Essays

Exploring New Frontiers in Personalized Criminal Responsibility through Neurotechnologies: Ethical Dilemmas and the Neuro-Rights Framework

Janna Da Nóbrega Souza*

ABSTRACT: In criminal law, adjudicating responsibility relies on two foundational capacities: the comprehension of the wrongful nature of the act and the ability to behave in accordance with this comprehension. These elements are subject to interpretation, often requiring the involvement of experts, although it is not uncommon the incidence of discrepancies among them. Brain-based mind reading technologies is principally aimed to enhance objectivity by analyzing brain metrics rather than observed behavior. While promising in addressing subjectivity, concerns remain about the limited reproducibility of the data and the brain's complexities. The application of artificial intelligence offers potential solutions to these challenges; however, it concurrently introduces profound ethical and legal dilemmas.

KEYWORDS: artificial intelligence; neuroscience; criminal responsibility; criminal law; neuro-rights

SUMMARY: 1. Introduction – 2. Towards a Novel Paradigm in Personalized Criminal Responsibility – 3. Advancements in Diagnostic Accuracy of Mental Disorders through Deep Learning Approaches: Implications for Forensic Assessments – 4. Navigating Ethical and Legal Frontiers: Protecting Mental Privacy, Preserving Psychological Continuity, Securing Consent, and Countering Manipulative Measures – 5. Conclusions.

1. Introduction

The legal provisions, in their stringent nature, fail to consider personalization. Human conduct is shaped by innumerable variations in temperament, intellect and education, all of which contribute to the uniqueness of individual actions.[1] In Criminal Law, however, the provisions appear to oppose this premise in the pursuit of meticulous and unequivocal individualization. The assessment of criminal responsibility rests on two fundamental individual capacities: the comprehension of the wrongful nature of the act and the ability to behave in accordance with this comprehension.[2] In judicial practice,

* *Assistant Professor of Criminal Law, University of Jaén. Email: jda@ujaen.es.* *The article was subject to a double-blind peer review process.*

[1] D. DENNO, *Neuroscience, and the personalization of Criminal Law*, in *Symposium Personalized Law*, 2, 2019, 366.

[2] Spain, Penal Code, Organic Law, 10/1995, 23 November 1995 stipulates that "exemption from criminal responsibility applies to: a) Those who, at the time of committing the criminal offence, due to any mental anomaly or alteration, are unable to comprehend the unlawfulness of the act or to act in accordance with that understanding; b) Those who, at the time of committing the criminal offense are in a state of complete intoxication due to the consumption of alcoholic beverages, toxic substances, narcotics, psychotropic substances, or other substances with analogous effects, provided that such intoxication was not sought with the purpose of committing the offense, or if

these elements are reflected in behavioral fragments deemed relevant at the time of the offense, typically interpreted with the assistance of expert professionals. Meta-analytic research consistently demostrates that forensic examiners often reach divergent conclusions when evaluating the same case.[3] For instance, a study by Gowensmith, Murrie, and Boccaccini reported unanimous consensus in only 55.1% of the assessments.[4]

Multiple factors account for this discrepancy. They include the evaluator's idiosyncrasies and the subjectivity inherent in interpreting diagnostic criteria for mental disorders.[5] In this context, subjectivity is conceived as that which emanates or originates from the mind itself, not as a product of external influences, but rather as an intrinsic attribute of the thinking subject, exerting precedence over the object of thought.[6] In instances where they arise, the occurrence of cognitive biases exerts an impact on the evaluation process by generating divergent decision thresholds.[7] In connection therewith, disparities in expertise have been noted, often associated with the racial and/or ethnic origin of the accused, as well as the political ideology and antecedent professional training of those entrusted with their assessment.[8]

The emergence of brain-reading seeks to enhance diagnostic objectivity and to detect feigned manifestations of psychopathology. Within this framework, objectivity refers to findings free from personal emotions, biased interpretations, or prejudices, and grounded in verifiable facts and empirical data.[9] Consequently, this technological development replaces reliance on behavioral observation with precise measurements of cerebral structure and/or function, thereby supporting a more impartial evaluation, free of subjective interference.[10] These components hold considerable significance and require careful consideration by the presiding Judge or the Court when determining criminal responsibility.[11] Human moral and

---

the commission of the offense was not foreseen or ought not to have been foreseen, or if they are under the influence of withdrawal syndrome due to dependence on such substances, which impedes their ability to comprehend the unlawfulness of the act or to act in accordance with that understanding; c) Those who, due alterations in perceptions from birth or childhood, have severely impaired awareness of reality (…)".

[3] L. Guarnera, D. Murrie, *Field reliability of competence and sanity options: a systematic review and meta-analysis*, in *Psychological Assessment*, 6, 2017, 795-818; L. Kacperska, J. Heitzman, *Reliability of repeated forensic evaluations of legal sanity*, in *International Journal of Law and Psychiatry*, 44, 2016, 24-49.

[4] W. Gowensmith, D. Murrie, M. Boccaccini, *How reliable are forensic evaluations of legal sanity?*, in *Legal and Human Behavior*, 2, 2013, 98-106.

[5] K. Mccallum, W. Gowensmith, *Tipping the scales of justice: the role of forensic evaluations in the criminalization of mental illness*, in *CNS Spectrums*, 2, 2020, e.g. 154-160; L. Curley, J. Munro, M. Lages, R. Mclean, J. Murray, *Assessing Cognitive Bias in Forensic Decisions: A review and Outlook*, in *Journal of Forensic Sciences*, 2, 2020, 354-360.

[6] S. Mohinder, M.C. Joshi, *Exploring subjectivity, objectivity, truth and reality in the context of forensic document examination: review and commentary*, in *International Journal of Forensic Science and Legal Medicine*, 1, 2021, 21-32.

[7] G. Parmigiani, G. Mandarelli, G. Meynen, S. Ferracuti, *Structured instruments for insanity defense evaluations: Opportunities and limitations*, in *Behavioral Sciences and the Law*, 41, 2023, 1-13.

[8] J. Hicks, Ethnicity, *Race, and Forensic Psychiatry: are we color-blind?*, in *The Journal of the American Academy of Psychiatry and the Law*, 1, 2004, 21-33; T. Neal, *Discerning bias in forensic psychological reports in insanity cases*, in *Behavioral Sciences and the Law*, 3, 2018, 338; D. Freedman, S. Zaami, *Neuroscience and mental state issues in forensic assessment*, in *International Journal of Law and Psychiatry*, 65, 2019, 2-3.

[9] S. Mohinder, M.C. Joshi, *op.cit.*, 21-32.

[10] G. Mecacci, P. Haselager, *identifying criteria for the evaluation of the implications of Brain reading for Mental Privacy*, in *Science and Engineering Ethics*, 25, 2019, 446-449.

