Maastricht University Leading in Learning!

Neurolaw

Edition 2014 | Volume V

MaRBLe Research Papers

Neurolaw

Maastricht Research Based Learning Project (MaRBLe)
Maastricht University, 2014
Faculty of Law

Authors:
Nina Koivula
Nina Ferreira
Petar Lozev
Franziska Böhlke
Birgit Thun
Janika Bockmeyer
Jan Smits

ISBN: 978–90–5681–427–4
NUR: 820

Preface

Jan M. Smits

In the spring semester of 2013 six students of Maastricht University participated in a course on Neuroscience and Law. This course was taught as a so-called MARBLE-project. Maastricht Research Based Learning-projects are part of the Dutch government's SIRIUS Program that aims to promote excellence in Dutch higher education, building upon the self-evident idea that talented students need to be challenged and encouraged to do their very best. At Maastricht University, MARBLE-projects funded through SIRIUS are open for third year bachelor students who want to engage in a research project and who belong to the top 20% of their class. The course on Neuroscience and Law attracted students from both the Faculty of Law and the Faculty of Arts and Social Sciences who worked intensively together during four months. The course started in February 2013 and ended on 30 May 2013 with a conference organised by the students.

Neuroscience and law (in brief 'Neurolaw') is one of the most exciting recent developments at the intersection of law and science. It is a rapidly emerging field that reveals all kinds of details about how the human brain works. The aim of the course was to explore how this new scientific knowledge can or should affect the law. This influence is potentially important. The main reason for this is that law is full of presumptions about how and why people act. These presumptions are increasingly questioned by neuroscientists, giving rise to what some have called a 'neuro-revolution' in our thinking about law. However, it is far from clear what the *exact* impact of neuro-scientific insights has to be. The main aim of the course was to explore this impact. This led us to explore a wide range of different issues:

a. a mapping exercise of the potential impact of neuroscience on law. This mapping exercise was the main activity during the first seven weeks of the project.
b. an exploration of a number of more specific topics. Students selected these topics during the first seven weeks and then wrote an individual paper during the second part of the project.

The activities that we undertook during the course are set out in the Introduction to this volume. The main part of this volume consists of the six papers that the students wrote. These papers deserve a broad audience. They are of high quality and testify of the enthusiasm and skills of their authors. All in all, the volume is evidence of the excellent research and editing skills of the six participating students.

Table of Contents

Introduction

This essay collection marks the conclusion of the MARBLE (Maastricht University Research Based Learning) for Excellence project on Neurolaw in 2013. Its objective was to address some of the complexities arising from the intersection of law and science. In particular, our attention was focused on the possible contributions that neuroimaging and neuroscience in general could yield to the field of law. The consequent research addressed the issue from a variety of perspectives – legal philosophy, criminal procedure, as well as human rights and rehabilitation.

This opportunity was generously funded by the Dutch Ministry Of Education, Culture and Science through its Sirius programme and supported by the Hague Institute on the Internationalisation of Law.

During the project we had the chance to visit the WODC (Wetenschappelijk Onderzoek – en Documentatiecentrum) under the Ministry of Security and Justice in The Hague. The appointment, led by Professor Frans L Leeuw, made clear that the 'real-world' inclusion of neuroscience was not the kind of distant scenario some of the academics in the field seemed to envisage – on the contrary, it was already taking place. The visit certainly provided additional impetus and motivation for the research process.

However, perhaps the most eye-opening experience of this project took place no further than the Faculty of Psychology and Neuroscience of Maastricht University. A discussion led by Professor Rainer Goebel illuminated the cutting-edge research into the human brain and behaviour that is taking place across the bridge. The Brains Unlimited scanner lab has at its disposal exceptional know-how and technology, and perhaps these could one day even be harnessed for inter-faculty neurolaw research.

A third event of importance was the Student Conference on 30th of May, which gave us the possibility to present our research to others and receive feedback and tips from accomplished experts. In this context, we would like to thank all the experts and guest lecturers for their insight and patience; Dr. Gary Low, Dr. David Roef, Drs. Liesbeth Vink, Prof. Hans Nelen, Dr. Hans Stauder and Dr. David Townend.

Lastly, we would like to express our heartfelt gratitude to Professor Jan M. Smits for making all this possible.

Nina Koivula
Nina Ferreira
Petar Lozev
Franziska Böhlke
Birgit Thun
Janika Bockmeyer

Rehabilitation of sexual offenders: clearing the stage for neuroscience?

By Nina Ferreira

Table of Contents

Abbreviations

ACC – anterior cingulate cortex
CT – Computed Tomography
fMRI – functional Magnetic Resonance Imaging
HEG – Hemoencephalography
PET – Positron Emission Tomography
TMS – Transcranial Magnetic Stimulation

1. Introduction

Sexual delinquency is among the most serious crimes constituting a significant societal problem and raising the most public attention (Hanson & Morton-Bourgon, 2005). They do so in particular when it concerns a re-offence committed by a forensic patient. The prevention of sexual recidivism through treatment and rehabilitation in correctional settings is crucial, given the irrefutable harm that these offences cause victims and the fear they generate in the community (CSOM, 2001). Such offences have grave physical, emotional, psychological and social impact on the victims (Boyd, 2011), which makes them incomparable and exceptional with regard to other criminal offences. The following two quotes by two victims of sexual assaults illustrate clearly the serious consequences for the victims:

> "Depression followed, as did lack of any self-care or self-worth. Though I am ashamed to admit this, yes, suicide did enter my mind on many occasions and thankfully I was blessed in my life by daughters because they were my reason for surviving and pressing on even when I could barely stand my existence. I am still frequented by many of these emotions and am now just beginning, through therapy and strong support system, to work through them." ("Summer" in Easteal & McCormond-Plummer, 2006, pp. 143-144)

> "The on-going violence throughout my years of marriage was mental and sexual. My urethra was so battered I became incontinent; my psyche was so battered I became a mental cripple. I finally got out and changed my name and city, and found myself again." (Unnamed victim in Easteal, 1994, p. 67)

With respect to the grave implications of sexual delinquency, much focus of research and the public is therefore on the issue of assessing the risk of recidivism among the group of sexual offenders. It is however challenging to find strategies that are most likely to achieve the goals of prevention of relapse (Levenson, 2009). Taking into account the significance to provide the community and more importantly victims of sexual offences a life without fear and to protect them from future assaults, rehabilitation systems could draw great benefit from a deeper engagement within the uprising neuroscientific dimension. The rapidly growing field of neurosciences and the technological and medical development offer new opportunities to understand the brain, and the correlation between damaged brain structures and misbehaviour of criminals. Sexual deviance and

violence are increasingly being associated with anomalies in the brain structure of the offenders. A better understanding of a sexual delinquent's brain and the rise of new different neuroscientific techniques provide new prospects for the rehabilitation of these delinquents.

