

특별호: 심리학과 4차 산업혁명 2

제4차 산업혁명 시대의 법심리학: 새로운 도전과 과제

박 광 배[†]

인류가 머지않아 진입하게 될 제4차 산업혁명 시대는 물리적 세계와 가상세계, 생물학적 영역과 디지털 영역의 경계를 부수고 융합하는 기술의 발전에 의해 주도될 것으로 예상된다. 새로운 기술의 발전은 장기적으로 인간에 대한 법의 전제와 법의 기능 및 역할에 변화를 가져올 것으로 예상된다. 단기적으로는, 기술의 발전이 사법시스템에 종사하거나 관련된 사람들을 조력, 지원, 인도하고, 나아가서 사람이 하던 일과 업무를 일부 대체하거나, 일부 민사 및 형사 절차가 온라인과 컴퓨터 시스템에 의해 이루어지는 등의 변화를 초래할 것으로 예상된다. 이러한 법환경의 변화는 사람과 법의 상호작용 과정에서 발생하는 문제를 파악하고 해결 방안을 모색하는 법심리학의 적응적인 진화도 촉진하게 될 것으로 기대된다. 본 소고에서는 최근에 발전된 기술들이 이미 사법영역에 유입되어 나타되기 시작한 미세한 변화들 예컨대, 인공지능에 의한 범죄예측, 등에 기초하여, 멀지 않은 미래에 법심리학이 포렌식평가, 배심원 선정, 과학적 증거, 사법 의사결정, 그리고 수사 영역에서 새롭게 당면하게 될 것으로 예상되는 도전과 과제를 점검하였다. 또한, 기술에 기반한 새로운 범죄와 피해자보호의 문제가 나타나면서, 법심리학이 범죄자와 피해자에 대한 연구에서 새롭게 모색해야 할 잠재적 이슈들에 관해 논의하였다.

주요어 : 제4차 산업혁명, 기술, 과학, 법심리학, 법

[†] 교신저자: 박광배, 법무법인 케이스앤피 법과인간행동연구소. 충북대학교 심리학과 (명예교수).
Email: kwangbai@chungbuk.ac.kr.

과학기술의 발전에 의해 산업과 인간활동이 또 다시 혁명적으로 도약하게 되는 시대가 예견되고 있다. Schwab(2016)은 소위 “제4차 산업혁명”으로 불리는 생산 수단과 방식 및 절차의 획기적인 변화는 인간이 살아가고, 일하고, 관계를 형성하고, 상호작용하는 행위와 활동뿐만 아니라, 생각하고, 느끼고, 이해하는 의식의 질과 양태를 근본적으로 바꿀 것이고, 그 변화의 규모와 수준, 그리고 복잡성은 인간이 지금까지 경험하지 못한 전혀 새로운 것이 될 것으로 예상하였다.

제4차 산업혁명 시대에 모든 정보와 생산물을 방대한 네트워크로 연결하여 고도의 복잡성을 가진 문제를 다루는 기술, 생물학적 기제와 공학적 구조를 물리적으로 융합하여 인간과 기계의 기능과 역량을 확장하는 기술 등에 의해 인간의 삶과 활동, 의식의 질과 양태가 근본적으로 변화하면, 사회과학과 인문과학에도 내용적, 방법적 변화가 생기게 된다. 사회의 질서와 가치가 순리적으로 유지, 보존될 수 있도록 조절하고, 통제하는 법의 기능과 역할 또한 변화해야 하고, 사람과 법의 상호작용 과정에서 발생하는 문제를 파악하고 해결방안을 모색하는 법심리학(Psychology and Law, Legal Psychology)과 법정심리학(Forensic Psychology)도 적응적으로 진화할 것으로 기대된다.

법심리학은 심리학의 이론과 임상실무를 사법제도와 사법절차에 관여되는 사람에게 적용하여 법이 사람의 행동에 미치는 효과와 사람의 행동이 법에 초래하는 영향을 과학적으로 탐구하는 연구분야다(Ogloff, 2002). 법심리학자는 인간행동에 관한 법의 가정을 평가하고, 사회적 변화에 대해 법이 반응하는 방식에 관심을 가진다. 예를 들어, 법은 의도적으로 허

위의 진술을 할 이유가 없고, 감각과 인지 기능에 명백한 결함이 없는 목격증인의 진술은 사실과 부합할 것으로 가정한다. 그래서 대부분의 재판에서 목격증인의 진술은 일반적으로 매우 강력한 증거로 인정된다. 그러나 심리학적 원리와 이론을 최초로 법 영역에 적용하여 법심리학을 태동시킨 인물로 알려진 휴고 뮌스터버그(Hugo Münsterberg)는 “증언대에서: 심리학과 범죄에 관한 에세이(On the Witness Stand: Essays on Psychology and Crime)”라는 저술(Münsterberg, 1908)을 통해 허위진술의 의도가 전혀 없는 경우에도 목격증인의 진술, 피고인의 자백 등이 사실과 괴리할 수 있고, 수사기관과 법정에서 발생하는 암시성(suggestibility)에 의해 진술인의 지각, 기억, 판단, 등이 왜곡될 수 있다는 것을 설명하였다. 이후, 엘리자베스 로프투스(Elizabeth Loftus)를 비롯한 많은 실험심리학자들이 인간의 기억이 외부에서 주어지는 암시, 잘못된 정보, 유도 질문 등에 의해 매우 쉽게 변형될 수도 있고 (Loftus & Palmer, 1974), 심지어는 존재하지 않았던 사실에 대한 가짜 기억이 만들어질 수도 있다는 것을 실증적으로 보여주었다. 게리 웰스(Gary Wells)와 그의 동료학자들은 수사기관에서 목격진술을 채증하는 절차에 따라 사실에 부합하지 않는 목격자의 기억과 진술이 만들어질 수 있다는 것을 입증하였고(Wells et al., 1979; Wells, 1978; Wells et al., 2015), 스테핀 쉐시(Stephen Ceci)는 범죄, 학대, 방임의 피해가 의심되는 아동의 기억, 피암시성, 증언능력에 관한 연구들을 통해 피해가 의심되는 아동의 경험과 기억을 파악하기 위한 면담 방식에 따라 아동의 기억과 진술이 매우 예민하게 왜곡되고 달라질 수 있다는 것을 보여주었다 (Ceci & Bruck, 1995). 이러한 실험연구 결과의

생태계 타당도를 의심의 여지없이 뒷받침하는 실제 사건들이 현실의 재판과정에서도 나타났다. 미국의 일부 형사재판에서 피고인측 변호인들이 변호인 중 한사람을 범정의 피고인석에 앉아있게 하였는데, 범죄 발생 당시 범인을 목격한 검사측 증인이 증언을 하면서 피고인석에 앉아있는 변호인을 자신이 본 범인으로 지목하였다(Illinois v. Simac, 161 Ill.2d. 297, 1994; United States v. Sabatar, 830F.2d 7(2d Cir), 1987; People v. Gow, 382N.E.2d 673, App. Ct., 1978).

목격자의 기억과 진술에 관한 초기 연구에 기초해서 지난 50년간 얼굴지목(face identification)의 정확성에 관해 1,000개가 넘는 연구들이 지속적으로 이루어진 바, 최근에 Lee(2019)는 그 대부분의 실증연구들을 메타분석하여 얼굴지목의 정확성에 관한 중요한 핵심 사실들을 정립하였다. 첫째로 얼굴지목의 정확성에 영향을 주는 것으로 알려진 38개의 변인들 중, 5개 정도의 변인들이 일관된 영향을 주는 안정적인 요인으로 파악되었고, 연구들의 미세한 방법적 특성(예. 실험자 지시)이 얼굴지목의 정확성에 관한 인지심리학적 연구들과 목격지목에 관한 법심리학 연구들 사이의 차이를 설명하는 것으로 판단되었다. 얼굴지목과 범인지목의 정확성과 신빙성에 관한 과학적인 실증연구들이 누적된 결과, 목격증인이 의도적으로 허위의 진술을 할 이유가 없는 경우에도 그 진술을 맹목적으로 신뢰해서는 안 된다는 인식이 일반화되었고, 사법절차에서 증거능력이 인정될 수 있는 목격진술의 요건이 구체화되었다(Wells & Quilivan, 2009; Wells et al., 2020).

목격진술의 정확성에 관한 연구들로 시작된 법심리학은 전통적으로 사법적 사실판단의 정

확성, 사법절차의 투명성과 공정성, 증거와 증언의 신빙성과 타당성에 관련된 과학적 지식을 발굴하고, 심리평가와 심리치료의 이론과 지식을 사법기관에 연관, 연루된 사람에게 적용하여 사법절차와 체계의 효율적인 운영을 조력하는 역할을 한다. 구체적으로는, 법의 제정, 적용, 운용(절차), 그리고 사법절차와 개인의 관계 및 상호작용에서 인간의 존엄성(dignity), 사생활보호(privacy), 정의(justice), 공정성(equality), 자율성(autonomy) 등의 심리적 가치가 보호, 보장되도록 끊임없이 모니터링하고, 그러한 심리적 가치의 손상이 예상되거나 발견되면 그 원인을 규명하고 대안을 모색하는 학문이다.

제4차 산업혁명 시대에 발전하는 혁신적인 기술로 인해 사법절차에 나타나는 미세한 변화는 점차로 크게 확대될 것이고, 그에 따라 법심리학도 지속적으로 새로운 연구 아젠다와 연구방법을 모색, 발전시킬 것으로 믿어진다. 예를 들어, 라인업에서 범인을 지목하는 목격자의 미세행동을 기계학습(machine learning) 알고리즘으로 분석하여 목격지목의 정확성을 예측하는 행동변인(e.g., 라인업에 제시된 각각의 얼굴을 목격자가 보는 횟수와 시간 등)을 모색하는 연구가 최근에 이루어지고 있다(Price et al., 2020). 기계학습 원리에 기반한 Support Vector Machine(SVM)은 유사한 목적으로 사용되는 다른 방법들 즉, 로지스틱회귀모형, 결정트리(decision tree), 인공신경망(artificial neural network)과 비교했을 때 정확성이 가장 높은 예측/분류 방법으로 알려졌다(Amini et al., 2016). 그러나 과학의 발전과 그에 따른 사회 및 인간의 변화와 함께 이루어질 심리과학의 변화와 발전을 장기적으로 예측하는 것은 불가능한 일이다. 본 소고는 최근에 발전된 기

술들이 이미 사법절차에 유입되어 나타되기 시작한 미세한 변화들에 기초하여, 향후 멀지 않은 미래에 법심리학이 새롭게 당면하게 될 것으로 예상되는 도전과 과제를 점검하기 위한 것이다.

기술의 혁신적 발전

산업혁명은 기술의 혁신적 발전에 의해 경계가 바뀌는 일련의 사건들로 정의된다(Mokyr, 2009). 현대문명의 발전과정에서 지금까지 기술의 발전에 수반되는 도약적인 변화는 심리학자 에이브라함 마슬로우(1908~1970)가 제안한 인간의 기본적인 욕구 피라미드(Maslow, 1943)의 보편적 실현을 단계적으로 이루는 방향으로 나타났다. 18세기 후반~19세기 전반에 걸쳐 영국에서 증기기관의 발명으로 촉진된 제1차 산업혁명은 생산수단을 기계화하였고, 19세기 중후반에 전기의 활용으로 이루어진 제2차 산업혁명은 대량생산 시대를 열었다. 생산의 기계화와 그에 따른 대량생산은 보편적 인간의 생리적 욕구(physiological needs)와 안전욕구(safety needs)를 충족할 수 있는 기회와 수단을 광범위하게 제공하여 세상에 새로운 변화를 가져왔다. 현재 진행 중인 제3차 산업혁명은 1970년대의 전자기술과 컴퓨터의 발전에 의해 주도되었다. 전자기술에 의해 생산이 자동화되고, 디지털 정보기술은 인간의 사회적 욕구(social needs)의 충족을 탈공간적이며 범인류적인 수준으로 확장하게 하였는데, 모든 개인이 인간의 역사상 유래 없는 방대한 인터넷 망으로 연결되고, 그 네트워크를 통하여 사람들이 경제활동에 참여하고 기여하는 방식에 획기적인 변화가 초래되었다.

전자기술과 정보기술의 발전에 의해 촉진되어 인간 사회가 머지않아 진입하게 될 제4차 산업혁명은 물리적 세계와 가상세계, 생물학적 영역과 디지털 영역의 경계를 부수고 융합하는 기술의 발전에 의해 주도될 것으로 예상된다. 사물인터넷(internet of things: IOT), 로봇틱스(robotics), 빅데이터(big data), 인공지능(artificial intelligence: AI) 혹은 기계학습(machine learning), 나노기술(nanotechnology), 바이오기술(biotechnology), 양자컴퓨팅(quantum computing) 등으로 회자되는 기술의 혁신적인 발전과 융합은 인간욕구의 최상위 단계라고 생각되는 자기실현욕구(self-actualization needs)를 자동화와 일반화의 사회적 양상으로 실현하는 기회를 제공할 것이다. 앞선 산업혁명과정에서 이루어진 대량생산은 동일하고 균일한 생산물을 많이 만들어서 널리 보급하는 시스템을 구축하여 가능하였으나, 인간의 자기실현욕구를 구성하는 존중욕구(esteem needs), 인지욕구(cognitive needs), 그리고 심미적 욕구(aesthetic needs)를 충족하는 산업기술은 개인의 특수한 요구와 필요가 고려된 맞춤형 생산을 대량으로 할 수 있는 “대량 맞춤형생산(mass customization)” 시스템을 구축하게 될 것으로 예상된다(Dombrowski & Wagner, 2014). 상위욕구의 보편적 실현에 집중되는 산업의 발전은 그 진행 속도가 기하급수적으로 가속하게 될 것으로 예상되고, 모든 국가의 모든 산업을 흔드는 광활한 범위, 그리고 생산, 관리, 지배, 통제의 모든 영역에 변화를 초래하는 조직적 충격을 가져온다는 점에서 앞선 세 차례의 산업혁명과 구별될 것으로 예상된다(Schwab, 2016).

“디지털 시대의 재도래(second coming of digital era)”로 불리기도 하는(Coldwell, 2019) 제4차 산업혁명은 정보의 유통과 교환을 촉진하

여 사람들 사이의 범문화적인 이해와 결속을 강화하고, 인류의 삶의 질을 획기적으로 향상시킬 것이며, 전세계적인 부의 증가를 가져올 것으로 기대되는 반면, 인간의 자연적인 능력을 능가하는 인공지능 기술의 발달은 인간의 실존적 위기를 초래할 것으로 우려된다. 예를 들어, 변호사 시험을 갓 통과하여 처음으로 로펌에 취직한 변호사가 로펌에서 수행하는 대부분의 업무는 인공지능(AI) 변호사가 훨씬 효율적으로 더 잘 할 수 있게 된다. 제4차 산업혁명을 주도할 것으로 기대되는 혁신적인 기술들의 발전은 사회와 개인의 안전과 기본권을 해칠 수 있는 새로운 위협 요인을 증가시킬 수도 있다. 또한 그러한 혁신적인 기술을 이용하여 사생활침해, 해킹, 지식재산권침해, 무기의 제조 등의 새로운 형태의 범죄가 나타날 수도 있다. 2010년을 기점으로 전세계 인구의 수보다 많아진 컴퓨터들(Gershenfeld & Vasseur, 2014)의 연결망을 통하여 근거 없는 부정적 정서가 무의식적으로 전염되어 퍼지고(Kramer et al., 2014), 성공적인 사람과 집단에 대한 비현실적인 기대와 가짜 정보가 유통되며, 비합리적인 사고와 극단적 이념이 배양되는 기회가 확장될 수 있다(Schwab, 2016). 산업, 사회, 문화의 모든 영역에서 기술 의존성이 가속화되어 인간의 소외와 사회적 불평등이 심화되고 고착될 가능성 또한 함께 우려된다. 예를 들면, 제4차 산업혁명에 의해 증가하는 부의 최대 수혜자는 물리적, 지적 자본을 소유한 혁신적 자본가, 주주, 투자자 등일 것으로 예상된다. 그래서 노동시장은 “저급기술/저임금(low-skill/low-pay)” 영역과 “고급기술/고임금(high-skill/high-pay)” 영역으로 양분화되고 그로 인한 사회적 긴장이 높아질 수 있다(Schwab, 2016). 고급기술을 가진 노동자에 대

한 수요는 증가하는 반면, 기술수준과 교육수준이 낮은 노동자에 대한 수요는 감소한다. 생산수단의 자동화에 의해 저임금 노동자에 대한 수요가 대거 사라지게 되므로 고급기술을 가지지 않은 노동자는 항시 실직의 위험과 직면하고, 그로 인한 지속적인 스트레스는 노동현장에서의 부정적인 행동과 부실한 생산성을 초래할 수 있다. 생산자가 얻는 이익이 노동시장으로 순환되지 않아서 자본과 노동의 분화와 격차가 커지고, 노동시장의 분포에서 소득의 중간계층이 점차 사라지게 되어 지금까지 높은 평균 소득을 향유하던 국가들의 소득수준이 정체되거나 오히려 감소한다. 이와 같은 중간계층의 몰락과 불만족은 민주주의의 위기를 초래하는 주요인이 될 수 있다.

