

# 欺骗的神经机制和测谎应用：来自 fMRI 研究的证据\*

崔茜 蒋军 杨文静 张庆林

(西南大学心理学部; 认知与人格教育部重点实验室(西南大学), 重庆 400715)

**摘要** 功能核磁共振成像技术在探讨欺骗机制和开发测谎技术中的应用是目前的研究热点, 前者使用多种范式来考察假装失忆、故意错误及人际交互中的欺骗行为的认知和神经机制, 而后者主要沿用传统的 GKT 和 CQT 范式检测欺骗行为。已有研究突显了前额叶和顶叶这些参与执行控制过程的脑区在欺骗中的核心作用, 但基于这些脑区信号进行测谎时, 得到的准确率不容乐观。未来研究应探索除执行控制过程外的欺骗的其它认知成分, 尝试从“记忆痕迹”和“欺骗结果”等方面寻找测谎的新线索。并应揭示“欺骗特质”的神经基础, 以启发“忠诚度”的测量方法。

**关键词** 欺骗; 测谎; fMRI

**分类号** B845; B849

## 1 引言

欺骗普遍存在于人类社会, 它是标志个体心理发展“正常”的一种认知能力(Spence et al., 2004)。欺骗具有多种形式或状态, 包括经济欺诈和骗局、掩饰、伪造、游戏中的假动作或策略、魔术、以及善意的谎言甚至玩笑等。一些欺骗行为具有自私和反社会性质, 它们将导致严重后果, 而另一些欺骗行为具有利他和亲社会性质, 它们作为社交过程中的润滑剂可促进人际交往。

Abe (2009)将欺骗定义为一种心理过程。通过这一心理过程, 个体在明知道某一事件或信念是错误或虚假的情况下, 有意试图说服他人相信这一事件或信念, 以获得某些好处或避免损失。Vrij (2000)认为欺骗是在没有预先警告的情况下, 某种成功或不成功的、给他人制造某种自己明知是虚假的信念的有意意图。Vrij (2004)再次强调了欺骗的无预警性和有意性特征。以上这些定义都同时指出了欺骗的两个重要特点: 首先欺骗不仅具

有非真实的属性, 更重要的是它必须是一种有意行为。因而病态性撒谎不能称之为欺骗, 因为患者不具备有意制造错误信念的主观意图, 仅仅是在客观上造成了错误信息的传达(Sip, Roepstorff, McGregor, & Frith, 2008)。其次, 欺骗通常是没有预警性的, 一方面, 在对方遭受欺骗之前, 不会收到任何的预警或提示; 另一方面, 欺骗行为是欺骗者自愿产生的, 而非受到来自外部的某种指导或提示信息的诱发(Derksen, 2012)。

识别欺骗对人类具有重要的适应意义(Lachmann & Bergstrom, 2004), 而人类先天在识别欺骗这一技能上具有缺失, 因此, 除了关注欺骗现象本身以外, 科学家们尤其专注于发展测谎技术。以生理多导仪为主的早期传统测谎技术通过测量非特异的、外周神经系统的自主唤醒, 推测欺骗的发生(Gamer, Verschuere, Crombez, & Vossel, 2008)。自主神经系统的唤醒表明个体产生了焦虑或恐惧情绪, 然而这些情绪并不一定由欺骗导致, 因此, 传统测谎方法的科学性受到质疑(Press, 2003)。随后研究者们发明了基于事件相关电位(Event Related Potential, ERP)的测谎技术, 该技术可以探测大脑活动, 通过判断受测者是否具有某些重要信息的先前知识(犯罪知识测试, Guilty Knowledge Test, GKT), 从而推知真相。这一方法虽然也不是直接检测欺骗过程, 但由于它

收稿日期: 2012-09-07

\* 国家自然科学基金资助项目(批准号: 31170983); 西南大学 211 工程国家重点学科建设项目(批准号: NSKD08002)。

通讯作者: 张庆林, E-mail: zhangql@swu.edu.cn

依赖客观反映刺激熟悉度的P300指标,因此获得了较为稳定的测谎效果。但是,ERP的低空间分辨力特征使其难以定位欺骗过程所涉及的大脑区域(Abe, 2011),使得对欺骗神经机制的探索受到阻碍。近年来,神经成像技术的发展,尤其是功能核磁共振成像(fMRI, Functional Magnetic Resonance Imaging)在认知科学中的应用,使得在高空间分辨力下直接检测与欺骗有关的大脑活动成为可能(Abe, 2011),也为寻找新的测谎线索创造了契机。

最初使用fMRI技术开展的欺骗研究是在先前ERP的欺骗研究基础上,使用迫选和提示反应的实验范式同时诱发欺骗反应和诚实反应,再将两种反应导致的大脑激活模式进行相减,从而揭示区分两种反应的神经信号(Langleben et al., 2002; Lee et al., 2002; Spence et al., 2001),这些研究发现诚实反应是一种基线式的反应,而欺骗反应则会导致参与认知控制功能的前额叶系统的更大激活(Spence, 2004)。在这些先驱研究的基础上,随后的研究根据目的分为两大部分:一部分研究关注欺骗的认知和神经机制,使用假装失忆范式(Browndyke et al., 2008; Lee et al., 2005; Liang et al., 2012)、故意错误任务(Abe et al., 2008; Bhatt et al., 2009; Fullam, McKie, & Dolan, 2009; Ganis, Morris, & Kosslyn, 2009; Ito et al., 2011; Kaylor-Hughes et al., 2011; Lee et al., 2009; Nunez, Casey, Egner, Hare, & Hirsch, 2005; Spence et al., 2004; Spence, Kaylor-Hughes, Brook, Lankappa, & Wilkinson, 2008)以及在具有人际交互的社会情境下诱发真实欺骗行为的多种新颖任务(Baumgartner, Fischbacher, Feierabend, Lutz, & Fehr, 2009; Bhatt, Lohrenz, Camerer, & Montague, 2010; Greene & Paxton, 2009; Kireev, Korotkov, & Medvedev, 2012; Sip et al., 2012; Spence, Kaylor-Hughes, Farrow, & Wilkinson, 2008),探讨不同类型的欺骗行为的神经机制差异及共性;在神经基础上寻找更为纯净的欺骗成分(如与记忆成分分离);并探索人际交互作用下自发欺骗的神经机制及其个体差异。而另一部分研究在探索欺骗的大脑机制的同时,将机制研究中发现的欺骗所诱发的神经信号作为测谎指标,进行个体诊断的尝试,这部分研究大多沿用传统的GKT和CQT(Control Question Test)范式,结合特殊的统计分析方法,进行个体水平上的鉴别分析(Davatzikos et al., 2005; Kozel,

Johnson, et al., 2009; Kozel et al., 2005; Kozel, Laken, et al., 2009; Kozel, Padgett, & George, 2004; Kozel, Revell, et al., 2004; Langleben, 2008; Langleben et al., 2005; Nose, Murai, & Taira, 2009)。

下文将以这两部分研究及其所使用的不同实验范式为主线,对现有研究的内容、方法及大致结果进行回顾,并根据目前研究存在的问题,结合其他领域的研究发现,提出未来的研究方向。

## 2 已有研究现状

### 2.1 欺骗机制研究

使用fMRI技术对欺骗神经机制的揭示,经历了从实验室情景下的简单欺骗反应到人际交互情境下的自发欺骗行为的探索。这些研究采用不同的实验范式从不同角度诱发欺骗行为,以逐步接近现实生活中的多种欺骗情景。以下我们对这些实验范式进行简要的总结和概括:

#### 2.1.1 假装记忆损伤范式

装病是一种欺骗行为,包括伪装或夸大某一生理、心理疾病或其他功能损伤的行为(Lee et al., 2002)。这一现象普遍存在于现实生活中,已经引起许多学科领域研究者的关注。由于人类在行为上伪装较为容易,因而使用行为范式对这一现象进行检测很难获得满意的效果。然而,人类几乎不可能控制其大脑激活模式,通过检测伪装行为导致的大脑信号来检测伪装已成为目前研究的热点。

假装记忆损伤的研究一般采用学习再认范式,先让被试学习一些材料,这些材料可能是字词或数字等,然后在实验过程中要求被试对学习过的材料进行“假装失忆”反应。“假装失忆”反应任务往往要求被试想象他们处于某种保险赔偿情境中,如果他们被诊断为记忆损伤,将会获得可观的赔偿金。被试的具体任务是使用策略、尽可能地伪装失忆,目的是在再认任务中表现出差的回忆成绩,但不能只是简单地做出完全相反的错误反应,因为这样很容易被发现在伪装。通常,为了让被试更加容易的理解假装失忆任务,从而更为自然娴熟地进行假装反应,实验任务中一般还会另外设置“正确反应”、“错误反应”以及“随机反应”3种任务。

