RETHINKING THE PRIVILEGE AGAINST SELF-INCRIMINATION IN TERMS OF EMERGING NEURO-TECHNOLOGY: COMPARING THE EUROPEAN AND UNITED STATES PERSPECTIVE[*1]

Marin Mrčela[**2] and Igor Vuletić[***3]

Abstract: This paper analyses new fact-finding methods in criminal proceedings, using state-of-the-art innovations in neuroscience and artificial intelligence (AI). It will outline the existing methods and explain their effects. Then it will address the criminal-law aspects of the use of such methods as evidence in criminal proceedings, with an emphasis on the assessment of their admissibility from the perspective of the right to a fair trial and the privilege against self-incrimination. This topic will be observed from the perspective of US and European law, highlighting the existing jurisprudence of the European Court of Human Rights (ECtHR) and the legal standards established by the court in relation to the privilege against self-incrimination. Based on this analysis, the authors will formulate a conclusion suggesting that the use of current AI technologies should be juxtaposed to the relevant benchmark of the privilege against self-incrimination as the requisite standard of the right to a fair trial.

Keywords: lie detector, fair trial, self-incrimination, truth, evidence, testimonial, artificial intelligence, neuroscience.

1 Introduction

The use of AI technologies for the enhancement of efficiency in criminal prosecution is entering a new era. Various forms of AI, which will most likely play a role as tools to assist law enforcement authorities in the performance of their duties, are currently being developed. However, despite the numerous advantages that such technology can provide, the question arises about whether the breakthrough of new scientific technologies, under the influence of the development of neuroscience and AI systems, could undermine the essence of the right to a fair trial.

---

[*] The research for this paper was conducted within 'Artificial Intelligence and Criminal Law (IP-PRAVOS-18)', a project funded by the Faculty of Law Osijek.
DOI: 10.3935/cyelp.19.2023.504.

[**] Justice of the Supreme Court of the Republic of Croatia, Marin.Mrcela@vsrh.hr. ORCID iD: orcid.org/0000-0002-7559-9543.

[***] Associate Professor of Criminal Law at JJ Strossmayer University in Osijek, ivuletic@pravos.hr. ORCID iD: 0000-0001-5472-5478.

During the interrogation of a defendant, law enforcement in some parts of the world use certain examining tests (or assessment techniques), such as the Guilty Knowledge Test (GNT), the Implicit Association Test (IAT) and the Timed Antagonistic Response Alethiometer Test (TARA). GNT is based on the application of technologies such as electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) and provides for the monitoring of blood flow in the brain. The scans obtained through these methods allow the recognition of certain physiological (not just psychological) reactions of the brain to the posed questions. The test is designed to determine whether or not the respondent recognises certain information connected to the criminal offence. If the information is recognised, then the brain will react differently than if the information is not recognised, and this will be visible on the screen.[1] TARA manufactures a situation in which, if respondents lie, they must perform two incompatible tasks, whereas if they tell the truth, they can perform two compatible ones. Both tasks involve repeatedly classifying target and control statements as true or false. The incompatible task combination, being more difficult, takes longer to complete correctly; hence, slower responses diagnose dishonesty.[2] IAT is a controversial assessment intended to detect subconscious associations between mental representations of objects in memory.[3] Some authors claim that it is suitable for detecting biases during *voir dire*.[4]

These techniques have the purpose of helping investigators distinguish truth from lies. At the same time, neuroscientists are developing supportive mechanisms for the purpose of enabling mind reading and improving accuracy in proving a distinction between truth and lies. After fMRI, the newest and most controversial of such mechanisms for mind reading is brain fingerprinting: an objective method to detect information stored in the brain by measuring EEG brainwaves through sensors placed on the scalp of the person who is being interrogated.[5]

These methods are already in use in different parts of the world, and courts have, from time to time, affirmed that a positive result in such a test can be used as proof of guilt in criminal proceedings. Although the described methodology can be very effective, it creates significant legal

[1] Gershon Ben-Shakhar and others, 'Trial by Polygraph: Reconsidering the Use of the Guilty Knowledge Technique in Court (2002) 26(5) Law and Human Behavior 529-530. See also Igor Vuletić and Tunjica Petrašević, 'Is It Time to Consider EU Criminal Law Rules on Robotics?' (2020) 16 Croatian Yearbook of European Law and Policy 234.

[2] Aiden Gregg, 'When Vying Reveals Lying: The Timed Antagonistic Response Alethiometer' (2007) 21(5) Applied Cognitive Psychology 621.

[3] See Madison Kilbride and Jason Iuliano, 'Neuro Lie Detection and Mental Privacy' (2015) 75 Md L Rev 163, 166-171.

[4] Dale Larson, 'A Fair and Implicitly Impartial Jury: An Argument for Administering the Implicit Association Test During Voir Dire' (2010) 3 DePaul Journal for Social Justice 139, 158.

[5] Lawrence A Farwell, 'Brain Fingerprinting: A Comprehensive Tutorial Review of Detection of Concealed Information with Event-related Brain Potentials (2012) 6(2) Cogn Neurodyn 115.

implications. One of these implications is that such interrogation methods violate the defendant's ability to exercise their right to remain silent. Another issue is the defendant's consent for the use of AI-based tests and the possibility of adverse inferences if the defendant refuses to give consent to the use of AI-based tests in criminal proceedings. These and other related legal issues have contributed to the establishment of a new branch of law: neurolaw.[6]

This paper will explore these issues more thoroughly. The specific contribution of this study lies precisely in a comparison of the European and American approaches to the subject matter. It is noteworthy that, in the existing literature, no such attempt has been made in this particular context. However, the primary purpose of the paper is not comparative, but instead normative in the sense of advocating a particular legal position on how AI technology should reflect on the privilege against self-incrimination. In this sense, the authors will endeavour to establish what they believe to be the most significant argument, which is that if the use of AI technology proves to be reliable and credible, then such a method must be applied in respect of the right to a fair trial as a well-established human rights standard in criminal proceedings.

