscience*, 2014, **26**(7): 1331-1346
- [12] Basu Mallick D, F. Magnotti J, S. Beauchamp M. Variability and stability in the McGurk effect: contributions of participants, stimuli, time, and response type. *Psychonomic Bulletin & Review*, 2015, **22**(5): 1299-1307
- [13] Stevenson R A, Zemtsov R K, Wallace M T. Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. *Journal of Experimental Psychology: Human Perception and Performance*, 2012, **38**(6): 1517-1529
- [14] Hairston W D, Wallace M T, Vaughan J W, *et al.* Visual localization ability influences cross-modal bias. *Journal of Cognitive Neuroscience*, 2003, **15**(1): 20-29
- [15] Wozny D R, Beierholm U R, Shams L. Probability matching as a computational strategy used in perception. *PLoS Computational Biology*, 2010, **6**(8): 1-7
- [16] Odegaard B, Shams L. The brain's tendency to bind audiovisual signals is stable but not general. *Psychological Science*, 2016, **27**(4): 583-591
- [17] Van Der Stoep N, Nijboer T C W, Van Der Stigchel S, *et al.* Multisensory interactions in the depth plane in front and rear space: A review. *Neuropsychologia*, 2015, **70**(1): 335-349
- [18] McGurk H, Macdonald J. Hearing lips and seeing voices. *Nature*, 1976, **264**(5588): 746-748
- [19] Shams L, Kamitani Y, Shimojo S. What you see is what you hear. *Nature*, 2000, **408**(6418): 788
- [20] Botvinick M, Cohen J. Rubber hands 'feel' touch that eyes see. *Nature*, 1998, **391**(6669): 756
- [21] Ghose D, Barnett Z P, Wallace M T. Impact of response duration on multisensory integration. *Journal of Neurophysiology*, 2012, **108**(9): 2534-2544
- [22] Meredith M A, Stein B E. Spatial determinants of multisensory integration in cat superior colliculus neurons. *Journal of Neurophysiology*, 1996, **75**(5): 1843-1857
- [23] Spence C. Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule. *Annals of the New York Academy of Sciences*, 2013, **1296**(1): 31-49
- [24] Samad M, Shams L. Visual-somatotopic interactions in spatial perception. *Neuroreport*, 2016, **27**(3): 180-185
- [25] Laing M, Rees A, Vuong Q C. Amplitude-modulated stimuli reveal auditory-visual interactions in brain activity and brain connectivity. *Frontiers in Psychology*, 2015, **6**(10): 1440
- [26] Lippert M, Logothetis N K, Kayser C. Improvement of visual contrast detection by a simultaneous sound. *Brain Research*, 2007, **1173**(1): 102-109
- [27] Crommett L E, Perez-Bellido A, Yau J M. Auditory adaptation improves tactile frequency perception. *Journal of Neurophysiology*, 2017, **117**(3): 1352-1362
- [28] Nozaradan S, Peretz I, Mouraux A. Steady-state evoked potentials as an index of multisensory temporal binding. *NeuroImage*, 2012,

