中国科学: 生命科学 2014 年 第 44 卷 第 9 期: 938 ~ 950
SCIENTIA SINICA Vitae www.scichina.com life.scichina.com
《中国科学》杂志社
SCIENCE CHINA PRESS

论 文

有无反馈对欺骗过程的神经机制的调控作用

崔茜[①②③], 雷旭[②], 杨文静[②], 王一峰[③], 张庆林[②*]

① 电子科技大学政治与公共管理学院, 成都 610054;
② 西南大学心理学部, 重庆 400715;
③ 电子科技大学生命科学与技术学院, 神经信息教育部重点实验室, 成都 610054
* 联系人, E-mail: zhangql@swu.edu.cn

收稿日期: 2014-02-03; 接受日期: 2014-05-05
国家自然科学基金(批准号: 30970892, 31170983)资助项目
doi: 10.1360/052013-59

摘要 欺骗行为会导致欺骗结果的产生, 欺骗结果又会直接影响欺骗行为的发生及其内在机制. 虽然有研究表明, 欺骗结果会对相应的欺骗过程产生调控作用, 但对这一调控作用的机制并不清楚. 本研究采用功能核磁共振技术, 对两组被试分别使用有、无反馈(欺骗结果)的 GKT 范式并记录两组被试在进行诚实反应和欺骗反应时的大脑激活模式. 结果发现, 有反馈组与无反馈组相比, 有反馈组的诚实反应和欺骗反应都导致了左侧顶叶皮层、左背部前扣带皮层、左侧脑岛、双侧视皮层和右侧小脑的更大激活; 对两组而言, 欺骗反应和诚实反应都导致了右腹外侧前额区域、双侧缘上回、左侧脑岛、右后内侧额叶、右侧颞中回和右侧纹状体的更大激活; 此外, 与无反馈组相比, 有反馈组的欺骗反应与诚实反应在双侧纹状体和左侧脑岛上的激活差异更加明显. 这些结果表明, 有无欺骗结果对欺骗过程的神经机制具有调控作用, 当需要面临欺骗结果时, 欺骗过程将更大程度地涉及到奖赏预期和风险厌恶过程的参与.

关键词
欺骗结果
欺骗过程
神经机制
功能核磁共振

欺骗这一人类社会普遍存在的现象一直以来都受到广泛学科领域的关注[1,2], 伴随着科学技术的革新, 欺骗研究也经历了漫长的发展. 最初的传统研究主要采用多导仪技术, 该技术通过记录自主神经活动(如皮肤电阻、血压、呼吸频率)的变化来标记欺骗行为. 但由于自主神经活动与欺骗并无直接联系, 这一方法饱受批评. 随着事件相关电位(event related potential, ERP)在认知科学中的应用, 研究者们逐渐开始关注欺骗相关的大脑活动, 但由于 ERP 的低空间分辨率特征, 欺骗的脑基础仍然很难明确. 近年来出现的功能核磁共振(functional magnetic resonance imaging, fMRI)技术同时具备无创性和高空间分辨率特征, 它可以直接探索与欺骗有关的大脑活动[1], 为揭示欺骗的神经机制提供了重要契机.

Langleben 等人[3]首次采用 fMRI 技术探索欺骗的神经机制. 他们使用改良后的 GKT(guilt knowledge test)范式诱发欺骗反应和诚实反应, 并比较两种反应的大脑激活模式. 结果发现, 与诚实反应相比, 欺骗反应导致了前扣带回和额上回等参与执行控制过程的脑区的更大激活, 证实了 fMRI 信号在

引用格式: 崔茜, 雷旭, 杨文静, 等. 有无反馈对欺骗过程的神经机制的调控作用. 中国科学: 生命科学, 2014, 44: 938–950
Cui Q, Lei X, Yang W J, et al. The existence of feedback modulates the neural mechanism underlying deception. SCIENTIA SINICA Vitae, 2014, 44: 938–950, doi: 10.1360/052013-59

辨别欺骗和诚实行为上的可行性. 与此同时, Spence 等人[4]和 Lee 等人[5]分别采用故意错误反应范式和假装失忆范式也证实了 fMRI 信号能够区分欺骗反应和诚实反应. 在这些研究的基础上, 后续研究开始考察不同类型的欺骗行为在神经机制上的异同[6~10]; 在神经基础上探索更为纯净的欺骗成分(如与记忆成分分离)[11~15]; 并尝试揭示自发欺骗行为的神经基础及欺骗行为的个体差异[16~19]. 这些研究的结论普遍认为, 诚实是一种基线式的反应, 而欺骗则需要额外的执行控制功能的参与, 因此会涉及大脑执行系统的激活, 主要包括前额叶、顶叶和前扣带回等[20]. 此外, 最近一些研究在人际交互的社会情境下考察欺骗行为, 结果发现, 与社会认知及情感过程(如奖赏预期、风险评估、道德判断和心理理论等)相关的脑区也参与欺骗[16,21,22], 说明在更为真实的社会情境下, 欺骗将涉及更加复杂多样的心理过程[23], 换言之, 欺骗的神经机制会受到社会情境因素的影响[20,23,24].

尽管基于 fMRI 的欺骗研究已经取得了一些成果, 但也存在许多问题. 例如, 虽然许多研究认识到了社会情境对欺骗(过程)具有重要影响[20,23,24], 但是, 关于它如何影响欺骗过程的具体机制并不清楚. 其中, 欺骗结果是组成欺骗的社会情境的必要因素, 它对欺骗过程的影响作用却受到很大的忽视. 在欺骗不需要面临结果的情境下, 根本无法获得有关欺骗的真实表征[23]. 目前已有少数研究采用具有欺骗结果的新实验范式(如有角色分工的游戏范式)探讨欺骗过程的神经机制[16,17,21], 结果发现, 除了经典的大脑执行系统有激活外, 还额外激活了与社会认知和情感过程相关的脑区[16,21,22]. 但这些研究都没有操纵欺骗结果的有无, 因此无法断定这些额外脑区的激活是否由于欺骗结果的加入所造成. 加之这些研究使用与以往欺骗研究完全不同的游戏范式, 所以也无法通过跨研究的比较分离出欺骗结果对欺骗过程的影响效应. 为了解决这些疑惑, 本研究直接操纵欺骗结果这一因素, 详细考察有、无欺骗结果对欺骗过程的神经机制的调控作用.

本研究采用一个“杀人游戏”模拟杀人情境, 然后对两组被试(“杀手”)分别使用有、无反馈的 GKT 范式进行测谎, 反馈是指测谎者对被试每次欺骗反应或诚实反应的判定结果(将这一反应判定为诚实或欺骗), 是通过加分或扣分的方式对实验过程中的欺骗反应或诚实反应的社会性评价(累积的最后总分直接决定测谎者对被试身份的最终判断), 在客观上形成了社会性的欺骗结果. 通过比较两组被试进行欺骗反应和诚实反应的大脑激活模式, 揭示欺骗结果对欺骗过程的神经机制的调控作用.

本研究预期: (ⅰ) 与无反馈组相比, 有反馈组被试需要持续跟踪测谎者的反馈并揣测其心理状态, 以及时更新印象管理的策略、促成成功欺骗, 其诚实反应和欺骗反应都将更大程度地卷入心理理论加工过程, 因此可能会导致相关脑区的更大激活, 如背部前扣带皮层(dorsal anterior cingulate cortex, dACC); (ⅱ) 与无反馈组相比, 在需要面临欺骗结果的有反馈组中, 欺骗过程将更大程度地涉及到奖赏预期和风险评估过程, 因此在欺骗过程所涉及的众多脑区中, 负责奖赏预期和风险评估的纹状体(striatum, STR)和脑岛(insula, INS)将可能受到有、无反馈(欺骗结果)因素的调控作用.