The analysis is based on a systematic approach for clarity, and the presentation of this problem is divided into four further main sections. The first will briefly outline new AI questioning technologies. The next two sections will address the standards of the privilege against self-incrimination, as a part of the right to a fair trial as developed under the relevant jurisprudence of the US Supreme Court and the ECtHR. In the final section, before the conclusion, the authors will elaborate their standpoint.

2 An overview of innovative neuroscientific methods in criminal proceedings

The concept of lie detection technologies appeared as early as the 18th century, when the measurement of thieves' pulse during interrogations was introduced for the first time. The concept expanded from the 1920s, when the first polygraph testing began.[7] The development of neuroscience and artificial intelligence, which has accelerated and attracted close attention during the past two decades, has led to the creation of new technologies which were inconceivable until recently and which replace some otherwise typical human activities. This phenomenon is present in many fields, so it should come as no surprise that it did not leave out the criminal legal system. Criminal law is one of the fields that is receptive to significant improvements through new technologies, because it allows the removal, or at least reduction, of the deficiencies that inherently flow

[6] Leda Tortora and others, 'Neuroprediction and AI in Forensic Psychiatry and Criminal Justice: A Neurolaw Perspective' (2020) 11 Frontiers in Psychology 1–9.

[7] Robert Bradshaw, 'Deception and Detection: The Use of Technology in Assessing Witness Credibility (2021) 37(3) Arbitration International 708–709.

from human reasoning and thought processes (such as subjectivity, bias and prejudice, naivety, etc). In simple terms, it is much more difficult (if not impossible) to deceive a mechanism based on artificial intelligence, unlike humans, as a fact-finder. Therefore, the use of such supporting mechanisms contributes to the discovery of the truth of a certain event, which carries significant weight in the field of criminal law, in light of the severity and significance of the legal consequences of a conviction or exoneration. In this sense, it is clear that such technology has become more popular and desirable over time, at least from the perspective of the law enforcement authorities. The following sections will present the key features of the most sophisticated mechanisms of modern science and technology in the field of fact-finding in criminal proceedings.

It is well known that the traditional (conventional) polygraph is essentially based on the measurement of certain physiological indicators of stress during questioning related to a specific event. The so-called conventional control question polygraph test (CQT) measures peripheral responses, such as the sweating of the skin, cardiovascular activity and blood pressure, as well as breathing, in an effort to detect reactions that are typical of lying.[8] The assessment of the truthfulness of an individual's statements is made based on the prominence of one or more such indicators. As reactions to stress are highly individual and differ from person to person, the results of such testing are insufficiently reliable. This is among the primary reasons why, over the past two decades, polygraph testing has become largely inadmissible as legitimate evidence in most jurisdictions in the world.[9] Therefore, scientists are striving to devise lie detection mechanisms based on reliable and science-based criteria in order to obtain a credible outcome.

The most dominant technology of this kind today is functional magnetic resonance imaging (fMRI), which enables a neural imaging procedure for observing changes in the cerebral blood flow. This method is also aimed at the monitoring of reactions that are usually linked to untruthful statements. Unlike the previously described traditional lie detector, fMRI does not measure stress reaction to the questioning, but it observes certain cerebral processes which are considered to be connected with lying and internal conflicts. These cerebral processes cannot be consciously controlled, unlike some elements that are used in the old-fashioned polygraph test. The fMRI technology follows the delivery of oxygenated blood to neurons that have recently fired. Based on such observations, scientists can conclude which parts of the brain react to certain stimuli. In this sense, fMRI is considered to be more precise than the older scalp-recorded event-related potentials (ERPs) technique, which was measured

[8] William G Iacono and David T Lykken, 'The Validity of the Lie Detector: Two Surveys of Scientific Opinion (1997) 82(3) J Appl Psychol 426–433.

[9] ibid. See also Timothy B Henseler, 'A Critical Look at the Admissibility of Polygraph Evidence in the Wake of Daubert: The Lie Detector Fails the Test' (1997) 46(4) Cath U L Rev 1247. See also Ed Johnston, 'Brain Scanning and Lie Detectors: Implications for Fundamental Defence Rights' (2016) 22(2) Eur J Current Legal Issues 4.

by an EEG. Observance of the ERPs could not localise their source in the brain, which is possible with fMRI.[10] Therefore, the aforementioned Guilty Knowledge Test (GKT), which was traditionally used during classical polygraph questioning and consisted of an examination of the details of a criminal offence which can only be known to the perpetrator, has been used in combination with fMRI technology in the past two decades to ensure more reliable and science-based results.[11] These developments have made fMRI a recognisable and commercially viable mechanism on the market, so certain US companies have specialised specifically in the development of this technology in the context of questioning for the purpose of criminal prosecution.[12]