- 60(1): 21-28
- [29] Driver J, Noesselt T. Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. *Neuron*, 2008, **57**(1): 11-23
- [30] Perrault T J, Vaughan J W, Stein B E, *et al.* Neuron-specific response characteristics predict the magnitude of multisensory integration. *Journal of Neurophysiology*, 2003, **90**(6): 4022-4026
- [31] Spence C. Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, 2011, **73**(4): 971-995
- [32] Parise C V, Spence C. Audiovisual crossmodal correspondences and sound symbolism: a study using the implicit association test. *Experimental Brain Research*, 2012, **220**(3-4): 319-333
- [33] Guo L, Bao M, Guan L, *et al.* Cognitive styles differentiate crossmodal correspondences between pitch glide and visual apparent motion. *Multisensory Research*, 2017, **30**(3-5): 363-385
- [34] Diederich A, Colonius H, Kandil F I. Prior knowledge of spatiotemporal configuration facilitates crossmodal saccadic response. *Experimental Brain Research*, 2016, **234**(7): 2059-2076
- [35] Wang Q, Keller S, Spence C. Sounds spicy: Enhancing the evaluation of piquancy by means of a customised crossmodally congruent soundtrack. *Food Quality and Preference*, 2017, **58**(1): 1-9
- [36] Ghazanfar A A, Schroeder C E. Is neocortex essentially multisensory?. *Trends in Cognitive Sciences*, 2006, **10**(6): 278-285
- [37] Dewson J H, Cowey A, Weiskrantz L. Disruptions of auditory sequence discrimination by unilateral and bilateral cortical ablations of superior temporal gyrus in the monkey. *Experimental Neurology*, 1970, **28**(3): 529-548
- [38] Belliveau J W, Kennedy D N, McKinsty R C, *et al.* Functional mapping of the human visual cortex by magnetic resonance imaging. *Science*, 1991, **254**(5032): 716 - 719
- [39] Kayser C. The multisensory nature of unisensory cortices: a puzzle continued. *Neuron*, 2010, **67**(2): 178-180
- [40] Bizley J K, Nodal F R, Bajo V M, *et al.* Physiological and anatomical evidence for multisensory interactions in auditory cortex. *Cerebral Cortex*, 2007, **17**(9): 2172-2189
- [41] Kayser C, Logothetis N K, Panzeri S. Visual enhancement of the information representation in auditory cortex. *Current Biology*, 2010, **20**(1): 19-24
- [42] Watkins S, Shams L, Tanaka S, *et al.* Sound alters activity in human V1 in association with illusory visual perception. *Neuroimage*, 2006, **31**(3): 1247-1256
- [43] Butler J S, Foxe J J, Fiebelkorn I C, *et al.* Multisensory representation of frequency across audition and touch: high density electrical mapping reveals early sensory-perceptual coupling. *Journal of Neuroscience*, 2012, **32**(44): 15338-15344
- [44] Kayser C, Petkov C I, Augath M, *et al.* Integration of touch and sound in auditory cortex. *Neuron*, 2005, **48**(2): 373-384
- [45] Amedi A, Malach R, Hendler T, *et al.* Visuo-haptic object-related activation in the ventral visual pathway. *Nature Neuroscience*, 2001, **4**(3): 324-430
