Neurotechnology in criminal justice: Key points for neuroscientists and engineers

S.H. Geukes, MSc[1], Dr. J. Bijlsma[2,3], Prof. dr. G. Meynen[2,3,4,7], Dr. M.A.H. Raemaekers[1], Prof. dr. N.F. Ramsey[1], Dr. M.A. Simon Thomas[5,6], Dr. D.A.G. van Toor[2,6], Dr. M.J. Vansteensel[1]

[1]University Medical Center Utrecht Brain Center, Department of Neurology and Neurosurgery, University Medical Center Utrecht
[2]Willem Pompe Institute for Criminal Law and Criminology, Department of Law, Faculty of Law, Economics, Governance and Organization, Utrecht University
[3]Utrecht Centre for Accountability and Liability Law, Department of Law, Faculty of Law, Economics, Governance and Organization, Utrecht University
[4]Department of Philosophy, Faculty of Humanities, VU Amsterdam
[5]Institute of Jurisprudence, Constitutional and Administrative Law, Department of Law, Faculty of Law, Economics, Governance and Organization, Utrecht University
[6]Montaigne Centre for the Rule of Law and Administration of Justice, Department of Law, Faculty of Law, Economics, Governance and Organization, Utrecht University
[7]GGZ inGeest, Amsterdam

Keywords: Neurotechnology, Law, Criminal justice, Ethics, Human rights, User-context, Intended use

Neurotechnology encompasses any device (including hardware, testing/stimulation paradigms and data analysis/ interpretation algorithms) that interfaces with the central nervous system to modulate or monitor neural activity. Driven by continuous scientific advances, neurotechnology is increasingly applied in research, clinical and consumer settings. The recent considerable investments in neurotechnology from private sector parties such as Blackrock, Neuralink and Kernel, and the growing interest of policy makers, are likely to further accelerate this trend.

One area of application for neurotechnology is the criminal justice system. Neurotechnology may provide legally relevant information about people's past, present and future behavior. In addition, treatment programs in criminal justice may apply neurotechnological risk assessment and interventions. Currently, the use of neurotechnologies in criminal justice is mostly limited to the diagnosis of neurological conditions such as epilepsy, brain trauma, and dementia, incidentally aiding the assessment of fitness to stand trial and legal insanity[1]. Neurotechnologies have also sporadically been used for other applications in criminal justice[2].

Here, we identify concerns and priorities for *responsible implementation* of neurotechnology in the criminal justice system beyond its current use (BOX1). For that, we first consider areas where neurotechnology has potential value. Then, we identify key requirements that have to be met for the techniques to be usable in criminal justice settings. These requirements often deviate from standards typically applied in neuroscientific research. Using some of the most prominent developments of neurotechnology as examples, we briefly describe to what extent emerging neurotechnologies currently match these requirements. Finally, we assess human rights issues related to the implementation of neurotechnology in criminal justice. We conclude that responsible real-world implementation requires neurotechnologies to be effective, reliable and compliant with human rights, and argue that research and development of neurotechnologies for use in criminal justice should focus on the user-context and prioritize validation for intended use.

Potential value of neurotechnology for the criminal justice system

There are various ways in which neuroscience could potentially be employed in criminal justice [2]. Here, we focus on three examples that relate to criminal proceedings.

First, in the early stages of criminal proceedings (*fact-finding*), after a suspect has been identified, an important question is whether or not the suspect has 'offender knowledge' or *guilty* knowledge. This may be concluded indirectly by investigating secondary sources (such as the suspect's mobile phone), but the only current source of first-hand information is interrogative questioning. Interrogative questioning, however, can introduce bias and statements by defendants are often unreliable. Thus, fact-finding could be served by novel tools that provide (more objective) information on the presence of offender knowledge. An example of a neurotechnological approach that could serve fact-finding is the so-called 'P300 response' [2]. The P300 response can be extracted from electroencephalography (EEG) recordings of brain signals, and is correlated with the recognition of items that may be relevant to the crime, such as the murder weapon in an array of five weapons.

Second, when a suspect has been demonstrated to be guilty of a crime, decisions about the verdict often take into account the *risk of reoffending*. Predictions of reoffending are used to determine whether or not an offender is eligible for parole. However, current risk assessment tools exhibit only poor to moderate performance [3]. Scientific models, designed to aid in risk assessment, may potentially be improved by supplementing them with neurobiological information. In an often-cited study, the activity in the anterior cingulate cortex, as measured with functional magnetic resonance imaging (fMRI), was associated with the likelihood of being re-arrested [4]. Others found that adding data from single photon emission computed tomography (SPECT) to typical risk factors, such as criminal history, improves the prediction of reoffending [5].

Third, in forensic settings, medication and psychotherapy is used to treat mental conditions that may result in violent behavior. Yet, such *interventions* are not always successful and may have severe side effects, and people may remain incarcerated because the risk of reoffending cannot be sufficiently reduced with available interventions. As such, neurotechnological interventions (neurointerventions) may contribute to reducing the risk of reoffending or serve broader rehabilitative applications. An example of a neurointervention is transcranial direct current stimulation (tDCS), which was recently reported to reduce aggression in a forensic population [6].