1　方法

1.1　被试

在某大学招募 31 名大学生参与该研究, 将其随机分配到无反馈组(15 人, 男性 6 人, 女性 9 人; 平均年龄(21.63±1.54)岁($\bar{x}$±SD, 下同))和有反馈组(16 人, 男性 8 人, 女性 8 人, 平均年龄(20.94±1.24)岁). 所有被试均为右利手, 没有任何生理或心理疾病史, 视力或矫正视力正常. 在获知整个实验过程的各项细节后, 每个被试自愿签订了知情同意书.

1.2　实验流程

(1) 模拟杀人阶段.　模拟杀人阶段的具体流程如图 1 所示. 首先, 实验者 A(主持人)组织 1 名被试和另外 3 名假被试一起抽签, 以确定他们在杀人游戏中的身份是“杀手”还是“无辜者”(图 1A). 主持人告之, 他们 4 人中将产生 1 名“杀手”而另外 3 人为“无辜者”. 其实所有的签都是“杀手”, 由此确保真被试抽到的总是“杀手”签, 并成为游戏中的杀手. 主持人还声明, 他们 4 人的身份除了自己和主持人知道以外, 其他 3 位参与者和另外的实验者都不知道, 每个参与者都需要避免暴露自己的身份. 身份确定以后, 4 名参与者都要完成一份问卷(图 1B). “杀手”(真被试)需要完成一份模拟杀人问卷(网络版附录1), 该问卷包括 7 个题项: 首先要求被试写下 2 个朋友的名字, 然后选

择其中一名作为“杀害”对象, 再依次选择作案时间、凶器、攻击的身体部位、装尸体口袋的颜色、以及弃尸地点. 除了朋友的名字需要被试自行填写以外, 其他的问题只需要被试从 2 个给定的选项中进行选择(被选出的 6 个选项组成“杀人”过程的杀人细节, 这些选项将作为 GKT 范式中的“探测刺激”出现在随后的测谎程序中). 而“无辜者”只需要填写一张记录个人基本信息的问卷. 填写问卷时, 4 名参与者的座位间保持足够距离, 使得他们无法看到其他参与者的问卷内容. 完成问卷以后, 4 名参与者被疏散到不同房间. 主持人来到“杀手”的房间, 要求“杀手”使用之前确定的 6 个杀人细节编造一个完整而生动的杀人过程, 并要求“杀手”不断完善和细化这一杀人过程, 使之形象鲜明、内容具体, 直到其变得印象深刻(可以非常容易的重复 6 个杀人细节)为止(图 1C). 然后主持人给予被试详细的指导语, 指导他如何在测谎过程中进行按键反应(网络版附录 2). 并让被试再一次背诵 6 个杀人细节. 最后, 主持人将被试带入核磁扫描室, 由另一位实验者 B(测谎者)对其进行测谎(图 1D).

(2) 测谎阶段. 测谎程序使用 GKT 范式, 具体的刺激编排和反应方式与以往多数研究一致[13,25~27], 包括 3 类刺激: (ⅰ) 探测刺激(probe, P). 这类刺激由模拟杀人阶段选择的杀人细节组成, 每个被试各不相同, 但都包括 6 个词语: 选择杀害的朋友名字(如“张三”)、作案时间(如“午夜”)、凶器(如“斧头”)、攻击身体部位(如“胸膛”)、装尸体口袋(如“黑口袋”)和藏尸地点(如“衣柜”). 被试需要假装没见过这些刺激, 做“没见过”的反应, 即欺骗反应; (ⅱ) 无关刺激(irrelevant, I). 这类刺激与模拟杀人过程无关, 由实验者事先指定, 对每个被试都相同. 它们在词语类型上与探测刺激匹配, 1 个探测刺激(杀人细节)配以 4 个无关刺激(网络版附录 3 和 4). 这些无关刺激对于杀手和无辜者来说都是之前没有见过的新词. 被试只需要按照实际情况对其进行反应, 做“没见过”的反应, 即诚实反应; (ⅲ) 目标刺激(target, T): 这类刺激与模拟杀人过程也没有任何关系, 也由实验者事先指定, 对每个被试都相同, 在词语类型上也与探测刺激匹配, 1 个探测刺激配以 1 个目标刺激(网络版附录 3). 不过, 被试需要在测谎前背诵这些词语, 然后在

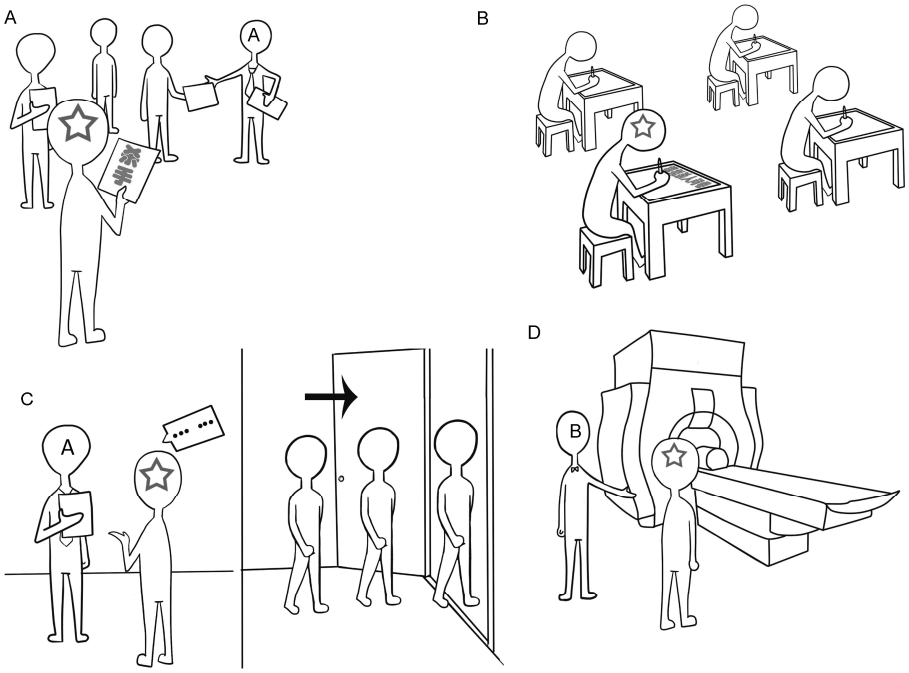

图 1 模拟杀人阶段的具体流程

A: 实验者 A 担当游戏主持人, 五角星标记的是真被试, 其余 3 人为假被试, 抽签决定“杀手”身份; B: “杀手”完成模拟杀人问卷; C: 4 名参与者疏散到不同房间, 杀手根据模拟杀人问卷中确定的凶案细节编造故事; D: 另外一名实验者 B 担当测谎者并使用核磁共振对被试进行测谎

测谎中按照实际情况对其做“见过”的反应, 即诚实反应. 由于对目标刺激的反应(“见过”)不同于其他两类刺激(均为“没见过”), 将目标刺激随机分散在整个刺激流里可起到维持被试注意力的作用[13,25,26]. GKT 范式的测谎原理是: 所有被试(包括杀手和无辜者)都对探测刺激和无关刺激进行相同的行为反应(回答“没见过”), 但只有杀手可以识别探测刺激, 对于他们而言, 探测刺激是小概率的、有意义的刺激, 对其做“没见过”反应需要涉及欺骗过程, 将诱发独特的大脑激活模式; 而无辜者无法区分探测刺激和无关刺激, 两类刺激将导致完全相同的大脑反应[26]. 因此, 通过观察被试对探测刺激和无关刺激的大脑反应是否存在差异, 便可得知被试是否撒谎, 判断他们是否“有罪”.