- [46] James T W, Humphrey G K, Gati G S, *et al.* Haptic study of three-dimensional objects activates extrastriate visual. pdf. *Neuropsychologia*, 2002, **40**(10): 1706-1714
- [47] Meyer K, Kaplan J T, Essex R, *et al.* Predicting visual stimuli on the basis of activity in auditory cortices. *Nature Neuroscience*, 2010, **13**(6): 667-668
- [48] Riedel P, Ragert P, Schelinski S, *et al.* Visual face-movement sensitive cortex is relevant for auditory-only speech recognition. *Cortex*, 2015, **68**(1): 86-99
- [49] Raj T, Ahveninen J, Lin F-H, *et al.* Onset timing of cross-sensory activations and multisensory interactions in auditory and visual sensory cortices: Onset timing of audiovisual processing for simple stimuli. *European Journal of Neuroscience*, 2010, **31**(10): 1772-1782
- [50] Liang M, Mouraux A, Hu L, *et al.* Primary sensory cortices contain distinguishable spatial patterns of activity for each sense. *Nature Communications*, 2013, **4**(1): 1979
- [51] Proulx M J, Brown D J, Pasqualotto A, *et al.* Multisensory perceptual learning and sensory substitution. *Neurosci Biobehav Rev*, 2014, **41**(8): 16-25
- [52] Collignon O, Voss P, Lassonde M, *et al.* Cross-modal plasticity for the spatial processing of sounds in visually deprived subjects. *Exp Brain Res*, 2009, **192**(3): 343-358
- [53] Werchan D M, Baumgartner H A, Lewkowicz D J, *et al.* The origins of cortical multisensory dynamics: Evidence from human infants. *Developmental Cognitive Neuroscience*, 2018, **34**(1): 75-81
- [54] Kupers R, Pietrini P, Ricciardi E, *et al.* The nature of consciousness in the visually deprived brain. *Frontiers in Psychology*, 2011, **2**(2): 19
- [55] Ricciardi E, Pietrini P. New light from the dark: what blindness can teach us about brain function. *Current Opinion in Neurology*, 2011, **24**(4): 357-363
- [56] Pietrini P, Furey M L, Ricciardi E, *et al.* Beyond sensory images-object-based representation in the human ventral pathway. *Proceedings of the National Academy of Sciences*, 2004, **101**(15): 5658-5663
- [57] Heimler B, Striem-Amit E, Amedi A. Origins of task-specific sensory-independent organization in the visual and auditory brain: neuroscience evidence, open questions and clinical implications. *Current Opinion in Neurobiology*, 2015, **35**(1): 169-177
- [58] Ricciardi E, Bonino D, Pellegrini S, *et al.* Mind the blind brain to understand the sighted one! Is there a supramodal cortical functional architecture?. *Neuroscience and Biobehavioral Reviews*, 2014, **41**(1): 64-77
- [59] Descartes R, Adam C, Tannery P. *Oeuvres de Descartes: Meditationes de Prima Philosophia*. Paris: Librairie Philosophique, 1973
- [60] James W. *The Principles of Psychology*. London: Macmillan, 2012
- [61] Dehaene S, Changeux J P. Experimental and theoretical approaches to conscious processing. *Neuron*, 2011, **70**(2): 200-227
- [62] Balduzzi D, Tononi G. Integrated information in discrete dynamical systems: motivation and theoretical framework. *PLoS*

