神经影像大数据与心脑关联: 方法学框架与应用

杨志*, 左西年*

中国科学院心理研究所行为科学重点实验室, 磁共振成像研究中心, 人脑功能连接组及其发展实验室, 北京 100101

* 联系人, E-mail: yangz@psych.ac.cn; zuoxn@psych.ac.cn

2014-08-01 收稿, 2014-09-11 接受, 2014-11-19 网络版发表

科技支撑计划(2012BAI36B01)、国家自然科学基金(81270023, 81278412, 81171409, 81000583, 81220108014)、中国科学院“百人计划”、中国科学院重点部署项目(KSZD-EW-TZ-002)、北京市精神疾病诊断和治疗重点实验室(2014JSJB03)和中国科学院心理研究所杰出青年人才基金(Y4CX062008)资助

摘要 心理科学成为一门成熟科学的必要条件之一是建立心理学概念与神经活动的联系, 即心脑关联. 当前心脑关联研究通常以心理学概念为出发点形成研究假设, 推断脑功能活动特征与心理学概念的关联. 心理学概念的形成依赖于观测者的主观经验和一致观念, 它们未必反映神经系统的功能组织结构. 经过近 20 年的心脑关联研究, 我们仍然难以建立明确而科学的心理学概念与脑活动特征之间的关系. 本文综述了当前心脑关联研究的瓶颈、分析其本质原因, 并提出以脑为中心、以神经影像大数据为基础的心脑关联研究框架. 随后, 举例综述新的心脑关联研究框架下神经影像大数据及其方法论. 最后, 讨论 2 个实际应用以说明新研究框架对发现心脑关联的独特作用.

关键词

神经影像
大数据
心脑关联
脑连接组
机器学习
数据挖掘

心理科学试图建立社会、意识和自然世界的联系. 心脑关联, 即心智、行为与神经系统的对应关系, 是心理科学试图回答的核心问题, 也是心理科学面临的巨大挑战^[1]. 心理学中许多概念(如思维、记忆、情绪、自动化加工、自我等)来自于生活经验, 它们的存在是由于一部分研究者一致认为它们能够有效描述心理活动, 而它们未必反映神经系统的功能组织结构. 当另一部分研究者一致地质疑这些概念的有效性、并提出新的概念体系时, 心理学中就出现了“流派”^[2].

出现“流派”现象是心理学成为一门成熟科学的重要障碍, 很多心理学概念的存在依赖于观测者的经验, 这就使观测者的主观经验影响了概念的构建. Barrett^[3,4]以情绪类别中的“愤怒”、“悲伤”及“恐惧”为例阐释了观测者经验对情绪分类的作用. 她发现, 我们从他人的行为表现中知觉到的“愤怒”、“悲伤”和“恐惧”是由于他人的行为表现诱发了存在于我们(观

测者)心中的不同情绪经验, 因而情绪的分类是被感知的, 而非独立于观察者的自然存在. 回顾大量研究的结果, Barrett^[1]发现人知觉某一类情绪时的个体变异与知觉不同情绪类别时的变异没有实质区别. 这些研究启示我们, 目前心理学中用于描述行为和心智的基本概念不完全独立于观测者.

建立心脑关联的目的是将依赖于观测者的心理学概念与独立于观测者的神经系统活动联系起来, 其意义在于使心理学概念与自然世界——而非研究者的经验和观念——建立牢固的联系, 进而推动心理学向一门成熟科学发展^[5]. 神经影像技术, 尤其是功能磁共振成像技术^[6]的飞速发展, 为建立心脑关联提供了最有效的途径. 利用神经影像技术, 研究者提出了“人脑功能图谱”的概念并开始探索各种心理现象在脑活动中的表征^[7-9], 从基本的视感觉^[10]到复杂的“共情”^[11]. 在经历了 20 余年研究后, 我们似乎很难找到一个还没有被“映射”到脑中的心理学术语.