자기실현욕구의 보편적 구현과 사회적 불평등과 부조리의 심화가 동시에 예상되는 제4차 산업혁명시대는 사법절차와 사법제도에서 개인의 안전과 기본권을 해칠 수 있는 새로운 위협 요인의 출현을 경계하고, 인간의 심리적 가치를 보호, 보장해야 하는 법심리학에 새로운 과제와 도전을 가져올 것으로 예상된다.

기술의 발전과 법심리학

새로운 기술의 발전은 3단계에 걸쳐 사법절차와 제도에 영향을 주고 사법시스템을 재구성하는 역할을 할 것으로 예상된다(Sourdin, 2015). 첫 번째 “사법지원기술(supportive justice technology)” 단계는 기술이 사법시스템에 종사하거나 관련된 사람들을 조력, 지원, 인도하는 것이다. 이 단계에서 새로운 기술은 주로 많은 사람들에게 사법에 대한 정확한 정보를 빠르게 제공하고, 사법서비스에 대한 접근을 용

이하에 하며, 여러가지 대안적인 분쟁해결방법을 제시하고 인도하는 역할을 한다. 두 번째 “사법대체기술(replacement justice technology)” 단계는 기술이 인간의 기능, 나아가서 존재를 대체하는 것이다. 이 단계에서는 여러 단계의 법조인이 하던 일과 업무를 인공지능 기술이 대체하고, 일부 민사 및 형사 공판절차가 소위 ‘산개된 법정(distributed court)’에서 온라인과 컴퓨터 시스템에 의해 이루어질 수도 있다(Rowden, 2015). 세 번째 “사법붕괴기술(disruptive justice technology)” 단계는 기술에 의해 사법절차와 사법시스템 자체가 바뀌어서 완전히 다른 형태의 사법제도를 형성하는 것이다(Sourdin, 2018).

김상준(2013)은 1995년부터 2012년까지 고등법원 형사 합의부 강력범죄 항소사건 중 고등법원에서 1심 유죄판결을 취소하고 무죄판결을 선고한 전체 사건 540건을 분석하였다. 이 사건들 중 513건(95%)은 최종적으로 무죄가 확정되었다. 이 사건들에서 1심 유죄판결은 오판(오류긍정)일 가능성이 높다. 그것이 오판이 아니라면 항소심 혹은 상고심의 무죄판결이 오판(오류부정)이다. 김상준(2013)은 심급간 유무죄 판단차이를 초래한 원인과 이유를 분석하였다. 분석된 요인은 크게 부실증거(피해자 허위/오인진술 49.3%, 정황증거의 문제 23%, 목격자의 오인 지목진술 20.7%, 허위자백 20.4%, 과학적 증거의 오류 13.9%, 공범의 허위자백 11.1%)와 법관의 사실인식 및 편향(e.g., 터널비전, “합리적 의심”에 관한 인식, 암묵적 유죄편향성)의 문제로 나타났다(Kim, 2013). 사물인터넷, 로봇틱스, 빅데이터, 인공지능과 기계학습, 나노기술, 바이오기술, 양자컴퓨팅 등의 기술의 발전은 증거의 질을 향상시켜 일차적으로 이러한 오판요인들의 상당수

를 줄이는데 기여할 것이다. 반면, 새로운 기술의 도입으로 인해 사법절차에 생기는 변화들 중에는 새로운 종류의 자료, 새로운 증거, 새로운 절차와 관련된 심리학적 문제가 잠재된 것들이 있을 것으로 예상된다. 세부적으로는 포렌식평가, 배심원선정, 과학적 증거, 사법사결정, 그리고 수사 영역에서 새로운 기술에 의한 변화와 그에 수반되는 범심리학적 연구의 필요성이 나타날 것이다. 더불어서, 기술에 기반한 새로운 범죄가 생겨나면서 범죄자와 피해자에 대한 새로운 범심리학 연구의 필요성이 조망되어야 한다.

포렌식평가

표준화 vs. 프로파일링

심리평가(psychological testing)는 행동표본의 객관적인 측정을 위한 표준화된 문항들(자극, 질문, 문제, 과제 등)로 구성된 심리검사를 사용해서 특정 요인이나 현상의 존재와 수준을 결정하고, 개인차(능력, 인성, 역량, 태도, 정서 등)를 점수화하는 표준화된 절차와 방법으로 정의된다(Cronbach, 1949; American Psychological Association, 2006; Anastasi & Urbina, 1997). 17세기 유럽에서 길이, 무게 등, 측정의 단위를 통일하기 시작하면서 과학과 산업이 도약적으로 발전하였듯이, 과학적인 심리평가의 핵심은 이론주도적인 표준화(standardization)다. 전통적인 심리평가에서 이루어지는 측정은 측정하고자 하는 바의 내용과 형식을 구체적으로 명시하는 이론적 개념 혹은 구성개념을 구축하고 그것을 객관적으로 측정할 수 있는 표준화된 절차에 의해 이루어진다.

그러나 디지털 기술의 발달로 사람들의 개인차를 평가하고 미래의 행동을 예측하는 새

로운 절차와 방법들이 등장하였다. 웨어러블(wearables), 소셜미디어(social media), 혹은 사회관계망서비스(social networking service) 등의 새로운 기술에 의해 생성되어 누적되는 행동자료 즉, 소위 빅데이터(big data)를 분석하여 개인차를 평가하고 미래의 행동을 예측하며, 집단관계를 파악하는 절차들이 나타나고 있다(Adjerid & Kelley, 2018; Bleidorn & Hopwood, 2018). 심리학 연구방법론 저널 중 하나인 Psychological Methods는 2016년에 특별판(Psychological Methods, 2016, Vol. 21, No. 4)을 기획하여 심리학 연구를 위해 빅데이터를 수집, 분석하고, 결과를 해석, 응용하는 구체적인 방법을 소개하는 10편의 논문을 발표하였다. 이미 소비자 심리학에서는 빅데이터를 분석하여 소비자의 심리적 특성과 심리적 상태를 예측하여 마케팅 전략을 수립하는 수단으로 활용하고 있다(Matz & Netzer, 2017). 또한 스마트폰과 기타 웨어러블 디바이스를 이용하여 실제 생활환경에서 순간적으로 경험하는 행동, 감정, 선택 뿐만 아니라 신체기능, 신체상태까지 실시간으로 측정해서 생태계 타당도가 높은 행동 데이터를 수집하고, 그 데이터에 기초해서 성격적인 특징을 파악하고 미래의 행동을 예측하는 “경험표집방법(experience sampling method)” (Kahneman & Krueger, 2006)과 “야전평가(ambulatory assessment)” (Timons et al., 2017)가 성공적으로 활용되고 있다. 이러한 데이터의 특징은 3V로 표현되는 규모(volume), 다양성(variety), 속도(velocity)다(Laney, 2001). 빅데이터는 양이 방대하고, 다양한 자료를 포함하며, 자료가 생성, 수집되는 속도가 매우 빠른 특징을 가진다.

심리측정과 관련하여 빅데이터는 3V 이외에 또 다른 중요한 특징을 가지는데, 전통적

인 심리과학에서 인간의 행동과 기능을 측정하고 예측하기 위해 사용되는 데이터와는 달리, 표준화되지 않은 절차와 방법에 의해 생성, 수집되는 특징이다. 예를 들어, 심리학은 범죄자의 재범여부를 예측하기 위한 목적으로 “고안된” 자료에 기초해서 재범예측을 하지만, 빅데이터는 특정 목적 없이 혹은 다양한 목적으로 수집된 자료에서 재범위험을 예측하는 인자들을 발굴하기 위해 사용된다. 만약 빅데이터에 편향성, 차별성, 왜곡 등이 구조적으로 존재하는 경우에는 그것들이 재범예측에 그대로 반영될 수 있다(Hannah-Moffat, 2018). 따라서 빅데이터에 기초해서 이루어진 분석과 결과의 신뢰성과 타당성은 항상 이론주도적인 기존의 심리측정방법으로 재확인되어야 한다(Stanley & Byrne, 2016). 즉, 현대의 심리과학은 조만간 도래할 제4차 산업혁명 시대의 표준화되지 않은 절차와 방법으로 불특정의 방대한 행동자료를 수집하고, 분석하여 개인차를 평가하고 개인의 행동을 예측하는 “기술주도적 프로파일링(technology-driven profiling)” 방법들도 과학적인 심리평가의 범주 안에 포괄될 수 있는지를 판단하고(Iliescu & Greiff, 2019; Qiu, Chan, & Chan, 2017), 그 포섭을 위한 구체적인 방법론(e.g., Stanley & Byrne, 2016; Chapman et al., 2016)을 모색해야 하는 도전에 직면하고 있다.

예측과 판별

2002년 영화 마이너리티 리포트(Minority Report)는 경찰이 인공지능 기술을 이용하여 미래의 범죄자를 미리 포착, 체포해서 범죄를 예방하는 사회를 묘사하였다. 범죄를 범하는 사람들을 과학적으로 판별하려는 노력은 19세기 유럽에서 시작되었는데(Pavlich, 2009), 예를

들어, 프란시스 갤턴(Francis Galton)은 많은 상습범죄자들의 얼굴을 조합하여 공통된 특징을 추출함으로써 범죄위험성 평가의 과학적인(인류학적인) 기준을 정립하고자 하였다(Galton, 1878, 1890). 현대의 범죄예측과 위험성평가도 범죄행위를 하는 사람들의 공통된 특징을 알면 그 특징들로 위험성을 평가할 수 있다는 전제 즉, '계리적 평가(actuarial assessment)'에 기초해서 이루어진다. 미국의 회사 노스포인트(Northpoint)는 폭력 범죄자들의 전과기록, 가족관계, 약물남용, 인터뷰 등의 자료를 분석하여 재범가능성을 예측하는 콤파스(Correctional Offender Management Profiling for Alternative Sanctions: COMPAS)라는 이름의 인공지능 알고리즘을 개발하여 그 알고리즘으로 형사피고인의 재범가능성을 예측하여 판사에게 양형 자료로 제공한다(Freeman, 2016; Brennan et al., 2009). 즉, 전통적으로 법심리학자(forensic psychologists)가 해오던 위험성평가, 폭력성평가, 재범가능성평가 등을 인공지능 기술이 대신 하는 것이다.

미국의 위스콘신 주법원은 2013년에 총기에 의한 범죄에 사용된 자동차를 운전한 것으로 알려진 에릭 루미스(Eric Loomis)라는 남자피고인에게 6년의 징역형과 5년의 보호관찰을 선고하였다. 재판의 양형단계에서 위스콘신 교정국은 에릭 루미스의 재범가능성에 관한 COMPAS 예측결과를 양형조사 자료의 하나로 판사에게 제출하였는데, COMPAS의 알고리즘은 기업비밀이라는 이유로 판사에게 알려지지 않았고 재범확률 추정치만 제출되었다. 예측 알고리즘이 공개되지 않았음에도 불구하고, 판사는 높은 재범가능성을 추정한 COMPAS 결과를 토대로 에릭 루미스에게 중형을 선고하였다. 이에 에릭 루미스는 법원이 COMPAS

의 재범예측에 근거하여 자신에게 중형을 선고한 것은 적법절차 위반이라는 이유로 상고하였다. 그의 상고이유는 COMPAS의 재범가능성 예측 알고리즘은 자신과 전혀 관련성이 없는 사람들의 데이터를 학습한 결과이고, 어떤 사람들의 어떤 데이터가 알고리즘의 기반이 된 것인지 알 수 없으며, 알고리즘 자체가 비공개 되었기 때문에 정확한 정보에 기초해서 형을 선고 받을 기본권리가 침해되었고, 재범가능성 예측에 성별이 인자로 고려되는 것이 위헌이라는 것이다. 그러나 위스콘신 대법원은 에릭 루미스의 상고를 기각하였다(State v. Loomis, 881 N.W.2d 749(Wis. 2016)). 다만, 대법원은 판결문에서 인공지능에 의한 재범예측을 양형에 고려하는 판사는 자신이 선고하는 형을 뒷받침하는 다른 정보들을 함께 적시해야 한다는 지침을 명시하였다.

Dresell & Farid(2018)는 COMPAS에 의한 재범예측의 타당도를 검증하기 위하여, 플로리다 주 브로워드 카운티에서 2013년과 2014년 사이에 체포된 7,000명 피고인의 재범가능성을 COMPAS로 예측한 결과, 137개의 예측인자를 사용한다고 알려진 COMPAS의 재범예측 정확성이 전체적으로 65%에 불과하여 매우 낮을 뿐만 아니라, 단지 2개의 인자(피고인의 연령과 전과)만으로 예측했을 때의 정확성과 거의 동일하였다. Dresell & Farid (2018)의 연구에서 드러난 COMPAS의 가장 중대한 결함은 재범예측에 심각한 인종차별이 나타난다는 것이다. COMPAS는 재범하지 않는 흑인피고인의 재범가능성을 백인피고인의 재범가능성보다 거의 2배 더 높게 예측하였는데, 재범하지 않은 흑인피고인 중 44.9%가 재범할 것으로 예측한 반면, 재범하지 않은 백인피고인의 23.5%가 재범할 것으로 예측하였다. 또한 재범한 백인

피고인 중에 47.7%가 재범하지 않을 것으로 예측된 반면, 재범한 흑인피고인은 28.0%가 재범하지 않을 것으로 예측되었다. 이러한 결과는 흑인피고인의 재범가능성 예측에서 백인피고인에 비해서 상당히 높은 오류긍정(false positive)과 상당히 낮은 오류부정(false negative)을 범하는 것을 뜻하고, 흑인에 대한 차별이 알고리즘으로 구조화되었다는 것을 의미하였다.

범죄행위를 하게 만드는 인간의 특성은 아직 과학적으로 밝혀진 것이 없다. 특정인이 과거 범죄자들이 공유하는 특징을 가지면 그 특정인이 미래에 범죄를 범할 위험성이 높다고 평가하는 계리적 평가에 의한 재범예측의 정확성은 높지 않다. Yang, Wong, & Coid(2010)는 1999년부터 2008년 사이에 발표된 재범예측 정확성 연구들을 메타분석하였는데, 그 연구들은 PCL-R(Psychopathy Checklist-Revised: Hare, 2003)을 비롯한 9개의 재범예측/위험성 평가도구들 중 적어도 2개 이상의 도구를 사용하여 계리적 평가에 의한 예측의 정확성을 검증한 연구들이었다. 메타분석에서 9개의 평가도구들은 재범여부에 대하여 공히 중간 정도의 효과크기(Cohen's d0.6)를 가지는 것으로 나타났고, 어떤 평가도구도 다른 것들보다 더 우수한 예측정확성을 보이지 않았다.