Requirements concerning effectivity and reliability

Given the potential value of neurotechnologies across the different stages of criminal justice, the question is which requirements have to be met for *responsible* real-world application in criminal proceedings. We highlight central issues that relate to efficacy and reliability and that are not always taken into account in neuroscientific research. We briefly discuss to what extent the neurotechnological tools mentioned above meet these requirements. Although we do not specifically address safety in this article, it is evident that only safe technologies should be used.

The first requirement for the application of neurotechnology in criminal justice is that it should meet standards for minimum efficacy and reliability and is sufficiently specific to its intended application. Importantly, the respective standards generally used for evaluation of isolated tools in the field of neuroscience do not necessarily concur with legal standards for proof. This difference is especially relevant for circumstances in which neurotechnology will be used as an investigational tool. During a police investigation, lower standards of proof may be sufficient – as police are merely looking for potential leads – compared to the court case where proof 'beyond reasonable doubt' is often the standard. Furthermore, courts consider the totality of the available evidence and not just a single source. So, even when neuroscientific evidence alone would not be fully convincing, *together with* other sources of information the legal threshold of proof may still be met. Yet, meta-analyses on the use of EEG-P300 approaches for identifying concealed information report that its accuracy is currently comparable to that of tools such as skin conductance, heart rate and breathing [7], which are not universally considered as sufficiently reliable and are typically excluded from evidence. Moreover, P300 results can be substantially affected by various parameters such as the specific context of the recording, interindividual differences of the response in both culprits and innocent witnesses of a crime, and the way in which the P300 response is computed [8]. Another concern relates to the number of trials required for typical P300 assessments: repeated exposure of an individual to a specific item may affect their memory to a significant extent (cf. [9]) and could potentially be a traumatic experience for a witness that needs to undergo the P300 assessment. Findings on the effectivity, reliability and specificity of fMRI and SPECT for risk assessment are mixed, and these topics clearly require further research.

Second, the efficacy and reliability of neurotechnology should be generalizable across individuals. Risk assessment, for example, may be used to inform decisions in which the liberty of an offender is at stake (e.g., parole decisions). When the reliability and validity of risk assessments are not the same across various groups in society (e.g., gender, race, age), or when unbiased (i.e., blind to the details of the crime) interpretation of neurotechnological data cannot be guaranteed, the application of neurotechnology may negatively affect the legal position of certain groups. This negative impact may conflict with the prohibition of discrimination protected by constitutions and human rights treaties [3]. Whereas generalizability is widely relevant, also for other neurotechnological applications, this topic has received only limited attention in neurotechnology research. For the EEG-P300 response, there are indications that individual traits can impact its amplitude [8], indicating that future research should focus more on evaluating the generalizability of neurotechnological tools in criminal justice.

A third requirement is ecological validity: the context in which neurotechnology is applied for criminal justice will entail many variables that are difficult to control, and will therefore differ substantially from that in a tightly controlled laboratory setting. For example, a suspect in a criminal investigation may benefit from manipulating the outcome of the experiment, while there is no such incentive during regular research or in clinical situations. Efficacy and reliability should thus not be substantially compromised by manipulation by an uncooperative suspect. Because the quality of EEG and fMRI data can be easily degraded by motion or other subject countermeasures [8], the real-world application of these techniques for fact-finding and risk assessment will require the development of novel approaches that minimize or correct for such confounds. For neurointerventions to reduce recidivism, ecological validity means that they actually reduce reoffending rates *in the real world*. A parole decision may be based on the assumption that risk of reoffending is 'manageable' through a certain intervention. The expected efficacy of the intervention should be taken into account in this decision, because in case of low real-world efficacy the risk of reoffending may actually remain too high. Regarding the use of tDCS as a neurointervention, more evidence is needed about its actual efficacy on behavioral outcomes [6].

From the above, it can be derived that more (neuroscientific and technical) research is required before neurotechnology can be applied in criminal justice in domains other than medical-diagnostic imaging. We recommend that this research should be guided by the context of its eventual use. It can also be inferred that application of neurotechnology in the criminal justice system eventually depends (also) on legal norms of evidence and decision-making.

Requirements concerning the protection of human rights

The future application of neurotechnology in criminal justice raises a range of new legal questions [10]. The most fundamental concerns exist vis-a-vis protection of human rights, as recently underscored by UNESCO and the United Nations Human Rights Counsel (A/HRC/RES/51/3) [11]. Human rights are protected by national constitutions, but also through international treaties such as the Universal Declaration of Human Rights and the European Convention on Human Rights. We highlight two important human rights issues specifically related to the application of neurotechnology in criminal justice.

Regarding fact-finding, EEG-P300 results may eventually complement interrogation of the suspect. The goal of EEG-P300 and interrogation is the same, namely to establish whether the suspect has knowledge of a certain crime. In an interrogation, the suspect can invoke the *right to remain silent* and thereby refuse to answer questions. In legal proceedings, it is in principle not allowed to use testimonies that are coercively obtained. Regarding EEG-P300 findings, however, it is as of yet not clear if they should be qualified as testimonial evidence, and therefore should receive the same protection as the spoken word [10].