- Computational Biology, 2008, **4**(6): e1000091
- [63] Tononi G. The integrated information theory of consciousness: an updated account. *Archives Italiennes de Biologie*, 2011, **150**(2/3): 56 – 90
- [64] Massimini M, Ferrarelli F, Huber R, *et al.* Breakdown of cortical effective connectivity during sleep. *Science*, 2005, **309**(1): 2228 – 2232
- [65] Palmer T D, Ramsey A K. The function of consciousness in multisensory integration. *Cognition*, 2012, **125**(3): 353-364
- [66] Alsius A, Munhall K G. Detection of audiovisual speech correspondences without visual awareness. *Psychological Science*, 2013, **24**(4): 423 – 431
- [67] Lunghi C, Binda P, Morrone M C. Touch disambiguates rivalrous perception at early stages of visual analysis. *Current Biology*, 2010, **20**(1): 143 – 144
- [68] Salomon R, Lim M, Herbelin B, *et al.* Posing for awareness: proprioception modulates access to visual consciousness in a continuous flash suppression task. *Journal of Vision*, 2013, **13**(7): 2
- [69] Kiesel A, Kunde W, Hoffmann J. Mechanisms of subliminal response priming. *Advances in Cognitive Psychology*, 2007, **3**(1-2): 307-315
- [70] Mudrik L, Faivre N, Koch C. Information integration without awareness. *Trends in Cognitive Sciences*, 2014, **18**(9): 488-496
- [71] Noel J P, Wallace M, Blake R. Cognitive neuroscience: integration of sight and sound outside of awareness?. *Current Biology*, 2015, **25**(4): R157-159
- [72] Arzi A, Shedlesky L, Ben-Shaul M, *et al.* Humans can learn new information during sleep. *Nature Neuroscience*, 2012, **15**(10): 1460-1465
- [73] 刘睿, 王莉, 蒋毅. 意识与多感觉信息整合的最新研究进展. *科学通报*, 2016, **61**(1): 2-11
Liu R, Wang L, Jiang Y. *Chinese Science Bulletin*, 2016, **61**(1): 2-11
- [74] Reber A S. Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, 1989, **118**(3): 219-235
- [75] Ashby F G, Alfonso-Reese L A, Waldron E M. A Neuropsychological theory of multiple systems in category learning. *Psychological Review*, 1998, **105**(3): 442-481
- [76] Maddox W T, Ing A D, Lauritzen J S. Stimulus modality interacts with category structure in perceptual category learning. *Perceptions & Psychophysics*, 2006, **68**(7): 1176-1190
- [77] Smith J D, Johnston J J, Musgrave R D, *et al.* Cross-modal information integration in category learning. *Attention, Perception, & Psychophysics*, 2014, **76**(5): 1473-1484
- [78] Kemeny F, Meier B. Multimodal sequence learning. *Acta Psychologica*, 2016, **164**(1): 27-33
- [79] Milne A E, Wilson B, Christiansen M H. Structured sequence learning across sensory modalities in humans and nonhuman primates. *Current Opinion in Behavioral Sciences*, 2018, **21**(1): 39-48
- [80] Rohlf S, Habets B, Von Frieling M, *et al.* Infants are superior in implicit crossmodal learning and use other learning mechanisms than adults. *Elife*, 2017, **6**(1): e28166
- [81] Frost R, Armstrong B C, Siegelman N, *et al.* Domain generality versus modality specificity: the paradox of statistical learning. *Trends in Cognitive Sciences*, 2015, **19**(3): 117-125
- [82] Kirkham N Z, Slemmer J A, Johnson S P. Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, 2002, **83**(2): B35-B42
- [83] Mitchel A D, Christiansen M H, Weiss D J. Multimodal integration in statistical learning: evidence from the McGurk illusion. *Frontiers in Psychology*, 2014, **5**(5): 407
- [84] Seitz A R, Kim R, Van Wassenhove V, *et al.* Simultaneous and independent acquisition of multisensory and unisensory associations. *Perception*, 2007, **36**(10): 1445-1453
- [85] Conway C M, Christiansen M H. Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 2005, **31**(1): 24-39
- [86] Conway C M, Christiansen M H. Seeing and hearing in space and time: Effects of modality and presentation rate on implicit statistical learning. *European Journal of Cognitive Psychology*, 2009, **21**(4): 561-580
- [87] Johansson T. Strengthening the case for stimulus-specificity in artificial grammar learning: no evidence for abstract representations with extended exposure. *Experimental Psychology*, 2009, **56**(3): 188-197
- [88] Mitchel A D, Weiss D J. Learning across senses: cross-modal effects in multisensory statistical learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 2011, **37**(5): 1081-1091
- [89] Ashton J E, Jefferies E, Gaskell M G. A role for consolidation in cross-modal category learning. *Neuropsychologia*, 2018, **108**(1): 50-60
- [90] Durrant S J, Cairney S A, Lewis P A. Cross-modal transfer of statistical information benefits from sleep. *Cortex*, 2016, **78**(1): 85-99
- [91] Kumaran D, Hassabis D, McClelland J L. What learning systems do intelligent agents need? Complementary learning systems theory updated. *Trends in Cognitive Sciences*, 2016, **20**(7): 512-534
- [92] Van Kesteren M T, Rijpkema M, Ruiter D J, *et al.* Consolidation differentially modulates schema effects on memory for items and associations. *PloS One*, 2013, **8**(2): e56155
- [93] Li X, Zhao X, Shi W, *et al.* Lack of Cross-modal effects in dual-modality implicit statistical learning. *Frontiers in Psychology*, 2018, **9**(2): 146
- [94] Dienes Z, Altmann G, Gao S J, *et al.* The transfer of implicit knowledge across domains. *Language and Cognitive Processes*, 1995, **10**(3-4): 363-367
- [95] Bratzke D, Seifried T, Ulrich R. Perceptual learning in temporal discrimination: Asymmetric cross-modal transfer from audition to vision. *Experimental Brain Research*, 2012, **221**(2): 205-210
- [96] Bratzke D, Schröter H, Ulrich R. The role of consolidation for

