Summary

Opportunities and Risks of the Application of Neurotechnology in Criminal Law

Mr. dr. J. Bijlsma, S.H. Geukes MSc, Prof. dr. G. Meynen, Dr. M.A.H. Raemaekers, Prof. dr. N.F. Ramsey, Dr. mr. M.A. Simon Thomas, Mr. dr. D.A.G. van Toor, Dr. M.J. Vansteensel

March 22, 2022

In recent years, neurotechnology has received significant attention. Neurotechnology concerns techniques that contribute to knowledge about the brain and/or that interact with the brain. The attention for these techniques partly results from continuous technological progress. Because neurotechnology (also) has relevance for application within the justice and security domain, the Scientific Research and Documentation Center (WODC) of the Ministry of Justice and Security has commissioned research into the opportunities and risks of such application of neurotechnology. This report is the result of that research.

This report focuses on the following research question:

'What opportunities and threats can be expected from neurotechnology for the domain of the Ministry of Justice and Security and what impact (legal, ethical and social) can neurotechnology have for policy?'

The research was conducted by researchers from UMC Utrecht (UMCU Brain Center) and researchers from the Utrecht University Law Department (UCALL and Montaigne Center).

In answering the research question, the researchers focused on the three most important applications of neurotechnology in the justice and security domain, being: 1*) investigation and fact-finding*; 2) *risk assessment*; and 3) *intervention*. A technique presents an opportunity if it is effective (and to some degree efficient) in achieving one or more goals that are central to the three applications. A risk is defined as tension with legal and ethical standards and possible unintended, negative side-effects of applying newly developed knowledge and technologies. The legal analysis of risks is limited to the human rights enshrined in the European Convention on Human Rights (ECHR). For this report, the researchers drew on a review of the medical, legal and ethical literature, interviews with fourteen academics with expertise in neuroscience, neurotechnology, neurolaw and neuro-ethics, and a proofreading of a draft version of the report by four other academics with relevant expertise. The findings are described in three sections: 1) technology; 2) law and ethics; and 3) synthesis.

In section 1, existing and emerging neurotechnologies, their application areas and relevant developments are mapped out. Within the range of available neurotechnologies, a distinction can be made between techniques that measure brain activity (EEG, MEG, fMRI, fNIRS, fTCD, PET, SPECT), techniques that can stimulate the brain (tDCS, TMS, TFUS and DBS), techniques that can be used for both measuring neural signals and stimulation of the brain (ECoG, sEEG, endovascular EEG, microelectrode arrays), and techniques that visualize the anatomical structure of the brain (CT, MRI, DTI). Techniques that measure brain activity can be divided further into methods that record electrical signals (EEG), or related changes in the magnetic signal (MEG), and methods (fMRI, fNIRS, fTCD, PET, SPECT) that target the metabolic processes of the brain: vascular reactions, such as changes in blood volume or blood oxygenation, in areas of the brain that are activated when people carry out a certain task.

Every neurotechnology has properties that, depending on the application, represent an advantage or a limitation. For example, measurements of the electrical signals have a high temporal resolution (they accurately measure rapid changes over time), while metabolic signals have a low temporal resolution

due to the multi-second delay between electrical signals and related vascular responses. There are also important differences between neurotechnologies in the spatial resolution and specificity and the spatial range. Microelectrodes, for example, allow the recording of signals from individual brain cells, but the dimensions of the 'arrays' in which these electrodes are organized limit the spatial range to a few millimeters. In contrast, EEG and fNIRS can map the signals of the surface of the entire brain, but per sensor a (relatively large) area of a few centimeters is covered. Other distinctive factors are the degree of invasiveness (for example, the necessity of brain surgery for the application of a certain technique), the possibility to reach deeper brain structures, the sensitivity to disturbances and the properties of the measuring device itself (size, degree of portability, price, etc.). In case neurotechnologies are used to induce changes in the neural signals (neuromodulation), there are important differences in the reversibility of the effects.

Many of the mentioned techniques have important scientific value and are (frequently) used in clinical practice, for example for the diagnosis of disorders such as epilepsy (EEG, MEG, ECoG, S-EEG), brain tumors (MRI, PET) or Alzheimer's disease (PET, SPECT, DTI), in the treatment of depression (TMS) and movement disorders such as Parkinson's disease (DBS), in the preparation for neurosurgical procedures (MRI, fMRI, fTCD) or for the assessment of neurological damage (CT, MRI). Other techniques are (almost) exclusively used in the neuroscientific domain, either because the clinical relevance and applicability of the technique in question is still under investigation (endovascular EEG, microelectrodes, tDCS, TFUS), or because it is mainly used to study fundamental neuroscientific questions in situations where other techniques are less suitable (fNIRS).