Lee等(2002)首次使用fMRI技术、在学习再

认范式下考察假装失忆的神经机制,发现假装失忆反应较之诚实反应导致了前额叶、左侧扣带回、顶叶和左侧尾状核等脑区的更大激活。随后的重复研究证实了前额叶-顶叶-皮层下基底神经节(尾状核和扣带回)这一神经连接在假装失忆中的重要作用,并发现这一神经机制不受被试性别和母语类型的影响(Lee et al., 2005),但受反应要求和错误类型的调控(Browndyke et al., 2008),证实了假装失忆激活的大脑反应不同于随机反应和错误反应,这意味着可能寻找到鉴别假装失忆行为的独特神经信号指标(Liang et al., 2012)。

假装失忆研究证实了前额叶-顶叶皮层在伪装这类欺骗行为中起着重要作用,并提出可以使用这一脑区的激活信号对假装失忆行为进行鉴别诊断。但目前这部分的 fMRI 研究还主要处于对机制的探讨和尝试寻找诊断指标的初始阶段,未来研究应该进行个体诊断的尝试。此外,目前研究中的假装行为仍然处于实验者的指导之下,并非由情景自然诱发,还没有研究探讨假装行为的决策过程,以及奖赏机制对该行为的调控作用。

### 2.1.2 故意错误反应范式

故意错误反应这类研究实际也采用学习再认范式,但要求被试在再认阶段故意执行错误的反应。一部分研究使用词语、数字或图片作为实验材料。而另外一部分研究,通常使用被试的个人相关信息或最近经历的事件或执行过的行为为实验材料。前类研究通常包括学习阶段和再认阶段,先要求被试在学习阶段学习一些实验材料(旧刺激),随后在再认阶段同时向被试呈现旧刺激和另外一些没有学过的新刺激,要求被试进行故意的错误再认反应,即将所有旧刺激有意判断为新刺激(假装漏报),以及将所有新刺激故意判断为旧刺激(假装误报)。后类研究由于是以被试个体所经历的事件作为实验材料,被试不再需要特定的学习阶段,但需要在实验前对每个被试进行访谈以搜集有关的个人信息。在这类研究中,被试的任务同样也是对发生、或没有发生在自己身上的经历进行故意错误的反应,即故意承认所有未发生过的经历和故意否认所有发生过的经历。

使用词语和图片作为材料的故意错误研究的主要目的是探讨故意错误(欺骗)与无意错误,或者与正确反应(诚实)所诱发的大脑激活模式的差异,从而更精确地把握与欺骗成分独特相关的大

脑激活模式。使用词语材料的研究证实了故意错误与诱饵词诱发的错误再认(Abe et al., 2008)、以及与无意错误(Lee et al., 2009)的大脑激活模式存在明显差异,这些研究发现故意错误作为欺骗的一种类型,涉及执行控制相关脑区,如前额叶皮层和扣带回等的参与,而错误记忆在记忆相关脑区,如海马上表现出激活。而以图片为材料的研究考察了图片的情绪成分(Ito et al., 2011)以及积极或消极的情绪效价(Lee, Lee, Raine, & Chan, 2010)对故意错误反应的影响,也有研究使用面孔图片为材料,模拟证人指证嫌疑犯这一司法审讯情景(Bhatt et al., 2009)。以上这些研究由于使用了不同材料,涉及不同的认知过程,因此每个研究具体诱发的大脑激活模式存在很大差异,但是,这些研究较一致的发现故意错误反应(欺骗)导致了认知控制和反应抑制相关的脑区的更大激活,这些脑区大致位于前额叶、顶叶皮层等。

以被试的个人相关信息或近期经历为实验材料的欺骗研究更多关注对不同信息进行不同类型的欺骗反应所导致的大脑激活模式,也有研究涉及对欺骗行为的个体差异及其神经机制的探讨。一些研究比较了个体对新近情景记忆(Spence et al., 2001)和长时情景记忆(Spence et al., 2004)、自传信息和常识信息(Nunez et al., 2005)、个人经历和他人经历(Ganis et al., 2009)进行欺骗反应时的大脑激活模式;探讨了基于编造的虚假经历(记忆的谎言)和一些临时提问(临时的自发欺骗)进行欺骗时所导致的大脑激活模式的差异。这些研究的核心结果仍然与欺骗研究的主要观点一致,同样发现欺骗较之诚实诱发了认知控制和反应抑制相关脑区(主要位于前额和顶叶区域)的更大激活,而当被试对个人相关信息进行欺骗时,这一控制过程表现得更为明显(Nunez et al., 2005)。另一些研究者利用这一范式对精神分裂症患者的欺骗行为和神经机制进行了考察,同样发现前额叶在欺骗过程中起到稳定而重要的作用(Fullam et al., 2009; Kaylor-Hughes et al., 2011),并发现这一区域的激活与精神病态特质质量表 PPI (Psychopathic Personality Inventory)的5个分量表——冷漠、无惧性、权谋中心性、领导性和压力免疫——上的得分相关显著,提示欺骗的神经机制上存在人格差异(Fullam et al., 2009)。此外,Spence等人(2008)使用这种故意错误范式,进行了个案诊断的尝试,

虽然结果不能用作司法凭证, 但为司法审判提供了启发和线索。

故意错误任务基于经典的学习再认范式, 它易于在实验室中进行操作, 便于控制无关变量, 并且可以提供机会了解欺骗和记忆的神经机制以及在欺骗过程中两者之间的联系和变化关系。这部分研究获得的结果再次证实了前额叶-顶叶皮层在欺骗中的重要作用; 发现欺骗和识别两个认知过程会诱发相互独立的神经信号; 发现假装行为中存在个体差异并初步揭示了其神经基础; 并且研究者已开始尝试利用这些结果对个体的故意错误行为进行诊断。未来研究需要注意增加这类欺骗行为的有意性并考虑社会情境因素可能对其产生的影响, 此外, 可以从这些结果中获得更多启发, 比如欺骗过程中稳定反映识别过程的神经信号也许可被用作新的测谎线索。

### 2.1.3 社会情境下的自发欺骗范式

有意性和无预警性是欺骗的两个重要成分(Vrij, 2000, 2004), 但在上述两类研究范式下诱发的欺骗行为却大多不具备这两个属性。在指导性的欺骗任务中, 欺骗导致的大脑激活模式主要反映抑制真实反应和产生欺骗反应这一认知过程, 虽然这一过程确实是欺骗的必要成分, 但现实生活中的欺骗远远不止是简单的产生一个错误的反应, 还涉及非常复杂的社会决策和人际交互过程, 包括情绪、注意、记忆、推理和决策等多种心理过程的参与(Kireev et al., 2012)。为了考察更为真实的欺骗行为, 一些研究通过各种巧妙的实验设计或新型范式, 在人际交互情境下诱发接近现实生活的欺骗行为, 并对其神经机制进行揭示。

在这些研究中, 一部分研究采用类似于指导性欺骗的范式, 它们仍然要求被试对经历过的事件或信息故意错误反应为不能识别, 只是放宽了被试的反应自由度, 使得他们可以自由选择是否以及何时进行欺骗反应, 结果发现, 这类自发欺骗行为也导致了前额叶皮层的更大激活(Spence, Kaylor-Hughes, Farrow, et al., 2008), 这与上述使用指导性欺骗范式的研究所发现的结果一致。但是, 这里的欺骗同样缺乏社会情境因素, 仍然停留在简单否认识别某些熟悉信息的形式上, 不具备现实欺骗行为所表现出的复杂性特点。为了模拟更为真实的欺骗行为, 另一部分研究采用游戏情景的方式, 通过与他人的交互作用过程诱发自