在刺激流程上, 无反馈组只有反应阶段, 要求被试对每个词语通过按键反应回答是否见过, 而有反馈组还包括额外的反馈阶段: 在对每个词语进行反应后, 被试会得到来自测谎者(计算机)的反馈. 反馈以“+2”或“−2”的形式出现, 分别代表加上或减去 2 分. 测谎者告诉被试, 针对目标刺激的反馈将根据他们是否正确反应来给出, 即如果正确回答“见过”则反馈“+2”(positive feedback following target, TP), 而错误回答“没见过”则反馈“−2”(negative feedback following target, TN). 最后的得分结果代表他们对目标刺激的记忆情况和测谎过程中进行反应的认真程度. 而探测刺激和无关刺激的反馈代表测谎者对其反应的判定结果, “+2”表示将他们的反应判定为诚实, 而“−2”代表将他们的反应判定为欺骗, 最后的得分越高越可能被判定为“无辜者”, 越低越可能被判定为“杀手”. 实际上, 除对目标刺激的反馈确实是基于被试是否正确反应给出外, 对探测刺激和无关刺激的反馈是事先设定好的. 对每种刺激而言, “+2”和“−2”出现的概率各占 50%, 由此形成 4 种类型的判定结果: 对探测刺激进行欺骗反应后被误判为诚实(探测刺激后的正性反馈(positive feedback following probe, PP))或被正确判定为欺骗(探测刺激后的负性反馈(negative feedback following probe, PN)); 对无关刺激进行诚实反应后被正确判定为诚实(无关刺激后的正性反馈(positive feedback following irrelevant, IP))或被误判为欺骗(无关刺激后的负性反馈(negative feedback following irrelevant, IN)). 针对每类刺激的反应方式及反馈设置请参见图 2A.

刺激流程通过 E-prime 2.0 呈现. 无、有反馈组的刺激流程见图 2B 和 C. 每个 trial 开始时, 屏幕中央出现一个持续 1 s 的词语, 随后出现一个持续 0.5 s 的反应提示信号“☆”. 要求被试在词语呈现阶段只做反应准备, 直到提示信号出现后才做反应. 被试的任务是判断是否见过该词语. 为了平衡手指的按键效应, 在每组被试中, 1/2 被试通过右手拇指按“1”键回答“见过”, 左手拇指按“3”键回答“没见过”, 而 1/2 被试使用相反的按键模式. 在无反馈组中, 提示信号消失后, 出现持续时间在 0.5/2.5/4.5 s 间随机的注释符号“+”, 随后进入下一个 trial; 而在有反馈组中, 提示信号消失后, 出现一个持续时间在 0.5/2.5/4.5 s 间随机的黑屏, 告知被试电脑此时正在分析他们的大脑信号, 黑屏消失后即出现持续 1.5 s 的反馈, 接着出现持续时间在 0.5/2.5/4.5 s 间随机的注释符号“+”, 随后进入下一个 trial. 如果被试进行了错误按键, 将直接反馈“−2”, 如果被试不能在 0.5 s 内做出反应, 则反馈“无反应”.

扫描前, 以代表作案时间的探测刺激及其相应的目标刺激和无关刺激作为实验材料让被试进行充分练习, 达到 90%以上的正确率后再进入正式实验.

正式实验包括 5 个实验组块, 每个组块对一个探

A

两组被试		有反馈组	
		反馈	
刺激	反应	概率	判定
探测	欺骗 (P)	50%	+2 (PP)
		50%	−2 (PN)
目标	诚实 (T)	正确	+2 (TP)
		错误	−2 (TN)
无关	诚实 (I)	50%	+2 (IP)
		50%	−2 (IN)

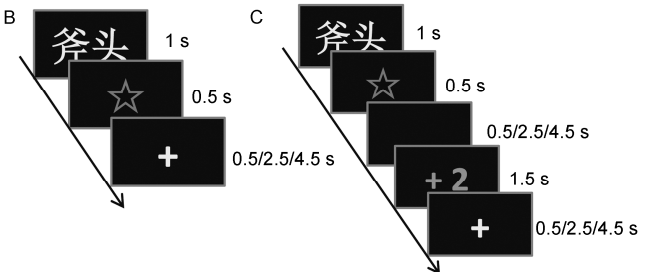

图 2　两组被试对 3 类刺激的反应类型及有反馈组中针对 3 类刺激的反馈设置(A)、无反馈组的刺激程序流程(B)和有反馈组的刺激呈现流程(C)

测刺激进行检测(用作练习材料的“作案时间”不参与检测), 5 个实验组块出现的先后顺序在被试间随机. 每个实验组块包含 1 个探测词语、1 个目标词语以及 4 个无关词语. 探测词语和目标词语分别重复呈现 14 次, 4 个无关词语中的 2 个词语分别重复 11 次, 而另 2 个词语分别重复 10 次, 因此每个实验组块包含 14 个探测 trial, 14 个目标 trial 和 42 个无关 trial, 共计 70 个 trial. 70 个 trial 在每个实验组块内随机出现.

(3) 事后主观报告. 扫描结束以后, 每个被试需要再次背诵探测刺激和目标刺激, 以确认他们对两类刺激的记忆情况. 随后, 他们需要完成一份事后主观报告问卷, 该问卷通过询问被试对逃过测谎(不被发现是杀手)的希望程度测量被试参与任务的“动机水平”; 通过询问被试看到每类刺激时的紧张程度测量被试对不同刺激的“主观卷入程度”; 并询问有反馈组被试对每类刺激后反馈的“关注程度”. 此外, 考察被试对本研究设计的整个背景故事的信任情况及策略使用情况. 通过搜集两组被试的这些主观体验, 了解欺骗结果调控欺骗过程的重要心理表现, 并为 fMRI 结果的解释提供指导.

1.3 fMRI 数据采集

脑成像数据使用西门子 3T 全身磁共振成像仪器(Siemens Magnetom Trio Tim, 德国)搜集, 配以 8 通道相控阵头线圈获取被试的大脑磁共振成像数据. 为了防止头动, 使用衬垫固定头部, 并告诉被试在实验过程中尽量保持头部静止. 视觉刺激通过磁共振仪器配置的专业投影设备呈现, 被试通过头部线圈上的倒视镜进行观看. 为了减少扫描仪产生的噪音对被试执行实验任务的影响, 每个被试佩戴专用的防噪音耳塞.

全脑功能像采用对血氧水平依赖(blood oxygenation level dependent, BOLD)具有敏感性的 T2*加权单次激发梯度回波的 EPI(echo planar imaging, EPI)序列获得, 无反馈组的每个被试共采集 880 个 TR 的全脑的功能影像, 而有反馈组的每个被试共采集 1360 个 TR(repeat time)的全脑的功能影像. 每个卷积包括 32 层, 采用隔层扫描的方式, 层厚 3 mm, 层间距 1 mm, 扫描方向平行于 AC-PC 平面. 扫描参数如下: TR=2000 ms, TE(echo time)=30 ms, 扫描视野 FOV(field of vision)=220×220, 扫描矩阵(matrix)为 64×64, 反转角(flip angle)为 90°. 每个 trial 呈现的时间通过 E-prime 控制以与扫描同步. 此外, 使用以下参数获取每个被试的结构像数据: TR=1900 ms, TE=2.52 ms, 层厚 1 ms, 扫描视野 FOV=256×256, 反转角为 90°, 最终每个体素的大小为 1×1×1 mm^3.

1.4 fMRI 数据分析

功能磁共振数据的预处理和分析采用 SPM8(Wellcome Department of Cognitive Neurology, 英国)工具包, 该工具包基于 MatLab(Math works, 美国) 平台运行. 在预处理阶段, 首先采用 slice timing 程序进行时间配准, 以对各扫描层在获取时间上的差异进行矫正, 再进行头动矫正(realigned). 然后, 将图像数据基于 SPM8 提供的 EPI 模版进行空间标准化处理(normalize), 同时将功能像数据重采样为 3×3×3 mm^3 分辨率大小. 最后, 使用 8 mm 的平滑核进行高斯平滑处理(smooth). 所有被试的头动小于 3 mm.