[11] G. Meynen, *Neuroscience-based Psychiatric assessments of Criminal Responsibility: Beyond Self-report?*, in *Cambridge Quarterly of Healthcare Ethics*, 3, 2020, 449-450.

legal judgements are inherently directed at decisions and actions. An in-depth understanding of the neurological mechanisms that underlie decision-making and the execution of human actions therefore has profound influence on such assessments.[12] In this regard, the observations of Da Nóbrega on the intrinsic interconnection between Neuroscience, Law, and Human behavior acquire particular relevance. If Legal norms are concerned with, and behavior emerges primarily from perception, information processing, emotional responses, deliberations, decisions, and other dynamically interacting states, then all theories of human conduct ultimately relate to the brain. For this reason, as research into these mechanisms advance, there will be a growing need to balance neuroscientific explanations those that carefully weigh the nature and severity of punishment to be applied.[13]

Paterson and Pardo highlight the valuable contribution that neuroscientific methods could offer in matters related to criminal responsibility, considering them one of the most feasible approaches through which neuroscience could enhance the legal field.[14] Gkotsi, Gasser, and Moulin emphasize this relevance emphatically, arguing that such techniques enable an objective evaluation of classical psychiatric conditions alleged affecting and potentially influencing their behavior.[15] Nadelhoffer and Sinnott-Armstrong endorse this perspective, contending that a comprehensive analysis of behavioral, neuropsychological, and/or neuroimaging evidence allows the construction of a detailed profile of brain dysfunction and its relation to the behavior under scrutiny.[16]

While the doctrinal perspectives previously outlined may endorse the integration of this novel paradigm, significant objections from prominent scholars remain. In this regard, Morse argues that the legal framework is grounded on a concept of personhood that transcends the confines of metaphysical free will. This concept functions as a fundamental pilar in understanding human beings as conscious agents and rational deliberators capable of directing their actions intentionally. According to Morse, this notion is deeply embedded in the normative foundations of human society and appears largely impervious to advancements in neuroscience. Therefore, only in the highly improbable scenario In which neuroscience could incontrovertibly establish that no individual is capable of resisting the impulses and desires that drive them, could one entertain the prospect of a profound reconfiguration of the legal construct of responsibility.[17] Darby, Macintyre, Cockerill, Stephens, Weinstock, and Darby reinforce their position by highlighting the limited reproducibility of neuroscientific investigations. They further note that variations in the processing of the neurological tool's software, along with subsequent statistical analyses of the collected data, may undermine the reliability of results, even when replicated datasets appear similar. Lastly, they emphasize the complex organization of the brain into interconnected networks, where certain diseases or clinical

[12] F. Zico Junius, et al., *The role of neuroprediction and artificial intelligence in the future of Criminal procedure support science: A new era in Neuroscience and Criminal Justice*, in *Yuridika*, 28, 2023, 593-619.

[13] J. Da Nóbrega, *Trastorno por déficit de atención e hiperactividad (TDAH)*, in *Neurociencia y Sistema de Justicia Penal. Madrid*, 2024, 15-16.

[14] M. Pardo, D. Patterson, *Mind, Brains, and the Law: the conceptual foundations of Law and Neuroscience*. New York, 2013, 121-147.

[15] G. Gkotsi, J. Gasser, V. Moulin, *Neuroimaging in criminal trials and the role of psychiatrics expert witnesses: a case study*, in *International Journal of Law and Psychiatry*, 65, 2019, 2.

[16] T. Nadelhoffer, W. Sinnott-Armstrong, *Neurolaw and Neuroprediction: potential promises and perils*, in *Philosophy Compass*, 7, 2012, 633-634.

[17] S. Morse, *New Neuroscience, Old problems: Legal implications of Brain Science*, in *Cerebrum: The Dana Forum, in Brain Science*, 6, 2004, 81-90.

symptoms may localize more accurately within a network rather than a discrete cerebral region, thereby increasing the potential for incongruity in findings.[18]

At this point, it is pertinent to underscore the significance of artificial intelligence (AI). In neuroscience, AI enables the integration of information from diverse sources, supporting the identification and categorization of patterns that reinforce conclusions regarding psychopathological phenomena. This capability holds particular relevance for Criminal Law, as it offers valuable insights into the complex relationship between the symptomatic manifestations of specific disorders and the resulting impairment of cognitive and volitional capacities in individuals who commit criminal acts or omissions. However, it is imperative to highlight that this situation engenders ethical-legal dilemmas of transcendent significance, initially pertaining to the preservation of mental privacy. Subsequently, discussions have broadened with the intention of exploring the issues of consent and the potential susceptibility of results to manipulation through the utilization of countermeasures. Lastly, albeit not directly, matters relating to psychological continuity are also implicated. The following sections will undertake a detailed exploration of these issues. Before doing so, it is essential to clarify certain foundational concepts that underpin the subsequent discussion.

2. Towards a Novel Paradigm in Personalized Criminal Responsibility

While it might be construed as a concern arising solely from contemporary scientific developments, the human aspiration to access the mental contents of others—that is, to engage in mind-reading—constitutes an ancient phenomenon, profoundly entrenched in the very foundations of social interaction. From time immemorial, humans exhibited a natural tendency to infer the beliefs, desires, intentions, and affective states of others. Initially, this interpretative endeavor relied on empirical and intuitive observation of behavior. Over time, it evolved through a complex trajectory towards systematized methodologies, culminating in recent decades with the sophisticated application of neurotechnology.

Within this trajectory, brain-reading methodologies have begun to assume the character of profoundly disruptive instruments, meticulously designed to unveil the internal architecture of mental states. These techniques identify and decode of neurobiological markers, allowing access to cognitive representations with remarkable specificity.[19] Haynes characterizes this procedural framework as a sophisticated modality for investigating the morphology and functional dynamics of cerebral activity, aimed at achieving deep insight into the psychological processes underlying human behavior.[20] Meynen emphasizes its potential relevance in criminal law, noting that these approaches would not only the anticipation of behavioral tendencies, but also their retrospective reconstruction based on empirical data derived from rigorous experimental studies of neurocognitive functioning.[21] This convergence has given rise to what may be termed an emerging neurolegal paradigm: an epistemic-normative framework intended to critically

[18] C. Darby, M. Macintyre, R. Cockerill, D. Stephens, R. Weinstock, R. Darby, *Ethics and Legal implications of emerging neuroscience technologies used for forensic purposes*, in L. Weiss Roberts (ed.), *Ethics and Clinical Neuroinnovation: Fundamentals, stakeholders, case studies, and emerging studies*, Cham, 2023, 178-179.

[19] G. Mecacci, P. Haselager, *op. cit.*, 446-449.

[20] J. Haynes, *Brain reading: decoding mental states from brain activity in humans*, in J. Illes, B. Sahakian (ed.), *Oxford Handbook of Neuroethics*, Oxford, 2012, 1-8.

[21] G. Meynen, *Neuroscience-based Psychiatric assessments of Criminal Responsibility: Beyond Self-report?, op. cit.*, 452.

reassess the foundational dogmatic categories of criminal law in light of insights from neurocognitive sciences. This paradigm extends beyond the simple inclusion of neurobiological evidence, advocating for a fundamental reconsideration of the ontological and epistemological foundations on which contemporary criminal theory rests.

The juridical rationality that structures and legitimizes the normative framework of criminal provisions is founded upon a particular anthropological conception of human behavior—one that is deeply anchored in fundamental categories such as volition, deliberation, freedom, and self-determination. These conceptual foundations construct an idealized image of the acting subject as an autonomous, rational, and free agent, fully capable of directing their conduct in accordance with normatively internalized imperatives. Similarly—though not without a degree of conceptual anachronism—judicial practice and traditional dogmatic formulations operate on comparable principles, reproducing a vision of the human being that presupposes coherent and self-regulating subjectivity. This subjectivity is assumed to select among courses of action with full awareness, deliberative capacity, and self-governance. If, however, it is accepted that the object of criminal law is human conduct—and that such conduct does not emerge *ex nihilo*, but rather constitutes the ultimate expression of a complex network of internal processes, including perception, information encoding, emotional experience, deliberation, decision-making, and other cognitive dynamics that interact in a reciprocal and intricately organized manner within the neural apparatus—then it becomes imperative to acknowledge that any legal theory aspiring to provide a rigorous and comprehensive account of human action must, in the final analysis, be grounded in its cerebral substrate.[22] This recognition must not be misconstrued as a concession to biological reductionism, nor should it be interpreted as an attempt to dissolve the notion of legal responsibility into a framework of neurobiological determinism. Rather, it advocates for a reconceptualization of penal dogmatics—one that embraces a more sophisticated, epistemologically enriched, and empirically grounded comprehension of the imputable subject. Such a paradigm seeks to integrate, with methodological rigor, empirical insights from contemporary cognitive neuroscience into the architecture of criminal responsibility. Ignoring the neural dimensions that condition—and, in certain circumstances, limit—human agency is, perpetuates a conceptual artifact whose epistemic legitimacy and normative coherence are increasingly challenged by advances in neuroscience.