계리적 평가 방법에 기초한 위험성평가는 특히 특정 개인의 위험성 혹은 재범 확률을 추정할 때 매우 큰 오류를 유발할 수 있다(Wollert, 2006; Hart, 2005). 범죄자들 속에서 어떤 특성과 범죄행동이 높은 상관관계를 가진다고 해서, 그 특성을 가진 특정인이 미래에 범죄행동을 할 것으로 예측할 수 없는 것이다. 계리적 평가 방법에 의한 위험성 평가의 정확성을 검증한 Hart & Cooke(2013)은 “폭

력을 유발하는 원인에 대한 이해가 없는 상황에서 계리적 평가 방법에 기초해서 미래의 폭력 가능성을 특정의 확률로 추정하는 것은 어떤 합리적인 수준의 정확성과 확실성도 전혀 담보하지 못한다”고 결론지었다. COMPAS와 같은 인공지능 알고리즘도 예외가 아니다. 그런데 인공지능을 이용하여 재범가능성을 평가하는 것은 사법의 적법절차원칙에 대해 더 심각한 위협이 될 수 있다. 영화 마이너리티 리포트에 묘사된 것처럼, 범죄예측이 광범위하게 이루어질 수 있기 때문이다. 법심리학은 과거범죄자들이 공유하는 특성에 기초한 위험성평가의 위험성을 지속적으로 평가하고, 과거 범죄자들이 공유하는 특성에서 범죄행위 혹은 위험성과 관련성을 가진 특성을 찾는 계리적 평가 이외에, 폭력유발 원인에 대한 이해에 기초한 평가 방법과 요인을 모색할 필요가 있다.

배심원선정

법에 관한 전문지식을 가지지 않은 일반인이 사실인정(유무죄 판결)을 하는 배심재판의 장점은 사법적 의사결정을 민주적이고 투명한(공정한) 방법으로 이룬다는 것이다. 배심재판 혹은 국민참여재판을 위해서는 법원이 관할지역에 거주하는 주민들 중 일정수의 배심원후보자를 무작위로 선정하여 배심원후보자 풀(pool) 혹은 소위 “비니어(venire)”를 구성한다. 비니어 구성에서 중요한 것은 관할지역에 거주하는 주민들의 인구학적 분포와 가치, 선호, 이념의 분포가 민주적으로 반영되어야 한다는 것이다. 비니어가 구성되면 배심원 선정절차를 통하여 배심원후보자들 중 재판의 피고 혹은 원고 일방에 편향되지 않는 사람들을 선정

한다. 배심원 선정절차에서는 피고측과 원고측이 각각의 배심원후보자에게 질문을 하고, 특정 후보자가 재판에서 불편부당하고 합리적인 판단을 하기 어렵다고 생각되거나, 자신에게 불리하게 편향될 것으로 판단되면 기피신청을 한다. 배심원 선정절차를 의미하는 라틴어 *voir dire*는 “진실을 말하다”는 의미다. 배심원 선정절차에서 배심원후보자는 주어진 질문에 대하여 솔직하게 응답하는 것이 중요하다. 따라서 배심원 선정절차에는 사생활침해의 위험이 존재한다. 선정절차에서 양당사자로부터 기피되지 않은 후보자들이 배심원으로 공판에 출석해서 증거를 보고 증언을 들은 후 평의를 거쳐 평결 혹은 사실인정을 도출한다.

배심원 선정절차의 목적은 지역사회 구성원들의 생각과 가치를 반영하고, 불편부당함(impartial) 판단을 할 수 있는 사람을 뽑는 것이다. 다시 말해, 지역사회를 대표하는 공정한 배심원을 선정하는 것이 배심재판의 가장 중요한 핵심이다. 그러나 지역사회를 대표하고 편향되지 않은 배심원을 선정하는 것은 결코 쉬운 일이 아니다. 우선 배심원후보자에 대한 많은 정보가 있어야 한다. 더 나아가서, 대표성과 공정성이 서로 상충할 수도 있고, 소송 당사자들은 공정한 배심원보다 자신들에게 호의적으로 편향된 배심원을 선정하려고 할 수도 있다. 따라서 지역사회 사람들의 생각과 가치를 대표할 수 있고, 특정 가치와 이념 등에 편향적이지 않은 사람을 정확히 알 수 있는 정보와 자료에 기초하여 배심원을 선정하는 소위 “과학적 배심원 선정(scientific jury selection)”의 중요성이 오래 전부터 인식되었다(Kovera et al., 2003; Lieberman, & Sales, 2007).

빅데이터는 공정하고 민주적인 배심재판을 구현하기 위한 과학적 배심원선정을 위하여

큰 잠재력을 가진다(Kasabian, 2015; Ferguson, 2016). 기계학습 기술을 이용하여 빅데이터를 분석하면 지역사회 구성원들의 가치와 선호를 정확히 파악하고, 각 개인의 행동경향성, 태도, 가치, 선호, 이념을 정확히 파악할 수 있는 가능성이 커진다. 소셜미디어(e.g., Facebook, Twitter)에 나타나는 자료를 이용하여 개인의 개방성, 양심성, 외향성 등의 성격프로파일을 도출할 수도 있다(Golbeck et al., 2011). 인터넷 환경에서 이루어지는 개인들의 행동에 대한 방대한 비구조화된 자료는 배심원후보자에 대해 풍부한 정보를 제공해줄 수 있다(Hoffmeister, 2012). 예를 들면, 특정 배심원후보자가 35세의 여성이고, 직업이 간호사라는 피상적인 정보를 넘어서, 육아관련 잡지를 구독하지만 시사잡지는 읽지 않고, 최근에 은행 대출을 받은 사실이 있으며, 종교단체에 정기적으로 기부를 하고, 과거에 배심원으로 재판에 참여한 경험이 있으며, 그 재판에서 어떤 결정을 했는지 등에 대한 정보는 그 배심원후보자가 어떤 사람인지에 대해 상당히 정확한 예측을 가능케하고, 인종, 성별 등의 피상적 정보에 기초한 고정관념(stereotype)에 의해 배심원후보자를 기피하거나 선정하는 대신, 성격과 선호에 대한 구체적인 정보에 바탕하여 배심원을 선정할 수 있는 가능성을 높여준다(Laitinen & Loynes, 2012). Ferguson(2016)은 배심원후보자를 보다 구체적으로 명확히 조망할 수 있다는 의미에서 빅데이터를 “밝은 데이터(bright data)”라고 지칭하였다.

빅데이터와 기계학습 기술의 발전은 지역사회를 대표하는 배심원후보자를 구성하고, 불편부당한 판단을 하는 배심원을 효과적으로 선정하는 절차를 과학화할 것으로 예상되지만, 위험성 또한 예상된다. 가장 우려되는 위험은

사생활침해다. 예를 들어, 배심원 선정절차에서 검사 혹은 변호인은 배심원후보자에게 과거의 선거에서 어떤 정치인에게 투표를 했는지 등의 질문을 할 수 없다. 만약 변호인이 그런 질문을 하면 판사의 제지를 받게된다. 그러나 변호사는 소셜미디어에서 특정 배심원 후보자의 과거 투표 경력에 관한 정보를 아무런 제지 없이 구할 수 있다. 배심원후보자의 정치적 성향뿐만 아니라, 건강상태, 성적 지향/취향, 전과유무 등도 빅데이터의 탐색과 분석에 의해 재판 당사자에게 노출될 수 있다. 소셜미디어에 공개된 개인정보를 수집하는 것 자체가 불법적인 사생활침해가 되지 않는 경우에도, 그러한 정보에 기초해서 배심원 선정 여부를 결정하는 것에 대한 반감이 생길 수 있다. 개인적인 영역의 정보가 본인의 의사와 상관없이 수집되고 이용되는 경우가 자주 발생하면 국민들이 배심원 봉사를 기피하게 되고, 배심재판제도의 운영이 어려워진다. 법심리학은 제4차 산업혁명 시대에 빅데이터와 기계학습 기술의 도움으로 이루어지게 될 배심원 선정절차의 과학성과 사생활침해의 가능성을 검토하고, 그 가치들의 균형을 유지하는 방안을 모색하는 연구에 관심을 가져야 한다.

과학적 증거

신경생리학

캐나다, 영국, 미국, 네덜란드 등에서 형사 사건의 법적 의사결정을 위하여 뇌파(EEG: electroencephalography)와 기능적 자기공명영상(functional magnetic resonance imaging: fMRI) 등의 신경과학 증거와 모노아민산화효소 A(monoamine oxidase A: MAOA)와 같은 행동유전학 증거가 점차로 많이 사용되고 있다(Fuss,

2016; Jones et al., 2013). 신경생리학 증거는 대부분의 형사사건에서 범죄행동이 피고인의 통제능력을 벗어난 것이라는 주장(“비정상적인 뇌가 피고인으로 하여금 그 행동을 하게 만들었다”는 주장)를 뒷받침하여 형사책임을 면하거나 형량을 낮출 목적으로 사용된다(Denno, 2009). 미국에서는 모든 살인사건 재판 중 약 5%에서 피고인이 신경생리학 증거를 사용한다(Farahany, 2016). 그러나 검사들도 피고인의 범죄행동이 갱생 혹은 교정이 불가능하다는 주장을 설득하여 형량을 높일 목적으로 신경생리학 증거를 사용하기도 한다(Farahany, 2016). 미국에서는 신경생리학 증거를 제시할 것을 판사가 검사 및 변호인에게 강력히 요구하기도 하고, 그러한 요구를 충족하지 않는 당사자에게 불이익을 주기도 한다(Denno, 2009).

행동유전학 연구들에 의하면 폭력적인 범죄 행동 변이의 약 50%를 유전적인 요소가 설명하는 것으로 알려졌다(Glenn & Raine, 2014). 그러나 범죄행동을 유발하는 특정의 유전자가 발견된 바는 아직 없다(Vassos et al., 2014). 가장 많은 관심을 모은 유전자는 MAOA다. 이 유전자에 돌연변이가 생기면 통제되지 않는 공격성과 충동성을 보이는 소위 “브루너 증후군(Brunner syndrome)”을 유발하는 것으로 알려졌다(Brunner et al., 1993). MAOA 변이형은 청소년의 주의력결핍과잉행동장애에서 동반되는 사이코패시 성향과 관련성을 가진다(Fowler et al., 2009). 또한 MAOA 변이형을 가진 남성들은 대뇌변연계(limbic structure)의 용적이 유난히 작고, 뇌기능에서 비정상적인 소견을 보이며(Meyer-Lindenberg et al., 2006), 최근 연구에서는 극단적으로 폭력적인 범죄자들(최소 10건 이상의 살인, 살인미수, 폭행범

죄)에서 폭력행동과 MAOA 변이형의 관련성이 크게 나타나는 것이 발견되었다(Tiihonen et al., 2015).

기술과 과학의 발전에 의해 법적 의사결정 영역에 점차 더 많이 유입될 것으로 예상되는 신경생리학 증거들에 대해서 변함없이 유념해야 할 점이 두 가지 있다. 첫째는 신경생리학 연구들이 보고하는 범죄행동의 신경학적, 유전학적 관련성은 많은 사례들로 이루어진 집단자료의 분석에서 얻어지는 결과다(Baker et al., 2006). MAOA 변이형이 폭력행동과 관련성을 가진다고 하더라도 특정 개인에서 그 관련성이 구현되기 위해서는 환경적인 요소들이 상호작용해야 한다(Kaplan, 2006). MAOA의 변이형을 가진 사람들 중 아동기에 학대적인 환경에서 성장한 사람들에서만 성인이 되어 반사회적인 행동을 하게 될 확률이 현저히 높게 나타났다(Caspi et al., 2005). 따라서 어떤 범죄자가 범죄행동과 관련되는 것으로 알려진 신경생리학적 특징을 가졌더라도, 그의 범죄행동이 그 특징에 의해서 유발되거나 영향을 받았다는 판단을 하는 것은 잘못된 판단이 될 가능성이 높다(Fuss, 2016). 둘째로, 신경학적, 유전학적 특징과 범죄행동은 1:1의 관계를 가지지 않는다. 범죄행동은 여러가지의 신경학적, 유전적 요인들이 복합적으로 작용하는 결과다. 따라서 신경생리학의 발달에 의해 점차로 증가하는 과학적 증거에 대하여 범죄행동에 대한 환경의 영향과 상호작용, 그리고 다양한 신경생리학적 요인들의 복합적 영향에 대해 지속적인 연구가 이루어져야 한다. 신경생리학의 발전은 법적 의사결정의 정당성과 합리성을 높이는데 큰 기여를 하게 될 것으로 예상된다. 법심리학은 사법절차에서 그러한 과학적 성과가 어떻게 활용되고 있는지, 그리

고 어떻게 활용되어야 하는지에 대해 관심을 가져야 할 것으로 생각된다.

형사재판에서 피고인의 범죄행동을 설명하기 위한 목적으로 사용되는 신경생리학 증거들은 양날의 칼이다. 동일한 증거가 범죄행동에 대한 피고인의 책임을 감하기 위한 목적으로 사용될 수도 있고, 처벌을 가중하기 위한 목적으로 사용될 수도 있다(Aspinwall et al., 2012). 양날의 칼로 쓰일 수 있는 신경생리학 증거들이 판사 혹은 배심원의 판단에 실제로 어떤 영향을 주는지도 법심리학적인 실증연구의 대상이다(Fuss et al., 2015).

브레인 디코딩

특정 정보 혹은 기억이 뇌에 존재하는지 여부를 EEG를 이용하여 측정해서 거짓말, 피병 등을 탐지하는 방법이 사용되고 있다(Ben-Shakhar & Elaad, 2002; Ben-Shakhar et al., 2002; Tardif et al., 2002; Vagnini et al., 2008; Synnott et al., 2015; Lee et al., 2018). 뇌파를 측정하여 뇌에 존재하는 정보를 탐지하는 방법은 뇌-컴퓨터 인터페이스 기술의 한 종류다. 뇌파는 사건 혹은 그림 등의 시각적 자극이 주어지면 시간에 따라 일정한 패턴을 보이며 나타나는데, 주어진 자극에 대해 천분의 1초 단위(ms)로 발생하는 뇌의 반응 혹은 활동을 반영한다. 자극이 주어지고 300~800ms에 나타나는 정적 정점(positive peak)을 P300으로 지칭하는데, 자극이 경험한 적이 있는 것이면 크게 나타나고, 경험한 적이 없는 것이면 작게 나타난다(Paller & Kutas, 1992). 따라서 자극 제시 후 크게 나타나는 P300 뇌파형은 피검자가 자극에 대한 기억을 가지고 있다는 것을 암시한다(Rugg & Yonelinas, 1993).

fMRI 영상을 이용하여 거짓말하는 뇌를 “보

는” 기술도 개발 중이다. fMRI는 뇌가 활동하는 순간에 활성화되는 영역을 시각적으로 표현하는 기술이다. 질문에 대하여 진실된 대답을 하는 사람의 뇌는 주로 기억이 저장되고 인출되는 영역에 국한되어 활성화되는 양상을 보이는 반면, 거짓된 대답을 하는 사람의 뇌는 계산과 인지적 통제를 담당하는 영역까지 확장된 활성화 양상을 보인다(Ganis et al., 2003; Nunez et al., 2005; Mohamed et al., 2006). fMRI에 의한 거짓말 탐지는 그 정확성이 아직 충분히 높지 않아서 법정에서 증거능력을 인정받지 못하고 있다. 미국의 뉴욕타임즈는 fMRI에 의한 거짓말탐지 기술이 향상되어 정확성이 90%를 상회하면 법원이 증거능력을 인정하게 될 것으로 예측한 바있다(Rosen, 2007). 그런데 그 정확성은 실험실 상황에서 이미 90%가 넘는 것으로 알려졌다(Woodruff, 2014). 많은 법학자들은 이 기술의 신뢰도가 조만간 과학적 증거의 법적인 증거능력을 판단하는 “다우버트기준(Daubert Standard)”을 충족할 정도로 발전할 것으로 예상한다(Amirian, 2016).