In light of the new insights neuroscience can provide, the question arises how neuroscience could be used for the rehabilitation of sexual offenders in order to reduce the possibility for recidivism. This paper will thus analyse the possible use of neuroscience in treatment of sexual offenders before release and reintegration to community. Therefore, the study intends to provide insight into the deficiencies of contemporary rehabilitation methods. Furthermore, it attempts to propose possible innovative approaches to safe reintegration of sexual offenders by using neuroscience and to examine in how far it would be possible with regards to ethical issues. The overall claim of the paper is that neuroscience could provide complementation to rehabilitation methods aiming at the reduction of recidivism among sexual offenders, though in limitation of the respect of certain rights of the delinquent.

On account of a proper understanding of the concept 'sexual offenders', it is defined in this study as "a highly heterogeneous mixture of individuals who have committed violent sexual assaults of strangers, offenders who have had inappropriate sexual contact with family members, individuals who have molested children, and those who have engaged in a wide range of other inappropriate and criminal sexual behaviours" (CSOM, 2001). Furthermore, 'recidivism' is defined as the relapse into crime by a person, who cannot be cured of criminal tendencies and is therefore a persistent offender (Hornby, 1974). However, it is important to consider that recidivism has different operational definitions. It can be either measured by a re-arrest, by subsequent conviction, or by subsequent incarceration. Moreover, it has to be decided, when measuring recidivism of sexual offenders, whether the commission of any crime is sufficient to be classified as a recidivating act or of only sexual delinquency will be regarded as recidivism. In this study, the focus lies on merely subsequent sexual delinquency as recidivating offence.

The discussion is structured as follows. The first section illustrates the academic relevance of this study by pointing out the deficiencies of existing studies and what it aims to add to the discussion. The following section addresses the problems of contemporary rehabilitation methods. It will do so by reviewing literature and studies on recidivism of sexual offenders and on the problems of current rehabilitation. By drawing general

conclusions from the findings and providing recidivism rates, the section provides the basis for the following discussion. Subsequently, the paper will discuss which insights neuroscience has brought so far on the brain structures of sexual offenders in order to determine the neuroscientific relevance. In the fifth section, the most relevant technologies of neuroscience for this study will be presented. An analysis of their possible use and new approaches for the rehabilitation of sexual offenders will be provided. Following that, the proposed approaches will be examined in the light of ethical concerns. The last section concludes by summarising the findings of this research paper and providing impetus for further research.

2. Academic Relevance

A precise reading of academic studies related to sexual recidivism, the effectiveness of the rehabilitation of sexual offenders and neuroscience enables to determine the academic relevance of this research paper. It revealed that not much has been researched on the intersection of neuroscience and the treatment of sexual delinquents. There are several meta-studies (Hall, 1995; Hanson & Bussière, 1998; Hanson et al., 2003; Lösel & Schmucker, 2005; Hanson, 2005; Hanson et al., 2009) on recidivism of sexual offenders, in particular of paedophiles and child molesters, which provide recidivism rates based on many different studies, which were selected on the basis of established criteria by these researchers. These meta-analyses offer an initial approach to the understanding of effectiveness of treatment. However, it is difficult to draw general conclusions from all of them, as they all use a different methodology, focus on distinct follow-up periods, and are based on different operational definitions of recidivism.

More importantly, so far there has been no literature on a potential improvement of recidivism prevention related to neuroscience. Until now, studies drawing a link between neuroscience and sexual offenders focus rather on their brain structure and the understanding of deviant interests and behaviour of sexual delinquents as a result of mental disorders (McKenna, 1999; Arnow et al., 2002; Joyal et al., 2007; Dreßing et al., 2007; Schiffer et al., 2007; Dreßing et al., 2008; Witzel et al., 2008; Gillespie et al., 2012). Witzel et al. (2008) suggest that neuroimaging, for example, could help in understanding recidivism of sexual offenders. They do not however state that neuroscience could potentially help in reducing future recidivism. The current paper tries to address this deficiency by examining the possible prospect application of neuroscience in the context of rehabilitation of sexual offenders.

With regard to the before mentioned limitations and the fact that no study has so far investigated the potential application of neuroscience, the current research paper attempts to provide an innovative approach to that particular intersection of neuroscience and the criminal justice systems. Moreover this research paper will add to the discussion an analysis of several neuroscientific methods that can address particular brain regions which are considered to be related to sexual deviance and behaviour. It will complement the research made to date by adding insights into the potential impact of these neuroscientific methods on the sexual offender during treatment, and how they could contribute in preventing sexual recidivism.

With regard to ethical concerns of neuroscience, Shaw (2012) and Vincent (2012) contribute considerably to this field. Both focus however on direct brain interventions and do not relate it to sexual offenders at all. Nevertheless, they make general conclusions, which are of considerable significance for this study. Other literature (Foster, 2006; Steven & Pascual-Leone, 2006; Ford & Henderson, 2006; Klitzman, 2006) also delivers meaningful insight into ethical problems regarding neuroscience by discussing of other neuroscientific techniques, such as the Transcranial Magnetic Stimulation, neuroimaging, and Deep-Brain Stimulation. Instead discussing merely one or two specific neuroscientific methods, this study will rather attempt to provide a discussion of the ethical concerns regarding the use of several neuroscientific techniques, as presented in section 5 of this paper, for sexual offenders in particular.

3. The Flaws of Contemporary Rehabilitation Methods

This section addresses the problems of current rehabilitation and treatment of sexual offenders. Terms such as rehabilitation and reintegration are used to refer to the psychological and social processes employed to assist criminals before or often even after their release (Seifert et al., 2003). It is employed as a sanction to protect the community from a relapse of serious sexual offenders. The main target of such a therapeutic treatment of sexual delinquents is to reduce the risk of re-offending sexually (Song & Lieb, 1994). Furthermore, treatment is provided to these criminals "to take responsibility for their behaviours, develop the necessary skills and techniques that will prevent them from engaging in sexually abusive and other harmful behaviours in the future, and

lead productive and prosocial lives" (CSOM, 2006, pp. 3-4). From a cognitive-behavioural perspective, correctional treatments intend to instil into criminals an understanding of the correlation between "thoughts, feelings, and behaviours, their impact on one's conduct, and then developing more healthy thinking patterns and appropriate ways of managing emotions" (ibid.). Furthermore, contemporary correctional programs are aimed at deviant sexual interests, distorted cognitive perception of offending, deficient social competence, lack of empathy, and impaired way of managing aggression (Song & Lieb, 1994).
In countries such as the Netherlands, Belgium, Germany, England and Canada delinquents with a mental disorder usually face an obligatory hospitalisation in forensic psychiatric clinics. In these states, national legislation arranges a risk assessment for individuals regarded as dangerous in view of recidivism. Moreover, it provides for the possibility of a prison sentence of indeterminate duration, to prolong the sentence or keep a criminal incarcerated as soon as the sentence has been served. Such a sentence is known in England as *life sentence, imprisonment for public protection* and *extended sentence;* in Germany it is called *Sicherheitsverwahrung*; in Canada *dangerous and long-term offender designation*; and in Belgium *terschikkingstelling van de regering*" (Kogel & Nagtegaal, 2006, p. 321).