In an increasingly sophisticated scientific and hermeneutic landscape, it becomes neccessary to reconsider the evaluation of traditional criteria for imputability through a more refine, dynamic, and empirically substantiated lens. While the contemporary penal paradigm already integrates the individualization of culpability by assessing the cognitive and volitional capacities of the defendant, this process remains predominantly reliant on clinical and psychodiagnostics tools, which, despite widespread use, remain partly subjective. Here, the concept of personalized criminal responsibility assumes particular relevance. It constitutes a framework that acknowledges and structurally incorporates the neurobiological idiosyncrasies of the acting subject in the determination of their imputability.[23] This approach allows for a precise understanding of the internal conditions of the agent and adapts of the penal judgment to their distinctive attributes, including the degree of neurocognitive maturation, intellectual capacity, or the ability for behavioral inhibition, the latter reflecting a tangible manifestation of self-regulation. Contemporary

[22] J. DA NÓBREGA, *Trastorno por déficit de atención e hiperactividad (TDAH)*, *op cit*., 15-16.
[23] D. DENNO, *op.cit*, 366.

neuroscientific research demonstrates that these characteristics have measurable correlates in brain morphology and functional connectivity.[24]

Regarding general intelligence (GI), the study conducted by Malpas, Genc, Saling, Velakoulis, Desmond, and O'Brien, employed diffusion tensor imaging combined with resting-state functional connectivity analysis, providing compelling evidence that GI significantly correlates with white matter organization across multiple cerebral regions. Moreover, variability in intelligence levels was associated with an extensive connectivity network spanning all cortical lobes, allowing for inferences regarding the underlying neural patterns that support superior cognitive abilities.[25] In terms of self-control, Krämer and Gruber used functional magnetic resonance imaging (fMRI) to examine the dynamic interactions among the amygdala, the *nucleus accumbens*, and the prefrontal cortex—regions central to understanding how emotions, desires, and rationality influence human behavior. They foung that increased connectivity facilitated approach behavior toward immediate rewards when emotional stimuli were presented. Furthermore, enhanced connections between the anteroventral prefrontal cortex, the amygdala and the nucleus accumbens were associated with reason-based decision-making in dilemmas. These findings corroborate prior studies, demonstrating that emotional signals from the amygdala and goal-directed information from the prefrontal cortex converge within the nucleus accumbens, thereby guiding decisions and reward-seeking behavior.[26]

These results do more than empirically corroborate prior work, they substantially advance scientific understanding of human behavior and delineate a legitimate epistemic and normative framework for integrating novel dimensions into the criminal law. The ethical and legal access to neurobiological markers can facilitate the development of a paradigm of personalized criminal responsibility—conceived not as a discretionary privilege or anomalous exception, but as a normative imperative anchored in a more equitable, accurate, and humanely calibrated conception of punitive justice. To disregard such neurobiological insights under the pretext of upholding an egalitarian abstraction may, paradoxically, give rise to a renewed manifestation of structural injustice.

Identifying and verifying neurobiological markers relevant to individual decision-making is particularly significant given the central role of response inhibition in behavior regulation and adaptation to dynamic environments. This executive function underlies the ability to delay immediate gratification in favor of long-term goals. Volitional control over one's behavior is essential for modulating impulsive tendencies and affective responses, both of with are cosely tied to the adherence to social norms. Individuals with reduced self-control demonstrate a heightened propensity to engage in criminal or deviant behavior due to increased reactivity to proximal stimuli and diminished capacity to anticipate or evaluate negative consequences.[27] With respect to the diminution of intelligence levels, it typically exerts an adverse influence

[24] G. Mecacci, P. Haselager, *op. cit.*, 446-449

[25] C. Malpas, S. Genc, M. Saling, D. Velakoulis, P. Desmond, T. O' Brien, *MRI correlates of general intelligence in neurotypical adults*, in *Journal of Clinical Neuroscience*, 24, 2016, 128-134.

[26] B. Krämer, O. Gruber, *Dynamic Amygdala influences on fronto-striatal brain mechanisms involved in self-control of impulse desires*, in *Neuropsychobiology*, 1, 2015, 37-45.

[27] J. Da Nobrega, *El papel de las disfunciones ejecutivas en el nuevo modelo de TDAH y consecuentes implicaciones en la delincuencia y responsabilidad penal*, in D. Morillas (ed.), *La responsabilidad penal de las personas con trastorno por déficit de atención e hiperactividad. Madrid*, 2019, 51.

on academic performance—an effect that has been empirically correlated with an elevated risk of engaging in criminal trajectories, whether initially or repeatedly.[28]

In light with these considerations, artificial intelligence (AI) emerges as a transformative tool for reconceptualizing neuroscientific understanding of mental states potentially linked to criminal conduct. Accurate delineation of an individual's bio-psycho-social profile can be significantly enhanced through the advanced deep learning methodologies. Deep learning, a specialized branch of artificial intelligence, aspires to replicate the operational dynamics of neural networks in the human brain, with the objective of facilitating the automated acquisition of knowledge and enabling the sophisticated processing and comprehension of data across diverse formats.[29] Within this framework, it is customary to employ two extensively validated methodological approaches in pursuit of the principal objective. The first entails a supervised learning paradigm, which relies on pre-classified datasets with the aim of precisely forecasting known outcomes (e.g., diagnostic confirmation). The second pertains to an unsupervised learning model, primarily utilized to discern patterns, trends, or the progression of a specific pathology in the absence of pre-labeled data. Concerning the data acquisition process—be it qualitative or quantitative—it encompasses the sequential phases of collection, exploration, and refinement, each of which is indispensable for the proper training of the system and the subsequent development of the algorithm required to execute a designated task. These stages are crucial for standardizing and enhancing the quality of input data for AI systems.[30]

As neural networks structures are further analyzed, they gain the capacity to learn and represent increasingly abstract and complex features.[31] Poldrack and Farah describe the process as follows: a brain monitoring device records neurological activity corresponding to specific behaviors. Statistical methodologies and deep learning techniques they analyze this data to generate a representative map linking mental states to neural events. These associations are learned and generalized, allowing the trained algorithm to classify and decode cognitive conditions from observed neural activity.[32] Deep learning surpasses automatic learning in two key ways: it eliminates subjectivity in feature selection, particularly when attribute dimensionality is high or prior knowledge is limited, and it can model highly complex information, making it an especially effective in detecting pathology through neurobiological markers.[33] Thus, deep learning represents not only a technical evolution in AI architecture but also an epistemological shift in accessing, decoding, and representing human psychological states. Its ability to detect complex patterns in extensive

[28] E. Von Wirth, et al., *Prediction of Educational Attainment and Occupational Functioning in young adults with a childhood diagnosis of ADHD: results from the Cologne adaptative multimodal treatment*, in *Journal of Attentional Disorders*, 7, 2022, 1018-1032; R. FABIO, et al., *Static and Dynamic Assessments of intelligence in ADHD subtypes*, in *Frontiers in Psychology*, 25, 2022, 2.