然的欺骗行为。Baumgartner 等人(2009)使用经济信任游戏(Economic trust game), 揭示了违背承诺这一欺骗行为的神经机制; Kireev 等人(2012)根据一个叫做“相信/不相信”的俄罗斯游戏, 考察了有意误导他人接受错误信息这一心理过程的神经机制; Sip 等人(2012)的研究关注欺骗的决策过程, 考察了当欺骗可能遭致受骗者的对质并遭受经济损失的情况下, 大脑如何进行欺骗决策。这部分的研究使用丰富多样的实验范式, 得出的结果差别很大, 难以统一, 也缺乏可比性。造成这一现象的原因包括以下两点: 一方面, 现实的欺骗行为本身就是千差万别, 在不同情景下具有完全不同形式和内容, 揭示这一几乎涉及所有心理过程的巨大课题, 势必需要复杂多变的实验任务; 另一方面, 目前对这一巨大课题的研究还处于最初的探索阶段, 采用发散式的研究思路, 开辟新的领域, 回答新的问题, 因此得到的结果非常零散。在目前的研究状况下, 更是无法进行个体诊断的尝试。尽管如此, 目前研究已经、并将继续获得许多有趣的发现, 这些发现将一步一步完善欺骗的认知和神经机制的理论框架, 更重要的是, 它们为未来研究启发了大量新思路, 预示其将成为一个很有生命力的课题。

## 2.2 测谎研究

### 2.2.1 GKT 范式

GKT 也称为 CIT (Concealed Information Test), 该范式被认为是在实验室情境下检测隐藏信息的黄金标准(Ben-Shakhar & Elaad, 2003)。测谎研究中通常使用的 GKT 范式由三类刺激的新异刺激范式(Oddball task)演变而来, GKT 范式包括以下三类刺激: 第一, 探测刺激(Probe)是和犯罪信息相关的刺激, 一般只有罪犯和职权者(警察)才知道这些信息, 而无辜者不知道这些信息。第二, 无关刺激(Irrelevant)是和犯罪行为没有任何关系的刺激, 无论是罪犯还是无辜者都无法识别这类刺激。第三, 目标刺激(Target)其实也是与犯罪完全无关的刺激, 但是要求被试对其执行某种不同于探测刺激和无关刺激的特定反应, 目的是维持被试的注意力(Ganis, Rosenfeld, Meixner, Kievit, & Schendan, 2011; Rosenfeld, Biroshak, & Furedy, 2006)。GKT 范式的测谎原理是: 虽然所有被试对探测刺激和无关刺激进行相同的行为反应, 但由于探测刺激对有罪被试具有特定意义, 会诱发

别反应。而对于无辜者而言, 他们无法区分探测刺激和无关刺激, 两类刺激将诱发相同的反应。因此, 以探测刺激和无关刺激诱发的电生理或神经反应的差异作为指标, 可以直接检测被试对这些信息的先前知识, 从而推测被试的欺骗行为。

GKT 早已被广泛运用于 ERP 的测谎研究中 (Winograd & Rosenfeld, 2011), Langleben 等人 (2002) 首次将 GKT 范式引入 fMRI 研究, 发现欺骗反应较之诚实反应导致了扣带前回 (ACC) 和额上回 (SFC, superior frontal cortex) 等参与执行控制的区域的更大激活, 说明可以使用 fMRI 信号对欺骗和诚实反应进行区分。后续的一些研究使用与其相似的范式, 针对以下几个问题进行了探讨: 第一, 根据 GKT 范式的测谎原理, 该范式会诱发被试对先前知识的识别反应 (Ben-Shakhar & Eyal, 2003), 因此在该范式下的欺骗行为所诱发的大脑激活模式其实是由欺骗和识别两个混淆成分共同导致的, 而并非单纯的欺骗过程所致 (Christ, van Essen, Watson, Brubaker, & McDermott, 2009)。针对这一问题, 出现了一系列后续研究以分离 GKT 范式所诱发的欺骗过程和识别过程的神经机制。其中, 一部分研究关注 GKT 范式所诱发的欺骗过程, 这些研究对 Langleben 等人 (2002) 所使用的实验范式进行改良, 比如在实验前发给被试两张扑克牌, 而在测谎阶段让其对一张扑克牌进行欺骗反应而对另一张扑克牌进行诚实反应, 通过这种方法排除识别过程对结果的影响, 以揭示单纯的欺骗过程的神经机制 (Davatzikos et al., 2005; Langleben et al., 2005; Luan Phan et al., 2005); 另一部分研究关注 GKT 范式所诱发的识别过程, 这些研究采用不需要被试进行欺骗反应的 GKT 范式, 结果发现探测刺激和无关刺激仍然导致不同的大脑激活模式, 这一激活模式反映两类刺激在记忆编码深度上的差异 (Gamer, Klimecki, Bauermann, Stoeter, & Vossel, 2012), 或反映刺激对个体的主观重要性 (显著性) (Furedy, 1986; Gamer et al., 2012)。此外, 这类研究还发现右腹外侧前额叶 (lateral ventral prefrontal cortex, VLPF) 参与对识别反应的抑制控制过程, 以该区域的激活信号为线索, 得到了较高的测谎准确率 (Nose et al., 2009)。第二, 使用 GKT 范式, 以欺骗反应和诚实反应之间存在差异的大脑激活模式为指标, 对欺骗行为进行个体诊断和个案研究 (Hakun et

al., 2009; Langleben et al., 2005)。第三, 结合反应时、多导仪和 fMRI 信号, 从多个渠道考察欺骗机制, 并探索几种信号各自反映的心理机制及其它它们之间的相互依赖关系 (Gamer, Bauermann, Stoeter, & Vossel, 2007; Gamer et al., 2012)。此外, Ganis 等人 (2011) 在以往的 ERP 研究 (Rosenfeld, Soskins, Bosh, & Ryan, 2004) 的基础上, 考察反策略的效应, 发现反策略大大降低了 fMRI 信号的测谎检出率, 从而指出了这一测谎方法可能存在的弊端。

### 2.2.2 CQT 范式

CQT (Comparison Question Test 或 Control Question Test) 也是传统测谎技术的一种标准范式。在 CQT 范式中, 刺激以问题的形式出现, 要求被试对每个问题进行“是”或“否”的回答。这些问题分为三种类型: 第一, 相关问题, 这类问题与要考察的事实有关, 用以诱发有罪被试的欺骗反应 (如, “你是否杀死了你的妻子?”); 第二, 比较问题 (控制问题), 无论对有罪还是无罪被试, 这类问题都可能诱发较强的心理反应 (如, “你曾经偷过东西吗?”); 第三, 无关问题, 相当于中性问题, 这类问题诱发的反应将作为基线 (如, “你现在坐在椅子上吗?”)。通常, 研究者用相关问题和控制问题所诱发的反应差异作为指标检测欺骗。也就是说, 被试对相关问题的反应如果持续强于控制问题, 则表明他/她可能在撒谎 (Wolpe, Foster, & Langleben, 2005)。

一些研究团队倾向于使用 CQT 范式。Kozel, Revell 等 (2004) 让被试从 10 个物体中寻找事先藏有现金的两个物体, 而在随后的测谎过程中要求被试对这两个藏有现金的物体分别进行诚实反应 (承认藏有现金) 和欺骗反应 (否认藏有现金), 结果发现欺骗反应激活了眶额皮层和扣带前回, 并且这两个脑区的激活强度和被试的皮肤电反应强度相关。随后, 该团队使用类似范式开展了一系列研究, 这些研究重复了 Kozel, Revell 等 (2004) 的结果, 但发现被试间的大脑激活模式存在很大的个体差异 (Kozel, Padgett, et al., 2004), 还有研究以欺骗的大脑激活模式为指标, 对欺骗行为进行了个体诊断 (Kozel et al., 2005; Kozel, Laken, et al., 2009), 并进一步验证了该方法对有罪和无辜被试的鉴别力度 (Kozel, Johnson, et al., 2009)。除了 Kozel 团队之外, Mohamed 等人 (2006) 在一个模

拟枪击情境后也使用 CQT 范式进行了测谎研究,他们以多导仪记录的生理指标对“有罪”和“无辜”被试进行区分,同时记录欺骗反应所导致的大脑激活模式。这些研究一致的观察到欺骗反应较之诚实反应导致了前额叶或扣带皮层区域的更大激活,使用大脑执行系统的激活信号对欺骗行为进行个体诊断时,可获得大概 90% 的敏感度,但特异度却低至 33% (Kozel, Johnson, et al., 2009)。