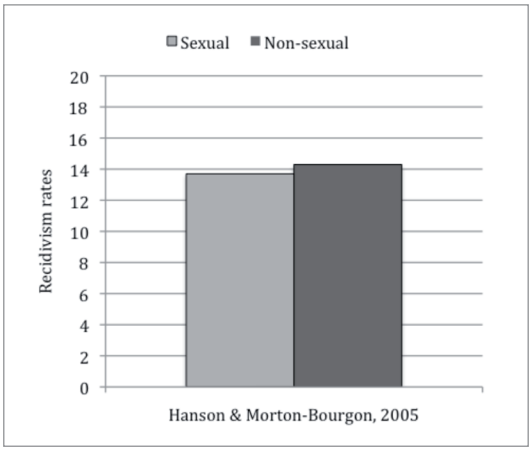
The rehabilitation of sexual offenders employs different treatment approaches – so far, these techniques are cognitive-behavioural, pharmacological, or psychotherapeutic (Motiuk, 2006). The cognitive-behavioural approach is designed to alter thinking schemes and patterns of arousal. Behaviour, beliefs, and attitudes, which are considered to increase the probability of sexual assaults, are addressed. Skills-based training is generally applied in order to strengthen empathy, self-control, and social competences (Seto, 2007; CSOM, 2001). Further, the pharmacological treatment aims at decreasing sexual arousal by means of medication. Lastly, the psychotherapeutic approach targets awareness and sense of responsibility for the offence and better concern for the victim (CSOM, 2001). Often these approaches are connected in practice. Sexual offender treatment also varies by location of the treatment – e.g. in prison, the community or a forensic clinic – and by the levels of intensity of treatment, i.e. the focus and duration (Motiuk, 2006; CSOM, 2001).

Recidivism might be much more likely among higher risk sexual offenders than lower risk ones. As Motiuk (2006) puts it, "although treatment may be more likely to reduce recidivism among higher risk sex offenders than their lower risk counterparts ... their higher risk level suggests that some of them will re-offend – even after treatment" (p. 21). In this paper it is held that every single case of victimization is one instance too much.

Therefore, we have the duty to reduce sexual offences as much as possible to prevent further victimization and hazards to society.

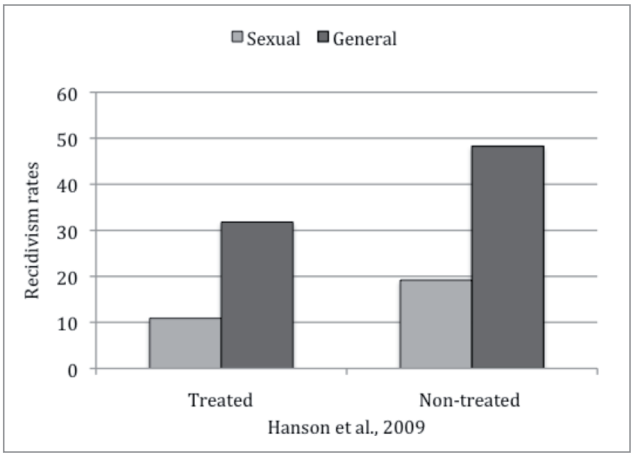
The measurement of the effectiveness of overall contemporary rehabilitation methods is rather difficult due to the different treatment approaches, the different location of treatments, the different levels of seriousness of the offenders' criminal history, and different degrees of self-reflection, i.e. whether it has been the own decision to participate or a forced one. Recidivism rates, however, can alternatively provide an approach to examine whether rehabilitation has been successful in reducing re-offence. Sexual offence recidivism rates can inform us about visible offenders, i.e. offenders who come into contact with the criminal justice systems. Although this seems to be a promising approach, there are crucial limitations. Firstly, recidivism rates "cannot tell us about hidden sexual assault such as intimate partner rape" (Stathopoulous, 2010). As Stathopoulous (2010) explains, "this type of offence, which may be repeated over years, may never come to the attention of police or end up in the justice system". Secondly, only one out of six known sexual assaults are reported, as announced by the Australian Bureau of Statistics in 2006 (ibid.). Consequently, recidivism rates can merely provide for a limited approach to understand the effectiveness of rehabilitation.

Studies on recidivism differ due to an application of distinct methods and follow-up periods, i.e. the time period until relapse. This research paper makes use of a meta-analysis by Hanson and Morton-Bourgon (2005), which reveals international results on recidivism. The analysis was based on 82 recidivism studies comprising in total 29,450 offenders.



It resulted in recidivism rates that are far lower than usually expected. By an average of 13.7 % (n = 19,267; 73 studies) sexual delinquents re-offend sexually; their violent non-sexual recidivism rate was 14,3 % (n = 6,928; 24 studies). On average, the follow-up time was 5 – 6 years (Hanson & Morton-Bourgon, 2005).

Besides this study, Hanson et al. conducted another meta-analysis in 2009. The meta-analysis was based on merely 22 recidivism outcome studies. It is particularly interesting that the study compared treated and non-treated delinquents. It included therefore 3,121 treated offenders and 3,625 untreated offenders in comparison groups. For the treatment group, the study revealed that the delinquents recidivated with an unweighted mean of 10.9 %, whereas non-treated offenders relapse with an unweighted mean of 19.2 % (Hanson et al., 2009). This illustrates clearly that rehabilitation methods do have a positive impact on sexual offenders and reduce sexual and non-sexual recidivism rates considerably.



Concluding from the foregoing findings by the international meta-analyses, most sexual offenders were not caught for another sexual offence – only 13.7 % were known for relevant recidivism. On average, sexual offenders are more likely to reoffend with a non-sexual offence than a sexual offence. Furthermore, sexual delinquents, who take treatment, are less likely to relapse than their comparison groups. However, it is important to stress that these figures underrate the actual recidivism since not all offences come to light. Nevertheless, we can learn from the recidivism rates over the effectiveness of correctional rehabilitation that not all interventions do effectively reduce recidivism. Given the grave consequences for victims of sexual offences, low recidivism rates are

nonetheless worrisome. Thus, there still remains considerable scope for improvement in sexual offender rehabilitation.

In respect of the effectiveness of rehabilitation for sexual delinquents and their deficiencies, one has to point out a crucial issue: there is never certainty over the outcome of the treatment. Even upon release a psychiatric expert can never make a prognosis over future behaviour and possible recidivism with a probability of 100%. There are always limits to each treatment and a remaining risk can hardly be assessed. Nonetheless, this cannot be a reason for keeping the delinquents arrested or in forensic clinics. In Germany, for example, a risk assessment is obligatory prior to discharge. Seifert et al. (2002) holds that "there is a noticeable trend in Germany towards over-predicting the risk posed by mentally ill offenders in an attempt to create a greater certainty" (pp. 62-63). In this respect, there is a need to improve risk assessment in order to avoid lifelong preventive detentions or clinical stay. It might be that some sexual offenders have to continue their treatment although they would probably not recidivate. Moreover, the effectiveness of risk assessment can be weakened when information fails to be correctly reported, or the experts focus on an irrelevant issue resulting in a biased impression of the offenders' character and behaviour (Belcher, 2008).