[29] T. Olugbemi, C. Adetunji, O. Adeyomove, A. Dare, M. Adeniyi, A. Enoch, *Clinical applications of Deep Learning in neurology and its enhancements with future directions*, in A. Abraham, S. Dash, S. Kumar Pani, L. García- Hernandez (ed.), *Artificial intelligence for Neurological Disorders*, London, 2023, 148-149; Y. Lecun, Y. Bengio, G. Hinton, *Deep Learning*, in *Nature*, 521, 2015, 436-444.

[30] T. Olugbemi, C. Adetunji, O. Adeyomoye, A. Dare, M. Adeniyi, A. Enoch, *op. cit.*, 148-149.

[31] T. Ching, D. Himmelstein, B. Beaullieu-Jones, A. Kalinin, et al., *Opportunities and Obstacles for deep learning in biology and medicine*, in *Journal of the Royal Society Interface*, 15, 2018, 1-47.

[32] E. Poldrack, M. Farrah, *Progress, and challenges in probing the human brain*, in *Nature*, 526, 2015, 371-379.

[33] T. Olugbemi, C. Adetunji, O. Adeyomoye, A. Dare, M. Adeniyi, A. Enoch, *op. cit.*, 215.

neurobiological and behavioral datasets offers unprecedented opportunities for evaluating pathological mental functioning, with profound implications for criminal law.

3. Advancements in Diagnostic Accuracy of Mental Disorders through Deep Learning Approaches: Implications for Forensic Assessments

In light of the most recent advancements within the realm of neuroscience, the integration of models predicated upon artificial intelligence—particularly those underpinned by deep learning methodologies—has ushered in a fundamentally transformative epistemic paradigm in the diagnostic delineation of mental disorders. This innovation transcends a mere incremental advancement, constituting rather a Copernican revolution, the far-reaching implications of which are poised to significantly and substantively reshape the domain of forensic evaluations of criminal responsibility. Within this emergent framework, a considerable body of scholarly work has rigorously investigated the deployment of deep learning techniques as a means of distinguishing between individuals clinically diagnosed with mental disorders and those within neurotypical control groups. Specifically, a range of advanced deep learning techniques has been deployed with the aim of achieving a more accurate and precise diagnosis of schizophrenia. In the initial instance, Plis and his colleagues devised a model utilizing a deep Bayesian network, applied to a comprehensive dataset comprising 198 patients and 191 healthy individuals. The model's precision, as evaluated through 10-fold cross-validation, attained an accuracy rate of 91% (2014, p. 1-11).[34] Similarly, Pinaya et al. employed a deep belief network (DBN) to analyze features extracted from brain morphometry data, aimed at differentiating between healthy controls (N=83) and diagnosed subjects (N=143). The DBN demonstrated an accuracy rate of 73.6% within the cohort under study (2016, p. 1-9).[35] Additionally, irregularities in resting-state neural activity have been documented, further underscoring the potential of such neural markers as reliable biomarkers for clinical diagnosis.[36]

The findings extend beyond the exclusive analysis of functional connectivity within the context of schizophrenia. Deshpande et al. accomplished a 90% certainty in distinguishing individuals with ADHD from those unaffected by the disorder, with the inattentive and combined presentations being differentiated with an accuracy rate of 95%.[37] To fully leverage the functional and structural data contained within the ADHD-2000 database, Zou et al. devised a multimodal convolutional neural network (CNN) architecture that integrated features derived from both fMRI and sMRI, achieving an accuracy of 69.15%.[38] In contrast to features artificially constructed, Mao et al. devised a spatiotemporal deep learning framework referred to as 4-D CNN, which is grounded in granular computing principles. In contrast to artificially generated attributes, Mao et al. introduced a spatiotemporal deep learning methodology known as the 4-D CNN,

[34] S. Plis, et al., *Deep Learning for neuroimaging: a validation study*, in *Frontiers in Neuroscience*, 8, 2014, 1-11.
[35] W. Pinaya, et al., *Using Deep Belief Network modelling to characterize differences in brain morphometry in schizophrenia*, in *Scientific Reports*, 6, 2016, 1-9.
[36] T. Olugbemi, C. Adetunji, O. Adeyomoye, A. Dare, M. Adeniyi, A. Enoch, *op.cit*., 150.
[37] G. Deshpand, et al., *Fully connected cascade artificial neural network architecture for attention deficit hyperactivity disorder classification from functional magnetic resonance imaging data*, in *IEEE transactions on cybernetics*, 12, 2015, 2668-2679.
[38] L. Zou, et al., *3D CNN based automatic diagnosis of attention hyperactivity disorder using functional and structural MRI*, in *IEEE Access*, 5, 2017, 23626-23636.

grounded in granular computing principles. This approach was trained utilizing changes in entropy and could be employed to compute granularity at a more generalized level through the overlay of layers. Evaluative studies confirmed that this model surpassed conventional methods within this dataset, attaining an accuracy of 71.3%.[39] In a subsequent investigation, Riaz et al. proposed a framework that processes preprocessed fMRI time series as input, producing diagnostic outcomes as output, with integral training via backpropagation, demonstrating an accuracy rate of 73.1%.[40]

Far from merely representing diagnostic achievements, as evidenced by the preceding discussion, these findings bear profound legal significance, offering highly dependable neurobiological anchors pertaining to brain dysfunctions localized in cortical regions closely tied to moral judgment, behavioral planning, and self-regulation. Consequently, clinical inferences derived from quantifiable data—such as cortical volume, functional connectivity, or alterations within the default mode network—attain forensic relevance, furnishing the adjudicator with an interpretative framework that surpasses conventional expert testimony predicated on clinical interviews and psychometric scales. A paradigmatic case that exemplifies this transition toward a neuro-informed justice system is *State of Florida v. Grady Nelson*. In this case, the defense presented evidence derived from quantitative electroencephalography (QEEG), which furnished a graphical representation of substantial dysfunction in the implicated brain structures. The jury, after considering these elements and their connection to the commission of the offense, opted to replace the originally sought capital punishment with a sentence of life imprisonment. This precedent unequivocally underscores that neuroscientific evidence has the potential to significantly influence the modulation of punitive responses, by introducing objectifiable factors that enable the precise calibration of the offender's degree of self-determination.[41]

At this juncture of analysis, it becomes imperative to elucidate the neurocognitive mechanisms through which a cerebral alteration may manifest as a total or partial incapacity to reconcile behavior with the imperatives of the legal framework. Within this paradigm, self-control is conceptualized as a higher-order executive function, underpinned by intricate neural circuits that establish connections between the prefrontal cortex and various subcortical structures. Disruption within these circuits impairs the inhibition of automatic responses, the anticipation of consequences, and the moral appraisal of decisions.[42] Ultimately, this neurocognitive architecture constitutes an essential component for the exercise of ethical agency and, by extension, forms the foundational basis for the adjudication of criminal responsibility.

In light of the preceding discussion, the phenomena of apathy and temptation emerge as clinical manifestations stemming from disruptions in the motivational equilibrium, which, in turn, directly influence the agent's capacity for self-regulation.[43] Within the context of assessing potential behavioral alternatives in

[39] Z. MAO, et al., *Spatio-temporal Deep learning method for ADHD fMRI classification*, in *Information Sciences*, 2, 2019, 1-11.