作为两种迫选范式, GKT 和 CQT 都具有简化分析过程的优点,因此目前在个体水平上检测欺骗的 fMRI 研究大多使用这两种范式。但其实两种范式基于完全不同的原理: CQT 范式诱发的是欺骗行为,它通过测量与欺骗过程相关的神经信号,直接检测欺骗行为。而 GKT 主要诱发的是对先前知识的识别过程,同时被试需要对探测刺激进行欺骗反应,由此检测到的神经信号可能来自两个心理成分,即对先前知识的识别过程和否认该识别过程的欺骗过程。那么,在实验室研究和实际应用中,哪一种范式更具优势?这一问题一直存在激烈的争议。在方法学上, GKT 范式诱发大脑激活信号的效应较之 CQT 范式更为明显 (Wolpe et al., 2005); 在测谎原理上, GKT 范式主要检测刺激对被试的主观意义性,而非检测欺骗,相比而言, CQT 更像是一种真正意义上的“测谎”范式。在实际应用的可行性上,两种范式在材料搜集上都存在困难。CQT 范式难以设置标准化的控制问题,而 GKT 范式必须获得一些与案件有关的、并且只有执法者和罪犯知道的信息(犯罪知识),这一要求在现实的断案过程中其实很难满足,因而使得 GKT 范式的实际应用受到限制。

### 3 已有研究的主要结论

综上所述,目前的欺骗和测谎的 fMRI 研究使用了大量不同的实验范式,得出的结果也不尽相同。基于这些繁杂的研究结果,以下对两部分研究获得的主要结论进行简要的概括。

#### 3.1 欺骗的神经机制

目前研究在不同的实验范式下诱发欺骗行为并考察其神经机制,由此获得了纷繁复杂的结果,其中,最具一致性的结果发现:在行为水平上,欺骗反应较之诚实反应表现出反应时的增长;在神经水平上,欺骗反应较之诚实反应导致了大脑的执行控制系统,主要包括前额叶-顶叶和扣带前回

等区域的更大激活,反之,诚实反应较之欺骗反应没有导致任何脑区的更大激活 (Spence et al., 2004)。这些结果说明诚实反应在人类的认知和交流过程中,可能属于一种“基线式”的反应方式,而欺骗反应类似于某种执行任务,它需要更高级的神经中枢的参与,并且会耗费更多的认知资源 (Abe, 2009, 2011; Christ et al., 2009; Gombos, 2006; Hughes et al., 2005; Langleben, 2008; Spence, 2004, 2008; Spence et al., 2004)。

然而,除了执行控制功能之外,人际交互情境下的欺骗行为还可能包含更多的社会认知过程。一方面,针对这类更具生态学效度的欺骗行为, Sip 等人 (2008) 对其的认知和神经机制进行了理论建构。他们认为人际交互情境下的欺骗主要涉及决策、心理理论和反应抑制这三大心理过程,因而这些心理过程所涉及的脑区也将参与到欺骗过程中。具体而言,欺骗的决策阶段需要风险评估和奖赏预期等心理过程的支持,这些心理过程将涉及内侧前额叶皮层、扣带前回、眶额皮层以及纹状体等脑区的参与;欺骗需要不断追踪受骗者的当前信念和心理状态,并需要执行反复的印象管理过程,这些过程都需要心理理论功能的支持,而心理理论过程将涉及腹内侧和背内侧前额皮层、扣带回的后部区域、以及颞上回和颞顶联合区等脑区的参与;欺骗需要抑制真实的反应倾向,以给欺骗反应提供便利,这需要反应抑制功能的支持,而反应抑制过程将涉及背外侧前额皮层和扣带回后部区域等广泛的执行控制系统的参与。以上三大心理过程所涉及的所有脑区构成欺骗行为的神经基础。另一方面,针对人际交互情景下的欺骗行为的实证研究获得了一些零散的、但是非常有趣的发现,这些发现在一定程度上为 Sip 等 (2008) 对欺骗机制的理论建构提供了实验证据。例如, Sip 等 (2012) 发现欺骗决策阶段的奖赏预期和风险厌恶过程涉及到尾状核和额下回的参与,而膝下扣带回 (subgenual ACC) 区域可能涉及欺骗决策中的道德判断过程。Baumgartner 等 (2009) 揭示了欺骗所涉及的情感成分的神神经机制,包括在欺骗的决策阶段中 ACC 和双侧脑岛所涉及的决策冲突和愧疚体验过程,欺骗的执行阶段中杏仁核所承担的对欺骗行为的愧疚厌恶、对欺骗被发现的恐惧情绪体验,以及纹状体所涉及的动机驱动过程,该动机由欺骗行为

将带来的利益导致, 通过奖赏预期过程诱发。而 Bhatt 等(2010)揭示了欺骗者在执行复杂的印象管理过程时所需要的心理理论过程, 该过程尤其涉及到颞顶联合区的参与。可见, 欺骗所涉及的认知和神经机制是复杂多样的, 而不同类型的欺骗行为可能对这些认知和神经机制有不同程度的涉及。对于人际交互情境下的复杂欺骗行为, 目前还远远未能完善对其神经机制的理论建构。

### 3.2 fMRI 测谎的目前成绩

如上所述, 虽然欺骗涉及到复杂多样的认知和神经机制, 但目前 fMRI 的测谎研究大多利用执行控制相关脑区的激活为指标, 在个体水平上对欺骗行为进行反推, 使用这一方法进行测谎可达到接近 90% 的准确率(Kozel et al., 2005; Kozel, Laken, et al., 2009), 不过也有研究发现这一方法对有罪者和无辜者的区分能力不容乐观, 容易造成阳性错误(Kozel, Johnson, et al., 2009)。此外, 少数研究关注欺骗中的记忆效应(Gamer et al., 2012; Nose et al., 2009), 这些研究基于个体对隐藏信息的记忆效应的神经信号间接推测欺骗行为, 获得了接近 80% 的准确率(Nose et al., 2009)。

## 4 已有研究的问题与未来研究思路

### 4.1 目前研究的问题

尽管使用 fMRI 技术的欺骗和测谎研究取得了丰硕成果, 但这些研究仍然存在一些问题, 并面临着困境和挑战。许多研究者已对该领域目前所面临的问题进行了总结, 按照本文的整体逻辑, 我们将这些观点按照欺骗的机制探讨和测谎应用两个方面进行简要概括。在欺骗的机制探讨方面, 包括实验情景的非真实性(Bles & Haynes, 2008; Sip et al., 2008; Spence et al., 2004), 欺骗种类的复杂性(Bles & Haynes, 2008; Sip et al., 2008; Spence, 2008; Spence et al., 2004), 个体差异对欺骗的认知和神经过程的影响等问题(Bles & Haynes, 2008; Sip et al., 2008; Spence et al., 2004); 在测谎技术方面, 包括从大脑激活模式反推欺骗行为的逻辑问题(Bles & Haynes, 2008; Langleben, 2008; Sip et al., 2008), 组分析结果用于个体诊断的方法学问题(Abe, 2011; Bles & Haynes, 2008; Spence et al., 2004; Spence & Kaylor-Hughes, 2008; Wolpe et al., 2005), 实验室测谎技术到司法应用的推广问题(Abe, 2011; Bles & Haynes, 2008;

Langleben, 2008; Sip et al., 2008; Spence et al., 2004; Spence & Kaylor-Hughes, 2008; Wolpe et al., 2005), 以及测谎可能导致的各种社会道德和伦理问题等等(Bles & Haynes, 2008; Langleben, 2008; Spence et al., 2004; Spence & Kaylor-Hughes, 2008; Wolpe et al., 2005)。这些观点在不同研究者的文章中、被不同程度的重复提及过, 属于目前该领域的公认观点, 本文不再赘述。以下, 我们针对机制探讨所忽视的几点具体问题和测谎研究所面临的困境, 提出未来的研究思路。

#### 4.1.1 欺骗机制研究较为忽视的几点问题

##### (1) 欺骗中的人格特质问题

近年来, 研究者开始逐渐关注欺骗行为中的个体差异的神经基础。Fullam 等(2009)发现欺骗所涉及的大脑激活模式与个体在精神病态特质量表 PPI 上的得分相关。M. A. Bhatt 等(2010)揭示了欺骗策略上的个体差异, 发现不同的欺骗策略导致了不同的大脑激活模式。Greene 等(2009)发现, 倾向于选择欺骗行为的个体与倾向于选择诚实行为的个体在行为的决策和执行阶段中都表现出不同的大脑激活模式, 暗示个体在欺骗特质上的差异可能存在其神经基础。但是, 这些研究都只关注了欺骗执行过程中的个体差异, 这是个体在欺骗行为上的状态性差异, 而非具有稳定倾向性的特质性差异。至今罕有研究考察欺骗特质的神经基础。Abe (2011)的综述中提到, 最近研究发现纹状体 D2 受体的易得性与个体在欺骗相关的人格特质上的得分呈负相关, 这可能揭示了欺骗/忠诚这一人格特质的神经基础(Cervenka, Gustavsson, Halldin, & Farde, 2010; Egerton et al., 2010; Huang et al., 2006; Reeves et al., 2007)。但是, 这些研究使用的欺骗特质问卷其实是测量社会赞许性的问卷, 这与个体在认知任务中表现出来的欺骗行为大不相同(Abe, 2011), 可以想见, 由此获得的结果将很难预测个体在现实情境下的自发欺骗行为。那么, 是否存在标志欺骗特质的神经信号? 这些信号能否预测个体在诱惑情景下的自发欺骗行为? 目前, 还没有研究能回答这些问题。