第一水平分析采用事件相关模式. 针对每个被试建立刺激矩阵, 对两组被试而言, 都包含反应阶段的 4 种条件(regressors): 被试对 3 类刺激(P, T 和 I)进行正确反应(对探测刺激的正确反应是指对其进行欺骗反应)的 trial 构成的 3 个感兴趣 regressor. 错误反应或无反应的 trail 的反应阶段构成另 1 个非兴趣 regressor(incorrect response, INR). 这 4 个 regressor 的矩阵开始时间对应于刺激呈现时间, duration 设置为 0. 对有反馈组而言, 还建立了额外 7 个反馈阶段的 regressors, 包括进行正确反应后出现的 6 类判定结果(PP, PN, TP, TN, IP 及 IN)构成的 6 个感兴趣 regressor, 以及由错误或无反应 trial 的反馈阶段构成的 1 个非兴趣 regressor(incorrect judgment, INJ). 这 7 个 regressor 的矩阵开始时间对应于反馈呈现时间, duration 设置为 0. 此外, 每个被试的 6 个头动参数也作为非兴趣 regressor 被纳入刺激矩阵, 以排除头动对信号的影响. 使用 GLM(general linear model)的 HRF(hemodynamic response function)函数进行个体水平的统计分析, 采用 128 s 的高通滤波消除低频噪音对数据的影响. 本研究关注有、无欺骗结果对欺骗过程的调控作用, 因此只考察反应阶段的大脑激活模式, 而不分析反馈阶段的数据. 此外, 由于设置目标刺激(T)仅在于维持被试对刺激的注意[25], 因此也不对目标刺激进行分析. 使用单样本 t 检验, 分别将 P(探测/欺骗)和 I(无关/诚实)条件与内隐基线相减, 并将 P 条件与 I 条件

相减, 获得 3 个相减(contrast)后的 contrast 脑和 beta 脑图像, 这些图像将用于随后的第二水平分析.

第二水平分析将第一水平分析中得到的 P 和 I 条件对应的 contrast 脑图像(con 文件)代入一个 2(无/有反馈)×2(探测/无关)的全因素模型(full factorial analysis). 主效应结果经过 *FWE*(family wise error rate)<0.001, 激活簇大于 50 个体素(K>50)的阈值校正. 为了探索两个因素的交互作用, 放宽阈值到 P<0.005, K>10, 虽然未经校正, 但 P<0.005 的阈值获得了社会神经科学领域的普遍认可[28,29], 在该阈值下获得的结果与以往研究具有可比性. 激活结果基于 AAL(anatomical automatic labeling)模板进行定位.

由于发现组别和反应类型之间的交互作用在双侧纹状体及左侧脑岛上显著, 根据交互作用的结果定义了 3 个功能兴趣区域(region of interest, ROI), 然后提取每个被试在 3 个 ROI 上探测刺激和无关刺激对应的激活强度参数(parameter estimates, PEs), 再将其代入一个 2(无/有反馈)×2(探测/无关)的重复测量方差分析, 以观察交互作用的具体模式.

2　结果

2.1　行为结果

(1) 正确率和反应时.　本研究设置目标刺激来维持被试对刺激的注意, 无、有反馈组被试对目标刺激的反应正确率都较高($\bar{x}$ 分别为(89.9±0.111)%和(95.6±0.030)%), 说明两组被试都认真执行了按键任务. 两组被试对探测刺激和无关刺激的反应正确率和反应时见表 1. 对两组的正确率进行 2(无/有反馈)×2(探测/无关)的重复测量方差分析, P 值采用 Greenhouse Geiser 法校正. 结果发现, 组别主效应不显著, F(1, 29)=1.855, P=0.184. 刺激类型主效应不显著, F(1, 29)=0.000, P=0.989. 两因素之间的交互作用也不显著, F(1, 29)=0.128, P=0.723. 对反应时进行同样的重复测量方差分析, 结果发现, 组别主效应不显著, F(1, 29)=2.548, P=0.121. 刺激类型主效应不显著, F(1, 29)=1.514, P=0.228. 组别和刺激类型两因素间的交互作用也不显著, F(1, 29)=1.259, P=0.271. 这些结果说明, 由于探测刺激和无关刺激需要完全相同的按键反应, 它们诱发了相同的行为反应模式.

(2) 事后主观报告.　扫描结束后, 所有被试都能准确无误地背诵目标刺激和探测刺激, 说明在整个实验过程中, 被试对两类刺激的记忆是明确和牢固的.

表 1　两组被试对 3 类刺激进行按键反应的正确率和反应时($\bar{x}$ ±SD)

行为变量	被试组别	探测刺激	无关刺激
正确率	无反馈组	0.901±0.098	0.897±0.124
	有反馈组	0.937±0.061	0.941±0.037
反应时	无反馈组	268.204±63.018	267.877±71.690
	有反馈组	301.053±33.613	293.970±28.998

事后主观报告的结果如表 2 所示. 整体而言, 两组被试希望通过测谎的动机水平都较高, 均大于 4(5 点评分), 且独立样本 t 检验的结果表明, 两组的动机水平没有显著差异, t(29)=1.158, P=0.256. 这说明本研究设置的模拟杀人情境诱发了所有被试较高的动机水平.

使用被试对刺激的主观卷入程度进行 2(无/有反馈)×2(探测/无关)的重复测量方差分析, P 值采用 Greenhouse Geiser 法校正. 结果发现, 组别主效应显著, F(1, 29)=18.081, P=0.000, 与无反馈组相比, 有反馈组对所有刺激的主观卷入程度都更高. 刺激类型主效应显著, F(1, 29)=32.592, P=0.000, 探测刺激比无关刺激导致了更高的主观卷入程度. 两因素间的交互作用显著, F(1, 29)=6.083, P=0.020, 简单效应分析的结果表明, 在两组中, 探测刺激都比无关刺激导致了更高的主观卷入程度(有反馈组 P=0.000, 无反馈组 P=0.027), 但是, 无关刺激在有反馈组中诱发的主观卷入程度明显强于无反馈组(P=0.000), 而探测刺激在两组中诱发的主观卷入程度没有显著差异(P=0.602).

使用有反馈组被试对探测刺激和无关刺激后反馈的关注程度进行配对 t 检验的结果表明, 被试对探测刺激后的反馈比对无关刺激后的反馈更为关注, t(15)=4.869, P=0.000, 说明被试不但可以区分探测刺激和无关刺激, 并且他们更关注对探测刺激进行欺骗反应的结果.

此外, 所有被试都报告他们相信实验中所使用的“测谎仪器”确实具备检测谎言的能力. 除了报告“见到杀人细节词语时尽量保持冷静”这一普遍策略外, 没有被试报告他们使用了任何其他的特殊策略.

2.2　fMRI 结果

全因素模型分析的结果如表 3 所示. 组别主效应

显著, 在反应阶段, 有反馈组比无反馈组导致了覆盖中央前、后回区域的左侧顶叶皮层、左侧 dACC、左侧 INS 以及双侧视觉皮层的广泛区域和右侧小脑的更大激活(图 3A); 而无反馈组较之有反馈组没有导致任何脑区的更大激活. 刺激类型主效应显著, 探测刺激比无关刺激导致了右侧眶额皮层(位于腹外侧前额区域, ventral lateral prefrontal cortex, VLPF)、双侧缘上回(supramarginal gyrus, SMG)(BA40)、左侧 INS、右后内侧额叶皮层(posterior medial frontal cortex, pMFC)、右侧颞中回(middle temporal gyrus, MTG)以及右侧 STR 的更大激活(图 3B); 而无关刺激较之探测刺激没有导致任何脑区的更大激活. 在更宽松的阈值下(P<0.005, K>10)发现, 组别和刺激类型两个因素的交互效应(有反馈组(探测>无关)>无反馈组(探测>无关))在双侧 STR 和左侧 INS 上显著(图 4A 和 B). 针对交互效应显著的 3 个激活区域定义功能 ROI, 并使用这 3 个 ROI 的激活强度参数(PEs)分别进行 2(无/有反馈)×2(探测/无关)的重复测量方差分析的结果显示, 3 个区域上确实存在显著的交互效应(左侧 INS: F(1, 29)=8.05, P<0.01; 左侧 STR: F(1, 29)=8.66, P<0.01; 右侧 STR: F(1, 29)=9.28, P<0.01); 进一步简单效应分析的结果表明, (ⅰ) 在有、无反馈两组中, 探测刺激在 3 个 ROI 上的激活都强于无关刺激(除在无反馈组中, 探测刺激和无关刺激在左侧 STR 上的激活差异边缘显著, P=0.067, 其余为 P<0.01); (ⅱ) 在探测和无关条件下, 3 个 ROI 在有反馈组中的激活都强于无反馈组(除左侧 INS 对无关刺激的激活强度在有、无反馈组间没有差异, P=0.81, 其余为 P<0.05); 但是, 探测刺激和无关刺激在这些脑区上的激活差异在有反馈组中比在无反馈组中更加明显. 也就是说, 在有反馈组, 对探测刺激进行反应时所需的欺骗过程更大程度的涉及到这 3 个脑区的参与.