However, this technology is also not without its flaws. There are warnings that the technology should be optimised to give results that reflect real-life circumstances, and not to create results under controlled conditions. One example is a study in which the respondents were asked to steal an object and put it in a closet, and then they were shown certain photographs related to the event to monitor cerebral activity. The respondents had to respond truthfully to the control and neutral questions, but they were instructed to deny the event. As a motivation, they were promised an additional monetary amount if they managed to deceive the examiner and fMRI. Scholarly literature notes that this situation significantly differs from reality, because the respondents in this situation know that they will not be punished, while in reality they would do their all to defend themselves. Furthermore, this study highlighted other deficiencies of the method, including the fact that not everybody has the same neurological reaction to lying and deceit.[13] Therefore, the use of fMRI as evidence in criminal proceedings is still largely in its early stages.[14]

Considering the outlined deficiencies in the described techniques, neuroscientists have continued to develop and optimise lie detection mechanisms. This has led to the development of the currently most controversial technology called brain fingerprinting. In essence, it differs from all previous technologies because it is not aimed at determining whether a person is lying or telling the truth, but whether or not a piece of information is stored in their brain.[15] Hence the term 'fingerprinting': 'fingerprinting' allows for the establishment of an objective and science-based link between the fingerprint at the crime scene and the finger of the per-

[10] Scott A Huettel, Allen W Song & Gregory McCarthy, *Functional Magnetic Resonance Imaging* (2nd edn, OUP 2009).

[11] For a detailed results analysis of the first published fMRI use report, see Daniel D Langleben and others, 'Brain Activity During Simulated Deception: An Event-Related Functional Magnetic Resonance Study' (2002) 15(3) Neurimage 727–732.

[12] One such example is No Lie MRI, a San Diego based company that has been producing fMRI-based lie detectors since 2006.

[13] Jonathan G Hakun and others, 'Towards Clinical Trials of Lie Detection with fMRI' (2009) 4(6) Soc Neuroscience 518–527.

[14] Johnston (n 9) 12.

[15] Farwell (n 5) 128.

petrator. Similarly, DNA fingerprinting proves the objective link between the DNA sample taken from the crime scene and the DNA sample taken from the suspect. The term 'brain fingerprinting' seeks to emphasise that this method provides objective and science-based evidence of the link between the images from the crime scene and the memories stored in the brain of the suspect. It is important to emphasise that brain fingerprinting cannot provide information on whether the memory is real or accurate, but simply whether it is stored in the brain (ie whether the suspect recognises a certain motive as something corresponding to their memory).[16]

The brain fingerprinting method functions under the principle of measuring the EEG (electroencephalographic) brain waves through non-invasive sensors which are placed on the head of the examinee. The examinee is then presented with certain words, phrases or images detailing a specific event on a computer screen, along with other irrelevant information. This technology measures the cerebral responses to the presented material and helps detect certain characteristic brain wave patterns. If a person recognises something as significant in a specific context, they will experience the so-called 'Aha!' effect, which will be visible as a specific pattern of brain waves, which are known in neuroscience as P300-MERMER. This test helps answer the question whether certain information is present or absent from the examinee's brain, while the system calculates the statistical reliability of the obtained results. If, however, the statistical processing cannot provide a sufficiently high percentage of reliability of the results, the system will show the result as 'indeterminate'.[17]

The accuracy tests of brain fingerprinting, conducted under controlled conditions by the FBI, CIA and the US Navy shows that, in 97% of cases, the system was able to assess whether the information was stored in the examinee's brain with 100% accuracy. Only in 3% of the examined cases was the system unable to provide a statistical confirmation and gave an 'indeterminate' result.[18] Despite this, none of the above-mentioned agencies provided recommendations for further investments in the development of this technology, ultimately finding it uneconomical.[19]

In the US, where this method originated, there are still no judgments based on evidence obtained through brain fingerprinting. In other parts of the world, there have been cases in which evidence was obtained by such means. For example, India took the lead in a recent case in which a suspected rapist was subjected to examination by means of the Brain

---

[16] Farwell (5) 128.

[17] ibid 115, 117, 128.

[18] ibid 139.

[19] See, for example, 'Investigative Techniques: Federal Agency Views on the Potential Application of "Brain Fingerprinting"' (GAO-02-22, 31 October 2001) 10 <www.govinfo.gov/content/pkg/GAOREPORTS-GAO-02-22/html/GAOREPORTS-GAO-02-22.htm> accessed 3 November 2022.

Electrical Oscillation Signature Profiling (BEOSP) method, which is a form of brain fingerprinting.[20] In some previous cases, Indian courts issued convictions on the basis of evidence obtained through brain scanning.[21]

It is clear from the analysis above that the combined development of neuroscience and AI-based technologies leads to the development of new supporting methods for the determination of the material truth in criminal proceedings. These technological advances should not be hindered, but should be subjected to clear legal criteria. These issues can also be observed from the perspective of the assessment of the authenticity of witness testimony. This paper will, however, focus only on the perspective of the deposition of the suspect and their right to the privilege against self-incrimination. The following analysis will study the scope and limitations of this privilege as a component of the right to a fair trial in cases of the application of the described technologies for the assessment of the truthfulness of the statements of the defendant in criminal proceedings. Both issues have been discussed in comparative literature and practice, but there has been no systematic comparison of US and European law. The following sections will present the legal standards developed under US and European law, shaped through the jurisprudence of the ECtHR. Based on this analysis, we will determine which of the two legal frameworks provides greater legal space for the future implementation of the described technologies, and others with similar features.