引用格式: 杨志, 左西年. 神经影像大数据与心脑关联: 方法学框架与应用. 科学通报, 2015, 60: 966–975

Yang Z, Zuo X N. Big neuroimaging data-informed mind-brain association studies: Methodology and applications (in Chinese). Chin Sci Bull, 2015, 60: 966–975, doi: 10.1360/N972014-00806

<https://engine.scichina.com/doi/10.1360/N972014-00806>

遗憾的是, 20 余年的“人脑功能图谱”研究似乎未能实现心脑关联的目的, 反而使得人脑功能与心理现象之间的关系变得扑朔迷离。例如, 前扣带回皮层的活动被关联于许多心理学概念(见文献[12]), 包括注意^[13]、错误监控^[14]、工作记忆^[15]、自动控制^[16]、情绪^[17]等。这些心理学概念在定义上似乎相去甚远。同样, 位于皮层下的海马也被关联到记忆^[18]、空间定位^[19]、行为抑制^[20]、情绪^[21]等似乎不相干的心理学概念。将不相干的心理学概念映射到同一脑区使我们难以解释该脑区的功能(即与心理学概念的关联)。更有研究发现, 简单的视觉刺激可以诱散发布于整个皮层的有规律的神经活动。这意味着可能整个皮层都对最简单的感觉刺激做出反应^[22]。同样的情况也出现在联系脑功能网络与心理概念的研究中。例如, 2001 年 Raichle 等人^[23]提出人脑功能网络中存在“默认网络”, 即该网络在不执行特定任务时活动增强。尽管有研究试图用心理概念解释该网络的功能, 但大量研究发现, 该网络的异常与几乎所有心理、精

神、神经疾病有关联^[24], 使我们很难用心理概念明确解释其功能。这种与心理概念的模糊对应关系普遍存在于所有脑功能网络中^[25], 因而我们几乎无法根据脑功能网络的异常准确预测心智和行为的变化, 难以建立明确的心理学概念与脑活动特征之间的对应关系, 使心脑关联研究进展缓慢。

1 发现心脑关联: 基于神经影像大数据的研究框架

传统心脑关联研究遇到的瓶颈来源于研究策略。我们通常从心理学概念出发, 检验心智、行为、症状方面的差异所关联的脑活动特征, 进而推断某个脑区或脑功能网络与某个心理学概念的关联(图 1(a))。例如, 我们比较不同智商的人群在脑功能影像信号的差异, 进而推断呈现显著活动差异的脑区与智商有关。然而, 如上所述, 心理学概念由生活经验提出, 未必反映神经系统的功能组织结构。纵观各种对心

图 1 两种心脑关联的研究策略。(a) 传统的心脑关联研究策略: 从现有心理学概念出发, 检验与心理学概念关联的脑活动特征; (b) 以脑为中心的心脑关联研究策略: 从脑功能网络出发, 检验与脑功能网络关联的心智、行为特征集合

Figure 1 Two opposing strategies for investigating brain-mind association. (a) Traditional research strategy for brain-mind association: Characterize brain activities based on the current psychological concepts; (b) brain-mind association research that focuses on the brain: Characterize mind and behavior feature sets based on functional brain networks

理学概念的神经影像研究,我们经常看到同一脑区与众多相去甚远的心理学概念的关联(图 1(a)),以致这些关联难以解释和应用。

为解决研究策略方面的瓶颈,我们可以借鉴美国国立精神健康研究所提出的以脑为中心的研究策略“研究领域准则(research domain criterion, RDoC)”^[26-28]。该研究策略的目的是建立基于脑机制的精神疾病分类系统。根据 RDoC 的思想,笔者提出以脑为中心、以神经影像大数据为基础的心脑关联研究框架(图 1(b))。该研究框架的思路与传统研究策略相反,强调以脑功能网络为出发点建立研究假设,推断与某一脑功能网络相关联的心智、行为特征。图 2 展示了心脑研究框架的组成要素及其关系。我们首先利用“数据驱动”方法在神经影像数据中选取存在个体间变异或个体内动态变化的脑功能网络,进而筛选能够最大程度地解释脑功能网络变异的心智和行为特征,提出某个脑功能网络与一组心智、行为特征的关联。随后,我们试图独立地验证新提出的心脑关联。

以脑为中心的研究框架的优势在于,我们有望依据脑的功能组织结构重组心理学中的概念,将与某一脑功能网络关联的一组心智、行为特征定义为一个新的心理学概念。这一新概念以神经系统活动为基础,与观测者独立。独立于观测者的心理学概念将使心理学与自然世界的联系更加牢固,从而避免概念体系(“流派”)间的争论。因此,以脑为中心的研究框架将可能重塑目前心脑关联研究,甚至重新定义心理学的概念和术语体系。

