

거짓 탐지와 뇌과학 : 기능적 자기공명영상을 활용한 거짓 탐지

이화여자대학교 뇌융합과학연구원,¹ 서울고등법원,² 이화여자대학교 대학원 약학과,³ 이화여자대학교 경영대학,⁴ 이화여자대학교 뇌·인지과학과⁵

최예라¹ · 김상준² · 도혜인³ · 신경식⁴ · 김지은^{1,5}

Detecting Deception Using Neuroscience : A Review on Lie Detection Using Functional Magnetic Resonance Imaging

Yera Choi, MS,¹ Sangjoon Kim, JD,² Hyein Do, B. Pharm,³ Kyung-Shik Shin, PhD,⁴ Jieun E. Kim, MD^{1,5}

¹Ewha Brain Institute, Ewha Womans University, Seoul, Korea

²Seoul High Court, Seoul, Korea

³College of Pharmacy, Graduate School of Pharmaceutical Sciences, Ewha Womans University, Seoul, Korea

⁴School of Business, Ewha Womans University, Seoul, Korea

⁵Department of Brain and Cognitive Sciences, Ewha Womans University Graduate School, Seoul, Korea

Since the early 2000s, there has been a continued interest in lie detection using functional magnetic resonance imaging (fMRI) in neuroscience and forensic sciences, as well as in newly emerging fields including neuroethics and neurolaw. Related fMRI studies have revealed converging evidence that brain regions including the prefrontal cortex, anterior cingulate cortex, parietal cortex, and anterior insula are associated with deceptive behavior. However, fMRI-based lie detection has thus far not been generally accepted as evidence in court, as methodological shortcomings, generalizability issues, and ethical and legal concerns are yet to be resolved. In the present review, we aim to illustrate these achievements and limitations of fMRI-based lie detection.

Key Words Lie detection · Brain · Functional magnetic resonance imaging.

Received: June 17, 2015 / Revised: June 23, 2015 / Accepted: June 26, 2015

Address for correspondence: Kyung-Shik Shin, PhD

School of Business, Ewha Womans University, 52 Ewhayeodae-gil, Seodaemun-gu, Seoul 03760, Korea

Tel: +82-2-3277-2799, Fax: +82-2-3277-2776, E-mail: ksshin@ewha.ac.kr

Address for correspondence: Jieun E. Kim, MD

Department of Brain and Cognitive Sciences, Ewha Womans University Graduate School, 52 Ewhayeodae-gil, Seodaemun-gu, Seoul 03760, Korea

Tel: +82-2-3277-6932, Fax: +82-2-3277-6932, E-mail: kjeun@ewha.ac.kr

서 론

살아있는 사람에서 뇌의 활성 상태를 비침습적으로 측정할 수 있는 기능적 자기공명영상(functional magnetic resonance imaging, 이하 fMRI)은 뇌내 신경세포들의 활동이 증가하면 해당 영역의 혈류량이 증가하고, 이에 따라 산화헤모글로빈 및 환원헤모글로빈의 농도가 변화한다는 사실에 착안하여 뇌의 활동을 측정하는 뇌영상 기법이다.¹⁾ fMRI가 개발된 이후 이를 활용한 연구의 수는 지속적으로 증가하고 있으며, 미국 국립의학도서관에서 제공하는 의생명 분야의 검색엔진인 PubMed에서의 검색결과에 따르면 최근 10년간

매년 수천여 건의 관련 논문이 출판되고 있다.

fMRI의 개발 및 활용도의 증가로 뇌의 활동과 감각 및 행동을 비교적 높은 공간 및 시간 해상도 수준에서 긴밀하게 연관 지을 수 있게 된 것과 때를 같이하여, 단순한 감각이나 행동을 넘어 고차원적이고 복합적인 감정이나 행동, 사회적 상호작용 등에 대한 뇌과학적 이해를 증진하고자 하는 사회 신경과학(social neuroscience) 등의 분야가 1990년대 초 이래 발전하게 되었다.²⁾ 이러한 신경과학 분야가 fMRI 등의 뇌영상 기법을 적극적으로 활용하여 뇌가 인간 행동 전반을 관장한다는 점을 밝히면서, 그 성과를 바탕으로 하여 도덕, 윤리, 법에 대해 새로운 관점에서 접근하고자 하는 신경윤리학