뇌파의 측정에 의한 기억탐지와 fMRI에 의한 거짓말탐지를 넘어서, 인공지능과 뇌과학이 결합하여 뇌에 대한 이해를 증진하고, 뇌 질환을 극복하는 방안을 모색하는 것은 물론, 뇌의 활성을 조절하고, 생물학적 뇌를 인공적으로 배양하는 등의 연구가 세계 여러 나라에서 활발히 진행 중이다. 그 중 브레인 디코딩(brain decoding) 혹은 뇌 해독 기술은 기능적 자기 공명 영상을 이용하여 여러 가지 자극이 뇌에서 부호화(encoding)되는 패턴을 파악한 후, 자극과 부호화 패턴의 관계에 기초하여 생각, 기억, 느낌, 감정, 등의 뇌 활동을 읽어낸다. 예를 들면, 시각적 경험을 컴퓨터가 재구성하

고(Nishimoto et al., 2011), 사람이 보고 있는 네발 달린 동물이 개인지, 표범인지를 아는 것과 같이, 시각적 경험이 해석되는 구체적 의미도 읽어내고(Huth et al., 2016), 마음속으로 하는 혼잣말을 재생하기도 한다(Wang et al., 2017; Anumanchipalli et al., 2019). 고정된 이미지 뿐만 아니라, 꿈을 꾸는 동안 그 꿈을 영상으로 기록하는 것도 가능하다(Horikawa et al., 2013). 아직은 기초적이고 실험적인 수준에 머물러 있지만, 뇌 신경과 컴퓨터를 수천 가닥의 극세 전극으로 연결하는 인터페이스 기술(Musk & Neuralink, 2019)이 발전하면, 생각만으로 컴퓨터에 글을 쓰거나, 컴퓨터를 통해 음성(말)을 만들어내거나, 기계와 장치를 조작하는 등의 작업이 가능해진다(Carmena et al., 2003; Zander & Kothe, 2011).

사람이 누군가의 얼굴을 상상할 때 뇌를 fMRI로 스캔하여 컴퓨터가 그 상상된 얼굴을 컴퓨터 화면에 실시간으로 제시하는 브레인 디코딩 연구가 진행 중이다(Lee & Kuhl, 2016). 뇌-컴퓨터 인터페이스 기술이 더 발전하면, 목격증인의 범인지목 방법에 대해 새로운 가능성을 전개할 잠재성을 가진다. 목격증인이 자신이 본 범인의 얼굴을 회상하면 목격증인의 뇌와 연결된 컴퓨터가 그것을 화상으로 재구성하고, 나아가서 방대한 범죄자 데이터베이스에서 그 얼굴의 신원을 곧바로 파악할 수도 있다. 일반적으로 얼굴에 대해 기억된 이미지는 해상도와 완성도가 낮은 이미지다. 모호한 형태로 회상된 이미지를 인공지능 기술에 기반한 얼굴 맵핑(facial mapping) 방법을 사용해서 데이터베이스에 이미 알려진 인물의 얼굴과 동일한 얼굴인지 판단한다(e.g., Edmond, Biber, Kemp, & Potter, 2009). 브레인 디코딩에 의해 재구성된 목격증인의 얼굴기억이 신빙성

이 있는 것인지, 그러한 기술에 의해 채증된 증거가 배심원과 판사에게 어떤 영향을 주는지(e.g., McCabe, Castel, & Rhodes, 2011; Smethurst, 2015), 그러한 기술의 신뢰도가 각기 다른 문화와 인종에서 균일하게 유지되는지 등에 대한 법심리학적 연구가 필요할 것으로 예상된다. 더불어서, 목격자의 범인얼굴에 대한 기억뿐만 아니라, 과거에 있었던 특정 사건에 대한 자전적 기억(Barnier & McConkey, 1999), 대화와 상호작용에 대한 기억, 특정 상황에서 가졌던 느낌에 대한 기억, 행동에 대한 기억, 속도에 대한 기억(Brooks et al., 2011), 등에 대해서도 브레인 디코딩 기술이 적용될 것으로 예상되고, 그러한 기억들의 신빙성을 좌우하는 요인들이 무엇인지에 대한 연구가 절실히 요구될 것으로 예상된다.

과학적 증거의 증거능력

모든 과학적 증거는 그것이 의미하는 바를 증명하는 것이 아니고, 다만 확률적으로 암시할 뿐이다(Gazzaniga, 2008). 따라서 정확성이 거의 완벽하다고 알려진 과학적 증거도 그것의 증명력이 사실판단자의 편견을 유발하거나, 혼동시키거나, 오도할 위험보다 클 때 비로서 증거능력이 인정된다(미국연방 증거규칙 제403조). 과학적 증거가 사실판단자의 편견을 유발한다는 것은 그 증거가 창출되는 과정이 복잡하거나 이해하기 어려워서 사실판단자가 합리적인 이성으로 증거의 내용을 이해하지 못하고 다만 그것의 전문성에 압도되어 증거의 내용과 타당도를 맹목적으로 인정하는 경우를 포함한다. 예를 들어, 뇌의 절단면이나 삼차원 형상을 재현해서 보여주는 fMRI 증거는 배심원의 편견을 유발할 위험이 크다는 이유로 증거능력이 배제되는 경우가 많다(e.g.,

Sinnott-Armstrong, Roskies, Brown, & Murphy, 2008). fMRI 영상이 보여주는 뇌의 모습은 뇌를 40,000개 정도의 영역(voxel)으로 구분한 후 각 영역의 활성화에 의해 영역내 혈류에서 포착되는 변화를 일련의 통계적 처리(약 40,000번의 독립적인 가설검증)를 거쳐서, 그 결과를 해부학적 뇌 그림판 위에 표시하는 것이고, 그 과정에서 많은 통계적 가정과 측정오차가 개입할 수 있다(e.g., Bennett, Miller, & Wolford, 2009). 그런데 fMRI 영상을 보는 사실판단자는 그 영상이 실제 뇌의 물리적인 모습을 있는 그대로 사진으로 촬영한 것과 같은 것으로 오해할 수 있다.

과학적 증거가 편견과 터널비전을 강화하고, 인지부조화에 의한 왜곡된 판단을 유발하는 또 다른 예는 DNA 증거에서 발견된다. 최근의 분석기법의 발달로 인하여 극히 미세한 생체흔적에서도 DNA검출이 가능해졌다. 심지어는 지문에서도 DNA(디옥시리보핵산)가 추출될 수 있고, 잠깐 찬 시계, 지폐 등 사람의 손이 거쳐간 모든 곳에서 DNA를 검출할 수 있으며, 물속에 10일간 잠겨있던 칼에서도 미토콘드리아 DNA(mtDNA)를 검출할 수 있다. 네덜란드 쉬담(Schiedam)시의 한 공원 숲속에서 2000년에 성폭행 후 살해된 10세 여아의 시신이 발견되었다. 얼마 후 목격자들의 진술에 기초하여 용의자가 체포되었는데, 피해자의 시신에서 발견된 타인의 DNA는 용의자의 DNA와 일치하지 않았고, 용의자의 DNA는 현장 어디에서도 발견되지 않았다. 그럼에도 불구하고 용의자가 범인으로 기소되었고, 2001년에 유죄판결을 받았으며, 2003년에는 네덜란드 대법원에서 유죄가 확정되었다. 그러나 다음 해에 자백을 하는 진범이 나타나서 용의자는 재심을 거쳐 2005년에 다시 무죄가 확정되었

으며, 네덜란드 정부는 누명을 쓰고 유죄를 확정받았던 그에게 600,000유로를 배상하였다. 그 무고한 용의자가 유죄판결을 받는데 기여한 한 가지 요인은 놀랍게도 피해자의 시신과 그 주변에서 용의자의 DNA가 검출되지 않았다는 사실이다! 그를 기소한 검사는 “피해자와 접촉한 사람은 누구든지, 어떻게 접촉했던, 반드시 DNA흔적을 남겨야 한다. 그런데 용의자의 DNA가 발견되지 않은 것은 그가 범행 후 용의주도하게 자신의 흔적을 모두 제거했기 때문이다.”는 주장을 하였고, 그 주장이 대법원에서까지 부정되지 않았다(Brants, 2013). 고도로 발달한 과학의 덕분으로 과학적 증거의 정밀성과 정확성이 크게 향상되는 반면, 그것이 판단자의 예단이나 편견, 그것을 확증하려는 동기와 연계되면 매우 강한 터널비전을 형성하여 잘못된 판단이 유도될 수 있다.

다른 모든 과학적 증거와 마찬가지로, 신경생리학 증거나 브레인 디코딩 증거가 재판에서 사용되는 경우, 그 증거의 증거능력과 증명력에 대한 검토가 선행되어야 한다. 법적인 의사결정과 사법판단 과정에 점차로 많이 유입될 것으로 예상되는 과학적 증거들에 대하여 법심리학은 새로운 과학적 증거들이 사실 판단자에게 미치는 영향 혹은 충격을 평가하고, 그 증거들의 증명력과 편견유발 위험을 비교할 수 있는 실증적 자료를 제공해야 할 필요가 있다.

수사

부인을 살해한 혐의를 받는 남편의 뇌에서 부인을 살해한 기억이 브레인 디코딩 기술에 의해 탐지될 수 있다면, 그 탐지된 살해기억은 남편의 살인혐의를 증명하는 증거로 사용

될 수 있을까? 대한민국 헌법 12조 2항은 자기부죄거부의 특권을 명시하였다. “모든 국민은 고문을 받지 아니하며, 형사상 자기에게 불리한 진술을 강요당하지 아니한다.” 미국의 수정헌법 제5조도 범죄혐의를 의심받는 사람이 자신에게 불리한 진술을 강요당하는 것을 금지한다. 체포되는 사람에게 고지되는 미란다 경고(Miranda warning)에 포함되는 진술거부권도 자기부죄거부특권의 한 종류다. 헌법은 피의자의 입에서 나오는 불리한 말 혹은 진술의 강요를 금지하지만, 입에서 나오는 침, 상피조직, 호흡기와 같은 물리적 증거의 채취를 강요하는 것을 금지하는 것은 아니다(Schmerber, 384 U.S. at 764; 헌법재판소 1997.3.27. 선고 96헌가11 결정). 즉, 자기부죄거부특권이 적용되는 증거의 범위에 대해 미국이나 한국의 법은 의사소통(communication)를 통해 변화할 수 있는 진술증거에 국한되고, 의사소통에 상관없이 고정된 신체적 상태를 나타내는 물리적 증거, 예를 들면, 지문, 혈액 등에는 자기부죄거부특권이 적용되지 않는다.

사람들은 과거의 경험, 대화, 약속, 행동들을 여러 가지 다양한 방식으로 “기억”한다. 회의 날짜를 달력에 기록하거나, 금전거래를 장부에 기록하거나, 최근 만난 사람과의 대화내용을 일기장에 기록하거나, 방문한 장소에 대한 정보를 인터넷 파일에 기록한다. 이메일, 통화녹음, 핸드폰문자, 은행입출금기록 등 기억을 위해 만들어진 기록물은 물리적 증거이고, 자기부죄거부특권이 적용되지 않는다. 그런데, 뇌에 저장된 기억도 ‘진술’의 형태를 가지지 않고 외부로 인출되면 수첩에 적힌 기록과 마찬가지로 자기부죄거부특권이 적용되지 않는 물리적 증거인가? 피의자의 기억이나 생각, 감정과 정서를 피의자의 입을 통하지

않고 피의자의 뇌로부터 직접 탐지한다면 (Rissman et al., 2010; Chadwick et al., 2010; Harrison & Tong, 2009), 탐지된 기억, 생각, 감정, 정서는 자기부죄거부특권이 적용되는 증거인가? 부인을 살해한 혐의를 받는 남편의 뇌에서 부인을 살해한 기억이 브레인 디코딩 기술에 의해 탐지될 수 있다면, 남편은 자기부죄거부특권을 이유로 그 기억의 탐지를 거부할 수 있을까? 그러한 마음증거는 피의자가 의식하지도 못하는 방식으로 채증될 수도 있는데(Tamietto & de Gelder, 2010), 그런 경우 남편은 자기부죄거부특권의 위배를 이유로 그 기억증거를 탄핵할 수 있는가?

새로운 기술의 발전에 의해 발생할 수 있는 새로운 법적 딜레마에 대해 Farahany(2012)는 자기부죄거부특권이 적용되는 ‘진술증거 vs. 물적증거’의 이분적인 구분을 수정할 것을 제안하였다. 수사기관이 피의자에게 증거를 구하기 위한 도발(provocation)을 했는지 여부와, 용의자 혹은 피의자가 증거의 현출과 내용을 주체적으로 통제할 수 있는지의 여부에 따라 자기부죄거부특권의 적용여부를 정해야 한다는 것이다. 용의자/피의자에 대한 수사기관의 도발(심문, 질문 등)이 없이 얻어진 기억 증거는 자기부죄거부특권의 적용 대상이 아니므로 용의자 혹은 피의자에게 불리한 증거로 사용 가능하다. 또한 DNA 혹은 지문 등의 증거와 마찬가지로, 용의자 혹은 피의자가 의식적으로 기억 증거의 내용을 변형, 조작하거나, 기억 증거의 인출과정을 통제하지 않는 경우는 자기부죄거부특권의 적용 대상이 아니다.

기술혁명 시대에 기억 혹은 기타 마음증거가 신경과학과 브레인 디코딩 기술의 발전으로 사법영역에 쏟아져 나오는 경우, 그 증거들의 현출 과정에 대한 이해가 매우 중요해질

것으로 예상된다. 예를 들어, 피의자에 대한 수사기관의 도발이란 어떤 상황에서 어떤 질문 혹은 심문을 의미하는지를 규명하는 심리학적인 연구가 필요하게 될 것으로 예상된다. 폴리그래프 검사와 EEG 혹은 fMRI를 이용해서 거짓말을 탐지하는 ‘죄지식검사(guilty knowledge test)’는 피검자가 증거의 현출을 의식적으로 조작하는 것이 가능한 것으로 알려졌다(Ben-Shakhar et al., 2002; Ben-Shakhar & Elaad, 2003). 그와 같이, 기억 혹은 기타 마음증거를 뇌-컴퓨터 인터페이스 기술로 취득할 때 해당 기술로 얻어진 증거의 내용과 현출 과정에 대해 의식적 통제가 어떤 상황에서, 어떤 정도로 가능한가를 파악하는 범심리학 연구들이 함께 이루어져야 한다. 또한 용의자 혹은 피의자가 의식적으로 기억 증거의 내용을 변형, 조작하거나, 기억 증거의 인출과정을 통제하지 않는 경우에도, 현출된 기억이 실제 경험과 일치하는 것인지, 변형되거나, 만들어진 기억일 가능성이 있는지에 대한 연구가 필수적인 것으로 생각된다.

사법의사결정

변호사와 법무사

대한민국 대법원은 2021년 시행을 목표로 빅데이터 분석을 통한 ‘인공지능(AI) 소송도우미’ 개발을 추진한다(법률신문, 2019. 9. 23). 민사소송에서 사건 성격에 맞는 소송제기 방법을 추천해 주고 그에 필요한 제출 서류와 소송 예상 종료시점 등을 안내하는 등, 변호사를 선임하지 않고도 소송할 수 있도록 돕는 것이 목표다. 소송 당사자가 관례나 법규 등을 물어보면 인공지능이 빅데이터를 분석해 답변을 내놓는 방식이다. 단순 안내에 그치지

않고 서류 작성을 지원하는 시스템도 갖출 계획이다. 즉, 일반적으로 변호사가 하는 소송업무의 상당 부분을 인공지능 로봇이 하는 것이다. 미국에서는 이미 인공지능 알고리즘으로 작동하는 로봇 변호사가 인간 변호사의 일을 대신 하고 있다. 세계 최초의 인공지능 로봇 변호사 ROSS가 2016년에 로펌(Baker Hostetler)에 파산전문 변호사로 채용되었는데, 만 1년만에 10개의 다른 로펌들에도 취업하였다(Nunez, 2017). ROSS는 자연어처리에 기반한 기계학습 알고리즘을 이용하여 엄청난 분량의 법률문서를 순식간에 검색하여 당면한 사건에 적용되는 법률, 그리고 가장 유사한 과거사건의 판례를 찾아 법리해석을 하고, 법률적인 조언을 제안하는 등의 일을 수행한다.