[40] A. RIAZ, et al., *Deep FMRI: end-to end deep learning for functional connectivity and classification of ADHD using fMRI*, in *Journal of Neuroscience Methods*, 1, 2020, 1-14.

[41] State of Florida v. Grady Nelson, F05-00846 (October 22, 2010).

[42] C. LUCIFORA, G. MARTINO, A. CURCURUTO, M. SALEHINEJAD, C. VICARIO, *How self-control predicts moral decision making: an exploratory study on health participants*, in *International Journal of Environmental Research and Public Health*, 18, 2021, 3840; K. MARTIN KRONKE, M. WOLFF, T. GOSCHKE, *Predicting Real-Life Self-control from Brain activity encoding the value of anticipated future outcomes*, in *Psychological Science*, 31, 2020, 268-279.

[43] T. CONNOR, *Self-control, willpower and the problem of diminished motivation*, in *Philosophical Studies*, 168, 2013, 783-796.

pursuit of normative obligations, apathy is characterized by a pronounced attenuation in responsiveness to rewards, coupled with an exaggerated perception of the exertion required for their attainment. Conversely, temptation is delineated by an excessive responsiveness to immediate rewards, accompanied by a deliberate disregard for the attendant costs of such actions. This fundamental divergence in the appraisal of alternatives gives rise to profound implications for decision-making: under the influence of apathy, the agent tends to engage in protracted and unproductive deliberation, whereas in the case of temptation, there occurs a reactivation of deliberations previously resolved, with the intent to reassess and, ultimately, subvert intentions that would otherwise favor the fulfillment of legal duties.[44]

Such affective-motivational states likewise produce distinct consequences at the level of behavioral execution. Specifically, individuals exhibiting apathy display a pronounced inhibition in initiating actions aligned with normative imperatives, whereas those subject to temptation demonstrate a heightened proclivity to engage in hasty behaviors whose immediate gratification stands in stark opposition to previously formed volitional intentions. Within this framework, it becomes analytically pertinent to correlate these dysregulations with the prototypical expressions of unlawful conduct. Apathy, by impeding the activation of the motivational drive requisite for the enactment of legally mandated behaviors, is primarily associated with criminally relevant omissions. Conversely, temptation frequently precipitates overt acts wherein the individual, compelled by the immediate salience of a reward, executes conduct that had formerly been rejected in accordance with internalized normative frameworks. In both instances—whether characterized by omission due to apathetic inhibition or by impulsive transgression under the influence of temptation—the violation of legal duty emerges as the product of a dysfunction within the neurocognitive substrates underpinning self-regulatory control.[45]

In light of the foregoing considerations, the potential conferred by deep learning technologies applied to the analysis of neurofunctional data acquires particular relevance in both academic and forensic contexts.[46] These advanced computational methodologies enable the precise identification of disruptions within executive control systems and support second-order inferences regarding their intensity, chronicity, and functional impact. This significantly enriches the evaluative framework for determining criminal responsibility from an interdisciplinary standpoint. It must be emphasized, however, that the discussion on the potential contribution of neuroscience to the assessment of criminal responsibility extends beyond the exclusive use of artificial intelligence tools. Neuroscience itself provides a rich body of knowledge and perspectives capable of critically challenging the traditional foundations of culpability. With this point clarified, the convergence of both domains fosters a novel analytical paradigm emerges wherein neurobiological evidence—far from subverting the foundational principles of criminal law— becomes a scientifically instrument for calibrating culpability according to the agent's demonstrable cognitive and volitional capacities. A heightened degree of precision in identifying mental disorders carries profound

[44] J. Bermudéz, S. Berthelette, A. Anaya, G. Fernández, D. Rodríguez, *Temptation and Apathy*, in S. Amaya, D. Shoemaker, M. Vargas (Ed.), *Oxford Studies in Agency and Responsibility*, New York, 2024, 10-17.

[45] J. Bermudéz, S. Berthelette, A. Anaya, G. Fernández, D. Rodríguez, *op.cit*., 10-17.

[46] In this regard, empathy is closely associated with consistent alterations in specific regions of the medial frontal cortex and subcortical structures, leading to the hypothesis that the dysfunction of a common system-level mechanism may constitute the underlying cause of its manifestation, irrespective of the nature of the condition from which it originates. See, for example, C. Heron, C. Holroyd, J. Salamone, M. Husain, *Brain Mechanisms Underlying Apathy*, in *Cognitive Neuroscience*, 90, 2019, 302–312.

implications for elucidating the extent of criminal responsibility among affected individuals, who, as a result, may be rendered wholly or partially incapable of acting in accordance with the law.

In line with this premise, it becomes imperative to undertake a succinct yet rigorously substantiated analysis of the relationship between neurobiological alterations and their manifestation in unlawful conduct. Such an investigation is essential for establishing the necessity of precise detection by appropriately qualified professionals, whose findings may serve as the foundation for a more equitable, coherent, and contextually appropriate assessment, fully aligned with the neuropsychological conditions of the individual under examination. In this context, the forensic evaluation of criminal responsibility necessitates the comprehensive reconstruction of the individual's mental state before, during, and after the offense, accompanied by an exhaustive scrutiny of the correlation between this mental state and the actions undertaken. To this end, a diagnosis is made to ascertain the potential presence of any psychological alteration, disorder, or deficit that the accused may have experienced at the time of the criminal act, supplemented by a meticulous observation of their behavior and the context in which the incident transpired. Furthermore, it is essential to perform a thorough evaluation of the individual's capacity to stand trial, with particular emphasis placed on assessing their level of dangerousness.

With regard to schizophrenia, this condition yields profound consequences on cerebral architecture and the daily functioning. It manifests through positive symptoms, affective disturbances, and cognitive deficits that result in considerable impairment of social and occupational capacities, sustained over a minimum duration of six months.[47] The diagnostic criteria require at least two symptoms, one of which must be disillusions, hallucinations, or disorganized speech. Markedly disorganized behavior or catatonia may also occur, alongside negative symptoms such as diminished emotional expression or significant avolition.[48] Consequently, those afflicted with this condition are deemed to possess an elevated propensity for engaging in aggressive behaviors towards others. Robust evidence indicates a significantly higher risk of comitting both violent and non-violent acts comparing to neurotypical populations.[49] A meta-analysis indicated that the prevalence of violence in the cases studied was approximately five times higher than in controls groups without psychopathological conditions. Witt and colleagues identified several contributing factors, including parental deviance, childhood maltreatment, and psychoactive substance abuse.[50] With respect to the heightened use of narcotics, it stands as a significant determinant, effectively doubling the likelihood of hostile behaviors.[51] A comprehensive longitudinal study reported similar findings,

[47] Concerning symptomatology, Iqbal et al., further elucidate that it encompasses the positive psychotic domain (delusions, hallucinations, disorganized or catatonic thoughts or behavior) as well as negative symptoms, which include avolition (diminished initiation or activities), alogia (reduced speech production), anhedonia (decreased capacity to experience pleasure) and asociality (lack of interest in social interactions). The diagnostic criteria for schizophrenia represent a constellation of symptoms associated with substantial cognitive, behavioral, and emotional dysfunctions (S. IQBAL, et al., *Schizophrenia and chronic psychotic disorders: update treatment options and new trends*, in *Psychiatric Annals*, 4, 2023, 155-159).

[48] E. LUVSANNYAM, et al., *Neurobiology of Schizophrenia: A comprehensive review*, in *Cureus*, 14, 2022, 1-7.