表 2　两组被试对探测刺激和无关刺激的主观报告结果($\bar{x}$±SD)

行为变量	组别	探测刺激	无关刺激
参与任务的动机水平	无反馈组	4.33±0.816	
	有反馈组	4.06±0.443	
对刺激的主观卷入程度	无反馈组	3.533±0.990	1.800±0.775
	有反馈组	3.688±0.602	3.000±0.516
对结果的关注程度	有反馈组	3.63±0.619	2.31±0.873

表 3　全因素模型分析下呈现显著激活的大脑区域 a)

比较条件	脑区	BA(Brodmann area)	t 值	体素值	x	y	z
组别主效应							
有反馈>无反馈	L parietal lobe	3/40/7	9.91	1116	−18	−54	60
	L dorsal anterior cingulate cortex	24/32	8.73	275	−3	6	42
	L insula	13	8.59	176	−42	0	9
	L/R occipital lobe	18/19	8.54	1150	−15	−81	30
	R cerebellum	–	7.51	73	3	−69	−15
刺激类型主效应							
探测>无关	R orbitofrontal cortex	47	10.91	564	33	27	−6
	R supramarginal gyrus	40	10.77	233	57	−45	33
	L insula	47/13	9.43	170	−30	21	−6
	R posterior medial frontal cortex	6/8	9.41	245	9	15	63
	L supramarginal gyrus	40	9.19	140	−57	−48	36
	R middle temporal gyrus	21	8.51	90	57	−27	−6
	R striatum	–	8.33	122	15	9	3
组别×刺激类型							
交互效应	R striatum	–	3.48	36	12	9	0
	L striatum	–	3.3	35	−12	9	0
	L insula	47	3.19	12	−27	24	−3

a) x, y, z 对应 MNI(montreal neurological institute)坐标, 脑区位置基于 AAL 模板进行定位; R: 大脑右半球; L: 大脑左半球. *: P<0.005, K>0, 其余结果经过 FWE<0.001, K>50 的阈值校正

3　讨论

本研究在一个模拟杀人情境后，对两组被试分别使用无、有反馈(欺骗结果)的 GKT 范式进行测谎，通过比较两组被试进行诚实反应和欺骗反应时的大脑激活模式发现，有无反馈因素调控了整个反应阶段(包括诚实反应和欺骗反应)的神经机制；在欺骗过程所涉及的众多脑区中，负责奖赏预期和风险评估过程的纹状体和脑岛的神经活动明显受到有无反馈因素的调控作用.

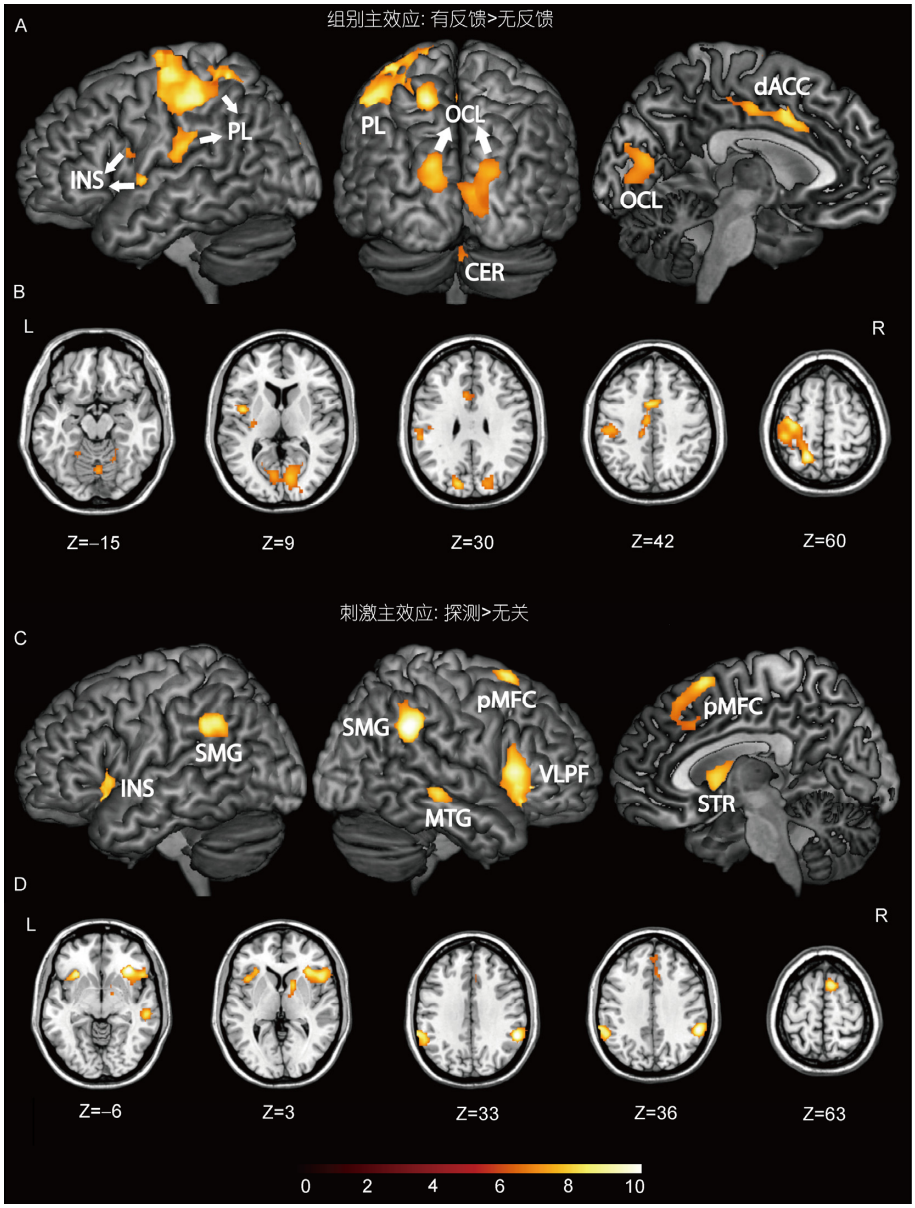

图 3　组别主效应(有反馈组>无反馈组)显著激活的脑区(A, B)与刺激类型主效应(探测刺激>无关刺激)显著激活的脑区(C, D)

先在立体大脑模型(render)上呈现所有激活显著的 cluster(A, C)，然后按照每个激活簇中最大激活点的坐标，进行横断面上的分层显示(B, D). 颜色标尺上的数字代表 t 值. 所有结果在 $FWE<0.001$, $K>50$ 的阈值上获得. L: 大脑左半球; R: 大脑右半球; INS: insula，脑岛; PL: parietal lobe，顶叶; OCL: occipital lobe，枕叶; CER: cerebellum，小脑; dACC: dorsal anterior cingulate cortex，背部前扣带皮层; SMG: supramarginal gyrus，缘上回; pMFC: posterior medial frontal cortex，后内侧额叶皮层; VLPF: ventral lateral prefrontal cortex，腹外侧前额区域; MTG: middle temporal gyrus，颞中回; STR: striatum，纹状体