神经影像大数据是以脑为中心的研究框架的基础。与传统研究策略中依赖现有理论提出研究假设

的做法不同,在以脑为中心的研究框架中,研究假设的提出源于对神经影像大数据的挖掘,即“数据驱动”的探索性分析。这是由于我们目前对脑功能的了解尚不足以形成明确、清晰的研究假设。自 2010 年 Biswal 等人^[29]提出利用大规模数据建立神经影像的“发现科学”以来,大规模神经影像数据发展迅猛,目前已经形成若干 1000 人以上的神经影像数据集^[30,31],美国(BRAIN/HCP, <http://www.nih.gov/science/brain>)和欧洲(HBP, <https://www.humanbrainproject.eu>)等发达国家和地区分别投入巨资发展人脑研究计划。这些大规模数据不仅可以使我们从神经科学角度重新认识人脑功能组织特点^[32],而且可以通过数据挖掘(见下文)提出新的心脑关联假设。与数据量的快速增长不相匹配,神经影像领域的各种方法学缺乏信效度和稳定性评估,特别是人脑功能连接组学领域^[33]。为推进人脑连接组学领域标准的建立,笔者联合全世界 18 个科研院所的 35 个实验室,最近发布国际首个神经影像重测信度数据库(CoRR, http://fcon_1000.projects.nitrc.org/indi/CoRR/html/index.html),共享超过 1600 人的重测数据,其中包含近万个多模态脑成像数据和 5000 余套静息态功能磁共振数据。结合丰富的行为测量数据,CoRR 将为检验心脑关联提供有力工具:我们可以利用一次测量的数据提出心脑关联,而利用重测数据独立地验证心脑关联。值得注意的是,心脑关联研究不仅需要大规模神经影像数据,还需要与影像数据对应的大规模行为测量数量,二者的结合构成适于心脑关联研究的神经影像大数据。以脑为中心的心脑关联发现框架的快速发展需要大样本脑与行为数据的积累,但大规模神经影像数据与多维度、大规模的行为数据在该框架中的地位不同:大规模神经影像数据是心脑关联发现框架的基础,多维度、大规模的行为数据是该框架的重要资源,但不是必要条件。

需要特别指出的是,以脑为中心的心脑关联发现框架并不否认传统的、从心理学概念出发的心脑关联研究的价值。行为测量数据的采集本身就是以心理学概念为基础的。新研究框架不否定行为数据蕴含的信息,而是致力于更好地整合众多心理学概念的科学价值和信息,使之与脑的工作机制具有更好的兼容性。在新研究框架中,行为数据位于被检验和重组的地位,而不是生成研究假设的依据。因此尽管行为数据的测量本身依赖于观测者的心理学概念,

图 2 (网络版彩色)以脑为中心的心脑关联研究框架的关键要素
Figure 2 (Color Online) Key elements in the brain-centered research strategy for brain-mind association

但心脑关联发现不是由其观测者依赖性驱动的. 新的研究框架与传统心脑关联研究具有较强的互补性. 我们在后文初步提出了定量化整合 2 种研究途径的思路(详见第 4 节).

2 发现心脑关联: 神经影像数据挖掘方法

如上文所述, 以脑为中心的心脑关联研究框架通过挖掘神经影像大数据提取存在明显个体间变异的脑功能网络, 进而建立该网络与心智、行为之间的关联. 研发数据挖掘算法对实现新的研究框架具有

关键作用. 除了常规的各种人脑功能连接组学方法(参见文献[33]), 笔者针对新研究框架的需要, 基于机器学习和数据挖掘理论, 提出了“数据驱动”的神经影像大数据挖掘方法, gRAICAR^[34~36](该工具已发布于互联网, <https://github.com/yangzhi-psy/gRAICAR>), 亦将作为人脑功能连接组计算系统(CCS, <http://lfc.d.psych.ac.cn/ccs.html>)^[37]的一个功能组件发布). 该方法为实现以脑为中心的心脑关联研究框架提供了核心方法. 如图 3 所示, gRAICAR 首先通过独立成分分析(ICA)获得反映脑功能网络的脑活动成分, 再依据

图 3 gRAICAR 方法的示意图. gRAICAR 依据空间分布相似性将来自于不同个体(S1~S3)的脑活动成分(S1_IC1, S1_IC2, ..., S4_IC4)划分为组水平成分(AC1~AC4), 并计算每一组水平成分中来自于不同个体成分的相似性矩阵. 这些矩阵分别反映不同脑功能网络在个体间的变异(或一致性). 参考文献[39]

Figure 3 Demonstration of the gRAICAR algorithm. gRAICAR groups the brain activity components (S1_IC1, S1_IC2, ..., S4_IC4) from individual subjects (S1, S2, S3) into group-level components (AC1~AC4). For each group-level component, gRAICAR computes a similarity matrix across subjects. These similarity matrices reflect inter-subject variability carried in the group-level components. See Ref. [39]