[49] S. HODGINS, M. PIATOSA, B. SCHIFFER, *Violence among people with schizophrenia: phenotypes and neurobiology*, in *Current Topics in behavioral neurosciences*, 5, 2013, 329-368.

[50] K. WITT, R. VAN DORN, S. FAZEL, *Risk factors for violence in psychosis: systematic review and meta-regression analysis of 110 studies*, in *Plos One*, 8, 2013, 1-15.

[51] M. ERONEN, P. HAKOLA, J. TIHONEN, *Mental disorders and homicidal behavior in Finland*, in *Archives in General Psychiatry*, 53, 1996, 497-501.

confirming that rates where three times higher among affected individuals.[52] Research also support correlations with adverse childhood experiences. It is suggested that an interaction between heightened stress sensitivity, emotional dysregulation, and reactivity (e.g. irritability, anxiety, general distress, and hypervigilance), combined with insufficient coping strategies, may increase susceptibility in individuals experiencing their first psychotic episode.[53]

Conversely, limitations have been documented in both objective and subjective assessments of executive functions, such as inhibition and working memory. These limitations are also associated with this trend.[54] A similar correlation occurs with attention deficit hyperactivity disorder. ADHD presents a range of challenges that hinder the capacity to inhibit behavior, sustain self-awareness, and maintain working memory.[55] This disorder results from the interaction of genetic and environmental factors during prenatal and early postnatal development, which enhance neurobiological susceptibility. These produce subtle alterations in cerebral systems, causing deficits across multiple neuropsychological domains. Research on structural impairments shows substantial reductions in gray matter volume within the basal ganglia, specifically in the putamen, globus pallidus, and caudate. These regions are key components of the frontal-striatal-thalamocortical circuits, which are somatotopically organized and crucial for optimal executive functioning.[56] Functional analyses reveal diminished blood flow in the prefrontal cortices and in the pathways connecting these areas to the limbic system.[57] Moreover, resting states studies show disruptions in neural networks beyond the prefrontal circuit, particularly in the default mode network.[58] These factors together significantly contribute to the risk of engaging in criminal acts and omissions.[59] This assertion is supported by evidence demonstrating a twofold increase in the likelihood of arrest, more than a threefold rise in the frequency of convictions, and nearly a threefold extension in the length of incarceration during both youth and adulthood.[60]

---

[52] S. Fazel, et al., *Schizophrenia and violence: systematic review and meta-analysis*, in *Plos Medicine*, 6, 1-15.

[53] J. Pozzo, J. Walsh-Messinger, D. Antonius, *The influence of childhood trauma on aggression and violent behavior in first episode psychosis: A critical review*, in *Aggression and Violent Behavior*, 61, 2021, 1-9.

[54] E. Luvsannyam, et al., *op. cit*., 1-7.

[55] J. Da Nóbrega, *TDAH y Sistema de Justicia Penal: un estudio desde la perspectiva del Neuroderecho*, *op. cit*., 370-372.

[56] T. Nakao, J. Radua, et al., *Gray matter volumen abnormalities in ADHD: voxel-based meta-analysis exploring the effects of age and stimulant medication*, in *The American Journal of Psychiatry*, 11, 2011, 1157-1159.

[57] P. Gustafsson, G. Thernlund, et al., *Associations between cerebral blood flow measured by single photon emission computed tomography (SPECT), electro-encephalogram (EGG), behavior symptoms, cognition, and neurological soft signs in children with attention-deficit hyperactivity disorder (ADHD)*, in *Acta Paediatrica*, 7, 2000, 830-855; S. Cortese, *Functional Decoding and Meta Analytic connectivity modeling in adult attention-deficit/hyperactivity disorders*, in *Biological Psychiatry*, 12, 2016, 896-904.

[58] P. Erika, et al., *Actividad funcional cerebral en estado de reposo: redes en conexión*, in *Revista de Neurología*, 1, 2011, 3.

[59] L. Sanchez, *ADHD, and crime: what is the relationship? Can implementation of Pharmacological and Psychosocial interventions facilitate offenders' rehabilitation?*, in *Submission to the Forensic Psychiatry Medical Students Price*, 2019, 2.

[60] J. Da Nóbrega, *TDAH y Sistema de Justicia Penal: un estudio desde la perspectiva del Neuroderecho*, *op.cit*., 171.

Statistical data support these findings. Meta-analyses across 42 prisons, using international clinical assessment interviews, indicate that 25.5% of the incarcerated population meet the criteria established for ADHD.[61]

This phenomenon can be understood as resulting from diminished self-control capacity. While this concept is not new – it was incorporated in the General Theory of Crime emphasizing socialization – recent research provides deeper insight into its biological basis, identifying specific dysfunctions that impair inhibitory systems controlling behavior.[62] This perspective expands the traditional interpretive paradigm by adding a neurobiological dimension. It enhances understanding of the phenomenon and allows for a rigorously grounded, individualized assessment of mental faculty impairment in each factual context.[63]

Once a correlation between specific neurobiological markers and criminal behaviors is established, it becomes crucial to evaluate mitigating circumstances through meticulous case-by-case analysis. This evaluation must determine the nature and extent of mental impairment at the time of the offense.[64] This requirement represents a profound and contested issue in the potential integration of neurotechnological methods into criminal responsibility assessments. While current-state irregularities do not prove their presence at the time of the criminal act, objective validation through deep learning methods can increase the likelihood of stablishing such presence. Observation of distinctive neurobiological patterns may indicate continuity, reinforcing this evidence. Criminal behaviors typically evolve in accordance with a logical progression, initiating with relatively trivial and defiant actions that progressively escalate into more severe transgressions. This persistence severity of the underlying mental disorder correlates with earlier onset, higher recidivism, and a broader manifestations.[65] Finally, it is imperative not to overlook other issues that constitute the core of vigorous doctrinal debates in the context of this potential application. Among these, particular attention must be paid to concerns regarding the preservation of mental privacy, the procurement of informed consent from the subject undergoing evaluation, and the inherent risk of manipulation of results through the use of countermeasures.

4. Navigating Ethical and Legal Frontiers: Protecting Mental Privacy, Preserving Psychological Continuity, Securing Consent, and Countering Manipulative Measures

The extraction of thoughts directly from the brain, independent of behavioral manifestations, raises profound ethical and legal concerns. Neural processes constitute an essential component of individual identity. The disclosure of neurobiological markers—whether conscious or unconscious—may therefore represent an encroachment on personal autonomy over cognitive states. The notion of cognitive freedom arises in this context. It seeks to broaden the understanding of freedom of thought and to promote its

[61] S. YOUNG, G. GUDJONSSON, P. CHITABESAN, *Identification and treatment of offender with attention/deficit hyperactivity disorder in the prison population: a practical approach based upon expert consensus*, in *BCM Psychiatry*, 18, 2018, 1-16.

[62] M. GOTTFREDSON, T. HIRSCHI, *A General Theory of Crime. Stanford*, 1990, 1-297; N. WAGER, *Psychobiology and Crime: ADHD*, in *Criminal Justice Matters*, 51, 2005, 21.

[63] N. VENABLES, et al., *Integrating criminological and mental health perspectives on low self-control: a multi-domain analysis*, in *Journal of Criminal Justice*, 56, 2018, 2.

[64] J. DA NÓBREGA, *TDAH y Sistema de Justicia Penal: un estudio desde la perspectiva del Neuroderecho*, *op. cit.*, 372.