3 Neuro-science vs the privilege against self-incrimination: the US perspective

The privilege against self-incrimination is very significant in the Anglo-American legal tradition. Its emergence is considered as one of the most consequential milestones in the development of criminal law and is connected to the expansion of the accusatory type of proceedings from the late 18th century. In the preceding period, from the mid-16th century, the purpose of criminal proceedings was to enable the defendant to speak (and not remain silent) in their case, thus providing them with the opportunity to raise their defence. During that time, the defendant could not retain legal counsel. These conditions drastically changed in the late 18th and especially in the second half of the 19th century, when defendants gained the right to be represented by legal counsel, as experts who could test and assess the hypothesis of the indictment.[22] This new understanding at the core of criminal proceedings soon led to the recognition of the privilege against self-incrimination (along with the 'beyond

[20] Vaibhab Yha, 'The Accused in Hatras Rape Case Will Undergo Brain Fingerprinting. What Is It?' <https://indianexpress.com/article/explained/hathras-rape-case-brain-fingerprinting-7070587/> accessed 19 March 2022.

[21] See, for example, *State of Maharashtra v Sharma*, CC No 508/07 Pune, 12 June 2008 (India).

[22] John T McNaughton, 'The Privilege against Self-Incrimination' (1960) 51(2) J Crim L Criminology & Police Sci 139.

reasonable doubt' standard and the exclusionary rule) as an integral and central component of common law criminal proceedings. This approach has been sustained to date.

In order for any examination technology to be applied on defendants before US criminal courts, it has to meet certain minimal conditions, which ensure that the results can be treated as reliable. In this sense, the admissibility standards for the use of certain evidence before the court in criminal proceedings were established in the landmark *Daubert* case. In this case, the Supreme Court established a clear set of criteria on the basis of which the court should assess the admissibility of certain scientific evidence in the proceedings. Thus, it was determined that the evidence can be admitted if it cumulatively meets the following five criteria: 1) that the technology on which it is based can be scientifically tested; 2) that it was subjected to review and scientific (public) publication; 3) that the percentage of the accuracy of its results is known; 4) that there are clear standards for oversight and control of the functions of the technology; and 5) that the technology is accepted in the scientific community.[23]

Under US law, the privilege against self-incrimination is regulated by the Fifth Amendment to the US Constitution.[24] Thus, the US Supreme Court formulated clear criteria under which certain evidence could fall under the scope of the Fifth Amendment if it is cumulatively incriminating, testimonial and compelled.[25] Regarding the latter criterion, the key decision of *Griffin v California*[26] should be mentioned. In this case, the US Supreme Court held that a comment from the prosecutor or judge to the jury that the defendant's silence should be treated as incriminating evidence represented a violation of the Fifth Amendment of the US Constitution. The Supreme Court affirmed this position in subsequent decisions.[27]

The privilege against self-incrimination is viewed from different perspectives in US judicial practice today,[28] but it can be noted that both the theory and practice focus more on the scope of this privilege in the context of new digital technologies and evidentiary potential that such technologies entail by literally converting evidentiary materials from the

---

[23] *Daubert v Merrel Dow Pharmaceuticals, Inc* [509 US 579] 1993.

[24] US Constitution, Fifth Amendment: 'no person [...] shall be compelled in any criminal case to be a witness against himself'.

[25] *Fischer v United States* [425 US 391, 408] 1976.

[26] *Griffin v California* [380 US 609] 1965.

[27] See, for example, *Mitchell v United States* [526 US 314] 1999.

[28] Debate on all potential aspects of privilege against self-incrimination is beyond the scope of this article. For a detailed insight on the issue of privilege against self-incrimination in the post-conviction phase, see Stephen Vance, 'Looking at the Law: An Updated Look at the Privilege Against Self-Incrimination in Post-Conviction Supervision' (2011) 75(1) Federal Probation Journal 33; for the issue of invoking this privilege due to the fear of foreign prosecution, see a detailed analysis in Gregory O Tuttle, '"Cooperative Prosecution" and the Fifth Amendment Privilege Against Self-Incrimination' (2010) 85(4) NYU L Rev 1348.

physical into the virtual sphere.[29] In this sense, increasing attention is being paid to the issue we are addressing in this paper.

Discussions on whether the use of examination technologies violates the defendant's privilege against self-incrimination mostly focus on the question of whether the results of neuro-testing are a 'testimonial' or 'physical' type of evidence. If the evidence is testimonial, under the condition that the defendant was directly or indirectly coerced, such evidence could not be used in the proceedings as it would fall under the scope of the privilege against self-incrimination. In the *Schmerber v California* decision, the court established the rule that the coerced drawing of blood from the defendant should be treated as physical evidence that would not violate their privilege against self-incrimination, and the same rationale should apply to all other types of testing.[30] However, the issue is that neither this decision nor those that followed established clear criteria for the distinction between physical and testimonial evidence, so practice in this area varies significantly.[31]

This issue is particularly complicated in relation to the use of sophisticated neuro-lie detection technologies. If it is the result of classical polygraph testing, then the answer is clear because these results can only be reached thorough an interaction with the defendant. Whenever there is direct communication, the evidence is testimonial, in accordance with the *Schmerber* standard,[32] which falls under the privilege. The situation is somewhat more complicated with technologies such as fMRI, since they are not based on classic verbal communication. However, given the fact that the examinee responds by pressing buttons on a device and that this is also a form of communication, such evidence may also be considered testimonial and thus protected by the privilege against self-incrimination.[33]

The biggest dilemma is created by sophisticated technologies such as brain-fingerprinting, because they do not require any communication with the examinee, so they cannot be subsumed under the classic concept of testimonial evidence. In this sense, and in light of the *Schmerber* criteria, such evidence would be inadmissible. If that is the case, some forms of neuro-technologies would be admissible in court, while others would not, which creates confusion. Some authors argue that all such technologies should be legally treated as images of cerebral waves, which would give them the same status as DNA evidence, which would no longer make it testimonial or covered by the privilege against self-incrimina-

---

[29] Diego Wright, 'The Right Against Self-Incrimination in the Digital Age' <https://proceedings.nyumootcourt.org/2021/09/the-right-against-self-incrimination-in-the-digital-age/> accessed 19 March 2022.