(neuroethics), 신경법(neurolaw)과 같은 분야가 태동되었다.³⁾

이러한 신경윤리학, 신경법 등의 신흥 학문은 인간의 자유 의지와 책임, 본성에 대한 여러 중대한 의문을 제기했을 뿐만 아니라, 뇌과학에 기반한 증거가 법정에서 어떻게 활용되어야 하는지에 대한 논쟁 또한 필연적으로 다루게 되었다. 특히 어조, 표정 등의 비언어적 요소와, 혈압, 땀, 호흡 등의 생리적 요소 등을 통해 증언의 참, 거짓 여부를 판별하기 위한 오랜 노력에 fMRI 등의 최신 뇌영상 기법이 가세하게 되면서, 뇌영상 증거를 법정에서의 거짓 판별에 사용하는 것에 대한 기대와 우려가 공존하고 있다.⁴⁾ 이에 본 논문에서는 fMRI를 활용한 거짓 판별에 대한 연구 결과들을 살펴보고, 그 유용성과 한계에 대해 논의하고자 한다. 본 논문에서는 PubMed 검색을 활용하여 제목 또는 초록에서 거짓과 관계된 단어(lying, deception, deceptive, deceive, dishonest, dishonesty)와 fMRI를 언급한 논문 중 관련성이 없거나 약한 논문을 제외한 58건을 논의 대상으로 삼았다.

본 론

거짓 행동에 관한 fMRI 연구 개괄

2000년대 초에 이르러, 최근 일화 기억(episodic memory)에 대해 참인 응답과 거짓인 응답을 했을 때 각각 활성화되는 부위를 비교한 Spence 등⁵⁾의 연구를 시작으로 거짓 행동과 관계된 뇌 부위를 파악하기 위한 초기의 fMRI 연구들이 보고되기 시작하였다.⁶⁻⁸⁾ 이러한 연구들은 기본적으로 참여자들에게 여러 질문을 제시하고, 이에 대해 진실된 반응을 보일 때와 거짓된 반응을 보일 때의 fMRI 반응 차이를 통해 거짓 행동에 특징적으로 관여하는 뇌 부위를 가려내고자 하였다. 이를 통해 전전두피질(prefrontal cortex, 이하 PFC),^{5,7)} 전대상회피질(anterior cingulate cortex, 이하 ACC^{a)}),⁶⁾ 두정피질(parietal cortex)^{6,7)} 및 피질하(subcortical) 영역⁷⁾ 등의 다양한 부위가 거짓 행동과 관계된 부위로 제시되었다. 하지만 이러한 연구들은 일관된 뇌과학적 증거를 보이는 데는 부족함이 있었고, 사용된 실험 패러다임이 거짓 행동이 이루어지는 실제 상황을 충분히 반영하지 못하였으며, fMRI 결과를 바탕으로 개인 수준에서 참, 거짓을 판별하려는 시도가 이루어지지 않았다는 한계점이 있었다.

이후 Kozel 등⁹⁾은 참여자들이 2개의 물건 중 하나를 가져가는 모의 절도를 수행하게 하고, 그들이 가져간 물건과 가져가지 않은 물건에 대한 질문들에 대한 fMRI 반응을 측정하

는 방식으로 실제 거짓 행동이 발생하는 상황과 더욱 유사한 실험을 수행하였다. 또한 선행 연구를 통해 드러난 관련 부위인 우측 ACC, 하안와전두피질(inferior orbitofrontal cortex), 하전두피질(inferior frontal cortex), 중간전두피질(middle frontal cortex) 및 좌측 중간측두피질(middle temporal cortex)을 관심영역으로 하여 모형수립군(model-building group)에서 해당 부위들의 활성을 확인하였으며, 거짓 반응과 관계되어 활성화된 부위들을 기반으로 모형수립군과 모형시험군(model-testing group)에서 각각 93%, 90%의 정확도(accuracy)로 참여자들이 가져간 물건을 특정하였다고 보고하였다.