A further deficiency of rehabilitation is that problems may arise with regard to the documentation of post-discharge progress, i.e. in the case of medication compliance (Seifert et al., 2003), which could result in an ultimate ineffective rehabilitation. Moreover, a self-assessment or self-report by the offenders forms a crucial part in rehabilitation treatments in forensic psychiatry (Hanson & Bussière, 1998). Hanson & Bussière (1998) set out that sexual offenders typically "deny recurrent deviant sexual interests or behaviour" (p. 348). Since such self-reports are easily exposed to bias, additional sources of information on the actual state of deviant interest and behaviour are necessary. Another limitation of rehabilitation methods is the fact that it is dubious whether these treatments are able to have an effect on the delinquents' behaviour even years after they have been released (Fortune et al., 2011).

Furthermore, preventive detention causes considerable costs in the long-term. In Germany, for example, the monthly costs for a single offender amount to 6,000 to 10,000 Euros (Thüringer Allgemeine, 2011). The question remains whether introducing complementary neuroscientific treatment measures at the stage of correctional treatment cannot prevent such costs. Another major weakness is the artificial setting of treatment. Accordingly, one

cannot prove whether improvements shown within this artificial setting will also remain in the real world (Belcher, 2008). In respect of all these deficiencies, the question arises whether neuroscientific methods could provide a solution to reduce even more recidivism rates. Could it be used to improve risk assessment? Or could one make use of it for correctional treatment? These questions are addressed in the following section.

4. Clearing the Stage for Neuroscience?

4.1. A View into Sexual Offenders' Brains

In light of the relevance of neuroscience for the rehabilitation of sexual offenders, we have to understand whether sexual offenders show unusual patterns of brain activity that are critical for sexual behaviour. McKenna (1999) argues that the brain is the 'master organ' responsible for sexual functioning. Sexual functioning involves both the cortical and subcortical structures of the brain – in particular its anterior regions (Joyal et al., 2007). The regulation of sexual drive, initiation, and activation takes place in the frontal and temporal lobes. Subcortical structures, such as the amygdala, the hippocampus, the septal complex and the hypothalamus, are associated with the mental regulation of sexual behaviours and genital reactions (ibid.). Further, neuroimaging data found that the frontal, temporal cingulate and subcortical structures are responsible for the modulation of sexual arousal (Arnow et al., 2002). Joyal et al. (2007) claim moreover that hyper sexuality and paraphilia have been observed to be associated with damage to subcortical structures. Frequent findings of neuropsychological and neuroimaging studies on sexual offenders hypothesise these brain parts to be involved in sexual deviance.

Deviant sexual interest is specific for sexual delinquents, who are usually attracted to sexual acts that are considerably abnormal and unlawful (Hanson & Morton-Bourgon, 2005). Joyal et al. (2007) argue that sexual anomalies represent "signs of an inherited diseased condition of the central nervous system" (p. 156). Additionally, Schiffer et al. (2007) find that the association between frontal and temporal cortex abnormalities and paedophilia support the hypothesis that there is a connection between damaged brain structures and sexual delinquency (also: Witzel et al., 2008). Furthermore, most paedophiles show alterations of the amygdala (Schiltz et al., 2007).

In addition to the foregoing, it is crucial to point out responsible brain structures of the abilities, which are lacking for sexual offenders. Dreßing et al. (2008) set out that "the ability to control impulsivity and to exhibit socially adequate behaviour as well as the ability for moral and ethical judgement seems to be linked to the function of the prefrontal cortex" (p. 8). Further, the occurrence of antisocial personality disorders among sexual delinquents is common. As a functional Magnetic Resonance Imaging (fMRI) study revealed, men with antisocial personality disorder possess at least 11% less prefrontal grey matter compared to the control groups (Raine et al., 2000). This decreased grey matter volume is clearly found in paedophiles, in particular in the orbitofrontal cortex, the ventral striatum and the cerebellum (Schiffer et al., 2007).

In light of the aforementioned, one can conclude that sexual offenders' brains display abnormalities in their structures. However, it is very important to stress the heterogeneity of the group of sexual offenders in which paedophiles, for example, have much more significant damages than delinquents in adult rape cases (Joyel et al., 2007; Motiuk, 2006). Nonetheless, we can argue from the foregoing that there is neuroscientific relevance for the treatment of sexual offenders.

4.2. Invasive and Non-invasive Neuroscientific Methods and their Potential Use for Rehabilitation

This section addresses the analysis of whether neuroscience could be used for the rehabilitation of sexual offenders. In respect of the deficiencies of contemporary rehabilitation methods, which were explained in section 3 of this paper, it is assumed that there is a need to search for other criteria for risk assessment and other methods for treatment. Therefore, the following paragraphs will discuss whether neuroscientific techniques could help in this context. Different techniques of neuroscience, from neuroimaging over neurostimulation to chemical castration, will be presented and their possible use will be discussed. In general, the neuroscientific dimension could contribute to rehabilitation by producing more clarity in predicting outcomes of treatment and future risk. Moreover, it could supply forensic institutions with more detailed information on the current state of the sexual offenders' interests and behaviour.

Neuroscience provides many different techniques, whose use and focus vary significantly. Hanson & Bussière (1998) argue that the strongest predictors of sexual recidivism are factors related to sexual deviance and thus deviant sexual interests. As result of

the introduction of neuroimaging methods, the assessment of sexual interests has substantially developed (Seto, 2007). Of the several neuroimaging techniques, this research paper refers exclusively to functional Magnetic Resonance Imaging (fMRI), Positron Emission Tomography (PET) and Computed Tomography (CT).

To begin with, fMRI is a technique used for the measurement of brain activity. It is based on the fact that brain areas spend more oxygen when they are active and blood flow rises to the active brain region to meet the growing demand of oxygen. fMRI reveals the alteration in blood oxygenation and blood flow resulting from the neural activity. It is a non-invasive technique, which does not expose the patient to radiation (Devlin, 2007). Relating it to sexual deviance, one can state that functional brain imaging of sexual deviance is still at the very beginning and therefore very few studies exist. Stoléru et al. (2003), for example, hold that there is a decrease in inhibition as a result of showing sexual stimuli leading to a reduction of orbitofrontal blood flow in men with normal libido. The case for men with hypoactive sexual desire is distinct, as the same brain area remains active. The maintenance of brain activation hence brings about constant higher levels of inhibition. Does the contrary, i.e. increased blood flow and even lesser inhibition, happen to men with hyper-sexuality? Brooks (2012) states that fMRI reveals an increased cerebral blood flow of hypersexual men when presenting them sexually arousing images. In addition, paedophiles display crucially higher activation in orbitofrontal brain areas and in the right amygdala when presenting them images of children. This has been shown in fMRI studies which contrasted paedophiles with non-paedophiliac control subjects, giving them images of men, women, girls and boys in the course of the imaging (Dreßling et al., 2007; Dreßing et al., 2008). A further study on homosexual paedophiles proved significant brain activation around the brain stem, the basal ganglia and the orbitofrontal cortex (Dreßing et al., 2007). One can conclude from the foregoing that the inhibition control of paedophiles is not adequately active after facing images of children (ibid.).