[65] *Ivi*, 372-373.

recognition within legal frameworks. Cognitive freedom comprises three essential dimensions: (a) the autonomy to alter one's thoughts or to change one's mind, together with the means through which such change may occur; (b) the safeguarding of individuals from interventions that could undermine mental integrity; and (c) the ethical and legal duty to actively foster and protect this freedom. These elements collectively define cognitive freedom as a right that embraces both positive and negative liberties. In its negative sense, it signifies the capacity to govern one's own cognitive domain without interference from governmental, corporate, or criminal entities. It also includes the right to preserve mental integrity against such incursions. Conversely, in its affirmative dimension, it denotes the authority to exercise complete control over the privacy of one's mental processes.[66] With regard to this latter aspect, it protects individuals from unauthorized intrusion of third parties into their brain data, preventing the collection and non-consensual dissemination of such information.[67] This guarantee is increasingly importante in a technological environment where traditional anonymization mechanisms are eroding. The weakening of such protections magnifies the risk of misuse and may expose neurobiological correlates of diseases, disorders, or general mental states. These exposures, when not mediated by explicit consent, produce serious ethical and legal consequences.[68] In this regard, the notion of cognitive biometric data, as delineated on page 5 of the *Initial Draft of the UNESCO Ethics of Neurotechnology standard*, offers an expanded theoretical framework that is essential for the comprehensive understanding of the intricacies and sensitivities inherent in these risks. This framework assumes particular importance, given that data pertaining to mental and cerebral states possess the potential to divulge profoundly intimate information.[69] From a doctrinal standpoint, it has been posited that the protection of such data may find implicit grounding in Article 8(1) of the *European Convention on Human Rights* (ECHR), which enshrines the right to privacy. In consonance with this, the Spanish Constitution recognizes personal intimacy as a protected sphere.

According to Ligthart, Douglas, Bublitz, Kooijmans, and Meynen, privacy should not be understood as an exhaustive concept. The authors draw an analogy between forensic neuroimaging to other investigative methods, emphasizing that the data obtained are intrinsically connected to individual's biological characteristics. DNA and fingerprints illustrate this connection, as they contain unique information tied to an identified or identifiable individual. They therefore qualify as personal data subject to legal protection. The same reasoning applies to data acquired through AI-assisted neuroimaging techniques, given that the brain's anatomy and connectivity of each person are unique relative to one another. However, two significant distinctions should be noted: first, these techniques are not primarily designed for identification. Second, they have the unique capacity to revel mental states, thereby probing the deeper dimensions of human subjectivity. It is precisely at this point that the parallels reach their limits.[70]

---

[66] P. Sommaggio, M. Mazzocca, A. Gerola, F. Ferro, *Cognitive Liberty. A first step towards a human neuro-rights declaration*, in *Biolaw Journal*, 3, 2017, 32.

[67] E. Gulyaeva, F. Farinella, *Human Neuro-rigths*, in *Revista Quaestio Iuris*, 15, 2022, 289-290; S. Ligthart, M. Ienca, G. Meynen, F. Molnar-Gabor, R. Andorno, C. Bublitz, P. Catley, L. Claydon, T. Douglas, N. Farahany, *Minding rights: Mapping ethical and Legal foundations of Neurorights*, in *Cambridge Quarterly of Healthcare Ethics*, 32, 2023, 461-481.

[68] P. Magee, M. Ienca, N. Farahany, *Beyond neural data: Cognitive biometrics and mental privacy*, in *Neuron*, 112, 2024, 3017-3028.

[69] UNESCO, First Draft of the Recommendation on the Ethics of Neurotechnology, 2024, 8-9.

[70] S. Ligthart, T. Douglas, C. Bublitz, T. Kooijmans, G. Meynen, *Forensic Brain-reading and Mental Privacy in European Human Rights Law: Foundations and Challenges*, in *Neuroethics*, 14, 2021, 191-203.

Of utmost significance is the issue concerning whether such an infringement may be deemed justifiable. In accordance with Article 8(2) of the *European Convention on Human Rights*, the collection of personal data through DNA testing and fingerprinting may be regarded as justifiable only if it meets the following criteria: 1) it is based on a legally foreseeable and accessible framework; 2) it serves a legitimate interest, and 3) it is both necessary and proportionate to achieve the intented objective, namely, the legitimate interest. The initial requirements do not necessarily impose a limitation for forensic neuroreading, provided that its application is governed by robust legislation, under which the information obtained may be utilized for the legitimate benefit of national security, the detection and prevention of crime, or the protection of rights and freedoms of third parties (e.g. contributing to the assessment of criminal responsibility). Concerning the evaluation of necessity and proportionality, the recording, retention, and utilization of neuroscientific data would be contingent upon the severity of the offense. The more profound the impairment, the more substantial the objective to be pursued must be. In this context, another important factor to consider would be the sensitivity of this data with respect to privacy. The determination of whether an individual suffers from a specific psychiatric disorder may reveal a wide array of supplementary information not confined to the purpose of investigation, thereby reinforcing the imperative for appropriate regulation.[71]

The acquisition of data without prior consent is not a recent issue, it has been the subject of scholarly debate since the latter part of the 19th century.[72] Consent plays a decisive role in safeguarding the autonomy of forensic subjects. Within the scope of this research, it offers protection against unauthorized and detrimental manipulations. For consent to be both effective and valid, the individual must understand the specific purpose for which it is granted and remain free from undue influence or coercion.[73] This requirement has direct practical implications. Certain applications of functional neuroimaging demand active engagement by the subject. Coercive methods are therefore unlikely to succeed in ineffective ensuring cooperation during tasks or in securing full attention to specific stimuli.[74]

In contemporary contexts, individuals may also attempt to compromise the reliability of measurements through countermeasures. These are systematically designed physical or cognitive strategies intended to alter or distort the validity of results. Empirical studies have shown that even minor movements, such as finger tapping, may reduce the accuracy of examinations by as much as 33%. This risk increases when subjects have a strong interest in influencing the outcomes.[75] Among the factors that exacerbate this vulnerability, delays in presentation are especially relevant, as they give participants time to deploy countermeasures. Same of these actions are visible to examiner, making them detectable. Others, such as mental

[71] *ivi*, 191-203.

[72] G. MECACCI, P. HASELAGER, *op. cit*., 454-455.

[73] C. DARBY, M. MACINTYRE, R. COCKERILL, D. STEPHENS, R. WEINSTOCK, R. DARBY, *op. cit*., 187-188.

[74] G. MEYNEN, *Brain-based mind reading in forensic psychiatry: exploring possibilities and perils, in Journal of Law and the Biosciences*, 2, 2017, 311-329.