- perceptual learning in temporal discrimination within and across modalities. *Acta Psychologica*, 2014, **147**(1): 75-79
- [97] Chen L, Guo L, Bao M. Sleep-dependent consolidation benefits fast transfer of time interval training. *Experimental Brain Research*, 2017, **235**(3): 661-672
- [98] Occelli V, Spence C, Zampini M. Audiotactile interactions in temporal perception. *Psychonomic Bulletin & Review*, 2011, **18**(3): 429-454
- [99] Grondin S, Ulrich R. Duration Discrimination Performance: No Cross-Modal Transfer from Audition to Vision Even after Massive Perceptual Learning. Berlin, Heidelberg: Springer, 2011
- [100] Lapid E, Ulrich R, Rammsayer T. Perceptual learning in auditory temporal discrimination: no evidence for a cross-modal transfer to the visual modality. *Psychonomic Bulletin & Review*, 2009, **16**(2): 382-389
- [101] Burr D, Banks M S, Morrone M C. Auditory dominance over vision in the perception of interval duration. *Exp Brain Res*, 2009, **198**(1): 49-57
- [102] Welch R B, Warren D H. Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, 1980, **88**(3): 638-667
- [103] McGovern DP, Astle A T, Clavin S L, *et al.* Task-specific transfer of perceptual learning across sensory modalities. *Current Biology*, 2016, **26**(1): R20-21
- [104] Alais D, Cass J. Multisensory perceptual learning of temporal order: audiovisual learning transfers to vision but not audition. *Plos One*, 2010, **5**(6): e11283
- [105] Barakat B, Seitz A R, Shams L. Visual rhythm perception improves through auditory but not visual training. *Current Biology*, 2015, **25**(2): R60-R61
- [106] Bruce C, Desimone R, Gross C G. Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology*, 1981, **46**(2): 369-384
- [107] Hikosaka K, Iwai E, Saito H, *et al.* Polysensory properties of neurons in the anterior bank of the caudal superior temporal sulcus of the macaque monkey. *Journal of Neurophysiology*, 1988, **60**(5): 1615-1637
- [108] Barraclough N E, Xiao D, Baker C I, *et al.* Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. *Journal of Cognitive Neuroscience*, 2005, **17**(3): 377-391
- [109] Yau J M, Deangelis G C, Angelaki D E. Dissecting neural circuits for multisensory integration and crossmodal processing. *Philos Trans R Soc Lond B Biol Sci*, 2015, **370**(1677): 20140203
- [110] Sereno M I, Huang R S. Multisensory maps in parietal cortex. *Current Opinion in Neurobiology*, 2014, **24**(1): 39-46
- [111] Akrami A, Kopec C D, Diamond M E, *et al.* Posterior parietal cortex represents sensory history and mediates its effects on behaviour. *Nature*, 2018, **554**(7692): 368-372
- [112] Bitzidou M, Bale M R, Maravall M. Cortical lifelogging: the posterior parietal cortex as sensory history buffer. *Neuron*, 2018, **98**(2): 249-252
- [113] Konen C S, Haggard P. Multisensory parietal cortex contributes to visual enhancement of touch in humans: A single-pulse TMS study. *Cerebral Cortex*, 2014, **24**(2): 501-507
- [114] Schlack A, Sterbing-D'angelo S J, Hartung K, *et al.* Multisensory space representations in the macaque ventral intraparietal area. *Journal of Neuroscience*, 2005, **25**(18): 4616-4625
- [115] Pasalar S, Ro T, Beauchamp M S. TMS of posterior parietal cortex disrupts visual tactile multisensory integration. *European Journal of Neuroscience*, 2010, **31**(10): 1783-1790
- [116] Azañón E, Longo M R, Soto-Faraco S, *et al.* The posterior parietal cortex remaps touch into external space. *Current Biology*, 2010, **20**(14): 1304-1309
- [117] Ghose D, Maier A, Nidiffer A, *et al.* Multisensory response modulation in the superficial layers of the superior colliculus. *Journal of Neuroscience*, 2014, **34**(12): 4332-4344
- [118] Meredith M A, Stein B E. Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, 1986, **56**(3): 640-662
- [119] Hall N J, Colby C L. Remapping for visual stability. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 2011, **366**(1564): 528-539
- [120] King A J. The superior colliculus. *Current Biology*, 2004, **14**(9): R335-338
- [121] Yu L, Xu J, Rowland B A, *et al.* Multisensory plasticity in superior colliculus neurons is mediated by association cortex. *Cerebral Cortex*, 2016, **26**(3): 1130-1137
- [122] Nishijo H, Ono T, Nishino H. Topographic distribution of modality-specific amygdalar neurons in alert monkey. *Journal of Neuroscience*, 1988, **8**(10): 3556-3569
- [123] Komura Y, Tamura R, Uwano T, *et al.* Auditory thalamus integrates visual inputs into behavioral gains. *Nature Neuroscience*, 2005, **8**(9): 1203-1209
- [124] Basura G J, Koehler S D, Shore S E. Multi-sensory integration in brainstem and auditory cortex. *Brain Research*, 2012, **1485**(1): 95-107
- [125] Naghavi H R, Eriksson J, Larsson A, *et al.* The claustrum/insula region integrates conceptually related sounds and pictures. *Neuroscience Letters*, 2007, **422**(1): 77-80
- [126] Niechwiej-Szwedo E, Chin J, Wolfe P J, *et al.* Abnormal visual experience during development alters the early stages of visual-tactile integration. *Behavioural Brain Research*, 2016, **304**(1): 111-119
- [127] Santangelo V, Van Der Lubbe R H, Olivetti Belardinelli M, *et al.* Multisensory integration affects ERP components elicited by exogenous cues. *Experimental Brain Research*, 2008, **185**(2): 269-277
- [128] Alsius A, Möttönen R, Sams M E, *et al.* Effect of attentional load on audiovisual speech perception: evidence from ERPs. *Frontiers in Psychology*, 2014, **5**(4): 727
- [129] Hyde D C, Jones B L, Flom R, *et al.* Neural signatures of face-voice synchrony in 5-month-old human infants. *Developmental Psychobiology*, 2011, **53**(4): 359-370