空间分布的相似性,将源于多个个体的脑活动成分划分为几类(称为组水平成分),使每一类内的成分尽可能相似,并且要求每一类中的成分分别来自不同个体.在每一个组水平成分中,我们用相似矩阵描述各个体的代表成分间的相似程度,称为个体间相似度矩阵.这些个体间相似度矩阵分别反映了不同脑功能网络在个体间的变异(或一致性).因此 gRAICAR 可以量化地反映各脑功能网络的个体间变异情况.

gRAICAR 之所以胜任以脑为中心的心脑关联研究框架,是由于它不假设个体在神经影像特征上的一致性或个体的分组.在未知个体分组的前提下, gRAICAR 能够基于脑功能网络的变异情况挖掘潜在的个体分组,并提取形成个体分组的脑活动特征.依据 gRAICAR 的结果,我们可进一步将体现个体间变异的脑活动特征与行为数据相关联,从而提出新的心脑关联假设.值得注意的是,由 gRAICAR 结果提出的心脑关联假设是基于神经影像特征的,而不是行为特征.这就实现了心脑关联研究框架的核心思想,使依赖于观测者心理学概念的行为测量成为被检验和重组的对象,而使神经影像数据表达的个体间变异成为心脑关联研究的驱动力量.

3 心脑关联发现框架的应用

以脑为中心的研究框架是发现心脑关联的有力工具.在本节中,我们讨论 2 个实际应用以说明该研究框架的独特优势.第一个研究将楔前叶-背侧扣带回网络与默认网络分离,建立脑功能网络与毕生发展的新关联^[38];第二个研究发现脑功能网络与临床症状模式之间的新关联^[39]. 2 个研究的共同特点是,以脑功能网络在个体间的变异为出发点发现新的心脑关联.值得注意的是,由于目前研究数据中的行为测量有限,以下 2 个研究实例应用新的心脑关联研究框架所发现的心脑关联受到行为数据维度的限制.然而如上文所述,有限的行为数据并不影响新的研究框架的实际应用.以下两个研究实例体现了以脑为中心的研究框架的核心价值:从脑的个体差异出发建立心脑关联.

3.1 毕生发展轨线区分默认网络和楔前叶网络

以往研究中,研究人员一直认为人脑的“默认网络”随人的毕生发展而呈现显著动态变化^[40-44].该网络一直被认为是老化过程的主要攻击目标网络,其

功能活动的衰退被各种认知老化疾病研究广泛报道(如老年痴呆症)^[45],然而“默认网络”的区域界定及其与其他脑网络的功能边界在学术界一直存在争议^[46-48].

根据以脑为中心的心脑关联研究框架,我们应用 gRAICAR 方法系统地挖掘了来自 126 名不同年龄(年龄范围: 7~85 岁)的健康志愿者的“静息态”脑功能影像数据^[31],以展示各脑功能网络在个体间的变异情况.我们发现 2 个在空间位置上存在交叠的脑功能网络,即“默认网络”(图 4(b))和“楔前叶-背侧扣带回网络”(precuneus-dorsal posterior cingulate cortex network, PCU-dPCC, 图 4(a)),显示出了截然不同的个体间变异情况.具体来说,“默认网络”的内部功能联系和空间分布模式在毕生发展中相对稳定,而“楔前叶-背侧扣带回网络”的内部功能联系和空间分布模式随年龄增长而降低(图 4).我们首先验证了多种对独立成分数的选择对于以上发现没有显著影响(详见文献^[38]),因而从技术层面确保了观测结果的可靠性.

根据以上观测结果,我们假设这 2 个脑网络在功能上是分离的,并体现出不同的毕生发展轨线.在后验分析中我们进一步发现,构成“楔前叶-背侧扣带回”网络的 2 个主要脑区随着年龄的增长对“默认网络”的贡献越来越明显.这些结果说明,这 2 个原本具有不同功能的脑网络随着年龄增长而逐渐合并为一个功能网络.这 2 个“功能枢纽”网络的动态整合可以解释脑老化进程中各种认知功能的不同变化,例如自我意识基本不随脑老化而消退,这可能与稳定的“默认网络”存在和保持有关;而空间认知等随脑老化有明显减退的功能,则可能与“楔前叶-背侧扣带回”网络的逐步解体有关.从人类毕生发展的角度,基于本研究的发现可以推测:以往观测到的与脑老化相关的“默认网络”活动变化可能来源于“楔前叶-背侧扣带回”网络,对这 2 个脑功能网络不加区分的方法,使以往研究将观测到的现象笼统地归因于“默认网络”功能衰退.