##### (2) 欺骗中的记忆痕迹问题

以往研究探讨了隐藏信息在大脑中的记忆效应。这类研究大多使用不需要被试进行欺骗反应的 GKT 范式, 以考察大脑是否会对隐藏信息产生独特的激活模式(Gamer et al., 2012; Nose et al.,

2009)。结果发现,即使当被试不需要对隐藏信息进行欺骗反应时,这些隐藏信息仍然会自动诱发广泛大脑皮层的更大激活,包括顶下小叶、腹外侧前额叶皮层、乃至运动辅助区的更大激活,研究者认为这些脑区参与记忆相关的过程,可能反映个体对刺激的编码深度(Gamer et al., 2012)。然而,这些区域其实在其他很多认知过程中都有涉及,比如顶下小叶实际上与刺激的频率或主观重要性更为相关(Downar, Crawley, Mikulis, & Davis, 2000, 2001),而腹外侧前额叶皮层更多时候被认为参与反应抑制和控制过程(Blasi et al., 2006; Dove, Manly, Epstein, & Owen, 2008; Lie, Specht, Marshall, & Fink, 2006; Ridderinkhof, van den Wildenberg, Segalowitz, & Carter, 2004)。可见,这些脑区可能并非直接反映个体对刺激的记忆效应,而是反映刺激对个体的主观重要性,以及个体在面临这些重要信息时的抑制控制过程(Nose et al., 2009)。但是,以往研究将所有这些区域的激活笼统的归结为记忆效应,而所谓的记忆效应是一个非常模糊的概念。因此,至今还没有研究直接揭示欺骗过程中标记记忆痕迹的神经线索,也没有研究直接利用这一线索对真相进行反推。

#### (3) 人际交互情景对欺骗执行阶段的影响问题

如上所述,目前对欺骗的研究主要考察了欺骗的执行阶段,而关于欺骗执行阶段的神经机制的结论还主要停留在对大脑执行控制系统的关注,对于欺骗所涉及的其他诸如风险评估、心理理论和情绪调节等过程的有关发现还相当匮乏。导致这一状况的原因一方面是因为大多数实验室所诱发的欺骗缺少足够的人际交互的社会情景,另一方面是因为大多数研究忽视了欺骗的后果在现实欺骗行为中所扮演的重要角色,而欺骗后果往往是欺骗所处的社会情景中的必要要素。

“在欺骗不需要面临后果的情境下,根本无法获得有关欺骗行为的真实表征(Sip et al., 2008)。”可见,是否需要面临后果将直接影响欺骗行为的本质属性,而目前还没有研究探讨欺骗的结果加工阶段对执行阶段的影响作用。揭示这一影响过程的认知和神经机制将有助于我们理解更为真实的欺骗行为。

#### (4) 欺骗中的结果加工问题

欺骗是一种目标定向行为,欺骗结果在欺骗过程中扮演非常重要的角色,它既是诱发欺骗行

为的动力也为修正欺骗策略提供了重要的反馈信息,欺骗结果的加工过程本身属于欺骗过程的一部分,而目前还没有研究揭示欺骗结果加工过程的认知和神经机制。Fang, Liu 和 Shen (2003)以及孙世月和罗跃嘉(2008)使用 ERP 技术在有反馈的范式下考察欺骗行为,得出了非常有趣的实验结果,但他们没有考察结果加工阶段的大脑反应。崔茜和张庆林(2013)同样使用 ERP 技术,虽然考察了欺骗的结果加工过程,但并没有探讨欺骗行为对结果加工过程带来的效应,也不能回答结果加工阶段的大脑激活信号可否用于测谎这一问题,并且这些研究都受到 ERP 的低空间分辨率特点的限制,因此对该阶段的神经基础的揭示非常有限。对欺骗行为后所伴随结果的加工机制是否异于对诚实行为后所伴随结果的加工机制?两者在认知和神经机制上有什么差别?目前还没有研究对这一问题进行探讨。

#### 4.1.2 fMRI 测谎研究面临的困境

目前 fMRI 的测谎研究主要基于执行控制过程以及记忆效应相关的 fMRI 信号推测欺骗过程,虽然一些研究报道使用此方法检测欺骗行为可以获得接近 90% 的正确率(Davatzikos et al., 2005; Kozel et al., 2005; Kozel, Laken, et al., 2009; Langleben et al., 2005; Nose et al., 2009),但也有研究报道该方法可能造成很高的阳性错误率(Kozel, Johnson, et al., 2009)。因此,对现有的 fMRI 测谎技术在司法中的应用前景很多研究者仍然持悲观态度。研究者们总结了造成该困境的原因,主要的观点认为这一困境是由于欺骗的复杂多变性导致的。首先,欺骗本身包含复杂多样的认知过程,如风险评估、奖赏预期、心理理论、反应抑制和认知控制以及情绪调节等等,并且这些认知过程本身已经具备相当的复杂性,它们中的任何一项都可以形成心理学领域的一大研究课题,加之欺骗还具有多种类型,不同欺骗行为可能涉及完全不同的认知和神经机制。由此可见,执行控制过程远远不足以刻画真实的欺骗行为,也很难寻找到某一简单的神经指标来对欺骗行为进行全面的刻画(Bles & Haynes, 2008; Sip et al., 2008)。为了解决这一困境,Sip 等(2008)提出绕过寻找欺骗的神经指标这一难题,而转入对其他更为具体的、可能与欺骗过程相关的认知过程入手的观点(Sip et al., 2008)。这正好与上述机制研究



存在的问题不谋而合, 如果能够解决以上提出的四点问题, 也许能够找到更为有效的测谎指标。基于以上四点问题, 以下提出未来的研究思路。

## 4.2 未来研究思路

### 4.2.1 欺骗特质的神经基础及应用

如果可以找到标记个体欺骗特质的神经信号, 基于这一信号便可能预测个体的欺骗行为。对该课题的探索可以启发未来研究发明“忠诚度”的测量工具, 从而从源头上减少欺骗行为造成的伤害。

揭示个体特质的神经编码是目前神经科学研究的热点和难点(Adelstein et al., 2011; Canli & Amin, 2002; de Young & Gray, 2009), 静息状态下的功能连接方法(rest-state functional connectivity, RSFC)为解决这一问题提供了契机。该方法可以在不需要任何认知活动的情况下揭示个体的神经网络的本质特征, 并有大量研究发现个体的 RSFC 模式可以反映个体的人格以及多种行为倾向性特征(Adelstein et al., 2011; Cox et al., 2010), 包括社会技能(Di Martino et al., 2009), 风险寻求(Cox et al., 2010), 工作记忆(Hampson, Driesen, Skudlarski, Gore, & Constable, 2006), 情景记忆(Ystad, Eichele, Lundervold, & Lundervold, 2010), 攻击性(Hoptman et al., 2010)和认知效能(Andrews-Hanna et al., 2007)等。RSFC 方法可有效地揭示大脑和行为之间的稳定关系(Allemand, Zimprich, & Hendriks, 2008; John, Naumann, & Soto, 2008; Shehzad et al., 2009; van Dijk et al., 2010; Zuo et al., 2010), 探索欺骗特质的神经基础的未来研究可以利用 RSFC 方法的这一优势, 并可尝试使用 RSFC 指标对个体的欺骗行为进行预测。

### 4.2.2 记忆痕迹的神经基础及应用

经历过的事物都会在大脑中留下痕迹, 而欺骗的目的正是有意向他人传递与自己所知的真相不符的错误信念(Abe, 2011; Vrij, 2000)。欺骗掩盖了事实真相, 而人们对欺骗的反感情绪以及对测谎技术的热衷正是源自于对还原事实真相的迫切希望。既然如此, 是否可以直接寻找大脑中储存事实真相的神经指标? Bles 和 Haynes (2008)同样提出了这一观点, 并认为从直接检测大脑的隐藏信息入手, 可能取得更好的测谎效果。然而, 虽然目前已有研究探讨了隐藏信息在大脑上的记忆效