3.1 有无反馈调控整个反应阶段的神经机制

fMRI 结果中的组别主效应(有反馈组>无反馈组)体现了反馈对整个反应阶段的调控作用. 组别主效应显示, 有反馈组比无反馈组在反应阶段导致了覆盖中央前、后回的左侧顶叶皮层、左侧 dACC、左侧 INS、双侧视觉皮层、以及右侧小脑的更大激活, 说明反馈对反应阶段产生了广泛的影响, 这一影响同时涉及诚实反应和欺骗反应, 使得两种反应都诱发了更为复杂的认知过程. 其中, dACC 作为心理理论系统的核心脑区已有大量报道[30~32]. dACC 参与认知性的心理理论任务[30,33], 如参与表征他人心理状态的错误信念任务[31]. 在错误信念任务中, 个体需要从他人的知识状况出发, 将其信念与客观事实进行区分, 最终获得对他人信念的觉知和理解. 然而, 表征他人信念必须基于一定的知觉线索(perception-based)[34~36], 即只有当他人的信念通过一定的外在线索表露出来、并被个体觉知到时, 个体才可能启动对他人现有心理状态的推论. 在本研究中, 反馈则充当了这一知觉线索的角色: 在有反馈组中, 不论被试做诚实反应还是欺骗反应, 反应后都会获得测谎者对这一反应的判定结果, 根据这一判定, 被试得以推论测谎者关于自己身份(杀手/无辜者)的当前信念, 并基于对测谎者当前信念的掌握情况调整随后的反应策略, 这是一系列反复推进的心理理论加工过程, 广泛存在于整个反应阶段和结果加工阶段; 然而, 在无反馈组中, 被试无法获得有关测谎者当前信念的任何线索信息, 上述心理理论加工过程无法或只能很小程度的进行. 因此, 两组被试对心理理论加工过程的认知卷入程度是有明显差异的, 这一差异最终体现在 dACC 的激活差异上.

有反馈组比无反馈组还导致了覆盖中央前后回的顶叶区域、小脑、以及视觉皮层的更大激活. 中央

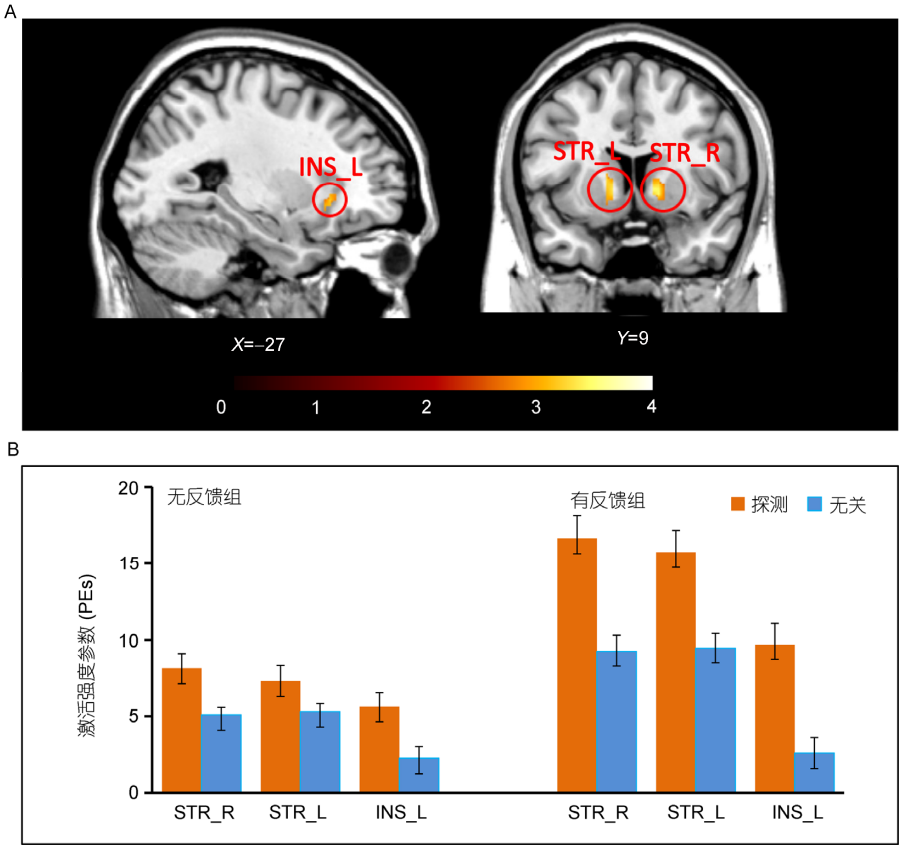

图 4 组别和刺激类型两因素的交互作用显著激活的 3 个脑区, 分别对应每个 cluster 的最大激活点坐标在矢状面或冠状面上进行显示(A)及在有、无反馈组中, 探测刺激和无关刺激在 3 个脑区上的激活强度参数(PEs)(B)

激活图在 $P<0.005$, $K>10$ 的阈值下获得, 颜色标尺上的数字代表 t 值. 误差线代表标准误(SE)

前后回和小脑区域属于运动皮层, 它们的更大激活可能说明有反馈情境下的反应动作诱发了更大程度的认知卷入[37,38]. 而属于注意系统的顶叶皮层、视觉皮层下部及舌回等区域在有反馈情境下的更大激活可能反映了主观卷入对注意过程的增强作用[39].

3.2　有无反馈调控欺骗过程的神经机制

fMRI 结果中的刺激类型主效应(探测>无关)反映出欺骗过程所涉及的大脑激活模式. 在两组中, 探测刺激比无关刺激导致了右侧 VLPF、双侧 SMG、左侧 INS、右侧 pMFC、右侧 MTG 以及右侧 STR 的更大激活, 这些脑区在以往欺骗研究中都已获得广泛的报告[3,18,20,25,40~43], 本研究的发现与已有研究保持一致, 证明本研究的实验设计有效地诱发了欺骗行为.

fMRI 结果中的组别和刺激类型的交互效应(有反馈(探测>无关)>无反馈(探测>无关))体现了反馈(欺骗结果)对欺骗过程的调控作用. 交互效应在双侧 STR 和左侧 INS 3 个区域上显著, 其中右侧 STR 和左侧 INS 与上述刺激主效应结果中的右侧 STR 和左侧 INS 重叠(peak 点坐标均<3 mm), 并且, 当降低阈值到 $FWE<0.005$ 时, 刺激主效应在左侧 STR 上也出现了显著(网络版附录 5), 意味着这 3 个区域都参与欺骗过程. 进一步简单效应分析的结果证实了这一主观观察: 在两组中, 3 个区域确实都表现出显著的刺激类型效应, 即对探测刺激的激活都强于对无关刺激的激活.