The *current* application of neurotechnology in criminal justice is limited to the diagnosis, through clinical-diagnostic methods, of neurological disorders such as frontotemporal dementia using techniques such as MRI and PET. These techniques can be applied, for example, in a pro Justitia behavioral assessment of a defendant. As such, neurotechnology may play a role in the assessment of legal insanity and the decision about a hospital order. In addition, four techniques are highlighted because of their possible *future* applicability in criminal justice. Within the domain of *investigation and fact-finding*, for example, research is being conducted into the use of 1) fMRI to identify deception (neuro-lie detection) and 2) EEG to detect offender knowledge via the so-called P300 response (neuro-memory detection). Neuro-memory and neuro-lie detection have significant potential within the realm of *fact-finding*. In order to determine whether a person is guilty, it is essential to establish what actually happened. A 'peek' into the memory of the suspect would provide valuable information in this respect. In the context of *risk assessment*, 3) fMRI is investigated to estimate the risk of recidivism. As such, neurotechnology may become of relevance for decisions on which sanction should be imposed on an individual offender. Furthermore, 4) the stimulation technique tDCS is being studied for its potential applicability as an *intervention* method within forensic psychiatry. Within the framework of various criminal sanctions, neurotechnology may play a role as an intervention to reduce the risk of recidivism and thereby contribute to the rehabilitation of offenders.

In a general sense, further developments of existing neurotechnologies seem particularly aimed at making non-invasive techniques more widely applicable and at the development and validation of fully implantable systems. With respect to techniques with a foreseeable applicability within the justice and security domain, developments are expected in increasing the spatial resolution of fMRI, in a better understanding and more reliable measurement of the EEG-P300 response, in improving the predictive value of neuroimaging by combination with other biological measures and in increasing knowledge about the effectiveness of the use of neurostimulation within forensic psychiatry. In addition, several recently developed techniques (e.g., fUSI, two-photon microscopy and optogenetics) may have relevance for the study of the human brain, but these techniques reside still (mainly) in the stage of animal research.

Knowledge about the structures and functioning of the brain has increased significantly in recent decades, as described above, partly as a result of technological developments that allow visualizing the brain in a living and active state. This development is already important for criminal (procedural) law, and this importance may develop further in the future. The question is, however, to what extent (future) technical possibilities are also legally permissible. For example, the suspect must be treated with *dignity* (Article 3 ECHR; this article prohibits the use of torture and inhuman and degrading treatment), he has the right to remain *silent* and to not be forced to (actively) cooperate to his own conviction (Article 6 ECHR) and the right of respect for his *private life* (Article 8 ECHR). The question is whether, and to what extent, new neurotechnologies can be applied in line with these human rights. In the legal part, it is assessed whether neurotechnologies can be applied in accordance with the law when they are used for 1) *investigation and fact-finding*; 2) *risk assessment*; and 3) *intervention*. This assessment evolves around the five applications of neurotechnology that arose from the technical part: 1) neuroimaging to diagnose neurological disorders; 2) the use of fMRI to identify deception; 3) the use of EEG-P300 as a means to identify offender knowledge; 4) the use of neurotechnology to estimate recidivism risk; and 5) the use of brain stimulation within forensic psychiatry.

The chapter on the assessment of the use of neurotechnologies for *investigation and fact-finding* describes that a general prohibition with regard to the use of these techniques does not follow, or is expected, from a human rights framework. The respect for human dignity and the prohibition of torture, the right to respect for privacy and the privilege against self-incrimination (i.e., the accused should not be forced to speak or otherwise cooperate in their own conviction) do not, in itself, prohibit certain investigative methods. A judge confronted with the results of a certain method – which may be a neurotechnological method – must specifically determine whether the use and implementation of the method is in accordance with applicable law. The main point of discussion here is whether the results of neurotechnologies can be compared to the spoken word or products of mental effort – for which a more extensive protection applies – or, in short, represent merely biological responses to stimuli. For this assessment, it is important to acknowledge that in criminal procedural law, the criminal law enforcement authorities may only act on a legal basis. In other words, the authority to use a particular method must be provided by law. Depending on the level of invasiveness of a method, such legal basis must provide certain guarantees, such as judicial review prior to implementation for the most drastic methods. With regard to neurotechnological methods, in particular neuro-memory detection that provides insight into the invisible memory, it must be noted that these are far-reaching methods that can only be allowed when there is a legal basis with strict guarantees. These safeguards would then regulate the decision to *use* neurotechnological methods.