应, 但还没有研究直接检测标志记忆痕迹的神经信号。

来自记忆领域的研究已经证实了内侧颞叶(MTL, medial temporal lobule)在记忆中的核心作用(Squire, Clark, & Bayley, 2004; Squire & Zola-Morgan, 1991)。在 Messinger, Squire, Zola 和 Albright (2005)的研究中, 他们让猴子对配对刺激进行学习, 再在测试阶段完成视觉匹配任务, 该任务要求猴子从两个刺激中选出目标刺激的正确配对刺激, 实验结果发现, 无论猴子是否做出正确的行为选择, 其颞下回(inferior temporal, IT)区域的神经元对正确的配对刺激(旧刺激)表现出更强的电位反应, 说明其大脑的记忆系统能够独立于行为反应、在客观上区分新旧刺激。随后, Daselaar 等人(2006)在人类大脑上也找到了能够标志刺激的客观新旧属性的 MTL 区域。他们发现, MTL 的后部区域(pMTL, posterior MTL)能够客观的区分词语的新、旧属性, 表现出对旧词有更大的激活反应, 哪怕被试在意识水平上并不能对新、旧词语进行正确的分类。Kirwan, Shrager 和 Squire (2009)的研究在更严谨的实验设计和更详尽的数据分析下进一步验证了 Daselaar 等人(2006)的发现, 他们的实验结果证实 MTL 确实能够客观标记个体对刺激的记忆痕迹。以上这些研究的发现都支持目前有关 MTL 认知功能的模型, 该模型认为: 在呈现记忆线索时, MTL 会对大脑中的记忆痕迹产生自动的激活反应(Eichenbaum, 2004; Moscovitch, 1995; Squire, Stark, & Clark, 2004), 因此表现出记忆痕迹效应。记忆研究领域的这些发现为测谎研究提供了重要启示: 是否可以使用 pMTL 的激活信号直接检测个体对某些重要信息的记忆痕迹, 从而直接推知真相? 至今还未有研究对这一想法进行考察。

### 4.2.3 欺骗结果对执行阶段的影响作用

欺骗执行阶段的神经信号已经获得了以往研究的充分关注(Spence, 2008), 其中, 涉及执行控制过程的大脑区域的激活信号是目前用于 fMRI 测谎技术的主要神经指标。但是, 这些研究大多使用不存在欺骗后果的欺骗范式, 而是否存在欺骗后果将直接影响欺骗行为的本质属性。为了揭示这一问题, 最简单的方法可能是直接比较在有、无欺骗后果情景下欺骗执行过程的大脑激活模式, 从而揭示欺骗后果对执行阶段的影响作

用。此外,还可以比较在有、无后果两种情景下欺骗执行阶段所诱发信号的测谎效力,以进一步理解该信号用于测谎时的适用条件,从而提高该指标在应用中的有效性。

#### 4.2.4 欺骗结果加工的神经基础及应用

中国俗语“做贼心虚”描述了人们在做坏事后因为害怕别人知道所表现出的不安心理(夏征农, 2009),此时个体会对外界针对自己的评价表现出异常的敏感性。欺骗作为一种不被社会认可的行为,执行欺骗行为后也应该导致这一心理。因此,加工欺骗行为所伴随的结果可能会诱发独特的大脑激活模式。未来研究可以在现有的欺骗范式中加入结果阶段,使用 fMRI 技术系统考察大脑如何加工欺骗行为后跟随的具有不同效价和不同意义的结果,从而详细揭示欺骗的结果加工过程的认知和神经机制。对这一问题的探讨同样可揭示结果加工阶段中与欺骗相关的神经信号,以间接推测欺骗行为。

#### 参考文献

- 崔茜,张庆林.(2013).测谎中反馈加工的认知神经机制:一项ERP研究.《心理科学》,36,61-66.
- 孙世月,罗跃嘉.(2008).欺骗任务中结果评价的FN效应.《心理学报》,40,693-700.
- 夏征农.(2009).《大辞海》.上海:辞书出版社.
- Abe, N. (2009). The neurobiology of deception: Evidence from neuroimaging and loss-of-function studies. *Current Opinion in Neurology*, 22, 594-600.
- Abe, N. (2011). How the Brain Shapes Deception. *The Neuroscientist*, 17, 560-574.
- Abe, N., Okuda, J., Suzuki, M., Sasaki, H., Matsuda, T., Mori, E.,... Fujii, T. (2008). Neural correlates of true memory, false memory, and deception. *Cerebral Cortex*, 18, 2811-2819.
- Adelstein, J. S., Shehzad, Z., Mennes, M., DeYoung, C. G., Zuo, X. -N., Kelly, C.,... Castellanos, F. X. (2011). Personality is reflected in the brain's intrinsic functional architecture. *Plos One*, 6, e27633.
- Allemand, M., Zimprich, D., & Hendriks, A. (2008). Age differences in five personality domains across the life span. *Developmental Psychology*, 44, 758-770.
- Andrews-Hanna, J. R., Snyder, A. Z., Vincent, J. L., Lustig, C., Head, D., Raichle, M. E., & Buckner, R. L. (2007). Disruption of large-scale brain systems in advanced aging. *Neuron*, 56, 924-935.
- Baumgartner, T., Fischbacher, U., Feierabend, A., Lutz, K., & Fehr, E. (2009). The neural circuitry of a broken promise. *Neuron*, 64, 756-770.
- Ben-Shakhar, G., & Eyal, E. (2003). The validity of psychophysiological detection of information with the Guilty Knowledge Test: A meta-analytic review. *Journal of Applied Psychology*, 88, 131-151.
- Bhatt, M. A., Lohrenz, T., Camerer, C. F., & Montague, P. R. (2010). Neural signatures of strategic types in a two-person bargaining game. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 19720-19725.
- Bhatt, S., Mbwana, J., Adeyemo, A., Sawyer, A., Hailu, A., & van Meter, J. (2009). Lying about facial recognition: An fMRI study. *Brain and Cognition*, 69, 382-390.
- Blasi, G., Goldberg, T. E., Weickert, T., Das, S., Kohn, P., Zolnick, B.,... Mattay, V. S. (2006). Brain regions underlying response inhibition and interference monitoring and suppression. *European Journal of Neuroscience*, 23, 1658-1664.
- Bles, M., & Haynes, J. D. (2008). Detecting concealed information using brain-imaging technology. *Neurocase*, 14, 82-92.
- Browndyke, J. N., Paskavitz, J., Sweet, L. H., Cohen, R. A., Tucker, K. A., Welsh-Bohmer, K. A.,... Schmechel, D. E. (2008). Neuroanatomical correlates of malingered memory impairment: Event-related fMRI of deception on a recognition memory task. *Brain Injury*, 22, 481-489.
- Canli, T., & Amin, Z. (2002). Neuroimaging of emotion and personality: Scientific evidence and ethical considerations. *Brain and Cognition*, 50, 414-431.
- Cervenka, S., Gustavsson, J. P., Halldin, C., & Farde, L. (2010). Association between striatal and extrastriatal dopamine D2-receptor binding and social desirability. *Neuroimage*, 50, 323-328.
- Christ, S. E., van Essen, D. C., Watson, J. M., Brubaker, L. E., & McDermott, K. B. (2009). The contributions of prefrontal cortex and executive control to deception: Evidence from activation likelihood estimate meta-analyses. *Cerebral Cortex*, 19, 1557-1566.
- Cox, C. L., Gotimer, K., Roy, A. K., Castellanos, F. X., Milham, M. P., & Kelly, C. (2010). Your resting brain CAREs about your risky behavior. *Plos One*, 5, e12296.
- Daselaar, S. M., Fleck, M. S., Prince, S. E., & Cabeza, R. (2006). The medial temporal lobe distinguishes old from new independently of consciousness. *Journal of Neuroscience*, 26, 5835-5839.
- Davatzikos, C., Ruparel, K., Fan, Y., Shen, D. G., Acharyya, M., Loughhead, J. W.,... Langleben, D. D. (2005). Classifying spatial patterns of brain activity with machine learning methods: Application to lie detection. *NeuroImage*, 28, 663-668.