此外, 简单效应分析结果还发现, 在探测刺激和无关刺激条件上, 3 个区域在有反馈组中的激活都强于无反馈组, 说明反馈普遍增强了这些区域在欺骗反应和诚实反应中的激活水平; 更为重要的是, 探测刺激和无关刺激在 3 个区域上的激活差异在有反馈组中比在无反馈组中更加明显, 说明有反馈情境下的欺骗过程更大程度地涉及到这些区域的参与. 已有大量研究证实 INS 参与风险厌恶过程[44,45], 而纹状体参与奖赏预期过程[46~50], 并发现 STR 对预期奖赏的主观价值敏感[50~52]. 结合以上发现, 本研究认为, (ⅰ) 当不存在欺骗结果时(无反馈组), 欺骗已经涉及一定程度的风险厌恶和奖赏预期过程, 即与诚实相比, 欺骗被视为更具风险, 与此同时, 与诚实后被他人正确判定为诚实(不被冤枉)的情况相比, 欺骗后却被他人误判为诚实(成功欺骗)这一社会性的预期奖赏对欺骗者具有更大的主观价值; (ⅱ) 当反应需要面临即刻的判定结果时(有反馈组), 被试对反应结果(反馈)的担忧和期待增强, 因此诚实反应和欺骗反应都涉及了更强的风险评估和奖赏预期过程; 但是, 这些心理过程的增强尤其倾斜于欺骗反应, 可能说明需要面对欺骗结果的有反馈组被试在欺骗时特别担心被发现, 同时又特别希望逃过检测而获得意外奖赏, 这些心理使得他们(与无反馈组被试相比)在欺骗过程中产生尤其强烈的风险厌恶体验并对成功欺骗的奖赏(哪怕只是社会性的奖赏)赋予了更高的主观价值. 来自有反馈组被试的主观报告结果在一定程度上印证了以上推论, 主观报告的结果表明, 有反馈组被试对探测刺激的欺骗反应能否成功的关注程度要显著高于对无关刺激的诚实反应能否被正确判定为诚实(不被冤枉)的关注程度, 说明欺骗后的反馈(欺骗结果)对被试的独特重要性, 因此不难理解对这些结果的预期加工也会尤其强烈.

除 INS 和 STR 外, 欺骗过程涉及的其他区域, 如双侧 SMG、右侧 VLPF、右侧 MTG 和右侧 pMFC 上没有出现组别和刺激类型的交互效应, 说明它们对欺骗的参与没有受到有无反馈(欺骗结果)的调控. 大量研究认为, 这些区域在欺骗中参与执行控制过程[1,20], 它们负责抑制诚实反应倾向, 以利于欺骗反应的产生[20]. 基于已有研究和本研究结果, 推测反应抑制过程在不同类型的欺骗中都是最为基础的认知成分, 因而它不容易受到欺骗结果这类社会情境因素的影响. 相反, INS 和 STR 所负责的风险评估和奖赏预期过程是欺骗中的社会认知和情感成分, 它们较容易受到社会情境的影响, 明显受到欺骗结果的调控.

4　结论

本研究使用无、有反馈(欺骗结果)的 GKT 范式, 意在从神经机制的角度考察欺骗结果对欺骗过程的调控作用. 研究结果发现, 与无反馈组相比, 有反馈组被试在欺骗反应和诚实反应上都更大程度地卷入了心理理论加工过程, 反映在 dACC 的更大激活上; 并且发现, 有反馈组被试比无反馈组被试在欺骗过程中更大程度地卷入了奖赏预期和风险评估过程, 反映在与这两个心理过程密切相关的脑区 STR 和 INS 的更大激活上. 欺骗结果对欺骗的神经机制具有明显的调控作用, 它主要对欺骗的社会认知和情感

成分产生影响.

5 本研究的不足

本研究采用的 GKT 范式是一种“指导性”的欺骗范式, 即要求被试在实验者的指导下对特定的刺激进行欺骗反应, 这与现实中个体自发的欺骗行为有一定差距. 指导性欺骗范式由于具有易于控制变量、容易控制不同实验条件的信号叠加次数的优点, 适用于探讨欺骗中的新变量效应、聚焦新问题的探索性研究, 因此仍然被大多数欺骗和测谎研究所使用[6,7,10,11,18,25,41,42,53~58]. 但是, 指导性欺骗范式毕竟缺乏生态学效度, 将来的研究有必要采用自发欺骗范式, 如将指导语改为“尽量逃避测谎”, 然后让被试自发的选择何时、对哪些刺激撒谎, 在更为真实的欺骗行为中考察欺骗结果对欺骗过程的调控作用.

参考文献

1 Abe N. How the brain shapes deception. Neuroscientist, 2011, 17: 560–574
2 Ford E B. Lie detection: historical, neuropsychiatric and legal dimensions. Int J Law Psychiat, 2006, 29: 159–177
3 Langleben D D, Schroeder L, Maldjian J A, et al. Brain activity during simulated deception: an event-related functional magnetic resonance study. NeuroImage, 2002, 15: 727–732
4 Spence S A, Farrow T F D, Herford A E, et al. Behavioural and functional anatomical correlates of deception in humans. Neuroreport, 2001, 12: 2849–2853
5 Lee T M C, Liu H L, Tan L H, et al. Lie detection by functional magnetic resonance imaging. Hum Brain Mapp, 2002, 15: 157–164
6 Ganis G, Morris R R, Kosslyn S M. Neural processes underlying self-and other-related lies: an individual difference approach using fMRI. Soc Neurosci, 2009, 4: 539–553
7 Abe N, Suzuki M, Tsukiura T, et al. Dissociable roles of prefrontal and anterior cingulate cortices in deception. Cereb Cortex, 2006, 16: 192–199
8 Abe N, Suzuki M, Mori E, et al. Deceiving others: distinct neural responses of the prefrontal cortex and amygdala in simple fabrication and deception with social interactions. J Cogn Neurosci, 2007, 19: 287–295
9 Ganis G, Kosslyn S M, Stose S, et al. Neural correlates of different types of deception: an fMRI investigation. Cereb Cortex, 2003, 13: 830–836
10 Nunez J M, Casey B, Egner T, et al. Intentional false responding shares neural substrates with response conflict and cognitive control. NeuroImage, 2005, 25: 267–277
11 Abe N, Okuda J, Suzuki M, et al. Neural correlates of true memory, false memory, and deception. Cereb Cortex, 2008, 18: 2811–2819
12 Lee T M C, Au R K C, Liu H L, et al. Are errors differentiable from deceptive responses when feigning memory impairment? An fMRI study. Brain Cogn, 2009, 69: 406–412
13 Gamer M, Klimecki O, Bauermann T, et al. fMRI-activation patterns in the detection of concealed information rely on memory-related effects. Soc Cogn Affect Neurosci, 2012, 7: 506–515
14 Nose I, Murai J, Taira M. Disclosing concealed information on the basis of cortical activations. NeuroImage, 2009, 44: 1380–1386
15 Kozel F A, Padgett T M, George M S. A replication study of the neural correlates of deception. Behav Neurosci, 2004, 118: 852–856
16 Baumgartner T, Fischbacher U, Feierabend A, et al. The neural circuitry of a broken promise. Neuron, 2009, 64: 756–770
17 Greene J D, Paxton J M. Patterns of neural activity associated with honest and dishonest moral decisions. Proc Natl Acad Sci USA, 2009, 106: 12506–12511
18 Fullam R S, McKie S, Dolan M C. Psychopathic traits and deception: functional magnetic resonance imaging study. Brit J Psychiat, 2009, 194: 229–235
19 Jung K H, Lee J H. Implicit and explicit attitude dissociation in spontaneous deceptive behavior. Acta Psychol, 2009, 132: 62–67
20 Spence S A, Hunter M D, Farrow T, et al. A cognitive neurobiological account of deception: evidence from functional neuroimaging. Philos Trans R Soc Lond B-Biol Sci, 2004, 359: 1755–1762
21 Sip K E, Skewes J C, Marchant J L, et al. What if I get busted? Deception, choice, and decision-making in social interaction. Front Neurosci, 2012, 6: 58
22 Bhatt M A, Lohrenz T, Camerer C F, et al. Neural signatures of strategic types in a two-person bargaining game. Proc Natl Acad Sci USA, 2010, 107: 19720–19725
23 Sip K E, Roepstorff A, McGregor W, et al. Detecting deception: the scope and limits. Trends Cogn Sci, 2008, 12: 48–53
24 Bles M, Haynes J D. Detecting concealed information using brain-imaging technology. Neurocase, 2008, 14: 82–92
25 Ganis G, Rosenfeld J P, Meixner J, et al. Lying in the scanner: covert countermeasures disrupt deception detection by functional magnetic