[30] *Schmerber v California* [384 US 761] 1966.

[31] Kilbride and Iuliano (n 3) 174.

[32] See n 30, 764.

[33] Kilbride and Iuliano (n 3) 176.

tion.[34] However, other authors oppose this view, claiming that this type of evidence analyses the content of the defendant's mind and is based on the knowledge of the defendant. As such, the evidence is testimonial and the denial of the privilege against self-incrimination in such cases would be unjustified.[35] This position is criticised for being contrary to prevailing judicial practice that does not treat evidence as testimonial if it does not contain any acts or statements by the defendant. Therefore, these authors suggest that the focus of the discission should be diverted from the Fifth to the Fourth Amendment and the defendant's right to privacy.[36] Other authors advocate the abandonment of the traditional division of evidence into physical and testimonial and propose four categories of evidence: identifying, automatic, memorialised, and uttered. These authors consider that only such a categorisation of evidence can protect the right to cognitive freedom and mental privacy, within the framework of the Fourth and Fifth Amendments to the US Constitution.[37] However, such views remain in the minority for now and the issue is still primarily observed from the perspective of the Fifth Amendment and the privilege against self-incrimination and the traditional categorisation of evidence into testimonial and physical.

Although neuro-scientific examination is broadly present in criminal proceedings in US courts, it still predominantly relates to the determination of certain mental and neurological disorders which may impact liability.[38] However, with regards to neuro-lie detection technologies, it can be validly concluded that there are still legal doubts about their scientific basis in the US legal system and this is probably one of the main reasons that their evidentiary value is still not recognised in judicial practice. The case law remains modest to date. A noteworthy case is *United States v Semrau*,[39] in which the defendant passed the fMRI test, but the court refused to admit it because it considered it scientifically unreliable. However, the court left the door open for the future application of fMRI if it is scientifically affirmed through further research.[40]

The fMRI technique is still not considered up to par with the *Daubert* standards,[41] although there are positions that this judgment is unfair

---

[34] Henry T Greely and Anthony D Wagner, 'Reference Guide on Neuroscience' in *Reference Manual on Scientific Evidence* (3rd edn, Federal Judicial Center 2011) 791 <www.nap.edu/read/13163/chapter/15#790> accessed 19 March.

[35] Michael S Pardo and Daniel M Filler, 'Neuroscience, Evidence, Legal Culture and Criminal Procedure' (2006) 33(3) Am J Crim L 316.

[36] Kilbride and Iuliano (n 3) 180–186, 193.

[37] Nita A Farahany, 'Incriminating Thoughts' (2012) 64 Stan L Rev 351.

[38] Darby Aono, Gideon Yaffe and Hedy Kober, 'Neuroscientific Evidence in the Courtroom: A Review (2019) 4 Cognitive Research: Principles and Implications 3.

[39] *United States v Semrau*, Court of Appeals for the Sixth Circuit, 693 F.3d 510 (2012).

[40] ibid 31.

[41] Isabella Sousa, 'fMRI v the Frye & Daubert Standards of Evidence: Re-searching for the Truth' (Columbia Undergraduate Law Review, 22 August 2021) <www.culawreview.org/journal/fmri-v-the-frye-amp-daubert-standards-of-evidence-re-searching-for-the-truth> accessed 19 March 2022.

since fMRI is no less reliable than some other technologies whose results are regularly admitted by the court, and that scientific confirmation can only be obtained through its broader use in the described context.[42] Some authors assert that the adoption of progressive procedural legislation in some states, such as Oklahoma, will enable a new approach to the interpretation of the *Daubert* criteria, which will create opportunities for the acceptance of fMRI as valid evidence for lie detection in the near future.[43]

In the USA, where the brain fingerprinting method was conceived, there have only been a few cases on record where evidence obtained in this manner was actually used in criminal proceedings (in the *James B Grinder* case, the *Terry Harrington* case, the *Jimmy Ray Slaughter* case). However, none of these cases was decided on the basis of evidence collected through the application of this method.[44] Therefore, it can be rightly concluded that the technique of brain fingerprinting is still far from accepted as a standard and as regular evidence of the veracity of testimony, at least in criminal proceedings.

4 Neuro-science vs the privilege against self-incrimination: the European perspective

The development and implementation of AI-based technologies in Europe is still significantly lagging behind the process in the US and Asia.[45] While the sight of AI-operated vehicles is commonplace in US and Asian cities, such innovations are exclusively experimental in Europe. However, Europe has recently tried to catch up with the competition, which has been reflected in strengthened legislative activity at the supra-national level in recent years, especially under the auspices of the EU.[46] It is worth mentioning the AI Act and the European Parliament Resolution of 6 October 2021 on artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters (2020/2016(INI)). In addition, some countries are developing technologies for the purpose of predicting crimes and reducing the crime rate. However, the actual prac-

---

[42] Justin Amirian, 'Weighing the Admissibility of fMRI Technology Under FRE 403: For the Law, fMRI Changes Everything – and Nothing' (2013) 41(2) Fordham Urban L J 769, 770.