이러한 초창기의 연구들 이래 여러 패러다임을 사용한 fMRI 연구들이 거짓 및 속임수와 관계된 뇌 영역을 판별하고, 나아가 이를 통해 거짓 행동을 효과적으로 판별하는 방법론을 확립하고자 하였다.^{4,10)} Christ 등¹⁰⁾의 구분에 따르면, 이러한 연구들은 1) 법과학(forensic sciences) 분야에서 범죄에 관련한 특징적인 상세정보를 인식하는지를 판단하기 위해 폭넓게 사용되어 온 guilty knowledge test(GKT)에 기반한 방법으로 물건의 소유 여부에 대해 거짓된 반응을 보이도록 하거나, 2) 과거의 자전적 정보 및 개인적 경험, 3) 최근에 수행한 활동, 또는 4) 최근에 획득한 지식과 관련한 거짓된 반응을 보이도록 하여 이를 진실된 반응 및 중립적 반응과 비교하는 패러다임을 주로 활용하고 있다. 이를 통한 성과는 상업화에까지 이어져, 2006년 이후 Langleben 등¹¹⁾의 방법론¹²⁾에 기반한 No Lie MRI와, Kozel 등¹³⁾의 방법론⁹⁾에 기반한 Cephos 등 fMRI를 활용한 상업적 거짓 탐지 서비스가 제공되기 시작하였다.⁴⁾

이렇듯 관련 분야의 연구가 계속해서 활발하게 진행되면서, 2009년 Christ 등,¹⁰⁾ 2014년 Farah 등⁴⁾이 현재까지의 관련 연구들을 정리하고 이를 바탕으로 한 수렴적 증거(converging evidence)를 제시하기 위한 메타분석(meta-analysis) 연구를 각각 수행하였다. Christ 등¹⁰⁾은 거짓 판별에 관한 대부분의 fMRI 선행연구들에서 실험 패러다임의 다양성에도 불구하고 실행 제어(executive control)에 있어 중요한 역할을 담당하는 영역인 PFC와 ACC가 지속적으로 관련 부위로 특정되어 왔다는 점을 지적하며, 이는 거짓 행동을 강도 높은 실행통제 능력을 요구하는 행동으로 보는 시각과도 잘 합치한다는 것을 강조하였다. 또한, 메타분석에서 더욱 정확한 뇌영역 특징을 가능하게 하는 활성 우도 추정(activation likelihood estimate, 이하 ALE) 기법을 활용하여, 양전자단층촬영(positron emission tomography) 및 fMRI를 사용한 관련 연구들에서 배외측(dorsolateral) PFC, 후두정피질(posterior parietal cortex), 복외측(ventrolateral) PFC, 전섬엽

a) ACC가 PFC에 포함되는 것으로 보는 시각과, ACC가 PFC와 구분되는 것으로 보는 시각이 모두 존재하나, 본 논문에서는 PFC와 ACC가 함께 언급된 경우 ACC를 개별적으로 명시하였다.

(anterior insula) 및 ACC 등의 부위가 거짓 행동과 연관되어 있으며, 이들이 각각 작업 기억(working memory), 억제 제어(inhibitory control), 작업 전환(task switching) 등 실행 제어의 여러 측면과 관련이 있음을 보였다.¹⁰⁾ 더욱 최근에 이루어진 Farah 등⁴⁾의 메타분석 연구 또한 ALE 기법을 활용하여, 여러 fMRI 선행연구들에서 공통적으로 배외측 및 복외측 PFC, 하두정피질(inferior parietal cortex), 전섬엽, 내측상전두피질(medial superior frontal cortex) 등이 거짓 행동과 관련된 것으로 나타났음을 보였다.