3.2 基于脑功能网络发现精神疾病亚型

在传统精神医学体系和临床实践中,精神疾病的诊断几乎全部依赖于心理行为症状表现.然而,基于症状学的诊断往往不能准确反映精神疾病的生物学机制,使得诊断与疗效脱节,许多患者需要尝试多

图 4 2 个存在空间交叠的脑功能网络显示不同的毕生发展轨迹。(a) “楔前叶-背侧扣带后回”网络; (b) “默认网络”。(c) 2 个网络的构成脑区及空间交叠示意图。构成这 2 个网络的脑区如(a)或(b)中左上角所示。这 2 个网络在受试者间的一致性或可重复性如左下角的相似性矩阵所示, 受试按年龄从低到高排序。基于相似性矩阵的 3 种受试间一致性指标(subject centrality, within-group consistency, fraction of significant subjects)显示, “楔前叶-背侧扣带后回”网络随受试年龄稳定性下降, 而“默认网络”相对稳定。参考文献[38]

Figure 4 Two spatially overlapping functional brain networks exhibit divergent life-span trajectories. (a) A “Precuneus-posterior cingulate cortex” network; (b) the “default mode network”; (c) overlapping regions of the two networks. Element regions for these networks are demonstrated in the upper-left corners of panels (a) and (b). Inter-subject consistency for the two networks are shown in the lower-left corners of panels (a) and (b). The subjects are ordered according to increasing age. Based on these similarity matrices, three different inter-subject consistency indices (subject centrality, within-group consistency, fraction of significant subjects) show that the consistency of the “precuneus-posterior cingulate cortex” network decreases with aging, while the “default mode network” is relatively stable. See Ref. [38]

种治疗方法和药物, 且大量患者终生无法治愈。传统的神经影像方法学因此严重限制了“以脑为中心”的精神疾病研究: 绝大多数研究首先根据症状定义病人组, 并比较病人组与控制组之间的平均脑活动差异, 因而使其在即使测量了脑活动的情况下, 本质上仍然是“以症状学为中心”的研究, 往往不具有对某种精神疾病的特异性, 导致其结果对精神疾病的临床实践改革意义有限。

根据以脑为中心的心脑关联研究框架, 在我们最近的一项研究中, 应用 gRAICAR 探索首次发作、未服药、早发精神分裂症患者的亚型。我们将 26 名精

神分裂症患者与 25 名健康对照组志愿者的“静息态”脑活动成分混合, 基于个体脑功能网络的相似性来检测在脑网络特征上具有高度一致性的患者亚组。我们发现了额下回-额上回网络与精神分裂症患者多维度症状模式的关联(图 5)。对于阳性症状高于阴性症状的患者, 该脑功能网络与健康对照组高度一致, 而在阴性症状大于阳性症状的患者中未发现该脑功能网络。因此, 根据这一脑功能网络特征, 我们可以将患者分为 2 个亚型。利用非参数检验、功能联接方法及一般线性模型, 我们验证了该分型的普遍性。该研究说明脑功能网络可以反映多维度的临床症状“模

图 5 基于脑功能网络相似性检测被试亚组. 利用 gRAICAR 的数据挖掘研究发现, 早发精神分裂症的阴性和阳性症状的相对强度与额下回-颞上回网络是否正常有关. 依据此脑功能网络可将患者分为 2 个亚组. 参考文献[39]

Figure 5 Discover subject communities based on similarity of functional brain networks. A data-mining study using gRAICAR reveals that the relative severity of positive and negative symptoms in early-onset schizophrenia is related to the abnormality of an inferior frontal gyri-superior temporal gyri network. According to the inter-subject variability of this network, early-onset schizophrenia patients can be categorized into two sub-groups. See Ref. [39]

式”，而非单一症状，提示在未来研究中需要对“多元症状”模式引起重视。更重要的是，该研究展示了利用数据挖掘算法实现“以脑为中心”的精神疾病研究的新途径。类似的基于脑与行为数据来进行疾病分组最近已经在国际上受到高度重视，Fair 等人^[49,50]就基于类似的方法对当前儿童多动症的亚型进行了研究。这些研究中发现的脑功能网络与多维症状模式之间的关联难以被以心理学、症状学概念为出发点的研究发现。