[43] Julie Elizabeth Myers, 'The Moment of Truth for fMRI: Will Deception Detection Pass Admissibility Hurdles in Oklahoma?' (2010) 6(1) Oklahoma J L & Tech 47.

[44] See details of those cases and their outcomes in Arpad Budahazi and others, 'The Options and Limitations of the Brain Fingerprinting Lie Detection Method in the Criminal Proceedings' (2018) 18(5) Magyar Rendészet 45-48.

[45] See more in Daniel Castro and Michael McLaughin, 'Who Is Winning the AI Race: China, the EU, or the United States? 2021 Update' (*Center for Data Innovation, January* 2021 <www2.datainnovation.org/2021-china-eu-us-ai.pdf> accessed 22 March 2022. See also Igor Vuletić and Tunjica Petrašević, 'Is It Time To Consider EU Criminal Law Rules On Robotics?' (2020) 16 Croatian Yearbook of European Law and Policy 227.

[46] Philipp Hacker, 'AI Regulation in Europe' (7 May 2020) <https://ssrn.com/abstract=3556532> or <http://dx.doi.org/10.2139/ssrn.3556532> accessed 22 March 2022.

tical efficiency of such technologies is still very questionable.[47]

Therefore, it is not surprising that the use of new fact-finding methodologies in criminal proceedings in European countries is somewhat more restricted than in the US. However, from the perspective of individual countries, there have been some interesting decisions related to different AI-related issues. For example, Italian courts have based their decisions on controversial scientific methods in two instances. In the first case, the court accepted the examination of the defendant's inability to form intent through the MMPI test and the Rorschach personality test, as well as the genetic test of predisposition for aggressive behaviour. The positive results of these tests were used as mitigating factors during sentencing. In the other case, the court verified the veracity of the testimony of a sexual harassment victim through IAT and TARA testing in order to determine the real memories. Since the results showed that her memories were natural, and the allegations of her manipulative character raised by the defence and attention seeking were scientifically unfounded, the court convicted the defendant.[48]

The use of modern neuroscientific technologies in European national legislation has so far focused on examinations of the defendant's personality and psycho-somatic capacities in order to enable a more favourable verdict for them. England and Wales, the Netherlands, Slovenia and Italy are leaders in this area.[49] In the context of this paper, it is particularly important to note a recent decision of the Supreme Court of the Netherlands which ruled that the defendant's privilege against self-incrimination was not violated because the defendant was forced to unlock his smartphone with his fingerprint.[50] The court based its decision on the interpretation of the ECtHR in the case of *Saunders v United Kingdom*, according to which there is a difference between evidentiary material which requires consent and that which can be collected independent of consent.[51] The latter material (such as fingerprints) can be taken by force,

---

[47] See for example Dominik Gerstner, 'Predictive Policing in the Context of Residential Burglary: An Empirical Illustration on the Basis of a Pilot Project in Baden-Württemberg, Germany' (2018) 3 European Journal for Security Research 115. See also Sunčana Roksandić, Nikola Protrka and Markus Engelhart, 'Trustworthy Artificial Intelligence and Its Use by Law Enforcement Authorities: Where Do We Stand? in K Skala (ed) *Proceedings of the 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO)* 23-27 May 2022 <www.bib.irb.hr/1196746> accessed 20 August 2023.

[48] See details on both cases in Armando Simbari, 'N 9965 Nota a Ufficio Indagini Preliminari di Cremona' (2012) 2 Rivista Italiana di medicina legale 749–758.

[49] For the situation in England and Wales, see Paul Catley and Lisa Claydon, 'The Use of Neuroscientific Evidence in the Courtroom by Those Accused of Criminal Offenses in England and Wales (2015) 14(2) J L & Biosciences 510. For the situation in Slovenia, see Miha Hafner, 'Judging Homicide Defendants by Their Brains: An Empirical Study on the Use of Neuroscience in Homicide Trials in Slovenia (2019) 6(1) J L & Biosciences 226. For the Netherlands, see Peggy ter Vrugt, 'A Pragmatic Attitude: The Right to Silence in the Netherlands (2021) 12(3) New J Eur Crim L 389.

[50] Supreme Court of the Netherlands ECLI:NL:HR:2021:202.

[51] *Saunders v United Kingdom* [GC] 19187/91, 17 December 1996.

if it is otherwise impossible, and this is not an infringement of the privilege against self-incrimination.[52] The ECtHR has shaped certain legal standards through years of jurisprudence for the scope and limitations of the privilege against self-incrimination. However, it is notable that these standards have not always been consistent. The following sections will outline some of the more significant cases in this area and which have shaped the position of the ECtHR and as such have had a significant impact on national legislation.