거짓 행동에 관한 fMRI 연구의 한계와 전망

앞서 소개된 바와 같이, 거짓 행동에 관여하는 뇌 부위를 파악하고자 시도한 여러 fMRI 연구들은 연구 방법론 및 결과의 변동성(variability)에도 불구하고 PFC, ACC 등의 부위를 공통적으로 제시하는 등 일정 수준의 수렴적 증거를 보이는 데 성공하였다. 그러나 Farah 등⁴⁾은 이에 사용된 방법론에 내재되어 있는 문제점과, 이러한 연구 성과를 현실의 상황, 특히 법정에서 적용하는 것의 한계를 분명히 파악할 필요가 있음을 시사하고 있다.

방법론적 측면에서, Farah 등⁴⁾은 기억 및 주의 교란변인(confound)의 영향을 주요한 문제점으로 언급하고 있다. 일례로 Hakun 등⁴⁾의 연구에서는 특정한 숫자를 선택하게 한 뒤 단순히 숫자를 관찰하게 한 경우와, 선택한 숫자에 대해 거짓 진술을 하도록 한 경우 모두에서 다른 숫자에 비해 선택한 숫자에 대한 외측 PFC와 두정피질에서의 활성이 증가하는 것으로 나타났다. 또한, Gamer 등¹⁵⁾은 특정 물건을 기억 하도록 한 뒤, 기억한 물건과 그렇지 않은 물건을 단순히 관찰하게 하며 fMRI를 촬영하였을 때에도 PFC와 전섬엽에서의 활성이 증가하였다고 보고하였다. 이는 거짓 행동에 관여하는 뇌 부위에 대한 기존의 여러 fMRI 연구 결과가 기억 및 주의 교란변인의 영향으로부터 자유롭지 못할 가능성을 시사한다. 이에 따라, 기존 연구에서 제시된 여러 영역들이 거짓 행동 그 자체가 아니라, 거짓 행동을 하는 데 요구되는 기억 능력이나 인지적 부하와 관계되어 높은 활성을 보인 것은 아닌지 고려해야 한다.

한편, 현실 적용의 측면에서, Farah 등⁴⁾은 먼저 fMRI를 통한 개인 수준에서의 거짓 판별의 정확도에 대한 의문을 제기하고 있다. 거짓 판별의 정확도는 거짓이 존재할 때 이를 옳게 판별할 확률인 민감도(sensitivity)와, 거짓이 존재하지 않을 때 이를 옳게 판별할 확률인 특이도(specificity), 그리고 전체 모집단(population) 내에서 거짓이 발생할 확률인 기저 확률(base rate) 등에 의해 영향을 받게 된다. 일례로 fMRI를 이용하여 거짓 판별을 시행하고 민감도와 특이도를 평가한

Kozel 등¹³⁾은 모의 범죄를 수행하고 거짓인 진술을 한 참여자들 중 100%에서 거짓을 거짓으로 옳게 판별하였으나, 모의 범죄를 수행하지 않고 참인 진술을 한 참여자들 중 67%에서 참을 거짓으로 잘못 판별하였다고 보고하였다. 이렇게 판별의 특이도가 낮은 경우, 특히 거짓 행동의 기저 확률이 낮을 때 참을 거짓으로 오판하는 거짓 양성(false positive) 판정의 확률이 크게 증가하는 문제가 발생하게 된다.⁴⁾¹⁶⁾

Farah 등⁴⁾은 이에 더해, 뇌의 구조 및 기능, 나이, 질환, 성격 등의 개인별 특질이 거짓 행동과 관계된 fMRI 반응에 현저한 영향을 미쳐 사이코패스(psychopath) 등 특정 집단에서 보편적 거짓 판별 기준을 적용할 수 없을 가능성과, 거짓이 강한 개인적 감정과 연관되어 있는 경우 거짓 판별에 이용되는 fMRI 지표에 변동이 생길 가능성, 그리고 거짓 판별 시험에 적발되지 않기 위한 대응 기술(countermeasure)의 사용이 fMRI를 이용한 판별의 정확도를 감소시킬 가능성 등을 극복해야 할 난점으로 언급하고 있다.