4 心脑关联发现框架面临的挑战和解决方案展望

同任何方法学一样，本文提出的基于神经影像大数据的心脑关联发现框架面临着多方面的挑战。笔者将新研究框架的局限性归纳如下，并展望潜在的解决方案。第一，由于当前的神经影像技术只能以较低的时间和空间分辨率间接地反映神经活动，我们能够获取的神经影像特征极为有限。而以有限的神经影像特征几乎不可能完全解释人的无限的行为。例如，我们通常以较大的、数量有限的解剖区域或数量有限的脑功能网络为单位提取脑活动特征，这些脑活动特征的变异与人类行为的巨大变异相比非常有限。这一情况是对所有心脑关联研究(包括传统的、从心理学概念出发的研究框架)最本质的挑战。除了改进神经影像技术以获取更高时空分辨率的神经活动信息外，笔者设想利用多变量模型建立心脑关联，即建立多维度神经影像特征与多维度行为特征之间的联系。利用神经影像的多维特征表达神经活动大大扩展了其解释行为变异的能力。“典型相关”(canonical correlation)^[51]是实现多变量心脑关联模型的有力工具。第二，先验知识不足。由于我们对大规模神经活动的知识和测量手段有限，我们对脑功能特征的先验知识亦有限。例如，我们尚不清楚脑功能网络是否存在明确的边界、最适于建立心脑关联的脑活动单元等重要问题。这对我们从脑的特征出

发建立心脑联系构成了挑战。贝叶斯方法可能是应对这一挑战、整合传统心脑关联研究成果和新的研究框架的有力工具。一个具体的设想是：以心理学概念反映的个体差异特征为个体差异的先验概率，以神经影像数据反映的个体差异特征为似然性，估计个体差异的后验概率。第三，行为数据收集不足。多维度的、大规模的行为数据测量数据使我们能发现更多心脑关联，是推动新的心脑关联，发现框架发展的重要资源。与大规模神经影像数据不同，多维度、大规模行为数据的采集和共享尚未快速发展。我们预期随着大规模神经影像数据的丰富，研究者会越来越重视与之匹配的行为测量数据的采集。同时，我们应该看到，尽管多维度的行为数据能推进新心脑关联发现框架的发展，但它并不是该研究框架发挥作用的必要条件。上文中介绍的 2 个应用实例即是在目前行为数据并不丰富的情况下对新研究框架的应用。同时，利用 gRAICAR 对个体间神经影像特征的相似性的测量目前只局限于脑功能网络的空间分布特性，是多种相似性量度中的一种。整合多种相似性量度以增进心脑关联的稳定性和有效性是未来研究的方向之一。

5 结论

心脑关联是心理科学研究中的关键性基础问题，将心理学概念与独立于观察者的神经活动紧密联系起来是心理科学发展为一门成熟科学的重要而不可或缺的条件之一。为突破传统心脑关联研究的瓶颈，我们提出以脑为中心、以神经影像大数据为基础的心脑关联发现框架。该研究框架强调以脑功能网络为出发点建立研究假设，推断与各脑功能网络相关联的心智、行为特征。该研究框架有望依据脑的功能组织结构重组心理学中的概念，使其与神经活动为基础、与观测者独立。未来神经影像学大数据和以脑为中心的数据挖掘方法论的发展将有力地推动以脑为中心的心脑关联研究。

参考文献

- 1 Barrett L F. The future of psychology: Connecting mind to brain. *Perspect Psychol Sci*, 2009, 4: 326–339
- 2 Ludin R W. *Theories and Systems of Psychology*. 5th ed. Lexington, MA: D. C. Heath & Company, 1996
- 3 Barrett L F. Feeling is perceiving: Core affect and conceptualization in the experience of emotion. In: Barrett L F, Niedenthal P M, Winkielman P, eds. *Emotions: Conscious and Unconscious*. New York: Guilford, 2005. 255–284
- 4 Barrett L F. Solving the emotion paradox: Categorization and the experience of emotion. *Pers Soc Psychol Rev*, 2006, 10: 20–46