Similarly to the fMRI, the PET is used to display brain activity and its functional processes (Demitri, 2007). It examines trace elements of short-lived radioactive material in the brain. For this purpose special liquid containing radioactive markers, which light the brain activity, are injected. The liquid indicator detects when the radioactive substance decays due to the emission of positrons. High radioactivity in specific cerebral areas is connected with high brain activity (ibid.). Though this technique is invasive, necessitating a small quantity of radioactive substance and therefore having minimal possible side effects, it provides a good resolution of the brain activity due to the positron emission, which

effuses from brain regions that consume the oxygen and burn glucose transported by the blood flow (Gunkelman & Johnstone, 2005).

Relating these findings to the rehabilitation of sexual offenders, one can argue that fMRI and PET have certain potential at this stage of the criminal justice systems. Firstly, it could contribute to the assessments on the sexual offenders' interests, behaviours and personal traits. It could be used to detect the specific sexual interests of the offenders in order to find the most appropriate therapeutic approach. By adopting a treatment, which is not only chosen on a subjective opinion of an expert but also on objective neuroscientific findings, it could result in far more effective rehabilitation.

Furthermore, it could not only be used to find the right treatment, but also to complement the assessment of the sexual offender's state during rehabilitation. Thereby, it could optimize the on-going therapy. Lastly, fMRI and PET results could be used as additional objective criteria in contrast to subjective views of forensic experts in order to reduce possible bias caused by the 'halo-effect'. Concluding from the above, fMRI and PET could particularly mark contribution to (risk) assessment of sexual offenders. Overall, its use could affirm the certainty of choices in treatment by the forensic experts. The fact that it is a non-invasive procedure, it allows for study reproducibility and there are no known risks which makes it a convincing and appealing technology (Illes et al., 2006). Nevertheless, the information gathered through the use of fMRI and PET should be cautiously dealt with as so far it has not been proved and tested sufficiently with sexual offenders. As already mentioned above, functional brain imaging of sexual deviance is still in its infancy.

CT scanning can further picture the brain on grounds of variable absorption of x-rays. CT scans can reveal the gross features of the brain. X-rays are well absorbed by bones and hard tissue, poorly absorbed by air and water, and soft tissue lies somewhere in between. Hence, the grey and white matter in the brain can be displayed. CT scans cannot however provide concrete findings on the brain activity (ibid.). Nevertheless, it is of significance for the underlying discussion since paedophiles feature, for example, an orbitofrontal and striatal reduction of grey matter (Dreßing et al., 2007). CT would offer the possibility of examining the volumetry of grey substance and could thereby provide a complementary check on the right therapy for individual sexual offenders. It could result in a better specification of the type of sexual offender and hence in a better and more appropriate therapeutic approach.

Besides neuroimaging, neurofeedback is a neuroscientific technology that allows for the recondition, training and self-regulation of brain waves (Masterpasqua &Healey, 2003; Hammond, 2008; Gunkelman & Johnstone, 2005). During a conventional neurofeedback session, electrodes are positioned on the scalp and typically also the earlobes. These electrode devices measure the person's brainwaves while watching a computer screen, listening to audio tone and doing specific tasks. The information revealed by the measurement of the brainwaves is read and transmitted to the computer, which translates it into audio-visual feedback (Hammond, 2008; Neurofeedback, n.d.). This technology enables the patient to train the alteration of brainwave patterns and to practice the maintenance of healthier brainwave patterns. As no electrical current enters the brain, this method is non-invasive (ibid.).

So far it has been effectively used for, *inter alia*, epilepsy, stroke, alcoholism, drug abuse, post-traumatic stress disorder, depression, obsessive-compulsive disorders, attention-deficits, hyperactivity disorder, autism, schizophrenia, Parkinson's tremor, tinnitus and Tourette Syndrome (Masterpasqua & Healey, 2003; Hammond, 2008). It is commonly argued that neurofeedback can potentially produce relief from any problems influenced by brainwaves. "This would include almost anything controlled by the brain including thinking abilities, motoric responses, behavioural, emotional, and social difficulties" (Neurofeedback, n.d., par. 4).
One tool commonly applied in neurofeedback sessions is real-time fMRI. Renaud et al. (2011) proposed a methodology for neurofeedback with regard to paedophilia. They suggest a neurofeedback technique targeting the activation of the anterior cingulate cortex (ACC). The ACC is associated with impulse control and sexual arousal. If paedophiles train to decrease the cerebral activation in the ACC, their urge to act upon the sexual attraction to children might be reduced (ibid.). The authors suggest a voluntary real-time imaging for paedophiles in which they present images of children on a screen and display the children's movements according to their ACC activation. When ACC activation diminishes, then the child displayed might raise its hands, and the patient would thus recognise that the ACC activity has been effectively reduced. As complementary strategy, the patients should permanently reflect on the severe consequences of acting on their sexual urges (ibid.).

In light of the fact that the frontal lobe is associated with sexual drive, initiation, and activation, Hemoencephalography (HEG) could provide an alternative neurofeedback technique useful for frontal lobe measurement (Neurofeedback, n.d.). One form is the Near

Infrared HEG, which examines the cerebral oxygenation. A headband, which comprises an infrared light source and two optodes, surrounds the head. When the infrared lights begin to blink, the brain oxygenation is measured upon the absorption of these lights and the ratio of this absorption received at the two optodes. The information gathered in the measurement is then transformed in feedback to enable the patient to train the cerebral functions (ibid). Another form of HEG is the Passive Infrared HEG. An infrared lens, which is used as 'brain thermometer', measures the temperature and from it derives cerebral metabolism. It is a technology used for a special neurofeedback to train the enhancement and regulation of cerebral function in the frontal lobe (ibid.).

Having discussed neurofeedback, one can state that there is so far poor experience with treating paedophiles, or sexual offenders in general. Nonetheless, as Renaud et al. (2011) suggest, there might be a potential use of real-time fMRI in the context of neurofeedback to treat this type of offenders in the future. Moreover, in respect of personality and psychological disorders, which are associated with sexual offenders, neurofeedback can provide an alternative to cognitive-behavioural treatment. Gillespie et al. (2012) state that "deficits in emotional regulation are particularly pertinent in sexual offenders" (p. 333). In particular, heightened levels of emotional loneliness, personal distress, and decreased levels of self-esteem have been found among sexual offenders. A person's capacity to keep sexual drives and behaviour under control can be impaired by such emotions states, and thus resulting *inter alia* in abnormal sexual fantasies (Gillespie et al, 2012).

The regions, which are particularly associated with these personality and psychological disorders, are the prefrontal cortex, the amygdala, and the ACC. The orbitofrontal cortex in the prefrontal cortex is a region associated with disorders such as the Post-Traumatic Stress Disorder, Panic Disorder, and Obsessive Compulsive Disorder (ibid.). Explaining the concrete potential of neurofeedback for sexual offenders' brains is difficult. Nevertheless, one can express the speculation that neurofeedback for the activity of these brain regions for rehabilitating sexual offenders might be possible in the future, as the aforementioned disorders are already being treated by neurofeedback. Concluding from the above, neurofeedback for the treatment of sexual offenders is rather at its methodological stage. Nonetheless, neurofeedback in brain regions such as the frontal lobe, which controls emotions, inhibition and impulsivity, could provide some further complementary tool to therapy to rehabilitate sexual offenders in the future.