대한민국 대법원이 개발하려는 ‘AI 소송도우미’나 미국의 ROSS와 같은 인공지능 시스템은 경험을 통한 학습에 의해 알고리즘이 완성되어가고 변화해가는 시스템이다. ROSS와 같은 인공지능 변호사는 스스로 기능을 향상시키고 정교화 할 것이고, 그 역할이 빠르게 확장될 것으로 예상된다. 문서검색, 빅데이터 분석, 논리적 추론을 비롯한 몇가지 영역에서 인공지능은 인간 변호사보다 효율적이고 정확하게 필요한 업무를 수행할 수 있고, 인간 변호사가 업무과중과 과로로 인해 경험할 수 있는 피로, 스트레스, 방임, 태만, 그리고 법/윤리적 이탈을 부추기는 유희과 탐욕으로부터 자유로울 수 있다. 미국의 이노센스 프로젝트(innocence project)는 오판에 의해 억울한 누명을 쓰고 교도소에 수감되어 있다가 뒤늦게 DNA 증거로 결백이 증명되어 석방되는 255개 사례에서 애초에 오판이 이루어진 원인을 분석한 결과, 54개 사례에서 피고인이 변호인의 무능 혹은 불성실 변호를 호소하였고, 그중 6

개 사례는 비효과적 변호를 이유로 항소심에서 무죄 판결이 이루어졌으며, 1개의 사례에서는 변호인을 바꾸어 재심리할 것을 결정하였고, 3개의 사례에서는 판결이 뒤집히거나 재심이 이루어지지 않는 않았으나 항소심이 하급심에서 변호인의 변호가 비효과적이었던 사실을 판결문에 명시하였다(West, 2010). 많은 오판 사례들에서 비효과적인 변호인의 조력이 중요한 오판요인 중 하나로 지적된다(Gould & Leo, 2010). 인공지능 기술이 발전하여 형사 분야에서 국선변호인(court-appointed defender)이나 공중변호인(public defender)의 업무를 보조하고, 더 나아가서 일부 업무에서 변호인을 대체할 수 있으면, 무능과 불성실에 의한 비효과적 변호 때문에 생기는 오판은 많이 줄어들 수 있을 것으로 예상된다.

반면에, 인공지능 변호사가 변호사로서의 직업윤리적인 판단을 할 수 있는지에 대한 의문이 존재한다(Nunez, 2017). 변호사는 의뢰인을 대신하여 법적인 행위를 하는 대리자(representative), 의뢰인에게 권리와 의무의 실질적, 법적 의미를 알려주는 충고자(advisor), 의뢰인의 권익을 성취하기 위해 투쟁하는 지지자(advocate), 의뢰인에게 이익이 되는 결과를 얻기 위해 중재와 타협을 하는 협상가(negotiator), 의뢰인의 법적인 입장과 상황을 정확히 평가하고 전달하는 평가자(evaluator) 등의 복합적인 역할을 해야 하고, 각 역할은 다양한 직업윤리적 원칙과 기준에 따라 수행되어야 한다. 인공지능 변호사가 역할에 따라서 상충할 수도 있는 직업윤리적 원칙들을 어떻게 적용하고 운영할지에 대한 논의와 연구가 필요하다. 인공지능 변호사에 대한 기대와 우려는 궁극적으로 그 알고리즘이 제공하는 법 서비스에 대한 의뢰인의 만족도에 귀착된다.

“AI 소송 도우미”와 ROSS와 같은 알고리즘에 대한 사용자 및 소비자의 만족도에 대한 심리학적 연구들이 나타날 것으로 예측된다.

판사와 배심원

인공지능은 변호사뿐만 아니라 사실인정과 사법적 의사결정을 하는 주체인 판사와 배심원을 대체할 가능성도 예측된다. 인간 판단자는 감정에 휩쓸릴 수도 있고, 설득 혹은 회유될 수도 있으며, 쉽게 지쳐서 무성의한 판단을 할 수도 있고, 휴리스틱에 의한 편향된 판단을 할 수도 있는 반면에, 일견 인공지능 기계는 인간에 비해서 더 불편부당하고 공정하며 논리적인 판단을 신속하게 할 수 있을 것으로 생각되기도 한다. 법관으로 16년간 사법 판단 경험을 한 정육도판사는 인간 판사는 이야기 혹은 “서사”에 관한 본능적 편향을 가지기 때문에 이해가 어려운 부조리한 현실을 직시하거나 인정하지 못하고 그럴듯하게 보이는 왜곡된 이야기를 구성하여 판단을 하는 반면, 생물로서의 본능이나 편향이 전혀 없는 AI 판사는 그럴듯한 이야기를 구성해야 한다는 압박없이 일체의 가용정보를 종합, 분석하여 판단하기 때문에 더 실제적 진실에 부합하는 판단을 할 것으로 추론하였다(Chung, 2020). 그러한 이유로 현재 사법판단에 관여하는 상당수의 사람들이 향후 20년 이내에 인공지능 기계로 대체될 것으로 예견되기도 하였다(Dolphin, 2015).

Aletras et al.(2016)은 자연어처리에 기반한 기계학습 알고리즘으로 인공지능 재판 프로그램을 개발하여 유럽인권법원(European Court of Human Rights)에서 이루어진 584건의 기존 재판결과(인권침해에 대한 가부판단)를 79%의 정확도로 예측하였다. 정형화되어 분쟁의 가

능성이 적은 사건, 소액사건, 이혼소송 등의 재판에서 법률, 증거, 확인된 사실, 판례, 법원칙, 그리고 정책적 고려에 기초해서 추론과 논증을 거쳐 이루어지는 판결을 판사 대신 인공지능 알고리즘, 소위 “AI 판사”가 하는 시스템이 이미 에스토니아, 호주, 중국 등의 여러 나라의 사법 실무에서 가동되고 있다. 미국의 연방대법원장 존 로버트 주니어(John G. Roberts Jr)는 2017년에 “인공지능이 사실인정과 사법적의사결정을 하게 될 때가 도래할 것으로 예상하는가?”라는 기자의 질문에 대해 “이미 그런 시대가 도래하였고, 사법부가 하는 일에 대해 심각한 긴장을 유발하고 있다”고 답변하였다(Liptak, 2017). 한국에서는 AI 판사 알고리즘이 개발되어도 그것이 실제 재판에서 인간 판사를 대체하여 판결을 하는 일이 생기는 어려울 것으로 생각된다. 대한민국 헌법 제27조 제1항은 법관에 의하여 재판받을 권리를 규정하고 있기 때문이다. 그러나 한국에서도 인간 판사의 판결 혹은 법원의 판단을 예측하는 인공지능 알고리즘은 개발될 가능성이 높다.

인공지능에 대한 상식적인 이해에 기초한 낙관적인 견해와는 달리, 인간에 의한 판단과 평가를 인공지능이 대체하면서 인간의 차별, 편파성, 불공정이 사법판단에 더 구조적으로 고착될 가능성이 크다. 위에서 소개한 재범가능성 알고리즘인 COMPAS도 미국의 사법통계 자료에 기초해서 학습한 인공지능인 까닭에 흑인에 대한 차별적 요소가 구조화되었다. 또 다른 예로, 미국 아마존이 2014년부터 인공지능을 이용한 신입직원 채용 시스템을 이용했는데 이때 개발된 인공지능 알고리즘은 과거 10년간의 지원자들 이력서 데이터를 기계학습하여 여성을 차별하는 패턴을 분석하였고 이

력서에 ‘여성’이란 단어가 있으면 감점을 부여 하였다. 아마존은 이 인공지능 채용 시스템을 폐기하였다. 사람에 의한 판단과 평가를 인공지능이 대신하면서 생길 수 있는 차별, 편파성, 불공정의 위험에 대해 미국대통령부의 과학기술정책담당관실(The United States White House’s Office of Science and Technology Policy)은 인공지능 기술이 정책결정의 자동화를 위해 사용되면 알고리즘이 기반하는 데이터에 존재하는 차별적 요소들 또한 학습하여 자동화된 결정에 사용될 가능성이 있으므로(“potential of encoding discrimination in automated decisions”), 사생활보호(privacy), 시민권(civil rights), 개인의 자유(individual freedom) 등의 기본권을 침해하는 위험요인이 될 수 있음을 경고한 바 있다 (Executive Office of the US President, 2014, p. 45).

기계학습의 원리로 구축되는 인공지능 알고리즘은 특정한 문화, 인종, 지역의 데이터에 기반한 패턴인식 메커니즘에 의해 판단하고 평가하는데, 그러한 인공지능 알고리즘은 학습의 기반이 된 데이터와 다른 문화, 인종, 지역의 사례에 대하여 편파적인 평가와 불공정한 판단을 할 수 있다(Surden, 2014). 따라서 만약 한국에 AI 판사 알고리즘이 도입된다면, 그 알고리즘은 한국에서 자체적으로 개발된 것이라야 한다. 필자가 아는 한, 아직은 한국에서 판사의 판결을 예측하는 인공지능 알고리즘이 개발되지는 않았지만, 머지 않아 자연어처리에 기반한 기계학습 알고리즘에 의해 과거 사건들에 대한 판결문을 학습하여 사법 판단을 예측하는 인공지능 시스템이 나타날 것으로 예상된다. 앞으로 나타날 수 있는 AI 판사에 대해서는 그 학습데이터와 알고리즘 자체의 정합성에 관한 면밀한 검토가 이루어

져야 한다. 앞서 언급한 김상준(2013)의 오판 요인에 관한 연구에서 나타난 요인들을 포함하는 판결문을 학습하여 판사의 판단을 예측하는 인공지능 시스템의 판단 알고리즘에는 오판을 초래하는 요인들이 그대로 구조화되어 고착될 수 있다.

범죄

인공지능은 “인간처럼 생각하는 기계”로 정의되는데, 이 정의를 튜링테스트(Turing test)로 부른다(Turing, 1950). 인공지능의 또 다른 정의는 “인간이 만들어내는 결과와 동일한 결과를 창출하는 기계”라는 것이다(Floridi et al., 2008). 아직 널리 인식되지는 않았지만, 인간과 상호작용하도록 프로그램된 인공지능은 한가지 중요한 잠재적 특징을 가진다. 그 특징은 상호작용하는 인간의 신뢰와 믿음을 얻어서 그 사람을 심리적으로 조정하는 능력을 가질 수 있다는 것이다.

인간은 사물에 인격을 부여하고 의인화(anthropomorphism)하여 세상과 환경을 이해하고 관계를 형성하려는 경향을 가진다(Epley et al., 2007). 어린 아이들은 인형과 대화하고, 성인들은 하늘의 구름, 바위, 동식물에 인격을 투사하여 상호작용한다. 인공지능을 인간처럼 생각하는 기계로 정의하는 튜링테스트도 기계를 의인화하는 한 예다(Watson, 2019; Proudfoot, 2011). 심지어는 신(God)과 같은 초자연적인 존재도 그 모습을 인간의 형상으로 조각하고 인간들이 서로 상호작용하는 방식으로 초자연적 존재와 상호작용하고자 한다 (Barrett & Keil, 1996).

의인화는 환경과 효과적으로 상호작용하고, 애착관계를 형성하려는 동기에 의해 추동된다

(Epley et al., 2008). Nass, Moon, & Green(1997)은 컴퓨터에 대한 사람들의 반응을 알아보는 한 실험연구에서 성 중립적인 컴퓨터가 사람의 목소리를 사용하면, 그 컴퓨터에 대해 사람들이 성고정관념에 의한 반응(sex-based stereotypic response)을 보인다는 것을 발견하였다. 사람들은 여성의 목소리를 사용하는 컴퓨터를 남성의 목소리를 사용하는 것보다 더 친근하게 느끼는 반면, 남성 목소리를 사용하는 컴퓨터가 해주는 칭찬에 더 많이 설득되고, 칭찬의 내용을 믿고 더 많이 공감한다는 것을 발견하였다. 자연어 처리에 기반한 인공지능을 갖춘 채터봇(chatterbot, 대화로봇)은 인간의 이러한 의인화 경향성을 이용하여 사용자와 사회적 유대를 형성하고, 사용자가 내면의 생각이나 사회적으로 민감한 정보, 혹은 개인적인 정보를 쉽게 노출하게 만드는 능력을 겸비할 수 있다(De Angeli, 2009; De Angeli, & Brahnam, 2008). 더 나아가서, 대화하는 인공지능은 사람의 경계심을 낮추고 신뢰를 획득하여 일반적으로 행하지 않을 성적인 행동, 위험한 행동 등을 하도록 조정하는 능력을 가질 수도 있다(Rehm, 2008; Veletsianos et al., 2008).

인공지능은 범죄를 예방하고 해결하는 유용한 도구가 될 수도 있고, 새로운 범죄를 만들어내고 기존의 범죄를 더 정교화하는 수단일 수도 있다. 인간이 초래하는 것과 동일한 결과를 만들어내는 기계는 당연히 범죄도 만들어낼 수 있다. 더군다나, 인공지능은 목적을 달성하기 위한 최적의 수단과 방법을 모색하고, 그 과정에서 도덕적, 윤리적 판단을 도외시할 수 있어서 범죄에 활용되면 개인과 사회의 안전에 잠재적으로 큰 위협이 될 수 있다. Martínez-Miranda, McBurney, & Howard(2016)는 가상의 주식 및 금융시장을 시뮬레이션하여

이익을 극대화하려는 가상 투자자들의 행동을 관찰하였다. 마치 인공지능 알고리즘 알파고(Alpha-Go)가 바둑게임에서 최적의 한 수를 찾듯, 가상의 투자자들은 기계학습기술을 사용하여 이익을 극대화하는 최적의 투자 전략을 찾아 투자하였다. 시뮬레이션 연구에서 이익을 극대화하는 투자전략을 기계학습 기술을 사용하여 수행하는 가상의 투자자들은 허위주문을 포함하는 시장조작, 불공정거래 방법을 쉽고 빠르게 학습한다는 것이 확인되었다. 사회심리학자인 Scymore & Tully(2016)는 기계학습에 기반한 인공지능을 피싱(Phising) 범죄에 활용하면 사람들이 얼마나 쉽게 범죄피해를 당하는지를 실험으로 보여주었다. 그들은 소셜미디어 사용자의 과거 행동과 공개된 정보 및 인터넷 사용 내용을 학습한 인공지능이 그 사용자에게 독특한 정보와 과거행동 및 개인적 경험을 포함하는 메시지(예를 들면, 거래 은행에서 발송한 것으로 위장한 메시지)를 만들고, 특정 링크를 방문하도록 유도하였다. 인공지능이 기계학습을 통하여 제작한 피싱 메시지를 받은 사용자들의 30%~66%가 유도된 링크를 방문하여 자신의 중요한 정보(예, 은행 계좌번호)를 노출하는 서류를 작성하였다. 일반적인 전화 혹은 인터넷을 통한 피싱범죄가 성공하는 비율이 5%~14%인 것(Jakobsson & Ratkiewicz, 2006)에 비교하면, 인공지능을 활용한 피싱은 그 심각성이 훨씬 클 것으로 예상된다.

King, Aggarwal, Taddeo, et al.(2020)은 인공지능의 발달로 인하여 새롭게 나타날 것으로 예상되는 범죄를 “인공지능범죄(artificial intelligence crime: AIC)”로 지칭하고, 그 예방 및 해결책을 모색하였다. AIC는 크게 두 종류로 구분될 것으로 예상되었다. 한 종류는 앞

서 소개한 Seymore & Tully(2016)가 상정하는 종류의 범죄 즉, 피싱범죄이고, 또 다른 한 종류는 상업과 거래 과정에서 발생하는 사기범죄다. 인간의 행동을 조정하는 능력이 발휘되는 AIC 범죄의 실제 예가 목소리 스푸핑(voice spoofing)이다. 특정인의 목소리가 가지는 억양, 톤, 리듬, 발음, 말버릇 등의 미세한 특징을 인공지능이 포착하여 다른 사람의 전화목소리를 디지털 기기로 변환시켜 사기를 치는 스푸핑범죄가 이루어지고 있다(Stupp, 2019). 2019년에 영국의 한 에너지 회사에서 어떤 헝가리 은행계좌로 거액의 회사돈을 송금하라는 사장의 전화 지시를 받은 직원이 220,000 유로를 송금하였다. 그런데 사장의 전화는 인공지능이 다른 사람의 목소리를 약간의 독일 액센트가 섞인 실제 사장의 목소리로 조작해 만든 사기전화였다. 인공지능이 사람의 경계심을 낮추고 신뢰를 획득하여 성적인 행동 혹은 위법한 행동을 하도록 유도하는 범죄는 아직 발견된 것이 없지만, 멀지 않은 미래에 나타날 것으로 믿어진다. 특히 인격형성이 완성되지 않은 청소년들이나 열악한 사회경제적 상황에 있는 사람들이 그러한 AIC의 피해자가 될 가능성이 높다.