- 5 Cleeremans A. The grand challenge for psychology: Integrate and fire! *Front Psychol*, 2010, 1: 1–2
- 6 Ogawa S, Lee T M, Kay A R, et al. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci USA*, 1990, 87: 9868–9872
- 7 Bandettini P A, Wong E C, Jesmanowicz A, et al. Spin-echo and gradient-echo EPI of human brain activation using BOLD contrast: A comparative study at 1.5 T. *NMR Biomed*, 1994, 7: 12–20
- 8 Friston K J, Holmes A P, Worsley K J, et al. Statistical parametric maps in functional imaging: A general linear approach. *Hum Brain Mapp*, 1994, 2: 189–210
- 9 Kwong K K, Belliveau J W, Chesler D A, et al. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc Natl Acad Sci USA*, 1992, 89: 5675–5679
- 10 Dale A M, Buckner R L. Selective averaging of rapidly presented individual trials using fMRI. *Hum Brain Mapp*, 1997, 5: 329–340
- 11 Fan Y, Duncan N W, de Greck M, et al. Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. *Neurosci Biobehav Rev*, 2011, 35: 903–911
- 12 Gasquoin P G. Localization of function in anterior cingulate cortex: From psychosurgery to functional neuroimaging. *Neurosci Biobehav Rev*, 2013, 37: 340–348
- 13 Bush G, Luu P, Posner M I. Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn Sci*, 2000, 4: 215–222
- 14 Mies G W, van der Molen M W, Smits M, et al. The anterior cingulate cortex responds differently to the validity and valence of feedback in a time-estimation task. *NeuroImage*, 2011, 56: 2321–2328
- 15 Haxby J V, Petit L, Ungerleider L G, et al. Distinguishing the functional roles of multiple regions in distributed neural systems for visual working memory. *NeuroImage*, 2000, 11: 380–391
- 16 Critchley H D, Mathias C J, Josephs O, et al. Human cingulate cortex and autonomic control: Converging neuroimaging and clinical evidence. *Brain*, 2003, 126: 2139–2152
- 17 Lindgren L, Westling G, Brulin C, et al. Pleasant human touch is represented in pregenual anterior cingulate cortex. *NeuroImage*, 2012, 59: 3427–3432
- 18 Bird C M, Burgess N. The hippocampus and memory: Insights from spatial processing. *Nat Rev Neurosci*, 2008, 9: 182–194
- 19 Howard L R, Javadi A H, Yu Y, et al. The Hippocampus and entorhinal cortex encode the path and euclidean distances to goals during navigation. *Curr Biol*, 2014, 24: 1331–1340
- 20 Scarf D, Millar J, Pow N, et al. Inhibition, the final frontier: The impact of hippocampal lesions on behavioral inhibition and spatial processing in pigeons. *Behav Neurosci*, 2014, 128: 42–47
- 21 Phelps E A. Human emotion and memory: Interactions of the amygdala and hippocampal complex. *Curr Opin Neurobiol*, 2004, 14: 198–202
- 22 Gonzalez-Castillo J, Saad Z S, Handwerker D A, et al. Whole-brain, time-locked activation with simple tasks revealed using massive averaging and model-free analysis. *Proc Natl Acad Sci USA*, 2012, 109: 5487–5492
- 23 Raichle M E, MacLeod A M, Snyder A Z, et al. A default mode of brain function. *Proc Natl Acad Sci USA*, 2001, 98: 676–682
- 24 Zhang D, Raichle M E. Disease and the brain's dark energy. *Nat Rev Neurol*, 2010, 6: 15–28
- 25 Smith S M, Fox P T, Miller K L, et al. Correspondence of the brain's functional architecture during activation and rest. *Proc Natl Acad Sci USA*, 2009, 106: 13040–13045
- 26 Insel T, Cuthbert B, Garvey M, et al. Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *Am J Psychiatry*, 2010, 167: 748–751
- 27 Insel T R. The NIMH Research Domain Criteria (RDoC) Project: Precision medicine for psychiatry. *Am J Psychiatry*, 2014, 171: 395–397
- 28 Simmons J M, Quinn K J. The NIMH Research Domain Criteria (RDoC) Project: Implications for genetics research. *Mamm Genome*, 2014, 25: 23–31
- 29 Biswal B B, Mennes M, Zuo X N, et al. Toward discovery science of human brain function. *Proc Natl Acad Sci USA*, 2010, 107: 4734–4739
- 30 Di Martino A, Yan C G, Li Q, et al. The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol Psychiatry*, 2014, 19: 659–667
- 31 Nooner K B, Colcombe S J, Tobe R H, et al. The NKI-Rockland sample: A model for accelerating the pace of discovery science in psychiatry. *Front Neurosci*, 2012, 6: 152
- 32 Kelly C, Biswal B B, Craddock R C, et al. Characterizing variation in the functional connectome: Promise and pitfalls. *Trends Cogn Sci*, 2012, 16: 181–188
- 33 Craddock R C, Jbabdi S, Yan C G, et al. Imaging human connectomes at the macroscale. *Nat Methods*, 2013, 10: 524–539