It is however important to point out that neurofeedback, as it stands now, would provide a complementary tool to therapy rather than being a therapeutic method by its own. Prof Dr R. Goebel (personal communication, April 23, 2013) argues that neurofeedback for sexual offenders would have merely indirect effects on subcortical functions. The more subcortical the deviance is, he says, the more difficult it is to correct. Nevertheless, as stated before, sexual arousal and sexual offenders' behaviour are often related to the cortical structures of the brain that would accordingly be addressed by neurofeedback. Neurofeedback is however limited to a certain compliance by the offender. The offender must have active thoughts and must focus on the treatment. Otherwise it is rather difficult to have a positive impact on the offender (ibid.). As a consequence, neurofeedback would have a rather uncertain effectiveness when used as single treatment technique.
Another advance in neuroscience, which has so far no use for the treatment of sexual deviance or delinquency, but could potentially be a complementary method for rehabilitation in the future, is neurostimulation. One neurostimulation tool is the Transcranial Magnetic Stimulation (TMS). It is a non-intrusive technology for the electrical stimulation of neural tissue and thus it has a moderate risk of serious consequences. Minimal side effects are, for example, migraines and minor skin injuries in the treatment area (López-Ibor et al., 2008). This technique "induces an electric current in the brain via application of a localized magnetic field pulse. The pulse penetrates the scalp and skull non-invasively and, depending on the parameters of stimulation, facilitates or depresses the local neuronal response with effects that can be long-lasting" (Steven & Pascual-Leone, 2006, p. 201). TMS finds applications in the study and treatment of movement disorders, epilepsy, depression, anxiety disorders, stuttering, Post-Traumatic Stress Disorder and schizophrenia (Walsh & Cowey, 2000; Kobayashi & Pascual-Leone, 2003; López-Ibor et al., 2008). The stimulation of neurons belonging to certain brain regions takes places through the supply of a current that builds a magnetic field (Walsh & Cowey, 2000). With regard to sexual offenders, the method could benefit rehabilitation by improving the correctional treatment as a therapy itself.

It is a method that activates cortical and subcortical structures in a safe, non-intrusive and relatively painless way, with minimal side effects reported (Kobayashi & Pascual-Leone, 2003; López-Ibor et al., 2008). Kobayashi and Pascual-Leone (2003) state that repetitive TMS can alter activity in the cortex with long term "behavioural effects, including visual, prefrontal, parietal cortex, as well as the cerebellum" (p.153). This offers a possibility to treatments applying repetitive TMS to regulate decreased or increased levels of cortical activity, by decreasing blood flow in the brain area stimulated which ultimately results

in a reduction of behavioural performance of tasks related to that region (ibid.; Walsh & Conwey, 2000). Relating it to sexual offenders, who have an increased cerebral blood flow, TMS could decrease blood flow in cortical and subcortical brain areas related to sexual arousal. As a consequence, the behavioural performance of sexual interests and urges might be diminished.

Lastly, this paper discusses chemical castration as neuroscientific method with potential use for the rehabilitation of sexual offenders. This form of treatment already takes place in several countries – in the United States, for example, it is used for relapsed paedophiles and rapists. Chemical castration is the use of medication, namely anti-androgenic drugs, to reduce male testosterone to pre-pubertal levels (Kutcher, 2010). With the aim of constraining the impact of male sexual hormones, anti-androgens block the receptors in the brain (Barret, 2008). The predominant anti-androgenic drugs used are cyproterone acetate, medroxyprogesterone, and gonadotrophin that secrete hormone agonists such as leuprolide, goserelin and tryptorelin (Grubin & Beech, 2010). The consequence of taking these drugs is the reduction of sexual drives, interest and performance. Logically it seems that such drugs might be an effective method to reduce the likelihood of sexual reoffending (ibid.). An offender with a decreased sexual interest and urges is considerably less likely to re-offence sexually (Sifferd, 2013). As most sexual offences follow from incapacity to control the sexual drives, chemical castration can be considered as a potentially effective method for rehabilitation. Grubin and Beech (2010) point out that such an intervention affecting the hormones is much more effective in decreasing recidivism rates than any other treatment approach.

The participation of offenders in psychological treatment, which might not have been previously possible or successful due to high testosterone levels, becomes possible with the taking of anti-androgenic drugs (Grubin & Beech, 2010). Consequently, these drugs enable to address the psychological qualities of sexual interest besides its hormonal part. The lowered testosterone levels not only reduce sex drive, but also sexual imagery, hence producing relief from obsessive fantasies (Berlin, 1997). Chemical castration however cannot address the psychological roots of offenders, who are not motivated by sexual drive and act as a result of personality disorders. Therefore, it is suggested that chemical castration should be used as complementary approach for rehabilitation besides cognitive-behavioural treatment or other neuroscientific methods. As it is not possible to monitor the compliance with oral drugs, they are mainly injected once a week (Berlin, 1997).

This method of dealing with the sexual interests and drives of offenders has the advantage of allowing for conditional release and enables a safe reintegration into society, provided the injection occurs in a regular and monitored manner. It offers the sexual offender the possibility to choose between longer arrest and an effective therapy for a faster release from prison and forensic psychiatry. Besides all these benefits, there are certain side effects resulting from chemical castration, for example osteoporosis, metabolic abnormalities, cardiovascular disease, and gynaecomastia (Grubin & Beech, 2010).

After having discussed the potential use of several neuroscientific methods, we must not disregard their limitations. These are addressed in the subsequent section.

4.3. Practical Limitations to Use of Neuroscience for Rehabilitation

Although neuroscientific methods seem to be promising for the rehabilitation of sexual offenders, one has to keep in mind that there are significant limitations to their application. Firstly, it has to be pointed out that early attempts to alter sexual interests have not been effective (Gordon, 2013). Moreover, there is an overall lack of knowledge of how the brain works and how to use neuroscience accordingly. Hence it seems that neuroscience has still a long way to go before being able to reveal much more information on the brain (Morse, 2011). Nonetheless, it has to be recognised that it could provide new approaches in the future if further investigation and the rise of innovative techniques provide a breeding ground for it.

Secondly, there is no proper understanding of the neurobiology of sexual interests and urges in general. Consequently, there is a need for further experiments and investigation in this field (ibid.). Thirdly, there are concerns regarding the validity of measures such as neurofeedback, as sexual offenders would not be exposed to a real life circumstances (ibid.; Morse, 2011). The situation of a sexual offender being connected to machinery and staring at a computer screen, on which images of children are shown, is certainly different from a real-life encounter with a child. Consequently, we cannot be certain about how exactly sexual offenders would behave in the real world, and whether neuroscientific advances could actually contribute to effective rehabilitation in real life, in spite of the positive results in laboratories.

Furthermore, there is a concern about the accuracy of predictions of sexual recidivism. The accuracy of neuroscientific methods in predicting recidivism is essential for rehabilitation in order to prevent any future harm caused to society (Greely, 2006). Greely (2006) further argues that

> "one crucial question would become just how much accuracy, and of what kind, should be required before relying on such predictive tests. In addition, how to calculate these accuracy rates will raise their own difficult problems: we would not want to test the accuracy of tests for predicting pedophilia by releasing tested suspects to see how often they molest children" (ibid.).