피싱범죄와 스푸핑범죄 이외에도 인공지능이 탑재된 악성코드(malware)에 의한 해킹 범죄, 무인수송수단(드론, 소형잠수정 등)을 이용한 마약 혹은 장물 운반, 소셜미디어에서 자동으로 댓글, 메시지 혹은 조작된 영상을 발송하는 소셜봇(social bot)을 이용한 괴롭힘과 스토킹 등의 범죄가 새롭게 나타날 수 있다. 악성코드에 의한 해킹의 대표적인 예는 이란의 나탄즈(Natanz) 핵시설을 무력화시킬 목적으로 미국과 이스라엘이 “올림픽게임작전(Operation Olympic Games)”으로 명명한 작전

에 사용된 사이버무기 스투스넷(Stuxnet)이다(Nakashima, 2012). 스투스넷과 같은 사이버무기에 인공지능 기능이 첨가되면 정확성과 스텔스 기능이 거의 완벽한 악성코드가 만들어질 수 있다. 미국의 IBM은 2018년에 그러한 인공지능을 갖춘 악성코드 덤록커(DeepLocker)를 공개하였다(Kirat et al., 2018). IBM이 미래에 예상되는 악성코드를 미리 파악하고 분석해서 생길 수 있는 피해를 미리 예방할 목적으로 개발한 덤록커는 공격대상을 정확히 포착할 때까지 악성의도를 완벽히 숨길 수 있는 인공지능 스텔스 기능을 갖춘 악성코드다. 사이버공간에는 물리적인 국경이 없기 때문에 인터넷을 통한 범죄는 광범위한 피해를 유발할 수 있다. 유럽의 국가들은 2001년에 “사이버범죄에 관한 유럽이사회(The 2001 Council of Europe Convention on Cybercrime)”를 개최하여 컴퓨터 기술과 사이버 공간을 통한 범죄의 정의, 범죄가 포착되었을 때 국가간 수사 및 사법처리 공조방안 등, 앞으로 예상되는 사이버범죄에 공동으로 대응하기 위한 대책을 논의하였다.

인공지능이 탑재된 드론이나 무인잠수정 등의 운반로봇은 다양한 상황에서 스스로 장애물을 피하고 최적의 코스를 선택하여 수송화물을 목적지까지 운반한다. 인공지능기능을 갖춘 운반로봇이 마약운반 등의 범죄에 이미 활용되고 있다(Europol, 2017; Sharkey et al., 2010). 범죄에 활용되는 운반로봇은 조정자 없이 스스로 운행하기 때문에 적발되더라도 그 로봇을 이용하는 범죄자가 누구인지 알기 어렵다(Gogarty & Hagger 2008,). 또한 로봇은 신원을 특정할 수 있는 지문이나 DNA 등의 흔적을 남기지 않는다. 더 나아가서, 로봇은 범죄행위 도중에 적발되면 스스로를 파괴하여

증거를 남기지 않도록 제작될 수 있다(Sharkey et al., 2010). 따라서 인공지능에 의해 기능하는 로봇을 이용한 범죄에 대응하기 위한 새로운 포렌식 과학과 프로파일링 기술이 개발되어야 한다.

기술의 혁신적인 발전에 의해 새롭게 나타나는 신종범죄에 대응하기 위하여 범심리학은 범죄가 이루어지는 영역, 상황 및 환경과, 범죄가 행해지는 과정을 이해하고 분석할 수 있는 새로운 패러다임을 개발해야 한다. 법의 일반적인 규범은 범죄의도(mens rea)를 가지고 범죄행위를 한 사람에게 범죄행위에 대한 책임이 귀속된다는 것이다. 그런데, 인간처럼 생각하는 기계 혹은 인간이 만드는 것과 동일한 결과를 만들어내는 기계로 정의되는 인공지능이 개입된 범죄의 경우에는 그 책임의 소재가 불명확할 수 있다(Čerka et al., 2015). 인공지능의 중요한 한가지 특징은 자율성(autonomy)을 가진다는 것이다. 자율성은 다양한 상황에서 인간의 개입 없이 스스로 예측, 판단, 선택, 결정하는 능력을 의미한다. 경험과 자료로부터 얻어진 학습된 정보에 기초하여 상향식 접근방식(bottom-up approach)으로 기능하는 인공지능은 그 기능과 발달이 인간의 관여 없이 이루어지기 때문에 기술을 처음 설계하고 제작한 사람이 의도하지 않은, 예측을 벗어난 영역으로 확장될 수 있다(Alaieri & Vellino, 2016). 인공지능이 목적을 달성하기 위한 최적의 수단과 방법을 자율적으로 모색하기 때문에, 기존에 알려지거나 제작자가 예측한 것과는 완전히 다른 새로운 범행수법(modus operandi)들이 창출될 수 있는 것이다. 따라서 일반적으로는 범죄에 사용된 인공지능의 사용자 혹은 제작자가 그 범죄에 대한 책임을 지게 되지만(Halley, 2010), 기계학습에 의하여

능력을 스스로 확장하는 인공지능은 애초의 제작자가 미리 예상하지 못했던 기능과 역량을 획득하여 자율적으로 판단, 결정, 기능하는 존재로 발전할 수 있고, 다른 법 인격과 독자적으로 상호작용하는 주체가 될 수도 있기 때문에, 인공지능의 제작자가 범의를 가지지 않은 경우에 자율적인 인공지능의 행위에 대해 어디까지 사용자 혹은 제작자가 책임을 질 수 있는가에 대해 사회적 판단 혹은 합의가 필요하게 된다(Scherer, 2016; Čerka et al., 2015). 범심리학은 그러한 사회적 판단과 합의가 합리적으로, 그리고 지속가능한 형태로 이루어질 수 있도록 기초적인 심리학적 자료와 과학적인 통찰력을 법 영역에 제공하는 역할을 해야 한다.

인공지능을 활용하는 범죄는 어떤 교육적, 지리적, 문화적 배경을 가진 사람이 어떤 상황에서 어떤 목적을 이루기 위하여 범하게 되는지, 그러한 범죄의 피해자는 대체로 어떤 사회인구학적 배경과 성격특질을 가진 사람들 인지 등을 파악하는 연구가 필요할 것으로 예상된다(King et al., 2020). 또한 인공지능의 심리적 조정 능력을 이용하는 피성범죄나 사기범죄, 성범죄에 대항하는 방법에 대한 연구도 필요할 것으로 예상된다. 인간의 의인화 경향성과, 그 반대로 사람에 대하여 인간성을 부정하고 물질이나 동물처럼 인식하는 비인간화 경향성(dehumanization. Haslam & Loughnan, 2014)을 이해할 수 있는 연구들을 통하여 범죄에 이용되는 인공지능의 기만과 속임수를 극복할 수 있는 방법을 모색할 필요가 있다.

피해자

뇌신경과학은 제4차 산업혁명을 주도하는

로보틱스, 인공지능, 나노기술, 바이오기술, 양자컴퓨팅 등과 융합하여 기하급수적인 속도로 발전하고 확장될 것으로 예상된다. 뇌의 활동(생각, 기억 등)을 탐지하는 기술, 기억의 내용을 선택적으로 조절하는 기술, 의식을 이식하는 기술, 줄기세포로부터 생물학적 뇌를 인공적으로 만들어서 키우는 기술(brain organoids) 등의 발전이 법에 미치는 영향을 선구적으로 연구하는 니타 페라하니(Nita Farahany) 교수는 뇌과학이 전개하는 지평의 확장에 의해 지금까지 법이 당연시 해온 인간에 대한 여러 가지 가정들이 새롭게 정의되어야 하고, 인간의 존엄성, 자유, 권리 등에 대한 법적 페러다임에 많은 문제가 제기될 것으로 예상한다(Farahany, 2012; Farahany, 2016; Farahany et al., 2018).

대한민국 대법원은 불법행위의 피해자에게는 그로 인한 손해의 확대를 방지하거나 줄이기 위하여 합리적인 노력을 해야 할 일반적인 손해경감의 의무(duty to mitigate)가 있음을 판결로 명시하였다(대법원 2003. 7. 25. 선고, 2003다22912, 판결). 피해가 커지지 않도록 합리적인 노력을 하지 않아서 발생하는 추가적인 손해와 피해에 대해서는 피해보상 혹은 배상을 청구할 권리가 인정되지 않는 것이다. 대법원이 명시한 손해경감의 의무가 금전적, 신체적, 물질적 손해뿐만 아니라 정신적, 심리적 손해에도 적용되는지는 법 분야에서 검토되어야 한다. 예를 들어, 성폭행 피해자는 폭행피해 당시 고통, 공포, 분노 등의 감정을 경험하게 되고, 그러한 감정에 의해 스트레스 호르몬인 에피네프린의 분비가 촉진되는데, 에피네프린은 경험하는 감정을 장기기억에 저장하고 오랫동안 유지될 수 있도록 고착시키는 역할을 한다(Cahill & Alkire, 2003). 따라서

성폭행 피해자가 피해 당시 경험하는 부정적 감정은 오랜 동안 기억되어 시간이 많이 지난 후, 외상 후 스트레스장애(post-traumatic stress disorder)로 악화될 수 있다(Chivers-Wilson, 2006). 그런데, 고혈압 등의 심혈관 질환의 치료에 사용되는 프로파놀롤(propranolol)을 성폭행 피해 직후의 피해자에게 투약하면, 이후에 외상 후 스트레스장애의 발생이 예방될 가능성이 있다(Pitman, Sanders, Zusman et al. 2002). 프로파놀롤이 뇌에서 정서 정보를 처리하는 편도체(amygdala)에 작용하여 공포, 두려움, 쇼크 등의 부정적인 감정 기억을 선택적으로 지우는 효과를 가지기 때문이다(Kindt et al., 2009). 프로파놀롤의 기억조절 효과는 아직 더 많은 연구들에 의해서 확실히 증명되어야 하지만, 만약 바이오기술의 발전에 의해 인간의 기억을 조절하는 것이 가능해져서 범죄피해자, 사고피해자, 기타 목격증인이 법적, 의학적인 이유로 외상적 경험에 대한 기억의 일부를 선택적으로 지우는 의료적 처치(Maren, 2011)를 받게 된다면, 피해경험에 대한 그들의 진술이 증거능력을 가질 수 있는지에 대한 의문이 생길 수 있다. 예를 들어, 성범죄 피해 직후에 범인의 인상착의를 경찰에 진술했던 피해자가 이후에 의사가 처방한 프로파놀롤을 복용하여 피해 및 범인에 대한 기억을 상당 부분 상실하여 재판에서 피고인을 알아보지 못한다면, 피해자의 진술을 근거로 피고인을 처벌할 수 있을까? 사법절차와 판단의 정확성을 위하여 손해경감의 의무를 정신적, 심리적 손해에 대해서는 배제해야 할 것인지(즉, 외상 후 스트레스 장애의 발병을 예방하기 위한 목적의 기억조절 기술의 사용을 불허해야 할 것인지), 아니면 법적 공정성과 의료윤리의 실현을 위해서 정신적, 심리적 손해에 대해서도 적용해야 할

것인지를 합리적으로 결정할 수 있는 과학적 근거를 법심리학이 모색해야 한다.

결론

제4차 산업혁명 시대에 예상되는 새로운 기술의 발전은 사법절차에 새로운 자료, 새로운 증거, 새로운 절차를 유입할 것이고, 사법시스템을 기술주도적 혹은 기술의존적인 시스템으로 조금씩 변화시킬 것이다. 우선 멀리 않은 미래에 포렌식평가, 배심원선정, 과학적 증거, 사법 의사결정, 그리고 수사 영역에서 새로운 기술에 의한 변화와 그에 수반되는 법심리학적 연구의 필요성이 나타날 것으로 예상된다. 기술주도적으로 변화해가는 사법시스템에서는 값비싼 고급기술을 사용하고 운영할 수 있는 재력과 능력을 가진 사람이 그렇지 않은 사람보다 항상 유리한 처우를 받게 되는 구조적 불평등이 심화될 수 있다. 따라서 새로운 기술의 유입에 의해 촉진되는 법심리학적 연구는 그러한 기술의 적용이 사법절차에서 인간의 존엄성, 사생활보호, 정의, 공정성, 자율성 등의 심리적 가치를 보호, 보장하는지 확인하고, 심리적 가치의 손상이 발견되면 그 원인을 규명하고 대안을 모색하는 노력에 집중되어야 한다. 이러한 심리적 가치를 보호, 보장하는 핵심은 최고의 정확성(accuracy)과 완전한 투명성(transparency)이 유지되는 형사절차와 사법절차를 구축하는 것이다(Simon, 2012). 빅데이터를 분석하여 개인의 행동을 예측하는 것은 정확한가? 기계학습에 기초한 인공지능 기술이 초래할 수 있는 사생활침해의 범위는 어떤 수준인가? 인공지능이 안정적으로 담당할 수 있는 사법판단의 영역과 범위는 어디까지

인가? 브레인 디코딩 기술이 만들어내는 새로운 과학적 증거에 기초해서 이루어지는 사법판단 및 결정은 정확한가? 그러한 증거가 판단자에 미치는 영향은 무엇인가? 등이 기술혁신 시대의 법심리학적 연구가 탐구해야 하는 물음들이 될 것으로 예상된다.

새로운 기술들에 의해 초래되는 물음들에 답하기 위한 법심리학적 연구 자체도 동일한 기술들에 의한 방법적인 혁신에 의해 이루어질 것으로 예상된다(e.g., Price, Bruer, & Adkins, 2020). 따라서 법심리학자는 제4차 산업혁명 시대를 주도할 것으로 예상되는 빅데이터, 인공지능, 기계학습 등의 혁신적 기술들에 대한 이해와 지식을 가져야 하고, 그 기술들을 스스로 사용 및 응용할 수 있는 능력을 갖추어야 한다. 법심리학이 제4차 산업혁명 시대에 본질적으로 변화하는 법적, 사회적 환경에 적응하고 유용한 역할을 유지하기 위해서는 미래의 법심리학자를 배출하는 교육이 심리학, 법학, 신경과학, 컴퓨터과학, 등 많은 분야의 협력에 의해 이루어져야 할 것으로 생각된다.