- Derksen, M. (2012). Control and resistance in the psychology of lying. *Theory & Psychology, 22*, 196–212.
- de Young, C. G., & Gray, J. R. (2009). *Personality neuroscience: Explaining individual differences in affect, behaviour and cognition* (pp. 323–346). New York: Cambridge University.
- Di Martino, A., Shehzad, Z., Kelly, C., Roy, A., Gee, D., Uddin, L.,... Milham, M. (2009). Relationship between cingulo-insular functional connectivity and autistic traits in neurotypical adults. *American Journal of Psychiatry, 166*, 891–899.
- Dove, A., Manly, T., Epstein, R., & Owen, A. M. (2008). The engagement of mid-ventrolateral prefrontal cortex and posterior brain regions in intentional cognitive activity. *Human Brain Mapping, 29*, 107–119.
- Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2000). A multimodal cortical network for the detection of changes in the sensory environment. *Nature Neuroscience, 3*, 277–283.
- Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2001). The effect of task relevance on the cortical response to changes in visual and auditory stimuli: An event-related fMRI study. *Neuroimage, 14*, 1256–1267.
- Egerton, A., Rees, E., Bose, S. K., Lappin, J. M., Stokes, P. R., Turkheimer, F. E., & Reeves, S. J. (2010). Truth, lies or self-deception? Striatal D<sub>2/3</sub> receptor availability predicts individual differences in social conformity. *Neuroimage, 53*, 777–781.
- Eichenbaum, H. (2004). Hippocampus: Cognitive processes and neural representations that underlie declarative memory. *Neuron, 44*, 109–120.
- Fang, F., Liu, Y. T., & Shen, Z. (2003). Lie detection with contingent negative variation. *International Journal of Psychophysiology, 50*, 247–255.
- Fullam, R. S., McKie, S., & Dolan, M. C. (2009). Psychopathic traits and deception: Functional magnetic resonance imaging study. *The British Journal of Psychiatry, 194*, 229–235.
- Furedy, J. J. (1986). *Lie detection as psychophysiological differentiation: Some fine lines* (pp. 683–701). New York: Guilford.
- Gamer, M., Bauermann, T., Stoeter, P., & Vossel, G. (2007). Covariations among fMRI, skin conductance, and behavioral data during processing of concealed information. *Human Brain Mapping, 28*, 1287–1301.
- Gamer, M., Klimecki, O., Bauermann, T., Stoeter, P., & Vossel, G. (2012). fMRI-activation patterns in the detection of concealed information rely on memory-related effects. *Social Cognitive and Affective Neuroscience, 7*, 506–515.
- Gamer, M., Verschuere, B., Crombez, G., & Vossel, G. (2008). Combining physiological measures in the detection of concealed information. *Physiology & Behavior, 95*, 333–340.
- Ganis, G., Morris, R. R., & Kosslyn, S. M. (2009). Neural processes underlying self-and other-related lies: An individual difference approach using fMRI. *Social Neuroscience, 4*, 539–553.
- Ganis, G., Rosenfeld, J. P., Meixner, J., Kievit, R. A., & Schendan, H. E. (2011). Lying in the scanner: Covert countermeasures disrupt deception detection by functional magnetic resonance imaging. *NeuroImage, 55*, 312–319.
- Gombos, V. A. (2006). The cognition of deception: The role of executive processes in producing lies. *Genetic, Social, and General Psychology Monographs, 132*, 197–214.
- Greene, J. D., & Paxton, J. M. (2009). Patterns of neural activity associated with honest and dishonest moral decisions. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 12506–12511.
- Hakun, J. G., Ruparel, K., Seelig, D., Busch, E., Loughead, J. W., Gur, R. C., & Langleben, D. D. (2009). Towards clinical trials of lie detection with fMRI. *Social Neuroscience, 4*, 518–527.
- Hampson, M., Driesen, N. R., Skudlarski, P., Gore, J. C., & Constable, R. T. (2006). Brain connectivity related to working memory performance. *Journal of Neuroscience, 26*, 13338–13343.
- Hoptman, M. J., D'Angelo, D., Catalano, D., Mauro, C. J., Shehzad, Z. E., Kelly, A. C.,... Milham, M. P. (2010). Amygdalofrontal functional disconnectivity and aggression in schizophrenia. *Schizophrenia Bulletin, 36*, 1020–1028.
- Huang, C. L., Yang, Y. K., Chu, C. L., Lee, I. H., Yeh, T. L., Chen, P. S., & Chiu, N. T. (2006). The association between the Lie scale of the Maudsley personality inventory and striatal dopamine D<sub>2</sub>/D<sub>3</sub> receptor availability of healthy Chinese community subjects. *European Psychiatry, 21*, 62–65.
- Hughes, C. J., Farrow, T. F. D., Hopwood, M. C., Pratt, A., Hunter, M. D., & Spence, S. A. (2005). Recent developments in deception research. *Current Psychiatry Reviews, 1*, 273–279.
- Ito, A., Abe, N., Fujii, T., Ueno, A., Koseki, Y., Hashimoto, R.,... Mori, E. (2011). The role of the dorsolateral prefrontal cortex in deception when remembering neutral and emotional events. *Neuroscience Research, 69*, 121–128.
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). *Paradigm shift to the integrative Big Five trait taxonomy: History,*

- measurement, and conceptual issues. New York: Guilford.
- Kaylor-Hughes, C. J., Lankappa, S. T., Fung, R., Hope-Urwin, A. E., Wilkinson, I. D., & Spence, S. A. (2011). The functional anatomical distinction between truth telling and deception is preserved among people with schizophrenia. *Criminal Behaviour and Mental Health, 21*, 8–20.
- Kireev, M., Korotkov, A., & Medvedev, S. (2012). Functional magnetic resonance study of deliberate deception. *Human Physiology, 38*, 32–39.
- Kirwan, C. B., Shrager, Y., & Squire, L. R. (2009). Medial temporal lobe activity can distinguish between old and new stimuli independently of overt behavioral choice. *Proceedings of the National Academy of Sciences, 106*, 14617–14621.
- Kozel, F. A., Johnson, K. A., Grenesko, E. L., Laken, S. J., Kose, S., Lu, X.,... George, M. S. (2009). Functional MRI Detection of Deception After Committing a Mock Sabotage Crime. *Journal of Forensic Sciences, 54*, 220–231.
- Kozel, F. A., Johnson, K. A., Mu, Q., Grenesko, E. L., Laken, S. J., & George, M. S. (2005). Detecting deception using functional magnetic resonance imaging. *Biological Psychiatry, 58*, 605–613.
- Kozel, F. A., Laken, S. J., Johnson, K. A., Boren, B., Mapes, K. S., Morgan, P. S., & George, M. S. (2009). Replication of functional MRI detection of deception. *The Open Forensic Science Journal, 2*, 6–11.
- Kozel, F. A., Padgett, T. M., & George, M. S. (2004). A replication study of the neural correlates of deception. *Behavioral Neuroscience, 118*, 852–856.
- Kozel, F. A., Revell, L. J., Lorberbaum, J. P., Shastri, A., Elhai, J. D., Horner, M. D.,... George, M. S. (2004). A pilot study of functional magnetic resonance imaging brain correlates of deception in healthy young men. *Journal of Neuropsychiatry and Clinical Neurosciences, 16*, 295–305.
- Lachmann, M., & Bergstrom, C. T. (2004). The disadvantage of combinatorial communication. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 271*, 2337–2343.
- Langleben, D. D. (2008). Detection of deception with fMRI: Are we there yet? *Legal and Criminological Psychology, 13*, 1–9.
- Langleben, D. D., Loughhead, J. W., Bilker, W. B., Ruparel, K., Childress, A. R., Busch, S. I., & Gur, R. C. (2005). Telling truth from lie in individual subjects with fast event-related fMRI. *Human Brain Mapping, 26*, 262–272.
- Langleben, D. D., Schroeder, L., Maldjian, J. A., Gur, R. C., McDonald, S., Ragland, J. D.,... Childress, A. R. (2002). Brain activity during simulated deception: An event-related functional magnetic resonance study. *NeuroImage, 15*, 727–732.
- Lee, T. M. C., Au, R. K. C., Liu, H. L., Ting, K., Huang, C. M., & Chan, C. C. H. (2009). Are errors differentiable from deceptive responses when feigning memory impairment? An fMRI study. *Brain and Cognition, 69*, 406–412.
- Lee, T. M. C., Lee, T. M. Y., Raine, A., & Chan, C. C. H. (2010). Lying about the valence of affective pictures: An fMRI study. *Plos One, 5*, e12291.
- Lee, T. M. C., Liu, H. -L., Tan, L. -H., Chan, C. C. H., Mahankali, S., Feng, C. -M.,... Gao, J. -H. (2002). Lie detection by functional magnetic resonance imaging. *Human Brain Mapping, 15*, 157–164.
- Lee, T. M. C., Liu, H. L., Chan, C. C. H., Ng, Y. B., Fox, P. T., & Gao, J. H. (2005). Neural correlates of feigned memory impairment. *NeuroImage, 28*, 305–313.
- Liang, C. Y., Xu, Z. Y., Mei, W., Wang, L. L., Xue, L., Lu, D. J., & Zhao, H. (2012). Neural correlates of feigned memory impairment are distinguishable from answering randomly and answering incorrectly: An fMRI and behavioral study. *Brain and Cognition, 79*, 70–77.
- Lie, C. H., Specht, K., Marshall, J. C., & Fink, G. R. (2006). Using fMRI to decompose the neural processes underlying the Wisconsin Card Sorting Test. *Neuroimage, 30*, 1038–1049.
- Luan Phan, K., Magalhaes, A., Ziemlewicz, T. J., Fitzgerald, D. A., Green, C., & Smith, W. (2005). Neural correlates of telling lies: A functional magnetic resonance imaging study at 4 Tesla. *Academic Radiology, 12*, 164–172.
- Messinger, A., Squire, L. R., Zola, S. M., & Albright, T. D. (2005). Neural correlates of knowledge: Stable representation of stimulus associations across variations in behavioral performance. *Neuron, 48*, 359–371.
- Mohamed, F. B., Faro, S. H., Gordon, N. J., Platek, S. M., Ahmad, H., & Williams, J. M. (2006). Brain mapping of deception and truth telling about an ecologically valid situation: Functional MR imaging and polygraph investigation-Initial experience. *Radiology, 238*, 679–688.
- Moscovitch, M. (1995). Recovered consciousness: A hypothesis concerning modularity and episodic memory. *Journal of Clinical and Experimental Neuropsychology, 17*, 276–290.
- Nose, I., Murai, J., & Taira, M. (2009). Disclosing concealed information on the basis of cortical activations. *NeuroImage, 44*, 1380–1386.
- Nunez, J. M., Casey, B., Egner, T., Hare, T., & Hirsch, J. (2005). Intentional false responding shares neural