As a consequence, it is necessary to find an appropriate way of foreseeing assuredly whether a sexual delinquent will reintegrate properly in society and act lawfully after release or not (Canli, 2006). A further crucial limitation is that neuroscientific methods are connected with enormous costs. These costs do not only result from the acquisition of the equipment, but also from the employment of experts who operate and maintain the neuroscientific techniques (Illes et al., 2006). Nevertheless, these costs have to be balanced with the expenses of actual incarceration, rehabilitation, re-arrest in case of re-offence, and more importantly the costs of new victimization and the therapeutic measures needed after a re-offence. Lastly and more generally, current neuroscience studies on humans take place only with a very little number of subjects. This complicates the setting up of approved methods at present (Morse, 2011). In the end, it is to stress that neuroscience as it stands now is largely in its infancy and still needs lots of development to play a significant role in the rehabilitation of sexual offenders.

All in all, one can conclude from the foregoing discussion that neuroscience could potentially contribute to the rehabilitation of sexual offenders in the future besides the practical limitations. It could provide devices that allow a better prediction of potential recidivism, and to adjust the treatment itself through better risk assessment. As additional data will be independently assessed on the basis of neuroscience, subjective mistakes by forensic experts could be avoided and there would be overall more clarity in predicting risk. In addition to that, experts would be able to develop individually adjusted treatments so that a 'one-size-fits-all' approach to rehabilitation would be avoided. Thereby, treatment could address the specific individual needs of each offender and thus improve the reduction of recidivism rates.

Further, the new strategies for rehabilitation in terms of neurofeedback, for example, allow an individual to have a better control of his behaviour after release. Besides this, the use of neuroscience in the context of rehabilitation and also of incarceration could not only be significant for the prediction of re-offence, but also for avoiding over-prediction, i.e. keeping sexual offenders incarcerated or in forensic clinics although they would not recidivate after release. Thus, an unjust confinement of successfully treated offenders would be prevented. In light of lifetime surveillance after release, as in the case of the German sexual offender Karl D.[1], a neuroscientific approach in terms of chemical castration or TMS for example could provide an alternative solution to such a situation. Lifetime surveillance is associated with immense costs. By conducting a cost-benefit analysis, one could decide whether a neuroscientific approach would not be the more desirable solution. A cost-benefit analysis would be similarly adequate for the decision to safe-keep an offender in preventive custody. Instead of investing in further preventive detention institutions, money could be put into neuroscientific investigations and techniques in order to enhance its future use in the field of sexual offender rehabilitation.

Moreover, neuroscientific technologies such as TMS could provide new future treatment approaches, which alter the sexual deviances, desires and associated disorders. Thus, there is certainly a potential for neuroscience in the context of the rehabilitation of sexual offenders. It has just to find its way through greater commitment to experiments, further investigation and addressing the practical limitations discussed above. Nonetheless, there are not only practical limitations to the use of neuroscientific methods. Just as crucial are the ethical concerns to their use. As these issues can hardly be reduced in the future and thus constitute a permanent limitation to the use of neuroscience, they deserve their own section in this research paper. The following section thus discusses these concerns and proposes a possible approach in handling them.

1 In the 1980s Karl D. brutally raped three girls. He was arrested until 2009 and was released although he was considered as highly risky, *inter alia* because he does not accept the most brutal rape as his own. Karl D. moved to the German town Heinsberg-Randerath nearby Aachen, where he was under permanent surveillance – 24 hours a day, 7 days a week. After his move to the city Gelsenkirchen, he still stands under lifetime surveillance.

5. Neuroscience and Rehabilitation of Sexual Offenders – an Ethical Perspective

Neuroethics is a field dealing with the benefits and dangers that neuroscientific advances on the brain provide, and debating the ethical, social and legal consequences of treating or manipulating the brain (CCLE, 2003). This section addresses the discussion of the ethical and legal challenges of neuroscience with regard to the rehabilitation of sexual offenders. The discussion revolves around concepts such as privacy, autonomy, choice, consent, personal identity and proportionality.

First of all, it is necessary to understand that sexual offenders are not merely 'sexual offenders'. They are individuals, who have offended sexually and who have nonetheless specific fundamental rights, like any other criminal. One can therefore not simply disregard their rights when applying neuroscientific methods in a rehabilitation context. These rights are at the core of what it means to be a free person. As neuroscience advances and is increasingly used for the observation, enhancement and manipulation of brains, it is of utmost importance to ensure that those rights are recognised and safeguarded.

To begin with, the Centre for Cognitive Liberty and Ethics (CCLE) (2003) emphasizes the importance of cognitive liberty. It defines cognitive liberty as "the right of each individual to think independently and autonomously, to use the full spectrum of his or her mind, and to engage in multiple modes of thought" (ibid.). Accordingly, cognitive privacy and autonomy have to be protected when applying methods that enable the control and alteration of cognition. Gordijn and Buys (2010) consider privacy to be "the power of a person or group to control information about themselves. Consequently, it involves the ability to prevent data becoming known to others than those whom one chooses to be informed" (pp. 295-296). As neuroscience is able to uncover intimate, private thoughts and mental attitudes of individuals, there are many concerns that these may jeopardize privacy (Fuchs, 2006, p. 601-602).

An approach, which would enable to overcome these concerns, is to ask for an informed consent of the offender at issue, to prohibit forced applications of neuroscience, and to provide proportional treatment (Greely, 2004). It is necessary to take account of the proportionality of the trade-offs between privacy and the use of neuroscientific methods. The use of neuroscience is justified with regard to the benefits they could provide for

offenders (Gordijn & Buys, 2010). To illustrate this conflicting issue, we can have a look at neuroimaging techniques. As has been argued before, neuroimaging techniques can provide many opportunities to ameliorate risk assessment of sexual offenders by observing their brain structures. However, as neuroimaging then uncover the offenders' psychological and mental states – thoughts, desires, and attitudes – it is clear then these cause concerns over privacy (Buller, 2006).

Another crucial point in this discussion is the impact of neuroscience on the personal identity and character. Vincent (2012), for example, argues that direct brain interventions severely impact on individual's character, personality and identity. Furthermore, Spalding (1998) states that also behavioural drugs profoundly interfere with the personality and character of a person. In this respect, Shaw (2012) sets out that interventions should not alter values, which lie at the core of an individual's personality. Neuroscience should rather enhance capacities of individuals, i.e. the capacity to better control the sexual urges in the case of sexual offenders. Even more worrying is the modification of features that are fundamental to who the person is (ibid.).