참고문헌

- Adjerid, I., & Kelley, K. (2018). Big data in psychology: A framework for research advancement. *American Psychologist*, 73(7), 899-917. <https://doi.org/10.1037/amp0000190>.
- Alaieri, F., & Vellino, A. (2016). Ethical decision making in robots: Autonomy, trust and responsibility. *Lecture Notes in Computer Science (LNCS)*, 9979, 159-168.
- Aletras, N., Tsarapatsanis, D., Preotiu-Pietro, D.,

- & Lampos, V. (2016). Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective. *PeerJ Computer Science*, 2, e93. <https://doi.org/10.7717/PEERJ-CS.93>.
- American Psychological Association (2006). Evidence-based practice in psychology. *American Psychologist*, 61(4), 271-285. <https://doi.org/10.1037/0003-066X.61.4.271>.
- Amini, P., Ahmadiania, H., Poorolajal, J., & Moqaddasi Amiri, M. (2016). Evaluating the high risk groups for suicide: A comparison of logistic regression, Support Vector Machine, Decision Tree and Artificial Neural Network. *Iranian Journal of Public Health*, 45, 1179-1187.
- Amirian, J. (2013). Weighing the admissibility of fMRI technology under FRE 403: For the law, fMRI changes everything - and nothing. *Fordham Urban Law Journal*, 41, 715-770.
- Anastasi, A., & Urbina, S. (1997). Psychological testing (7th ed.). Prentice Hall/Pearson Education.
- Anumanchipalli, G. K., Chartier, J. & Chang, E. F. (2019). Speech synthesis from neural decoding of spoken sentences. *Nature* 568, 493-498. <https://doi.org/10.1038/s41586-019-1119-1>.
- Aspinwall, L. G., Brown, T. R., Tabery, J. (2012). The double-edged sword: does biomechanism increase or decrease judges' sentencing of psychopaths? *Science*, 337, 6096, 846-849. <https://doi.org/10.1126/science.1219569>.
- Baker, L. A., Bezdjian, S., & Raine, A. (2006). Behavioral Genetics: The science of antisocial behavior. *Law and Contemporary Problems*, 69, 7-46 (Winter).
- Barnier, A. J., & McConkey, K. M. (1999). Autobiographical remembering and forgetting: what can hypnosis tell us?. *The International Journal of Clinical and Experimental Hypnosis*, 47(4), 346-365. <https://doi.org/10.1080/00207149908410041>
- Barrett, J. L. & Keil, F. C. (1996). Conceptualizing a nonnatural entity: Anthropomorphism in God concepts. *Cognitive Psychology*, 31, 219-247.
- Bennett, C., Miller, M., & Wolford, G. (2009). Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: an argument for multiple comparisons correction. *NeuroImage*, 47, S125. [https://doi.org/10.1016/s1053-8119\(09\)71202-9](https://doi.org/10.1016/s1053-8119(09)71202-9)
- Ben-Shakhar, G., Bar-Hillel, M. & Kremnitzer, M. (2002). Trial by polygraph: Reconsidering the use of the guilty knowledge technique in court. *Law and Human Behavior*, 26, 527-541. <https://doi.org/10.1023/A:1020204005730>.
- Ben-Shakhar, G., & Elaad, E. (2002). Effects of questions' repetition and variation on the efficiency of the Guilty Knowledge Test: A reexamination. *Journal of Applied Psychology*, 87(5), 972-977. <https://doi.org/10.1037/0021-9010.87.5.972>.
- Ben-Shakhar, G., & Elaad, E. (2003). The validity of psychophysiological detection of information with the Guilty Knowledge Test: A meta-analytic review. *Journal of Applied Psychology*, 88(1), 131-151. <https://doi.org/10.1037/0021-9010.88.1.131>.

- Bleidorn, W., & Hopwood, C. J. (2019). Using Machine Learning to Advance Personality Assessment and Theory. *Personality and Social Psychology Review, 23*(2), 190-203. <https://doi.org/10.1177/1088868318772990>.
- Brants, C. (2013). Wrongful convictions and inquisitorial process: The case of the Netherlands. *University of Cincinnati Law Review, 80*(4), 1069-1114.
- Brennan, T., Dieterich, W., & Ehret, B. (2009). Evaluating the predictive validity of the Compas risk and needs assessment system. *Criminal Justice and Behavior, 36*(1), 21-40. <https://doi.org/10.1177/0093854808326545>.
- Brooks, K. R., Morris, T., & Thompson, P. (2011). Contrast and stimulus complexity moderate the relationship between spatial frequency and perceived speed - Implications for MT models of speed perception. *Journal of Vision, 11*(19), 1-10. <https://doi.org/10.1167/11.14.19>.
- Brunner, H. G., Nelen, M., Breakefield, X. O., Ropers, H. H., & van Oost, B. A. (1993). Abnormal behavior associated with a point mutation in the structural gene for monoamine oxidase A. *Science, 262*, 5133, 578-580. <https://doi.org/10.1126/science.8211186>.
- Cahill, L., & Alkire, M. (2003). Epinephrine enhancement of human memory consolidation: Interaction with arousal at encoding. *Neurobiology of Learning and Memory, 79*, 194-198. [https://doi.org/10.1016/S1074-7427\(02\)00036-9](https://doi.org/10.1016/S1074-7427(02)00036-9).
- Carmena, J. M., Lebedev, M. A., Crist, R. E., O'Doherty, J. E., Santucci, D. M., Dimitrov, D. F., Patil, P. G., Henriquez, C. S., & Nicolelis, M. A. (2003). Learning to control a brain-machine interface for reaching and grasping by primates. *PLoS biology, 1*, 2, E42. <https://doi.org/10.1371/journal.pbio.0000042>.
- Caspi, A., McClay, J., Moffitt, T. E., et al. (2002). Role of genotype in the cycle of violence in maltreated children. *Science, 297*, 5582, 851-854. <https://doi.org/10.1126/science.1072290>.
- Ceci, S. J., & Bruck, M. (1995). Jeopardy in the courtroom: A scientific analysis of children's testimony. American Psychological Association. <https://doi.org/10.1037/10180-000>.
- Čerka, P., Grigiene, J., & Sirbikyte, G. (2015). Liability for damages caused by artificial intelligence. *Computer Law & Security Review, 31*, 376-389. <https://doi.org/10.1016/j.clsr.2015.03.008>.
- Chadwick, M. J., Hassabis, D., Weiskopf, N., & Maguire, E. A. (2010). Decoding individual episodic memory traces in the human hippocampus. *Current Biology, 20*(6), 544-547. <https://doi.org/10.1016/j.cub.2010.01.053>.
- Chapman, B. P., Weiss, A., & Duberstein, P. R. (2016). Statistical learning theory for high dimensional prediction: Application to criterion-keyed scale development. *Psychological Methods, 21*(4), 603-620. <https://doi.org/10.1037/met0000088>.
- Chivers-Wilson K. A. (2006). Sexual assault and posttraumatic stress disorder: a review of the biological, psychological and sociological factors and treatments. *McGill Journal of Medicine: An*

- International Forum for the Advancement of Medical Sciences by Students*, 9(2), 111-118.
- Chung, U. D. (2020). Will trials be enlightened?: Looking into the past, the present and the future of trials through the lens of the comparison with mythology. In Yoon, J.S., Han, S.H., Ahn, S.J. (Eds.), *Law's dilemmas*, (pp 324-337). Bobmunsa. ISBN: 9788918911106(8918911106)
- Coldwell, D. A. L. (2019). Negative influences of the 4th Industrial Revolution on the workplace: Towards a theoretical model of entropic citizen behavior in toxic organizations. *International Journal of Environmental Research and Public Health*, 16, 2670. <https://doi.org/10.3390/ijerph16152670>.
- Cronbach, L. J. (1949). *Essentials of psychological testing*. Harper.
- De Angeli, A. (2009). Ethical implications of verbal disinhibition with conversational agents. *Psychology Journal*, 7(1), 49-57.
- De Angeli, A., & Brahnam, S. (2008). I hate you! Disinhibition with virtual partners. *Interacting with Computers*, 20(3), 302-310. <https://doi.org/10.1016/j.intcom.2008.02.004>.
- Denno, D. W. (2009). Behavioral genetics evidence in criminal cases: 1994-2007. In Nita A. Farahany (Eds.) *The Impact of Behavioral Sciences on Criminal Law*, pp. 317-354. Oxford University Press. Fordham Law Legal Studies Research Paper No. 1089171.
- Dolphin, T. (Eds, 2015). *Technology, Globalisation and the Future of Work in Europe: Essays on Employment in a Digitised Economy*. Institute for Public Policy Research, 45.
- Dombrowski, U. & Wagner, T. (2014). Mental strain as field of action in the 4th industrial revolution. *Proceedings of the 47th CIRP Conference on Manufacturing Systems*, 100-105.
- Dressel, J., & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. *Science Advances*, 4,1. <https://doi.org/10.1126/sciadv.aao5580>.
- Edmond, G., Biber, K., Kemp, R. I. & Porter, G. (2009). Law's looking glass: Expert identification evidence derived from photographic and video images. *Current Issues in Criminal Justice*, 20(3), 337-377. <https://doi.org/10.1080/10345329.2009.12035817>.
- Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4), 864-886. <https://doi.org/10.1037/0033-295X.114.4.864>
- Epley, N., Waytz, A., Akalis, S., & Cacioppo, J. T. (2008). When we need a human: Motivational determinants of anthropomorphism. *Social Cognition*, 26, 143- 155. <https://doi.org/10.1521/soco.2008.26.2.143>
- Europol. (2017). *Serious and organized crime threat assessment*. <https://www.europol.europa.eu/socta/2017/>.
- Farahany, N. A. (2012). Incriminating thoughts, *Stanford Law Review*, 64, 351-408.
- Farahany, N. A. (2016). Neuroscience and behavioral genetics in US criminal law: An empirical analysis. *Journal of Law and the Biosciences*, 2, 485-509.

- <https://doi.org/10.1093/jlb/lsv059>.
- Farahany, N. A., Greely, H. T., Hyman, S., Koch, C., Grady, C., Paşca, S. P., Sestan, N., Arlotta, P., Bernat, J. L., Ting, J., Lunshof, J. E., Iyer, E., Hyun, I., Capestany, B. H., Church, G. M., Huang, H., & Song, H. (2018). The ethics of experimenting with human brain tissue. *Nature*, 556, 7702, 429-432.
<https://doi.org/10.1038/d41586-018-04813-x>.
- Ferguson, A. G. (2016) The big data jury. *Notre Dame Law Review*, 91(3), 935-1006.
- Fowler, T., Langley, K., Rice, F., et al. (2009). Psychopathy trait scores in adolescents with childhood ADHD: the contribution of genotypes affecting MAOA, 5HTT and COMT activity. *Psychiatric Genetics*, 19(6), 312-319.
<https://doi.org/10.1097/YPG.0b013e3283328df4>.
- Freeman, K. (2016). Algorithmic injustice: How the Wisconsin Supreme Court failed to protect due process rights in State v. Loomis. *North Carolina Journal of Law & Technology*, 18, 75-106.
- Fuss, J. (2016). Legal responses to neuroscience. *Journal of Psychiatry & Neuroscience*, 41(6), 363-365. <https://doi.org/10.1503/jpn.160147>.
- Fuss, J., Dressing, H., & Briken, P. (2015). Neurogenetic evidence in the courtroom: a randomized controlled trial with German judges. *Journal of Medical Genetics*, 52, 730-737.
<https://doi.org/10.1136/jmedgenet-2015-103284>
- Galton, F. (1878). Composite portraits. London: Harrison & Sons, St Martin's Lane.
www.galton.org.
- Galton, F. (1890) Criminal anthropology. *Nature*, 42, 75-6. www.galton.org.
- Ganis, G., Kosslyn, S. M., Stose, S., Thompson, W. L., & Yurgelun-Todd, D. A. (2003). Neural correlates of different types of deception: an fMRI investigation. *Cerebral cortex*, 13(8), 830-836.
<https://doi.org/10.1093/cercor/13.8.830>
- Gazzaniga, M. S. (2008). The law and neuroscience. *Neuron*, 60(3), 412-415.
<https://doi.org/10.1016/j.neuron.2008.10.022>
- Gershenfeld, N. & Vasseur, J. P. (2014). As objects go online: The promise (and pitfalls) of the Internet of Things. *Foreign Affairs*, March/April.
- Glenn, A. L., & Raine, A. (2014). Neurocriminology: Implications for the punishment, prediction and prevention of criminal behaviour. *Nature Reviews Neuroscience*, 15, 54-63.
<https://doi.org/10.1038/nrn3640>.
- Gogarty, B., & Hagger, M. (2008). The laws of man over vehicles unmanned: The legal response to robotic revolution on sea, land and air. *Journal of Law, Information and Science*, 19, 73-145.
<https://doi.org/10.1525/sp.2007.54.1.23>
- Golbeck, J., Robles, C., Edmondson, M., & Turner, K. (2011). Predicting personality from Twitter. IEEE Third International Conference on Privacy, Security, Risk and Trust (PASSAT) / IEEE Third International Conference on Social Computing (SocialCom), Boston, MA, 149-156.

- <https://doi.org/10.1109/PASSAT/SocialCom.2011.33>
- Gould, J., & Leo, R. (2010). One hundred years later: Wrongful convictions after a century of research. *The Journal of Criminal Law and Criminology*, 100(3), 825-868.
- Hallevey, G. (2010). The criminal liability of artificial intelligence entities - from science fiction to legal social control. *Akron Intellectual Property Journal*, 4(2), 171-201.
- Hannah-Moffat, K. (2019). Algorithmic risk governance: Big data analytics, race and information activism in criminal justice debates. *Theoretical Criminology*, 23(4), 453-470. <https://doi.org/10.1177/1362480618763582>.
- Hare, R. D. (2003). *The Hare Psychopathy Checklist-Revised* (2nd ed.). Toronto: Multi-Health Systems.
- Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, 458, 7238, 632-635. <https://doi.org/10.1038/nature07832>.
- Hart, S. D. (2005). The "State of the Art" in Risk Assessment. Presentation to the International Conference on the Management and Treatment of Dangerous Offenders in York, England, 28-30 September, 2005.
- Hart, S. D. & Cooke, D. J. (2013). Another look at the (im-)precision of individual risk estimates made using actuarial risk assessment instruments. *Behavioral Sciences and the Law*, 31(1), 81-102.
- Haslam, N. & Loughnan, S. (2014). Dehumanization and inhumanization. *Annual Review of Psychology*, 65(1), 399-423.
- Hoffmeister, T.A. (2012). Investigating jurors in the digital age: One click at a time. *Kansas Law Review*, 60, 611-648. <https://doi.org/10.17161/1808.20186>.
- Horikawa, T., Tamaki, M., Miyawaki, Y., & Kamitani, Y. (2013). Neural decoding of visual imagery during sleep. *Science (New York, N.Y.)*, 340, 6132, 639-642. <https://doi.org/10.1126/science.1234330>.
- Huth, A. G., Lee, T., Nishimoto, S., Bilenko, N. Y., Vu, A. T., & Gallant, J. L. (2016). Decoding the semantic content of natural movies from human brain activity. *Frontiers in Systems Neuroscience*, 10, Article 81. <https://doi.org/10.3389/fnsys.2016.00081>.
- Iliescu, D., & Greiff, S. (2019). The impact of technology on psychological testing in practice and policy: What will the future bring? *European Journal of Psychological Assessment*, 35(2), 151-155. <https://doi.org/10.1027/1015-5759/a000532>.
- Jakobsson, M. & Ratkiewicz, J. (2006). Designing ethical phishing experiments: a study of (ROT13) rOnl query features. In Proceedings of the 15th International Conference on World Wide Web (Edinburgh, Scotland, May 23-26). WWW '06. ACM Press, New York, NY, 513-522. <https://doi.org/10.1145/1135777.1135853>.
- Floridi, L., Taddeo, M. & Turilli, M. (2009). Turing's imitation game: Still an impossible challenge for all machines and some judges - An evaluation of the 2008 Loebner Contest. *Minds & Machines*, 19, 145-150. <https://doi.org/10.1007/s11023-008-9130-6>.