With regard to *implementation*, all discussed human rights set boundaries to the level of coercion that may be applied. In other words, if the use of a certain authorization is legally possible, this does not mean that all acts of implementation are *ipso facto* lawful. For example, fixating an individual with significant (unnecessary) violence for the purposes of neuro-memory detection is unlawful. This means that the authorities carrying out a neurotechnological method must behave carefully, in the sense that they only use *lawful coercion*.

Risk assessment is associated with the same tensions with human rights as discussed in the context of *investigation and fact-finding*. Brain scans that are used to estimate the risk of recidivism may not be in conflict with Articles 3 and 8 ECHR, similar to scans that are used in the context of *investigation*. The respective frameworks for this assessment do not fundamentally differ. Where the use of enforced brain scans in the context of *investigation* may conflict with the privilege against self-incrimination laid down in Article 6 ECHR, such conflict is not evident for the use of this type of scans for *risk assessment*. Indeed, it is unclear whether or not Article 6 ECHR offers protection against mandatory cooperation during brain scans for the purposes of diagnosis and risk assessment. Another important topic is that risk assessments should not lead to unjustified unequal treatment based on group characteristics.

In highly exceptional cases, when there is an acute danger to health, forced application of neurotechnology may be permitted under Dutch law and the ECHR. However, Dutch criminal law offers no ground for forced neuro-interventions to reduce the risk of recidivism, and forced *interventions* are virtually unthinkable in the light of Articles 3 and 8 ECHR. Neuro-interventions aimed to inflict suffering qualify as degrading and inhumane treatment within the meaning of Article 3 ECHR. In the light of Articles 3 and 8 ECHR, a difficult question is to what extent neuro-interventions that are imposed, for example, as a special condition in the context of a conditional sentence, are permissible. In that case, a convicted person is free to refuse the intervention, but the consequence is that a prison sentence will be executed. Shouldn't such interventions be considered enforced, and therefore, in principle, inadmissible? This question cannot be definitively answered on the basis of the current state of the case law of the European Court of Human Rights (ECtHR). Although, in general, the intervention would not qualify as involuntary by the ECtHR, there are indications that under certain circumstances involuntariness applies, such as when the convicted person is in a particularly vulnerable position. Conversely, the right to freedom, as laid down in Article 5 ECHR, may require certain categories of offenders – in particular those sentenced for life and those placed under a hospital order – to be offered rehabilitative treatment so that they have the opportunity to be released.

Brain scans that can be used for *interventions* may affect the negative right to freedom of expression (Article 10 ECHR): the right to *refrain* from disseminating opinions, ideas and information. Whether or not the information obtained with brain scans in fact relates to opinions, ideas or information within the meaning of Article 10 ECHR is uncertain. If neuro-interventions change the brain's processes, freedom of thought and conscience may also come into play (Article 9 ECHR). This right cannot be infringed. It is currently unclear, however, whether neuro-interventions affect the psyche in such a way as to constitute an infringement of this right.

Finally, a framework is outlined concerning factors that are relevant to answering the question of whether neuro-interventions are permissible. Factors that play a role here are the aim of the intervention, its level of invasiveness (including side effects and risks), the degree of pressure that is applied and in which context this takes place, the availability of alternatives and the effectiveness of the intervention. It should be borne in mind that the actual application of neuro-interventions may lead to a new dynamic in ECtHR jurisprudence, because of the new questions this technology generates. It is conceivable, for example, that the little pronounced right to freedom of thought and conscience will start to play a more prominent role.

It was briefly discussed that literature shows that the existing human rights framework may be inadequate and that proposals have therefore been made to create new fundamental rights, such as a right to mental integrity.