이러한 과학적 측면에서의 한계점들과, 결론이 법적 강제성을 갖게 되는 법정이라는 공간의 특수성으로 인해, fMRI 기반 거짓 탐지 데이터를 법적 증거로 활용하는 것은 아직까지 어려운 실정이다. 미국의 경우 2010년 이래 3건의 재판에서 fMRI 기반 거짓 탐지 데이터를 증거로 제출하려는 시도가 있었으나, 이는 과학계에서 일반적으로 통용되는 근거가 아니라는 이유로 무산되었다.⁴⁾ 한편, 이에 대해 비록 과학적으로 충분히 확립되지 못한 증거일지라도 더 나은 증거 수단이 없다면 이를 활용하는 것이 법적 관점에서는 더 유익할 수 있다는 실용주의적 견해도 제시되고 있다.¹⁷⁾ 우리나라의 경우에도 fMRI 증거의 법정 활용 가능성은 법학계¹⁸⁾ 및 재판 실무자들로부터 지대한 관심의 대상이 되고 있다. 미국의 Daubert¹⁹⁾ 판례법리를 수용한 우리나라 대법원 2011.9.2. 선고 2009다52649 전원합의체 판결²⁰⁾에 의한다면, fMRI 증거는 앞으로 과학적 증거의 신빙성(또는 허용성) 인정 요건 중 오차율(특이도) 및 일반적 승인 등의 요건을 충족할 수 있어야 법정에서 신뢰할 수 있는 과학적 증거로 받아들여질 수 있을 것이다.^{b)} 향후 관련 연구의 발전 및 뇌과학적, 법학적 논의를 통해 fMRI가 법정에서 의미 있는 증거로 활용될 수 있을지 귀추가 주목된다.

b) 대법원 2011.9.2. 선고 2009다52649 전원합의체 판결은 과학적 증거의 신뢰성 여부의 인정방법으로 “그 이론이나 기술이 실험될 수 있는 것인지, 이론이나 기술에 관하여 관련 전문가 집단의 검토가 이루어지고 공표된 것인지, 오차율 및 그 기술의 운용을 통제하는 기준이 존재하고 유지되는지, 그 해당 분야에서 일반적으로 승인되는 이론인지, 기초자료와 그로부터 도출된 결론 사이에 해결할 수 없는 분석적 차이가 존재하지는 않는지 등을 심리·판단하는 방법”에 의할 것을 제시하고 있다.

결 론

fMRI를 활용한 거짓 탐지는 2000년대 초 태동한 이래 여러 방법론을 사용한 연구와 메타분석 등을 통해 PFC, ACC, 두정피질, 전섬엽 등 거짓 행동에 관여하는 여러 뇌 부위를 수렴적으로 밝히는 데 이르렀다. 그러나 한편으로는 기억 및 주의 교란변인으로 인한 방법론적 문제점과, 거짓 탐지의 정확도에 영향을 미칠 수 있는 낮은 특이도와 기저 확률, 개인적 특질의 차이, 거짓과 관계된 감정 요인, 대응 기술 사용 등으로 인한 어려움, 그리고 법정이라는 특수한 환경에 따르는 제약 등으로 인한 여러 한계점을 안고 있다. 향후 fMRI 기반 거짓 탐지 데이터가 법정에서 의미 있는 증거로 활용되기 위해서는 더욱 정교하고 창의적인 연구와, 뇌과학적 관점과 법학적 관점을 포괄하는 심도 있는 논의가 지속적으로 이루어져야 할 것이다.

중심 단어: 거짓 탐지·뇌·기능적 자기공명영상.

Acknowledgments

이 논문은 2013년도 정부재원(교육부)으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2013S1A3A2054667).

Conflicts of interest

The authors have no financial conflicts of interest.