Although it is not explicitly mentioned, it is held that the privilege against self-incrimination is a constitutive element of the right to a fair trial codified in Article 6 of the European Convention on Human Rights. Its purpose is to protect the defendant from the coercion of state authorities and to contribute to the realisation of the goals of Article 6.[53] Whether the statements of the defendant are coerced or not is subject to a case-by-case assessment, but the ECtHR has established some criteria for orientation. In this sense, a statement will be considered coerced if: 1) the defendant is threatened by consequences if they fail to give a statement;[54] 2) the defendant is subjected to physical or mental coercion for the purpose of soliciting the statement;[55] 3) the statement was obtained by the insertion of a notification that the law enforcement authorities were unable to collect in examination.[56]

The practice of the ECtHR shows that this privilege is not of an absolute nature because negative inferences can be derived from the defendant's silence in some instances if the circumstances are such that they clearly require their pleading.[57] In such cases, the court will have to weigh the interest of protecting the defendant and of a breakthrough in the criminal proceedings, taking into account the nature and degree of coercion faced by the defendant and the purpose of certain evidentiary material.[58] This exception to the privilege against self-incrimination will carry particular weight in jurisdictions with jury trials, because the courts there will have to give very clear and precise instructions to the jury about the possibility of drawing negative inferences from the defendant's silence.[59] However, it should be noted that, from the perspective of the national legislation of the member states of the Council of Europe, the

---

[52] ter Vrugt (n 49) 394.

[53] *Bykov v Russia* [GC] no 4378/02, 10 March 2009, § 92; *John Murray v United Kingdom* [GC] no 18731/91, 8 February 1996 § 45; *JB v Switzerland*, no 31827/96, 3 May 2011 § 64.

[54] *Saunders v United Kingdom* (n 51).

[55] *Jalloh v Germany* [GC] no 54810/00, 11 July 2006; *Gäfgen v Germany* [GC] no 22978/05, 1 June 2010.

[56] *Allan v United Kingdom*, no 48539/99, 5 November 2002.

[57] *John Murray v United Kingdom* § 47.

[58] *Jalloh v Germany* § 101; *O`Halloran and Francis v the United Kingdom* [GC] no 15809/02 and 25624/02, 29 June 2007 § 55; *Bykov v Russia* § 104; *Ibrahim and Others v United Kingdom* [GC] no 50541/08, 13 September 2016 § 269.

[59] *O'Donnell v United Kingdom* no 16667/10, 7 April 2015 § 51.

permissibility of drawing negative inferences from the defendant's silence is possible only in the United Kingdom. This is not possible under the laws of most other countries of continental Europe, which relativises the above-mentioned standard of the ECtHR.[60]

Similarly, the privilege does not apply to evidentiary material which is obtained by force, but is in essence independent from the consent of the defendant (such as fingerprints, blood samples, urine samples, breathe, etc, for DNA testing).[61] There are scholarly debates in this regard on whether, in the European legal context, information collected through modern neuro-scientific methods is a type of evidence that can legally be equated to DNA samples and thus be taken without the defendant's consent. Specifically, could the results of tests such as fMRI or brain fingerprinting be used as evidence against the defendant, despite their objection and despite forceful collection? It should be noted that the situation cannot be fully equated with the collection of fingerprints or DNA samples, because the types of neuro-scientific testing that are analysed here require the collaboration of the examinee, and the results can be obstructed (or prevented) by resistance.[62] This is why, in this context, coercion relates to mental force (ie legal coercion) which includes the prospect of negative legal consequences for the refusal to cooperate or the promise of different legal benefits in exchange for cooperation.

Some authors claim that this possibility is realistic in the existing framework and it is thus necessary to establish a new fundamental human right to mental privacy. It is their position that the special nature of the information obtained through the reading of cerebral waves implies intrusion into the privacy of an individual, and the methods for the collection of such information require the development of a new fundamental human right. According to them, the existing law on the protection of privacy and personal data is insufficient to address the technological advances.[63] Others consider it unnecessary to introduce a new fundamental human right because sufficient protections can be drawn from Article 6 (the right to a fair trial), Article 8 (the right to the respect of private and family life) and Article 9 (the freedom of thought, conscience and religion) of the Convention.[64] The latter authors hold that there are significant parallels between cerebral waves and DNA material, which lead to the conclusion that the existing Convention protections should suffice.[65] It

[60] John D Jackson, 'Silence and Proof: Extending the Boundaries of Criminal Proceedings in the United Kingdom' (2001) 5(3) Int J of Evid and Proof 145.

[61] *Saunders v United Kingdom* § 69; *O`Halloran and Francis v the United Kingdom* § 47.

[62] Sean Kevin Thompson, 'The Legality of the Use of Psychiatric Neuroimaging in Intelligence Interrogation' (2005) 90 Cornell Law Rev 1601, 1624.

[63] Marcello Ienca and Roberto Andorno, 'Towards New Human Rights in the Age of Neuroscience and Neurotechnology (2017) 13(5) Life Sci Soc Policy 14–15.

[64] Sjors LTJ Ligthart and others, 'Forensic Brain-Reading and Mental Privacy in European Human Rights Law: Foundations and Challenges' (2021) 14(2) Neuroethics 191–203.

[65] ibid. See also Sjors LTJ Ligthart, 'Coercive Neuroimaging, Criminal Law, and Privacy: A European Perspective' (2019) 6(1) J L& Biosciences 289.

is our position that cerebral waves (or the content of thoughts) and DNA material cannot be equated, because the latter is purely physical in nature, and can be collected without the collaboration of the subject (ie by force), which is impossible for cerebral waves.

5 Reflections on the scope of application in both legal environments

Based on the previous elaborations, it can be concluded that the question of the scope and limitations of the use of modern neuro-scientific technologies in criminal proceedings is currently relevant for criminal law, despite the fact that it has not yet received its final confirmation in judicial practice. It can be assumed that the described mechanisms will be recognised and applied in the near future. This issue has captured the attention of both US and European authors, who discuss many of its controversial aspects, an analysis of which would exceed the scope of this paper.[66] Instead, our focus is placed on the issue of using modern technologies in the context of the privilege against self-incrimination, as one of the fundamental rights of defendants in criminal proceedings.