Neurotechnological interventions aimed at reducing the risk of reoffending may be applied in a context in which the offender will lose his or her freedom when refusing to submit to the intervention. This raises the question whether the offender has *consented freely* to the intervention. Free consent to medical treatment is protected by the right to private life enshrined in *i.a.* article 8 of the European Convention of Human Rights [10].

Other human rights, such as the right to freedom of thought and mental privacy, are also relevant and much is still unclear about the human rights issues attached to the application of neurotechnology in criminal justice [10, 11]. These issues, as well as topics such as data protection, have to receive careful attention before neurotechnology can be applied in a criminal justice setting. Significant benefits can be achieved when legal scholars inform neuroscientists about the relevant legal framework and neuroscientists inform lawyers about the specifics of the techniques that they are using.

Conclusion

The use of neurotechnologies within criminal justice is associated with specific requirements and norms, which often differ from those typically applied by neuroscientists. We therefore argue that research on and development of neurotechnological tools for criminal justice need to zoom in on the 'user-context', moving away from demonstrations of the proof-of-principle, to address topics, several of which are discussed in this manuscript, that relate to the real-world application of the techniques (e.g., efficacy and reliability, specificity, generalizability and ecological validity). The continued research and development of neurotechnologies (including hardware, testing or stimulation paradigms and data analysis and interpretation algorithms) for settings of criminal justice, as well as the responsible implementation of these techniques in this domain, require a multidisciplinary approach in which neuroscientists, behavioral experts, legal and ethical scholars, and end-users should join forces. These efforts should lead to the definition of an 'intended use' for each neurotechnology that is applied in criminal justice, with validation of the technology for the intended use, similar to current regulatory requirements for medical devices. This approach will eventually ensure that a developed neurotechnology adequately serves its desired purpose within the criminal justice domain, and is not used for other purposes, for which it was not designed and for which it may therefore be inappropriate.

BOX1: Key recommendations for responsible application of neurotechnology in criminal justice

- Focus research and development of neurotechnologies for criminal justice on topics that are critical for real-world implementation, including efficacy, reliability, specificity, generalizability and ecological validity
- Engage eventual neurotechnology end-users of the criminal justice domain to ensure the developed neurotechnological tools meet their requirements
- Develop guidelines for application of neurotechnology in criminal justice settings, including the definition of the intended use and validation for that intended use

Declaration of Interest

MJ Vansteensel is member of the advisory council of G3ICT and board member of the International BCI Society.

Funding and Acknowledgements

This manuscript is based on research commissioned by the Research and Documentation Centre of the Ministry of Justice and Security of the Netherlands. GM is funded by the project Law and Ethics of Neurotechnology in Criminal Justice (LENC), financed by the Dutch Research Council (NWO, project number VI.C.201.067).

References

1. Farahany, N. A. Neuroscience and behavioral genetics in US criminal law: an empirical analysis. *Journal of Law and the Biosciences* **2**, 485-509 (2015).

2. Murphy, E. R. D. & Rissman, J. Evidence of memory from brain data. *Journal of law and the biosciences* **7**, KK1-lsaa078 (2020).

3. Tortora, L., Meynen, G., Bijlsma, J., Tronci, E. & Ferracuti, S. Neuroprediction and A.I. in Forensic Psychiatry and Criminal Justice:A Neurolaw Perspective. **11** (2020).

4. Aharoni, E. *et al*. Neuroprediction of future rearrest. *Proceedings of the National Academy of Sciences - PNAS* **110**, 6223-6228 (2013).

5. Delfin, C. *et al*. Prediction of recidivism in a long-term follow-up of forensic psychiatric patients: Incremental effects of neuroimaging data. *PLoS ONE* **14** (2019).

6. Sergiou, C. S., Santarnecchi, E., Franken, I. H. A. & van Dongen, J. D. M. The effectiveness of Transcranial Direct Current Stimulation as an intervention to improve empathic abilities and reduce violent behavior: A literature review. *Aggression and violent behavior* **55**, 101463 (2020).

7. Meijer, E. H., Selle, N. K., Elber, L. & Ben-Shakhar, G. Memory detection with the Concealed Information Test: A meta analysis of skin conductance, respiration, heart rate, and P300 data. *Psychophysiology* **51**, 879-904 (2014).

8. Leue, A. & Beauducel, A. A meta-analysis of the P3 amplitude in tasks requiring deception in legal and social contexts. *Brain and cognition* **135**, 103564 (2019).

9. Wixted, J. T., Wells, G. L., Loftus, E. F. & Garrett, B. L. Test a Witness's Memory of a Suspect Only Once. *Psychol Sci Public Interest* **22**, 1S-18S (2021).

10. Ligthart, S. in *Coercive Brain-Reading in Criminal Justice* (Cambridge University Press, Cambridge, 2022).

11. Ligthart, S. *et al*. Minding Rights: Mapping Ethical and Legal Foundations of 'Neurorights'. *Cambridge quarterly of healthcare ethics*, 1-21 (2023).