- substrates with response conflict and cognitive control. *NeuroImage*, 25, 267–277.
- Press, N. A. (2003). *The Polygraph and Lie Detection. Committee to Review the Scientific Evidence of the Polygraph.*: Division of Behavioral and Social Sciences and Education. Washington, D. C.: National Academies Press.
- Reeves, S. J., Mehta, M. A., Montgomery, A. J., Amiras, D., Egerton, A., Howard, R. J., & Grasby, P. M. (2007). Striatal dopamine (D2) receptor availability predicts socially desirable responding. *Neuroimage*, 34, 1782–1789.
- Ridderinkhof, K. R., van den Wildenberg, W. P. M., Segalowitz, S. J., & Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: The role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain and Cognition*, 56, 129–140.
- Rosenfeld, J. P., Biroshak, J. R., & Furedy, J. J. (2006). P300-based detection of concealed autobiographical versus incidentally acquired information in target and non-target paradigms. *International Journal of Psychophysiology*, 60, 251–259.
- Rosenfeld, J. P., Soskins, M., Bosh, G., & Ryan, A. (2004). Simple, effective countermeasures to P300-based tests of detection of concealed information. *Psychophysiology*, 41, 205–219.
- Shehzad, Z., Kelly, A. C., Reiss, P. T., Gee, D. G., Gotimer, K., Uddin, L. Q.,... Biswal, B. B. (2009). The resting brain: Unconstrained yet reliable. *Cerebral Cortex*, 19, 2209–2229.
- Sip, K. E., Roepstorff, A., McGregor, W., & Frith, C. D. (2008). Detecting deception: The scope and limits. *Trends in Cognitive Sciences*, 12, 48–53.
- Sip, K. E., Skewes, J. C., Marchant, J. L., McGregor, W. B., Roepstorff, A., & Frith, C. D. (2012). What if I get busted? Deception, choice, and decision-making in social interaction. *Frontiers in Neuroscience*, 6, 58.
- Spence, S. A. (2004). The deceptive brain. *Journal of the Royal Society of Medicine*, 97, 6–9.
- Spence, S. A. (2008). Playing Devil's advocate: The case against fMRI lie detection. *Legal and Criminological Psychology*, 13, 11–25.
- Spence, S. A., Farrow, T. F. D., Herford, A. E., Wilkinson, I. D., Zheng, Y., & Woodruff, P. W. R. (2001). Behavioural and functional anatomical correlates of deception in humans. *Neuroreport*, 12, 2849–2853.
- Spence, S. A., Hunter, M. D., Farrow, T., Green, R. D., Leung, D. H., Hughes, C. J., & Ganesan, V. (2004). A cognitive neurobiological account of deception: Evidence from functional neuroimaging. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359, 1755–1762.
- Spence, S. A., & Kaylor-Hughes, C. J. (2008). Looking for truth and finding lies: The prospects for a nascent neuroimaging of deception. *Neurocase*, 14, 68–81.
- Spence, S. A., Kaylor-Hughes, C. J., Brook, M. L., Lankappa, S. T., & Wilkinson, I. D. (2008). 'Munchausen's syndrome by proxy' or a 'miscarriage of justice'? An initial application of functional neuroimaging to the question of guilt versus innocence. *European Psychiatry*, 23, 309–314.
- Spence, S. A., Kaylor-Hughes, C. J., Farrow, T. F. D., & Wilkinson, I. D. (2008). Speaking of secrets and lies: The contribution of ventrolateral prefrontal cortex to vocal deception. *NeuroImage*, 40, 1411–1418.
- Squire, L. R., Clark, R. E., & Bayley, P. J. (2004). *Medial Temporal Lobe Function and Memory.* Gazzaniga: MIT Press, 691–708.
- Squire, L. R., Stark, C. E., & Clark, R. E. (2004). The medial temporal lobe. *Annual Review of Neuroscience*, 27, 279–306.
- Squire, L. R., & Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science*, 253, 1380–1386.
- van Dijk, K. R., Hedden, T., Venkataraman, A., Evans, K. C., Lazar, S. W., & Buckner, R. L. (2010). Intrinsic functional connectivity as a tool for human connectomics: Theory, properties, and optimization. *Journal of Neurophysiology*, 103, 297–321.
- Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and the implications for professional practice.* Chichester, England: Wiley.
- Vrij, A. (2004). Guidelines to catch a liar. In: *The detection of deception in forensic contexts* (pp. 287–314). Cambridge: Cambridge University Press.
- Winograd, M. R., & Rosenfeld, J. P. (2011). Mock crime application of the Complex Trial Protocol (CTP) P300-based concealed information test. *Psychophysiology*, 48, 155–161.
- Wolpe, P. R., Foster, K. R., & Langleben, D. D. (2005). Emerging neurotechnologies for lie-detection: Promises and perils. *The American Journal of Bioethics*, 5, 39–49.
- Ystad, M., Eichele, T., Lundervold, A. J., & Lundervold, A. (2010). Subcortical functional connectivity and verbal episodic memory in healthy elderly-A resting state fMRI study. *Neuroimage*, 52, 379–388.
- Zuo, X. -N., Kelly, C., Adelman, J. S., Klein, D. F., Castellanos, F. X., & Milham, M. P. (2010). Reliable intrinsic connectivity networks: Test-retest evaluation using ICA and dual regression approach. *Neuroimage*, 49, 2163–2177.

## Neural Mechanism of Deception and Its Application to Lie Detection: Evidences from fMRI Studies

CUI Qian; JIANG Jun; YANG Wenjing; ZHANG Qinglin

*(Faculty of Psychology, Southwest University;*

*Key Laboratory of Cognition and Personality (SWU), Ministry of Education, Chongqing 400715, China)*

**Abstract:** It is a hot research spot that investigating the neural mechanism of deception using functional magnetic resonance imaging and detecting lie based on its findings. Various experimental paradigms are adopted for revealing the cognitive and neural mechanism of feigned memory impairment, intentional error, and interpersonal deception. Furthermore, traditional GKT and CQT paradigms are used for detecting lies. Previous studies have emphasized the core roles of the prefrontal and parietal cortices implicating in cognitive control process in deception, however, lie detection based on the activities in these regions poses insufficient accuracy rates. Aside from highlighting executive control process, future research should focus on other cognitive components of deception and seek new indications of lie detection from memory traces and deception outcome, and reveal the neural basis of deception trait.

**Key words:** deception; lie detection; fMRI