The nature of modern neuro-examination technologies is such that successful results imply a level of collaboration of the person subjected to such testing. One might even say that the results depend on the person's cooperation in such testing. Therefore, it is clear that any discussion of coercion in such proceedings (such as fixing the head or body, or the use of any tranquilising substances) is moot. We may safely say that there is a universal standard which forbids medical interventions that would influence the defendant's will to testify. The same goes for force, threat or similar means to obtain the defendant's testimony. Bearing this in mind, to apply neuro-examination technologies, the defendant's consent should be necessary.

[66] However, it should be noted that the use of AI in criminal proceedings is recognised in the EU. The European Parliament adopted the resolution of 6 October 2021 on artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters (2020/2016(INI)). The resolution: 'reaffirms that all AI solutions for law enforcement and the judiciary also need to fully respect the principles of human dignity, non-discrimination, freedom of movement, the presumption of innocence and right of defence, including the right to silence, freedom of expression and information, freedom of assembly and of association, equality before the law, the principle of equality of arms and the right to an effective remedy and a fair trial, in accordance with the Charter and the European Convention on Human Rights; stresses that use of AI applications must be prohibited when incompatible with fundamental rights'. Besides, 'any AI tools either developed or used by law enforcement or the judiciary should, as a minimum, be safe, robust, secure and fit for purpose, respect the principles of fairness, data minimisation, accountability, transparency, non-discrimination and explainability, and that their development, deployment and use should be subject to risk assessment and strict necessity and proportionality testing, where safeguards need to be proportionate to the identified risks; highlights that trust among citizens in the use of AI developed, deployed and used in the EU is conditional upon the full fulfilment of these criteria'.

A comparison of modern neuro-examination technologies with classic fingerprinting or DNA sampling is inadequate. Brain fingerprinting or fMRI is not typical physical evidence. Such methods are rather *sui generis* evidence which cannot be collected without the collaboration of the defendant. Therefore, the issue of coercion in this case is connected to the indirect coercion that comes from suffering negative consequences for rejecting the collaboration. The previous elaboration undoubtedly shows that the legal framework of England and Wales will be most receptive to such practices, because of its rule that the silence of the defendant can be used as an adverse inference. However, adverse inference in the US and in continental Europe criminal procedure is not possible. In the latter systems, the (valid) consent of the defendant will be necessary for such evidentiary methods.

Besides, in most countries of continental Europe, there will be a need for certain legislative amendments, which will provide for new methods of neuro-examination as evidentiary means in criminal proceedings. Such methods ought to be introduced in legal texts mainly in criminal proceedings laws in order to be applied in practice and used as evidence. If there is a new method of evidence gathering, continental legal systems of criminal proceedings require their provisions in the law. The same was true when DNA testing became a new scientific method in criminal proceedings. The application and interpretation of modern neuro-examination technologies require specific (medical) knowledge and skills. This, in fact, is a new type of expertise. New rules or the adjustment of existing rules regarding the findings and opinions of such expert witness testimony are therefore indispensable.

6 Conclusion

The acceleration of scientific development in recent years will inevitably lead to the implementation of new fact-finding methodologies in criminal proceedings. Criminal law and criminal proceedings must keep up with the new times, but they must also maintain satisfactory standards with regards to fundamental human rights and the right to a defence. Therefore, it is important to observe and discuss the application of such technologies in the context of the privilege against self-incrimination, as a constituent part of the right to a fair trial.

The aim of this paper was to compare the US and European interpretation of the privilege, in order to make an informed conclusion on which legal system is more receptive to the use of this technology. It can be concluded that, although they come from different vantage points, the US Supreme Court and the ECtHR are largely taking the same position on the issue, preventing any form of coerced taking of evidence. While physical coercion is excluded by the nature of neuro-examination which requires the collaboration of the defendant, legal coercion is excluded by the modern standards of the defendant's right to a fair trial. Deviation

from this standard exists in England and Wales, where the jury has discretion to draw adverse inference from the defendant's silence.

When discussing new evidentiary methods in criminal proceedings, such as the use of AI, one must start from a basic premise. In order for an evidentiary method to be acceptable in criminal proceedings, it must be reliable and credible. This criterion was not satisfied in the case of the old-fashioned lie detector and therefore such a method was not accepted as evidence in criminal procedure. This is why the use of AI technology must first of all be reliable and credible.

If the use of AI technology proves to be reliable and credible, then such a method must respect well-established human rights standards in criminal proceedings, among which is the right to a fair trial and all its components, which is a central principle of criminal procedure. This paper has shown how the use of AI as an evidential method should be viewed through one of the components of the right to a fair trial, which is the privilege against self-incrimination.

In any case, reliability depends not only on technology, but also on the interpretation of the use of such technology. Interpretation can only be given by experts' findings and opinions. The credibility of these findings and opinions is and should be assessed only by the court in a procedure in which the right to a fair trial is respected. These are sensitive thought and cognitive processes inherent in the human brain. This is why the determination of someone's guilt must not be left to a machine.

Suggested citation: M Mrčela and I Vuletić, 'Rethinking the Privilege Against Self-Incrimination in Terms of Emerging Neuro-Technology: Comparing the European and United States Perspective' (2023) 19 CYELP 207.