resonance imaging. NeuroImage, 2011, 55: 312–319
26 Ben-Shakhar G, Elaad E. The validity of psychophysiological detection of information with the Guilty Knowledge Test: a meta-analytic review. J Appl Psychol, 2003, 88: 131–151
27 Rosenfeld J P, Biroschak J R, Furedy J J. P300-based detection of concealed autobiographical versus incidentally acquired information in target and non-target paradigms. Int J Psychophysiol, 2006, 60: 251–259
28 Wager T D, Lindquist M, Kaplan L. Meta-analysis of functional neuroimaging data: current and future directions. Soc Cogn Affect Neurosci, 2007, 2: 150–158
29 Lieberman M D, Cunningham W A. Type I and Type II error concerns in fMRI research: re-balancing the scale. Soc Cogn Affect Neurosci, 2009, 4: 423–428
30 Abu-Akel A, Shamay-Tsoory S. Neuroanatomical and neurochemical bases of theory of mind. Neuropsychologia, 2011, 49: 2971–2984
31 Sommer M, Döhnel K, Sodian B, et al. Neural correlates of true and false belief reasoning. NeuroImage, 2007, 35: 1378–1384
32 Swanner J K, Beike D R. Incentives increase the rate of false but not true secondary confessions from informants with an allegiance to a suspect. Law Hum Behav, 2010, 34: 418–428
33 Poletti M, Enrici I, Adenzato M. Cognitive and affective theory of mind in neurodegenerative diseases: neuropsychological, neuroanatomical and neurochemical levels. Neurosci Biobehav Rev, 2012: 2147–2164
34 Carruthers P. How we know our own minds: the relationship between mindreading and metacognition. Behav Brain Sci, 2009, 32: 121
35 Williams D M, Lind S E, Happe F. Metacognition may be more impaired than mindreading in autism. Behav Brain Sci, 2009, 32: 162–163
36 Happé F. Theory of mind and the self. Ann NY Acad Sci, 2003, 1001: 134–144
37 Cheng Y, Meltzoff A N, Decety J. Motivation modulates the activity of the human mirror-neuron system. Cereb Cortex, 2007, 17: 1979–1986
38 Lee W, Reeve J. Self-determined, but not non-self-determined, motivation predicts activations in the anterior insular cortex: an fMRI study of personal agency. Soc Cogn Affect Neurosci, 2013, 8: 538–545
39 Ivanov I, Liu X, Clerkin S, et al. Effects of motivation on reward and attentional networks: an fMRI study. Brain Behav, 2012: 741–753
40 Kozel F A, Johnson K A, Mu Q, et al. Detecting deception using functional magnetic resonance imaging. Biol Psychiat, 2005, 58: 605–613
41 Kozel F A, Laken S J, Johnson K A, et al. Replication of functional MRI detection of deception. Open Forensic Sci J, 2009, 2: 6–11
42 Kozel F A, Johnson K A, Grenesko E L, et al. Functional MRI detection of deception after committing a mock sabotage crime. J Forensic Sci, 2009, 54: 220–231
43 Langleben D D, Loughead J W, Bilker W B, et al. Telling truth from lie in individual subjects with fast event-related fMRI. Hum Brain Mapp, 2005, 26: 262–272
44 Paulus M P, Rogalsky C, Simmons A, et al. Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. NeuroImage, 2003, 19: 1439–1448
45 Rudorf S, Preuschoff K, Weber B. Neural correlates of anticipation risk reflect risk preferences. J Neurosci, 2012, 32: 16683–16692
46 Beck A, Schlagenhauf F, Wüstenberg T, et al. Ventral striatal activation during reward anticipation correlates with impulsivity in alcoholics. Biol Psychiat, 2009, 66: 734–742
47 Dichter G S, Felder J N, Green S R, et al. Reward circuitry function in autism spectrum disorders. Soc Cogn Affect Neurosci, 2012, 7: 160–172
48 Hsu M, Krajbich I, Zhao C, et al. Neural response to reward anticipation under risk is nonlinear in probabilities. J Neurosci, 2009, 29: 2231–2237
49 O'Doherty J P. Reward representations and reward-related learning in the human brain: insights from neuroimaging. Curr Opin Neurobiol, 2004, 14: 769–776
50 Rademacher L, Krach S, Kohls G, et al. Dissociation of neural networks for anticipation and consumption of monetary and social rewards. NeuroImage, 2010, 49: 3276–3285
51 Bromberg-Martin E S, Hikosaka O. Midbrain dopamine neurons signal preference for advance information about upcoming rewards. Neuron, 2009, 63: 119–126
52 Morris G, Nevet A, Arkadir D, et al. Midbrain dopamine neurons encode decisions for future action. Nat Neurosci, 2006, 9: 1057–1063
53 Gamer M, Bauermann T, Stoeter P, et al. Covariations among fMRI, skin conductance, and behavioral data during processing of concealed information. Hum Brain Mapp, 2007, 28: 1287–1301
54 Hakun J G, Ruparel K, Seelig D, et al. Towards clinical trials of lie detection with fMRI. Soc Neurosci, 2009, 4: 518–527
55 Kaylor-Hughes C J, Lankappa S T, Fung R, et al. The functional anatomical distinction between truth telling and deception is preserved

among people with schizophrenia. Crim Behav Ment Health, 2011, 21: 8–20
56 Ito A, Abe N, Fujii T, et al. The role of the dorsolateral prefrontal cortex in deception when remembering neutral and emotional events. Neurosci Res, 2011, 69: 121–128
57 Lee T M C, Lee T M Y, Raine A, et al. Lying about the valence of affective pictures: an fMRI study. PLoS One, 2010, 5: e12291
58 Bhatt S, Mbwana J, Adeyemo A, et al. Lying about facial recognition: an fMRI study. Brain Cogn, 2009, 69: 382–390

The Existence of Feedback Modulates the Neural Mechanism Underlying Deception

CUI Qian[1,2,3], LEI Xu[2], YANG WenJing[2], WANG YiFeng[3] & ZHANG QingLin[2]

1 School of Political Science and Public Administration, University of Electronic Science and Technology of China, Chengdu 610054, China;
2 Faculty of Psychology, Southwest University, Chongqing 400715, China;
3 Key Laboratory for NeuroInformation of Ministry of Education, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 610054, China

Deception will result in some consequences (deception consequence), which will in turn affect the mechanism underlying deception process. Previous studies have demonstrated that the deception consequence modulates the corresponding deception process; however, the neural mechanism underlying this modulation effect remains unclear. The present study investigated how the neural signals elicited by deceptive response would be modulated by whether response would be judged by the lie-detector about its facticity. After a mock murder situation, BOLD responses in the brains of 31 healthy volunteers were recorded while they underwent two versions of guilt knowledge test (GKT) paradigm. In the GKT paradigm for the feedback group, a feedback indicating whether the response was judged honest or deceptive was given by the lie-detector after each response of the participant, while in the GKT paradigm for the no-feedback group, there was no feedback after their responses. The results showed that the feedback group showed stronger acitivities in the left dorsal anterior cingulate cortex, insula, parietal lobe, the bilateral occipital lobe, and the right cerebellum than the no-feedback group, irrespective of whether the participants did deceptive responses or honest responses. While deceptive responses elicited stronger activities in the right ventral lateral prefrontal cortex, posterior medial frontal cortex, middle temporal gyrus, the bilateral supramarginal gyrus, and the left insua, relative to honest responses, regardless of which group the participants belonged to. Most importantly, the left insula and the bilateral striatum were selectively more responsive for the deceptive response versus honest response in the feedback group than in the no-feedback group. These results initially demonstrate the important influence of deception consequence on the neural mechanism of producing deception, suggesting that deception would arouse stronger risk evaluation and reward predication processes when the deceiver had to face some consequences after telling lies than when there were no consequences following their lies.

deception consequence, deception production, neural mechanism, functional magnetic resonance imaging

doi: 10.1360/052013-59