- [130] Spreckelmeyer K N, Kutas M, Urbach T P, *et al.* Combined perception of emotion in pictures and musical sounds. *Brain Research*, 2006, **1070**(1): 160-170
- [131] Lu X, Ho H T, Sun Y, *et al.* The Influence of visual information on auditory processing in individuals with congenital amusia: an ERP study. *Neuroimage*, 2016, **135**: 142-151
- [132] Thelen A, Cappe C, Murray M M. Electrical neuroimaging of memory discrimination based on single-trial multisensory learning. *Neuroimage*, 2012, **62**(3): 1478-1488
- [133] Wang W, Hu L, Cui H, *et al.* Spatio-temporal measures of electrophysiological correlates for behavioral multisensory enhancement during visual, auditory and somatosensory stimulation: A behavioral and ERP study. *Neurosci Bull*, 2013, **29**(6): 715-724
- [134] Gui P, Ku Y, Li L, *et al.* Neural correlates of visuo-tactile crossmodal paired-associate learning and memory in humans. *Neuroscience*, 2017, **362**: 181-195
- [135] Kutas M, Federmeier K D. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 2011, **62**: 621-647
- [136] Liu B, Wu G, Meng X. Cross-modal priming effect based on short-term experience of ecologically unrelated audio-visual information: An event-related potential study. *Neuroscience*, 2012, **223**: 21-27
- [137] Ohara S, Lenz F, Zhou Y D. Sequential neural processes of tactile-visual crossmodal working memory. *Neuroscience*, 2006, **139**(1): 299-309
- [138] Ohara S, Wang L, Ku Y, *et al.* Neural activities of tactile cross-modal working memory in humans: an event-related potential study. *Neuroscience*, 2008, **152**(3): 692-702
- [139] Drijvers L, Ozyurek A, Jensen O. Alpha and beta oscillations index semantic congruency between speech and gestures in clear and degraded speech. *J Cogn Neurosci*, 2018, **30**(8): 1086-1097
- [140] Senkowski D, Schneider T R, Foxe J J, *et al.* Crossmodal binding through neural coherence: implications for multisensory processing. *Trends in Neurosciences*, 2008, **31**(8): 401-409
- [141] Mishra J, Martinez A, Sejnowski T J, *et al.* Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *Journal of Neuroscience*, 2007, **27**(15): 4120-4131
- [142] Kumar G V, Halder T, Jaiswal A K, *et al.* Large Scale functional brain networks underlying temporal integration of audio-visual speech perception: An EEG study. *Frontiers in Psychology*, 2016, **7**: 1558
- [143] Bosman C A, Aboitiz F. Functional constraints in the evolution of brain circuits. *Front Neurosci*, 2015, **9**: 303
- [144] Merker B. Cortical gamma oscillations: the functional key is activation, not cognition. *Neuroscience & Biobehavioral Reviews*, 2013, **37**(3): 401-417
- [145] Belardinelli M O, Sestieri C, Di Matteo R, *et al.* Audio-visual crossmodal interactions in environmental perception: an fMRI investigation. *Cognitive Processing*, 2004, **5**(3): 167-174
- [146] McCormick K, Lacey S, Stilla R, *et al.* Neural basis of the crossmodal correspondence between auditory pitch and visuospatial elevation. *Neuropsychologia*, 2018, **112**: 19-30
- [147] Peiffer-Smadja N, Cohen L. The cerebral bases of the bouba-kiki effect. *NeuroImage*, 2019, **186**: 679-689
- [148] Noppeney U, Ostwald D, Werner S. Perceptual decisions formed by accumulation of audiovisual evidence in prefrontal cortex. *Journal of Neuroscience*, 2010, **30**(21): 7434-7446
- [149] Mayer A R, Ryman S G, Hanlon F M, *et al.* Look hear! The prefrontal cortex is stratified by modality of sensory input during multisensory cognitive control. *Cerebral Cortex*, 2017, **27**(5): 2831-2840
- [150] Foxe J J, Wylie G R, Martinez A, *et al.* Auditory-somatosensory multisensory processing in auditory association cortex: An fMRI study. *Journal of Neurophysiology*, 2002, **88**(1): 540-543
- [151] Beauchamp M S. See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. *Current Opinion in Neurobiology*, 2005, **15**(2): 145-153
- [152] Beauchamp M S, Nath A R, Pasalar S. fMRI-Guided transcranial magnetic stimulation reveals that the superior temporal sulcus is a cortical locus of the McGurk effect. *Journal of Neuroscience*, 2010, **30**(7): 2414-2417
- [153] Marques L M, Lapenta O M, Merabet L B, *et al.* Tuning and disrupting the brain-modulating the McGurk illusion with electrical stimulation. *Frontiers in Human Neuroscience*, 2014, **8**(1): 533
- [154] Bolognini N, Rossetti A, Casati C, *et al.* Neuromodulation of multisensory perception: a tDCS study of the sound-induced flash illusion. *Neuropsychologia*, 2011, **49**(2): 231-237
- [155] Marchant J L, Ruff C C, Driver J. Audiovisual synchrony enhances BOLD responses in a brain network including multisensory STS while also enhancing target-detection performance for both modalities. *Human Brain Mapping*, 2012, **33**(5): 1212-1224
- [156] Noesselt T, Rieger J W, Schoenfeld M A, *et al.* Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. *Journal of Neuroscience*, 2007, **27**(42): 11431-11441
- [157] Macaluso E. Modulation of human visual cortex by crossmodal spatial attention. *Science*, 2000, **289**(5482): 1206-1208
- [158] Macaluso E, Driver J. Multisensory spatial interactions: a window onto functional integration in the human brain. *Trends Neurosci*, 2005, **28**(5): 264-271
- [159] Ehrsson H H, Spence C, Passingham R E. That's my hand! Activity in premotor cortex reflects feeling of ownership of a limb. *Science*, 2004, **305**(5685): 875-877
- [160] Olive I, Tempelmann C, Berthoz A, *et al.* Increased functional connectivity between superior colliculus and brain regions implicated in bodily self-consciousness during the rubber hand illusion. *Human Brain Mapping*, 2015, **36**(2): 717-730
- [161] Seirafi M, De Weerd P, Pegna A J, *et al.* Audiovisual association learning in the absence of primary visual cortex. *Frontiers in Human Neuroscience*, 2015, **9**(1): 686
- [162] Bolognini N, Olgiati E, Rossetti A, *et al.* Enhancing multisensory