[75] In this context, Kanwisher and Yovel asserts that data derived from functional magnetic resonance imaging becomes invalid if the individual displaces more than a few millimeters. This challenge is not confined merely to head or finger movements: even seemingly inconsequential actions, such as sifting the tongue within the mouth or closing the eyes, can profoundly undermine the integrity of the imaging data. In N. KANWISHER, G. YOVEL, *The fusiform face area: a cortical region specified for the perception of faces*, in *Philosophical Transactions of The Royal Society B – Biological sciences*, 361, 2109-2128.

arithmetic, are more difficult to identify.[76] While recent literature indicates that technological advances have reduced these risks, they remain significant obstacles to the generalization of findings real-world applications. The technique employed must therefore be either resistant to such tactics or allow their detection, so that evaluators can determine when results are distorted.[77]

Notwithstanding the fact that it is not directly, but rather indirectly associated with psychopathological identification in the adjudication of criminal responsibility, the issue regarding the potential disruption of psychological continuity resulting from the prospective application of neurotechnological tools in the rehabilitation of individuals – should they be declared exempt from criminal responsibility or granted a sentence reduction due to the presence of such an impairment – warrants considerable emphasis.[78] This safeguard serves as a critical measure for the preservation of mental integrity by shielding against unauthorized external alteration, thereby ensuring the stability of recurrent thoughts, preferences, and choices through the protection of the underlying neural processes.[79] The prolonged absence of such experimentation hinders a full self-understanding of our nature, given that we conceive ourselves as personal entities and agents capable of generating coherent and consistent attitudes. From this premise emerges a discernible correlation with the formation of personal identity, which is conceived as a process encompassing a constellation of attributes and characteristics that facilitate individualization within the social domain.[80] These qualities are expressed through both biological and cultural elements, and, according to Sessasego, are integrated into physical and cognitive integrity along with a sense of self-awareness, that is, the capacity to make autonomous choices regarding one's actions.[81]

The right to personal identity may be extrapolated from the provisions set forth in Articles 22 and 29 of the *Universal Declaration of Human Rights*, which pertain to the development of personality in a manner that is both free and comprehensive. Likewise, Article 10.1 of the *Spanish Constitution* refers to this inherent freedom, recognizing it as a fundamental right. The question arises as to whether such references are sufficiently effective in addressing the potential threats to psychological continuity, as this safeguard involves not only the preservation of external identity but also the protection of a prior dimension: the mental processes that form the foundation upon which self-comprehension and the autonomous exercise

[76] H. Bowman, et al., *Countering Measures: detecting identity lies by detecting conscious breakthrough*, in *Plos One*, 3, 2014, 1.

[77] G. Meynen, *Brain-based mind reading in forensic psychiatry: exploring possibilities and perils*, *op. cit.*, 320.

[78] In relation to the concept of exemption from criminal responsibility, article 101 of the Spanish Penal Code prescribes the following: A person who is declared exempt from criminal responsibility pursuant to paragraph 1 of Article 20 may, if deemed necessary, be subject to internment for medical treatment or special education in a facility appropriate to the type of mental anomaly or disorder identified, or any other measures delineated in paragraph 3 of Article 96. Such internment shall not exceed the duration of the custodial sentence that would have been imposed had the individual been adjudicated responsible; the Judge or Tribunal shall establish this maximum duration in the sentencing decision. Regarding scenarios of diminished responsibility, Article 104 of the Spanish Penal Code allows for the imposition, in addition to the corresponding penalty, of measures specified in Articles 101, 102, and 103. Nonetheless, the measure of internment is only applicable when the penalty imposed is custodial, and its duration shall not exceed the term prescribed by the Code for the specific offense.

[79] M. Ienca, R. Adorno, *Towards new human rights in the age of neuroscience and neurotechnology*, in *Life Sciences, Society and Police*, 13, 2017, 21.

[80] P. Tiedemann, *Identity and human rights. Considerations on a human right to identity right to identity*, in *Stuttgart*, 2016, 30.

[81] C. Fernández Sessasego, *El Derecho a la identidad personal*, in *Comparazione e Diritto Civile*, 1992, 1-46.

of consciousness are constructed. While the issue remains unsolved, current doctrinal perspectives suggest that such a determination may be inadequate, as psychological stability, by its nature, encompasses emergent situations that do not necessarily require the occurrence of tangible harm. Ienca and Adorno argue that, to guarantee this level of protection, a new right is required – one that ensures the preservation of mental processes against abusive exogenous disturbances[82].In close accordance with the preceding line of reasoning, and recognizing the structural limitations of current legal frameworks, there is a growing imperative to construct a normative architecture that reflects the neuro-specificity of the human condition. A doctrinal and political consensus is emerging in favor of new legal safeguards to counter the risks created by rapid advances in brain and mind technologies. The Chilean experience stands out as particularly paradigmatic, as its legislative initiative has culminated in the inclusion of Article 19.1 within its constitutional framework, thereby establishing an unprecedented legal precedent in the safeguarding of the neurobiological integrity and self-determination of the individual. This development not only attests to the legal plausibility of reconfiguring the paradigm of fundamental rights, but also highlights the urgent necessity of reimagining contemporary legal structures through a transdisciplinary lens, one that integrates both the progressive advancements of neuroscience and the ethical imperatives concomitant with such innovations.[83]

Uncertainties remain, and these currently overshadow the potential benefits. Yet the primary objective must not be overlooked: to provide courts with effective tools for determining criminal responsibility. Article 5.1 (d) of the *Artificial Intelligence Act* envisions this scenario by providing that such technologies may be utilized to augment human assessments of the subject's involvement in criminal activities, provided that these evaluations are grounded in directly verifiable objective facts directly related to the criminal conduct in question.[84] This matter assumes particular significance in criminal judicial proceedings, where fundamental liberties are exposed to heightened risk and the accuracy of evidence carries utmost importance.

On the other hand, the accurate identification of psychopathological conditions may require the use of neurotherapeutic tools for rehabilitative purposes. This opens a range of possibilities that, while promising, also pose significant challenges. The foregoing shows that the uncertainties explored in this section extend far beyond the issues already described. Additional concerns of equal relevance arise in practice. Access to neuroscientific assessments remains severely limited. The evaluations demand highly trained professionals. The process involves conducting the scans, interpreting the data, and applying AI-based methods when required. Costs are substantial and prohibitive. These practical and technical constraints produce, as noted, give rise to serious consequences. The legal implications associated with these

[82] M. Ienca, R. Adorno, *op. cit*., 21.

[83] Republic Of Chile, *Political Constitution of the Republic of Chile* (1980), as amended by Law No.21383, Official Gazette, 25 de October de 2021.

[84] The European Parliament and The Council, Regulation (EU) 2024/1689, 13 June 2024, article 5.1 (d): The placing on the market, the putting into service for this specific purpose, or the use of the AI system for making risk assessments of natural persons in order to assess or predict the risk of a natural person committing a criminal offence, based solely on the profiling of a natural person or on assessing their personality traits and characteristics; this prohibition shall not apply to AI systems used to support the human assessment of the involvement of a person in a criminal activity, which is already based on objective and verifiable facts directly linked to a criminal activity.

dilemmas are intricate and far-reaching. They deserve careful and detailed examination in future work, where the full scope of challenges and opportunities can be explored.

5. Conclusions

In conclusion, neuroimaging technologies provide an unprecedented opportunity to explore brain activity, enabling the extraction of a wide range of information that, until recently, remained beyond our understanding. These tools, with their considerable potential, represent an invaluable resource for forensic assessments of criminal responsibility. However, the risk of inconsistencies in the findings highlights the urgent need to refine methods and improve precision to ensure reliability and coherence. At this stage, it is essential to emphasize the role of Artificial Intelligence as fundamental pillar in this process. AI adds the integration of information by identifying and classifying patterns that reinforce conclusions about psychopathological phenomena. Its implementation may reduce subjectivity in diagnoses and improve detection of psychopathological simulations. Nevertheless, it rases significant ethical and legal dilemmas. These include the protection of mental privacy, the informed consent of the individual being assessed, and the potential manipulation of results through countermeasures; furthermore, they are indirectly related to psychological continuity.

Conducting supplementary research is crucial to advance the standardization, understanding, and accuracy of these technologies. Research also supports the development of neurolegislation. These processes are crucial for practical application and underscore the importance of close collaboration between Neuroscience and Criminal law to address the concerns empirically in this field.