REFERENCES

- 1) Huettel SA, Song AW, McCarthy G. Functional magnetic resonance imaging. 2nd ed. Sunderland: Sinauer Associates, Inc.;2009.
- 2) Cacioppo JT, Amaral DG, Blanchard JJ, Cameron JL, Carter CS, Crews D, et al. Social Neuroscience: Progress and Implications for Mental Health. *Perspect Psychol Sci* 2007;2:99-123.
- 3) Aggarwal NK. Neuroimaging, culture, and forensic psychiatry. *J Am Acad Psychiatry Law* 2009;37:239-244.
- 4) Farah MJ, Hutchinson JB, Phelps EA, Wagner AD. Functional MRI-based lie detection: scientific and societal challenges. *Nat Rev Neurosci* 2014;15:123-131.
- 5) Spence SA, Farrow TF, Herford AE, Wilkinson ID, Zheng Y, Woodruff PW. Behavioural and functional anatomical correlates of de-

- ception in humans. *Neuroreport* 2001;12:2849-2853.
- 6) Langleben DD, Schroeder L, Maldjian JA, Gur RC, McDonald S, Ragland JD, et al. Brain activity during simulated deception: an event-related functional magnetic resonance study. *Neuroimage* 2002;15:727-732.
- 7) Lee TM, Liu HL, Tan LH, Chan CC, Mahankali S, Feng CM, et al. Lie detection by functional magnetic resonance imaging. *Hum Brain Mapp* 2002;15:157-164.
- 8) Ganis G, Kosslyn SM, Stose S, Thompson WL, Yurgelun-Todd DA. Neural correlates of different types of deception: an fMRI investigation. *Cereb Cortex* 2003;13:830-836.
- 9) Kozel FA, Johnson KA, Mu Q, Grenesko EL, Laken SJ, George MS. Detecting deception using functional magnetic resonance imaging. *Biol Psychiatry* 2005;58:605-613.
- 10) Christ SE, Van Essen DC, Watson JM, Brubaker LE, McDermott KB. The contributions of prefrontal cortex and executive control to deception: evidence from activation likelihood estimate meta-analyses. *Cereb Cortex* 2009;19:1557-1566.
- 11) Langleben DD, Loughhead JW, Bilker WB, Ruparel K, Childress AR, Busch SI, et al. Telling truth from lie in individual subjects with fast event-related fMRI. *Hum Brain Mapp* 2005;26:262-272.
- 12) Davatzikos C, Ruparel K, Fan Y, Shen DG, Acharyya M, Loughhead JW, et al. Classifying spatial patterns of brain activity with machine learning methods: application to lie detection. *Neuroimage* 2005;28:663-668.
- 13) Kozel FA, Johnson KA, Grenesko EL, Laken SJ, Kose S, Lu X, et al. Functional MRI detection of deception after committing a mock sabotage crime. *J Forensic Sci* 2009;54:220-231.
- 14) Hakun JG, Seelig D, Ruparel K, Loughhead JW, Busch E, Gur RC, et al. fMRI investigation of the cognitive structure of the Concealed Information Test. *Neurocase* 2008;14:59-67.
- 15) Gamer M, Klimecki O, Bauermann T, Stoeter P, Vossel G. fMRI-activation patterns in the detection of concealed information rely on memory-related effects. *Soc Cogn Affect Neurosci* 2012;7:506-515.
- 16) Wolpe PR, Foster KR, Langleben DD. Emerging neurotechnologies for lie-detection: promises and perils. *Am J Bioeth* 2005;5:39-49.
- 17) Schauer F. Can bad science be good evidence? Neuroscience, lie detection, and beyond. *Cornell Law Rev* 2010;95:1191-1220.
- 18) Hong SO. The accuracy and legal effects of fMRI lie-detection evidence. *Seoul Law J* 2011;52:511-540.
- 19) Cornell University Law School Search Cornell [homepage on the Internet]. Daubert V. Merrell Dow Pharmaceuticals (92-102), 509 U.S. 579 (1993) [cited 2015 Aug 12]. Available from: <https://www.law.cornell.edu/supct/html/92-102.ZS.html>.
- 20) Supreme Court of Korea en banc Decision 2009Da52649 Decided September 2 2011.