- spatial orienting by brain polarization of the parietal cortex. *European Journal of Neuroscience*, 2010, **31**(10): 1800-1806
- [163] Jones J A, Callan D E. Brain activity during audiovisual speech perception: an fMRI study of the McGurk effect. *Neuroreport*, 2003, **14**(8): 1129-1133
- [164] Alais D, Morrone C, Burr D. Separate attentional resources for vision and audition. *Proceedings of the Royal Society of London B: Biological Sciences*, 2006, **273**(1592): 1339-1345
- [165] Stevenson R A, Siemann J K, Schneider B C, *et al.* Multisensory temporal integration in autism spectrum disorders. *Journal of Neuroscience*, 2014, **34**(3): 691-697
- [166] Hahn N, Foxe J J, Molholm S. Impairments of multisensory integration and cross-sensory learning as pathways to dyslexia. *Neuroscience & Biobehavioral Reviews*, 2014, **47**(1): 384-392
- [167] Beker S, Foxe J J, Molholm S, *et al.* Ripe for solution: delayed development of multisensory processing in autism and its remediation. *Neuroscience & Biobehavioral Reviews*, 2017, **84**(1): 182-192
- [168] Paraskevopoulos E, Kraneburg A, Herholz S C, *et al.* Musical expertise is related to altered functional connectivity during audiovisual integration. *Proceedings of the National Academy of Sciences*, 2015, **112**(40): 12522-12527