- 34 Dong G Z, Huang Z R, Yang Z, et al. Enhance fMRI data analysis by RAICAR. 2009 3rd International Conference on Bioinformatics and Biomedical Engineering, Vols 1-11, 2009. 1990–1993
- 35 Yang Z, LaConte S, Weng X C, et al. Ranking and averaging independent component analysis by reproducibility (RAICAR). *Hum Brain Mapp*, 2008, 29: 711–725
- 36 Yang Z, Zuo X N, Wang P P, et al. Generalized RAICAR: Discover homogeneous subject (sub)groups by reproducibility of their intrinsic connectivity networks. *NeuroImage*, 2012, 63: 403–414
- 37 Zuo X N, Xu T, Jiang L, et al. Toward reliable characterization of functional homogeneity in the human brain: Preprocessing, scan duration, imaging resolution and computational space. *NeuroImage*, 2013, 65: 374–386
- 38 Yang Z, Chang C, Xu T, et al. Connectivity trajectory across lifespan differentiates the precuneus from the default network. *NeuroImage*, 2014, 89: 45–56
- 39 Yang Z, Xu Y, Xu T, et al. Brain network informed subject community detection in early-onset schizophrenia. *Sci Rep*, 2014, 4: 5549
- 40 Damoiseaux J S, Beckmann C F, Arigita E J, et al. Reduced resting-state brain activity in the “default network” in normal aging. *Cereb Cortex*, 2008, 18: 1856–1864
- 41 Esposito F, Aragri A, Pesaresi I, et al. Independent component model of the default-mode brain function: Combining individual-level and population-level analyses in resting-state fMRI. *Magnet Resonance Imag*, 2008, 26: 905–913
- 42 Koch W, Teipel S, Mueller S, et al. Effects of aging on default mode network activity in resting state fMRI: Does the method of analysis matter? *NeuroImage*, 2010, 51: 280–287
- 43 Mevel K, Landeau B, Fouquet M, et al. Age effect on the default mode network, inner thoughts, and cognitive abilities. *Neurobiol Aging*, 2013, 34: 1292–1301
- 44 Sala-Llonch R, Arenaza-Urquijo E M, Valls-Pedret C, et al. Dynamic functional reorganizations and relationship with working memory performance in healthy aging. *Front Hum Neurosci*, 2012, 6: 152
- 45 Jagust W J, Mormino E C. Lifespan brain activity, beta-amyloid, and Alzheimer’s disease. *Trends Cogn Sci*, 2011, 15: 520–526
- 46 Andrews-Hanna J R, Reidler J S, Sepulcre J, et al. Functional-anatomic fractionation of the brain’s default network. *Neuron*, 2010, 65: 550–562
- 47 Zhang S, Li C S. Functional connectivity mapping of the human precuneus by resting state fMRI. *NeuroImage*, 2012, 59: 3548–3562
- 48 Zuo X N, Di Martino A, Kelly C, et al. The oscillating brain: Complex and reliable. *NeuroImage*, 2010, 49: 1432–1445
- 49 Fair D A, Bathula D, Nikolas M A, et al. Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. *Proc Natl Acad Sci USA*, 2012, 109: 6769–6774
- 50 Fair D A, Nigg J T, Iyer S, et al. Distinct neural signatures detected for ADHD subtypes after controlling for micro-movements in resting state functional connectivity MRI data. *Front Syst Neurosci*, 2012, 6: 80
- 51 Knapp T R. Canonical correlation analysis: A general parametric significance-testing system. *Psychol Bull*, 1978, 85: 410–416

Big neuroimaging data-informed mind-brain association studies: Methodology and applications

YANG Zhi & ZUO Xi-Nian

Key Laboratory of Behavioral Sciences, Magnetic Resonance Imaging Research Center, Laboratory of Functional Connectome and Development, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China

Difficulty in linking psychological concepts with neural mechanisms prevents the field of psychology from being considered a natural science. Previous mind-brain association studies were usually driven by psychological concepts and used to infer brain activity characteristics. Because psychological concepts are formed by observer-dependent experience and consensus, they may not correspond to the functional organization of neural networks in the brain. Even after two decades of mind-brain association research, we are still unable to establish clear and objective associations between psychological concepts and brain function. In this article we first review the challenges in linking psychological concepts with neural mechanisms and then analyze these difficulties’ causes. Based on these analyses, we propose a novel mind-brain association discovery methodology driven by big neuroimaging data. We discuss the advantage of this neuroimaging data-driven discovery methodology and introduce an algorithm, Graicar, to implement it. Finally, we demonstrate the usefulness of the methodology and the gRAICAR algorithm with two detailed applications

neuroimaging, big data, mind-brain association, brain connectome, machine learning, data mining

doi: 10.1360/N972014-00806