From an *ethical* perspective, the application of neurotechnology in the justice and security domain affects at least privacy, autonomy, physical and mental integrity, and human dignity. Privacy is obviously a central theme when information from the brain is registered. There is, however, some discussion about the extent to which brain data should be considered 'unique' compared to, for example, DNA data. Autonomy is relevant in (at least) three ways. Firstly: is the consent of a suspect or convicted person really a free, autonomous choice? Or is there a risk of accepting 'an offer you cannot refuse'? Second, neurotechnology that changes the brain may influence a person's decision-making process, which may threaten/undermine the individual's autonomy. Third, if neurotechnology helps people, in the long run, to organize their lives as they desire, it in fact supports their future autonomy. Mental and physical integrity are particularly important in relation to neuro-interventions. It has been argued that new human rights should be used to better protect mental integrity against neurotechnological interference than is currently the case. This immediately raises the question of whether such protection should be absolute, or whether infringements should be possible under certain circumstances. Human dignity seems to play an overarching – or foundational – role in the considerations mentioned above. In other words, in order to respect human dignity, we need to consider the implications of neurotechnology for privacy, autonomy and mental/physical integrity.

In section 3, the researchers reflect on the findings from sections 1 and 2. They conclude that a number of important steps must be taken before new neurotechnology can be implemented responsibly for *investigation and fact-finding*, *risk assessment* and *intervention*.

First, further *research* is needed into the *effectiveness and reliability* of the use of neurotechnologies for application in criminal justice. A better understanding is needed of, for example, the *predictive value* and *specificity* of brain measurements. A relevant question in that respect is for example: Is the occurrence of a certain brain signal specifically related to a lie or could it also be associated with another process? Because criminal justice often takes place at the level of the individual suspect/convict, future research into effectiveness and reliability will have to take the level of the *individual* into account. This requires a different approach than the correlation analyses and group comparisons that are common in neuroscience research. Related to this, it is important to determine to what extent statements about effectiveness and reliability can be *generalized*, or whether certain personal characteristics influence these measures. This is also important to avoid the risk of unjust unequal treatment of suspects. Finally, it needs to be determined to what extent neurotechnologies are vulnerable to *manipulation* of the outcome or usability of the data by uncooperative suspects.

A second important *research* topic is the *safety* of neurotechnologies. This topic concerns the *risks* of the application of neurotechnology itself, but also the possible physical or psychological *side effects*. Further research is especially needed for techniques that require brain surgery and techniques that could have a long-term or lasting effect on the brain.

In addition to conducting research, *implementation* of neurotechnology in criminal justice requires a *clarification of the legal and ethical frameworks*. As described above, much is still unclear about the admissibility of neurotechnology in criminal law. Because the legal context differs per country, it is important that The Netherlands develop their own view on the application of neurotechnology in the justice and security domain, which is also tailored to Dutch criminal law. This process preferably takes place already while research into, and development of, neurotechnology is conducted, because this offers opportunities to tailor the developed technology to, for example, requirements for reliability and effectiveness and legal safeguards. The development of such view can be supported by the ethical debate about the interface between neurotechnology in the justice and security domain and the topics of privacy, autonomy, physical and mental integrity, and human dignity. In this respect, the researchers have identified three important topics that require further reflection in particular.

First: The minimum requirements for reliability. While neuroscientists often regard a technique as an isolated tool and therefore impose high standards on the required reliability of the results of research, lawyers argue that information obtained with neurotechnology in the administration of justice will often be combined with other means of evidence and that the entirety of evidence will be used to demonstrate a criminal offense 'beyond reasonable doubt'. Because this concerns the answer to a *legal* question, where evidence is assessed in conjunction, it is necessary to develop specific requirements for application within criminal law.

Second: Further reflection on specific legal questions raised by the application of neurotechnology in criminal law. For example, for all applications, the question has arisen under what circumstances neurotechnology may be used against the will of the subject. In the context *of investigation and fact-finding* and *risk assessment*, this question arises in the light of the right to remain silent. In the context of neuro-*interventions*, it is clear that these may not be enforced and that the convicted person must choose the intervention voluntarily, but the question is to what extent undergoing a neuro-intervention under threat of deprivation of liberty can be called voluntary. Further reflection on these and other legal questions is needed. The ethical debate on these topics may be helpful in this regard.

Third: Provision of information to the judge. The researchers argue that it is essential that judges are adequately informed about the effectiveness, reliability and safety of neurotechnologies when they are applied in the practice of criminal justice.