- Jones, O. D., Marois, R., Farah, M. J., & Greely, H. T. (2013). Law and Neuroscience. *Journal of Neuroscience*, 33(45), 17624-17630.
<https://doi.org/10.1523/JNEUROSCI.3254-13>.
- Kahneman, D. & Krueger, A. B. (2006). Developments in the measurement of subjective well-being. *Journal of Economic Perspectives*, 20(1), 3-24.
<https://doi.org/10.1257/089533006776526030>.
- Kaplan, J. (2006). Misinformation, misrepresentation, and misuse of human behavioral genetics research, *Law and Contemporary Problems*, 69, 47-80 (Winter).
- Kasabian, A. (2015). Litigating in the 21st Century: amending challenges for cause in light of big data. *Pepperdine Law Review*, 43, 1, 173-212.
- Kim, S. J. (2013). Verdict of Innocence and Judicial Fact-finding. Kyungin Publishing. ISBN: 9788949910017(8949910012)
- Kindt, M., Soeter, M., & Vervliet, B. (2009). Beyond extinction: erasing human fear responses and preventing the return of fear. *Nature Neuroscience*, 12(3), 256-258.
<https://doi.org/10.1038/nn.2271>.
- King, T. C., Aggarwal, N., Taddeo, M. et al. (2020). Artificial intelligence crime: An interdisciplinary analysis of foreseeable threats and solutions. *Science and Engineering Ethics*, 26, 89-120.
<https://doi.org/10.1007/s11948-018-00081-0>
- Kirat, D., Jang, J., & Stoecklin, M. P. (2018). DeepLocker: Concealing targeted attacks with AI locksmithing. Blackhat. IBM Research.
- Kovera, M. B., Dickinson, J. J., & Cutler, B. L. (2003). Voir dire and jury selection: Practical issues, research findings, and directions for future research. In A. M. Goldstein (Ed.), *Handbook of psychology: Vol. 11. Forensic psychology* (pp. 161-175). Wiley.
- Kramera, A. D. I., Guilloryb, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *PNAS*, 111(24), 8788-8790.
- Kugler, L. (2018). AI judges and juries. *Communications of the ACM*, 61(12), 19-21.
<https://doi.org/10.1145/3283222>.
- Laitinen, S. P. & Loynes, H. J. (2012). A new “Must Use” tool in litigation? For The Defense, August.
- Laney, D. (2001). 3D data management: Controlling data volume, velocity, and variety. META Group.
- Lee, J. W. (2019). *Facial identification: A Meta-Analysis of 50 Years of Research*. [Doctoral Dissertation], The City University of New York.
- Lee, H., & Kuhl, B. A. (2016). Reconstructing perceived and retrieved faces from activity patterns in lateral parietal cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 36(22), 6069-6082.
<https://doi.org/10.1523/JNEUROSCI.4286-15.2016>.
- Lee, M-H., Jeon, H., Sohn, J-H., Eom, J-S., (2018). Detection of Feigned Memory Impairment using Brain Oscillation and Event-related Potential. *Journal of the Ergonomics Society of Korea*, 37(2), 205-217.
<https://dx.doi.org/10.5143/JESK.2018.37.2.205>.
- Lieberman, J. D., & Sales, B. D. (2007). Scientific

- jury selection. Washington, DC: American Psychological Association.
- Liptak, A. (2017). Sent to Prison by a Software Program's Secret Algorithms. *The New York Times* (online), 1 May.
- Loftus, E. F. & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, 13(5), 585-589.
- Maren S. (2011). Seeking a spotless mind: extinction, deconsolidation, and erasure of fear memory. *Neuron*, 70(5), 830-845.
<https://doi.org/10.1016/j.neuron.2011.04.023>
- Martínez-Miranda, E., McBurney, P., & Howard, M. J. W. (2016). Learning unfair trading: A market manipulation analysis from the reinforcement learning perspective. 2016 IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS), Natal, Brazil, pp. 103-109.
<https://doi.org/10.1109/EAIS.2016.7502499>.
- Maslow, A. H. (1943) A theory of human motivation. *Psychological Review*, 50, 370-396.
- Matz, S. C. & Netzer, O. (2017). Using Big Data as a window into consumers' psychology. *Current Opinion in Behavioral Sciences*, 18, 7-12.
- McCabe, D. P., Castel, A. D., & Rhodes, M. G. (2011). The influence of fMRI lie detection evidence on juror decision-making. *Behavioral Sciences & the Law*, 29(4), 566-577.
<https://doi.org/10.1002/bsl.993>.
- Meyer-Lindenberg, A., Buckholz, J. W., Kolachana, B., et al. (2006). Neural mechanisms of genetic risk for impulsivity and violence in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 103(16), 6269-6274.
<https://doi.org/10.1073/pnas.0511311103>
- 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 (2), 679-688.
<https://doi.org/10.1148/radiol.2382050237>
- Mokyr, J. (2009). Intellectual property rights, the industrial revolution, and the beginnings of modern economic growth. *American Economic Review: Papers & Proceedings*, 99(2), 349-355.
- Münsterberg, H. (1908). *On the Witness Stand: Essays on Psychology and Crime*. The McClure Company.
- Musk, E., & Neuralink (2019). An integrated brain-machine interface platform with thousands of channels. *Journal of Medical Internet Research*, 21(10), e16194.
 DOI: 10.2196/16194.
- Nakashima, E. (2012). Stuxnet was work of U.S. and Israeli experts, officials say. *The Washington Post*, 2 June.
- Nass, C., Moon, Y., & Green, N. (1997). Are machines gender neutral? Gender-stereotypic responses to computers with voices. *Journal of Applied Social Psychology*, 27, 864-876.
- Nishimoto, S., Vu, A. T., Naselaris, T., Benjamini, Y., Yu, B., & Gallant, J. L. (2011). Reconstructing visual experiences from brain

- activity evoked by natural movies. *Current Biology*, 21(19), 1641-1646.
<https://doi.org/10.1016/j.cub.2011.08.031>
- Nunez, C. (2017). Artificial intelligence and legal ethics: Whether AI lawyers can make ethical decisions. *Tulane Journal of Technology and Intellectual Property*, 20, 189-204.
- Nuñez, J. M., Casey, B. J., Egner, T., Hare, T., & Hirsch, J. (2005). Intentional false responding shares neural substrates with response conflict and cognitive control. *NeuroImage*, 25(1), 267-277.
<https://doi.org/10.1016/j.neuroimage.2004.10.041>
- Ogloff, J. R. P. (Eds., 2002). *Taking Psychology and Law into the Twenty-First Century*. Springer.
- Paller, K. A., & Kutas, M. (1992). Brain potentials during memory retrieval provide neurophysiological support for the distinction between conscious recollection and priming. *Journal of Cognitive Neuroscience*, 4(4), 375-391.
<https://doi.org/10.1162/jocn.1992.4.4.375>
- Pavlich, G. (2009). The subjects of criminal identification. *Punishment and Society*, 11(2), 171-190.
<https://doi.org/10.1177/1462474508101491>
- Pitman, R. K., Sanders, K. M., Zusman, R. M., Healy, A. R., Cheema, F., Lasko, N. B., Cahill, L., & Orr, S. P. (2002). Pilot study of secondary prevention of posttraumatic stress disorder with propranolol. *Biological Psychiatry*, 51(2), 189-192.
[https://doi.org/10.1016/s0006-3223\(01\)01279-3](https://doi.org/10.1016/s0006-3223(01)01279-3)
- Price, H. L., Bruer, K. C., & Adkins, M. C. (2020). Using machine learning analyses to explore relations between eyewitness lineup looking behaviors and suspect guilt. *Law and human behavior*, 44(3), 223-237.
<https://doi.org/10.1037/lhb0000364>
- Proudfoot, D. (2011). Anthromorphism and AI - Turing's much misunderstood imitation game. *Artificial Intelligence*, 175, 950-957.
- Qiu, L., Chan, S. H. M., & Chan, D. (2017). Big data in social and psychological science: Theoretical and methodological issues. *Journal of Computational Social Science*, 1(1), 59-66.
- Rehm, M. (2008). "She is just stupid"-Analyzing user-agent interactions in emotional game situations. *Interacting with Computers*, 20(3), 311-325.
- Rissman, J., Greely, H. T., & Wagner, A. D. (2010). Detecting individual memories through the neural decoding of memory states and past experience. *Proceedings of the National Academy of Sciences of the United States of America*, 107(21), 9849-9854.
<https://doi.org/10.1073/pnas.1001028107>
- Rosen, J. (2007). The Brain on the Stand- How Neuroscience is Transforming the Legal System. N.Y. TIMES, Mar. 11.
- Rowden, E. (2018). Distributed courts and legitimacy: What do we lose when we lose the courthouse? *Law, Culture and the Humanities*, 14(2), 263-281.
<https://doi.org/10.1177/1743872115612966>
- Rugg, M. D., & Yonelinas, A. P. (2003). Human recognition memory: a cognitive neuroscience perspective. *Trends in Cognitive Sciences*, 7(7), 313-319.
[https://doi.org/10.1016/s1364-6613\(03\)00131-1](https://doi.org/10.1016/s1364-6613(03)00131-1)

- Scherer, M. U. (2015). Regulating artificial intelligence systems: Risks, challenges, competencies, and strategies. *Harvard Journal of Law & Technology*, 29(2), 353-400. <https://doi.org/10.2139/ssrn.2609777>.
- Schwab, K. (2016). The Fourth Industrial Revolution: what it means, how to respond. World Economic Forum. January 14.
- Seymour, J., & Tully, P. (2018). Generative models for spear phishing posts on social media. ArXiv, abs/1802.05196.
- Sharkey, N., Goodman, M., & Ross, N. (2010). The coming robot crime wave. *IEEE Computer Magazine*, 43(8), 6-8.
- Simon, D. (2012). In Doubt: The Psychology of the Criminal Justice Process. Cambridge, Harvard University Press.
- Smethurst, A., Wilson, C., & Collins, K. (2015). The influence of fMRI lie detection evidence on jury decision making following post trial deliberations. *Applied Psychology in Criminal Justice*, 11(3), 147-161. <https://doi.org/10.1002/bsl.993>.
- Sourdin, T. M. (2015). Justice and technological innovation. *Journal of Judicial Administration*, 25(2), 96-105.
- Sourdin, T. (2018). Judge v robot? Artificial intelligence and judicial decision making. *University of New South Wales Law Journal*, 41(4), 1114-1133.
- Stanley, C., & Byrne, M. D. (2016). Comparing vector-based and Bayesian memory models using large-scale datasets: User-generated hashtag and tag prediction on Twitter and Stack Overflow. *Psychological Methods*, 21(4), 542-565. <https://doi.org/10.1037/met0000098>.
- Stupp, C. (2019). Fraudsters used AI to mimic CEO's voice in unusual cybercrime case. The Wall Street Journal, 30 August.
- Surden, H. (2014). Machine learning and law. *Washington Law Review*, 89, 87-115.
- Synnott, J., Dietzel, D., & Ioannou, M. (2015). A review of the polygraph: History, methodology and current status. *Crime Psychology Review*, 1(1), 59-83. <https://doi.org/10.1080/23744006.2015.1060080>.
- Tamietto, M., & de Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. *Nature Reviews Neuroscience*, 11(10), 697-709. <https://doi.org/10.1038/nrn2889>
- Tardif, H., Barry, R. J., & Johnstone, S. (2002). Event-related potentials reveal processing differences in honest vs malingered memory performance. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, 46(2), 147-158. [https://doi.org/10.1016/S0167-8760\(02\)00090-9](https://doi.org/10.1016/S0167-8760(02)00090-9)
- Tiihonen, J., Rautiainen, M., Ollila, H. et al. (2015). Genetic background of extreme violent behavior. *Molecular Psychiatry* 20, 786-792. <https://doi.org/10.1038/mp.2014.130>.
- Timmons, A. C., Baucom, B. R., Han, S. C., Perrone, L., Chaspari, T., Narayanan, S. S., & Margolin, G. (2017). New frontiers in ambulatory assessment: Big Data methods for capturing couples' emotions, vocalizations, and physiology in daily life. *Social Psychological and Personality Science*, 8(5), 552-563.

- <https://doi.org/0.1177/1948550617709115>.
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 49, 433-460.
- Vagnini, V. L., Berry, D. T., Clark, J. A., & Jiang, Y. (2008). New measures to detect malingered neurocognitive deficit: applying reaction time and event-related potentials. *Journal of clinical and experimental neuropsychology*, 30(7), 766-776.
<https://doi.org/10.1080/13803390701754746>.
- Vassos, E., Collier, D. A., & Fazel, S. (2014). Systematic meta-analyses and field synopsis of genetic association studies of violence and aggression. *Molecular Psychiatry*, 19(4), 471-477.
<https://doi.org/10.1038/mp.2013.31>.
- Veletsianos, G., Scharber, C., & Doering, A. (2008). When sex, drugs, and violence enter the classroom: Conversations between adolescents and a female pedagogical agent. *Interacting with Computers*, 20(3), 292-301.
- Wang, J., Cherkassky, V. L., & Just, M. A. (2017). Predicting the brain activation pattern associated with the propositional content of a sentence. *Human Brain Mapping*, 38(10), 4865-4881. <https://doi.org/10.1002/hbm.23692>.
- Watson, D. (2019). The rhetoric and reality of anthropomorphism in artificial intelligence. *Minds & Machines* 29, 417-440.
<https://doi.org/10.1007/s11023-019-09506-6>.
- Wells, G. L. (1978). Applied eyewitness-testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology*, 36(12), 1546-1557.
<https://doi.org/10.1037/0022-3514.36.12.1546>.
- Wells, G. L., Leippe, M. R. & Ostrom, T. M. (1979). Guidelines for empirically assessing the fairness of a lineup. *Law and Human Behavior*, 3, 285-293.
<https://doi.org/10.1007/BF01039807>.
- Wells, G. L., & Quinlivan, D. S. (2009). Suggestive eyewitness identification procedures and the Supreme Court's reliability test in light of eyewitness science: 30 years later. *Law and Human Behavior*, 33(1), 1-24.
<https://doi.org/10.1007/s10979-008-9130-3>.
- Wells, G. L., Steblay, N. K., & Dysart, J. E. (2015). Double-blind photo lineups using actual eyewitnesses: An experimental test of a sequential versus simultaneous lineup procedure. *Law and Human Behavior*, 39(1), 1-14.
- Wells, G. L., Kovera, M. B., Douglass, A. B., Brewer, N., Meissner, C. A., & Wixted, J. T. (2020). Policy and procedure recommendations for the collection and preservation of eyewitness identification evidence. *Law and Human Behavior*, 44(1), 3-36.
<https://doi.org/10.1037/lhb0000359>.
- West, E. M. (2010). Court findings of ineffective assistance of counsel claims in post-conviction appeal among the first 255 DNA exoneration cases. Innocence Project IAC Report.
- Wollert, R. (2006). Low Base Rates Limit Certainty When Current Actuarials Are Used to Identify Sexually Violent Predators: An Application of Bayes' Theorem. *Psychology, Public Policy, and Law*, 12, 56-85.
- Woodruff, W. A. (2014). Evidence of lies and rules of evidence: The admissibility of fMRI-based expert opinion of witness

- truthfulness. *North Carolina Journal of Law & Technology*, 16(1), 105-252.
- Yang, M., Wong, S. C., & Coid, J. (2010). The efficacy of violence prediction: a meta-analytic comparison of nine risk assessment tools. *Psychological bulletin*, 136(5), 740-767.
<https://doi.org/10.1037/a0020473>
- Zander, T. O., & Kothe, C. (2011). Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general. *Journal of Neural Engineering*, 8(2), 025005.
<https://doi.org/10.1088/1741-2560/8/2/025005>.
- 1차원고접수 : 2020. 10. 07.
2차원고접수 : 2020. 12. 28.
최종게재결정 : 2020. 12. 29.

Psychology and Law in the Era of the Fourth Industrial Revolution: New Challenges and Problems

Kwangbai Park

Law and Human Behavior Research Institute, KS&Partners
Department of Psychology, Chungbuk National University (Professor Emeritus)

The humanity stands on the brink of the fourth industrial revolution that will be characterized by a fusion of technologies that is blurring the lines between the physical, digital, and biological spheres (Schwab, 2016). Technological developments in the long run will bring about fundamental changes in the assumptions and functions of the law in respect to individuals and societies. In the short term, new technologies will assist, support, and guide people functioning in legal industries and justice systems. Further developments of the revolutionary technologies will accelerate the replacements of human works by machines, and some civil and criminal proceedings will be conducted online and by computer systems. These changes in the legal environment are expected to promote the adaptive evolution of legal psychology and forensic psychology to identify problems arising in the process of human-law interaction and seek solutions. Based on the microscopic changes that have begun to appear as recently developed technologies have already flowed into the judicial field, for example, crime prediction by artificial intelligence, etc., this review evaluated the novel challenges and problems with which legal psychology and forensic psychology will face in the near future in the areas of forensic evaluation, jury selection, scientific evidence, judicial decision-making, and crime investigation. In addition, as new technology-based crimes and issues of victim protection emerged, potential problems that should be newly addressed in research on criminals and victims were discussed.

Key words : *The Fourth Industrial Revolution, Technology, Science, Legal Psychology, Forensic Psychology, Law*