Cross-modal Learning and Its Cognitive and Neural Mechanisms*

SUN Xun-Wei^{1,2)}, SUN Ying^{1,2)}, FU Qiu-Fang^{1,2)}**

⁽¹⁾State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China;

⁽²⁾Department of Psychology, University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract Cross-modal learning refers to learning that involves obtaining information from multiple modalities and then integrating and utilizing it. Multisensory integration is an important basis of cross-modal learning. Although cross-modal conditions are more like the real-life environment of human learning, most studies still use single-modal stimuli, and the results of cross-modal learning are still somewhat messy and unsystematic. In order to better summarize the characteristics of cross-modal learning and its related mechanisms, the current review first introduces multisensory integration effects and the factors that influence them, as well as the experimental and theoretical researches on the modality non-specificity of primary cortexes. Then, we sum up the researches on consciousness, representation type, and transfer effect in cross-modal learning, and research progress for neural mechanism of cross-modal learning by adopting the techniques such as neuronal recording, ERP and fMRI. Finally, we summarize the current research results of cross-modal learning, and examine prospects for the potential applications of these research findings and future research issues in this area.

Key words cross-modal learning, multisensory integration, cognitive mechanism, neural mechanism

DOI: 10.16476/j.pibb.2018.0321

* This work was supported by grants from The National Natural Science Foundation of China and the German Research Foundation (NSFC 61621136008/DFG TRR-169) and The National Natural Science Foundation of China (61632004).

** Corresponding author.

Tel: 86-10-64845395, E-mail: fuqf@psych.ac.cn

Received: April 15, 2019 Accepted: May 6, 2019