Similarly, Vincent (2012) is afraid that changing an offender's character and central values might challenge his authenticity, or might even alter his personality in so far as he would be transformed into a different 'self'. In addition, Gordijn and Buys (2010) emphasize this loss of authenticity. They set out that a person "must be 'true to herself' and feel that her experience and feelings are 'her own'" in order to be authentic (p. 294). In addition, they argue that a person distances him or herself from his or her own emotions and experiences (ibid.). Such a change of a person's character traits and personality is considered to be an inexcusable violation of his or her autonomy rights (Sifferd, 2013). One would have to ensure that a person remains herself and authentic. In the case of sexual offenders, this is however rather difficult as the change or control of sexual interests and urges automatically alter attributes which are part of their identity. Sexual offenders would be alienated from their original feelings and experiences, and would thus not be the same person as before. This is therefore a striking concern regarding the use of neuroscience for rehabilitation.

A further right linked to cognitive liberty is the right to autonomy. It is the power of independence and self-rule over one's own cognition, as the CCLE defines it (2003). It remains an individual's decision whether or how to change his or her mental processes (ibid.). With regard to the discussed neuroscientific methods, it is to say that they are

compatible with the right to autonomy as long as they are reversible. This would be *inter alia* the case for chemical castration. The right to autonomy is further connected to the concept of choice, which implies that individuals have the freedom to choose. With regard to neuroscience, this means consequently that individuals should have the right to choose whether neuroscientific methods are applied or not, and to what exact extent (Sententia, 2004). Relating it to chemical castration, for example, sexual offenders should have the right to choose a treatment with anti-androgenic drugs, if they wish to voluntarily control their sexual urges (Berlin, 1997). As Berlin (1997) puts it, "a liberalized society calls for individuals to be able to make independent choices regarding their own lives, free of government interference" (p.195). The concept of choice should not merely be limited to the question whether and to what extend neuroscience is used, but also comprise the choice of the actual treatment method.

The idea of informed choice relates to the concept of informed consent. The voluntary use of neuroscientific tools for the rehabilitation of sexual offenders with an informed consent of the sexual offender would minder the legal and ethical concerns. What is necessary for consent to be informed? Bosmann-Larsen (2011, in Shaw, 2012) argues that the offer of treatment must be suitable in order for the consent to be well founded. Therefore, it needs firstly to be a genuine offer of treatment, without any threat. Secondly, the treatment must be restrictedly directed at the behaviour for which the delinquent was condemned (Ford & Henderson, 2006).

Furthermore, the offenders must be confident of the forensic expert and understand the possible risks and side effects of the treatment (ibid.). It might be difficult for the delinquents to assess the potential of a change in their personalities, or any other complication of intervention, it is crucial to provide as much information to the offender as possible. Thus, it is less difficult for them to reach a decision on the application of neuroscience. Additionally, informed consent involves the offender's realising and accord to advance the particular treatment. As it does not simply comprise an initial authorisation of the treatment, informed consent implies the voluntary wish and constant accord to continue the treatment. In order to encourage the offenders to make informed decisions, clear limits to the use of neuroscience should be set and they should be assisted to become aware of the possible consequences. For this purpose a safe, acceptable environment, in which the delinquents can simply recognise that more can be gained than lost from the use of the particular neuroscientific method, must be created.

Furthermore, informed consent involves the *voluntary* intent to participate in a neuroscientific treatment. Hence, the use of such technologies or drugs should not be forced (Spalding, 1998). Sententia (2004) puts it somewhat differently by stating that individuals should not be coerced to use technologies or to take drugs, which have an impact on their brain, as long as their future behaviour does not cause danger to others. Relating Sententia's idea to sexual offenders, he would argue that sexual offenders might be compelled to take anti-androgenic drugs or to participate in correctional treatment based on neuroscientific advances, as they constitute a danger for the society upon their release if not effectively rehabilitated. Limiting the offender's options can thus be justified by the need to protect society and by the value of rehabilitating the offender and restoring him to the community. In the context of sexual delinquency, it is overall difficult to say whether it is possible to make an informed decision as a sexual offender as there is always some external control and pressure, given the particular legal status of the offender. Nonetheless, Bosmann-Larsen (2011, in Shaw, 2012) states that "despite the coercive circumstances, an offender can sometimes give a valid consent to behaviour treatment when this is offered as a condition of early release from prison".

As a consequence from the foregoing, it is arguable that there is a need to follow the main purpose of using neuroscience for the rehabilitation of sexual offenders, namely the prevention of further harm to society. The different interests, i.e. the sexual offenders' rights and the society's protection, have to be balanced. It is necessary to determine in each context whether the traditional treatment is the safer option – as Shaw puts it: "what is safer for the offender may not be safer for the public" (2012). Sexual offenders do not only have rights, which have to be protected to a minimum, but also obligations to others in society, resulting from their harm committed to their victims and the community. As they have forfeited their rights as a result of their offence (Scott & Holmberg, 2003), it could be justified to use neuroscience in a non-excessive manner in order to help the sexual offender in improving his or her sexual deviance and behaviour. The question is then, to what extent it is permissible to constrain sexual delinquents' rights for the greater good of society.

Besides the discussion of constraining certain rights of sexual offenders, it is nonetheless crucial to point out the chance of greater exercise of the right to liberty. The exercise of this right could be possible through a treatment based on neuroscience that allows for parole or probation. Therefore, it is arguable that it would be a mutual give and take between society and offender in this context. Levenson and D'Amora (2005) state that the sexual

delinquent's treatment does not only benefit society, but is also a patient's best interest by reducing his or her suffering. "Community safety and alleviation of client distress are goals that can be mutually rewarding for patients and society" (ibid, p.146).

Overall, one can conclude by stating that the use of neuroscience for the rehabilitation of sexual offenders must be proportional, i.e. the benefits of treatment through neuroscientific technology or anti-androgenic drugs must exceed their risks and ethical concerns. Moreover, transparency should be increased, thus allowing the sexual offender to take an informed decision on whether to participate in neuroscientific treatment methods. Finally, an explicit scheme of continuous informed consent should be adopted to minder the effects of neuroscience on the sexual offenders' rights relating to privacy, autonomy and his personality. Nevertheless, it must not be disregarded that the delinquent has forfeited his rights by acting criminally. Thus an application of neuroscience, which constrains the sexual offenders' rights in a non-excessive way, could be permissible to some extent, however merely as last resort when there is no chance for any improvement in the foreseeable future.

6. Conclusion

Recidivism rates of sexual offenders and several problems of current rehabilitation systems, which weaken the overall effectiveness of rehabilitation, show that there is a need for improvement in order to prevent further sexual relapses. Neuroscience provides more and more insight on how damaged brain structures are connected with misbehaviour of criminals, and more specifically how unusual patterns of brain activity are related to sexual deviance. This research paper set out to examine how neuroscience could be used for rehabilitating sexual delinquents in order to reduce future sexual relapse. Therefore, it analysed the possible use of neuroscientific methods from neuroimaging, over neurostimulation and neurofeedback, to chemical castration. It is beyond me to explain the concrete possible use of neuroscience from a medical perspective. However, this paper aimed to provide a preliminary approach to this topic and to raise further discussion with the aim of raising the neuroscientific awareness in this specific context.
This study concludes that there is in fact a potential for neuroscientific methods to be used as complementary tools to therapy in the future. It could contribute in allowing for a better prediction of sexual recidivism and assessment – assessment of the progress made by the sexual offender, and assessment of the remaining risk. Further, it could provide more

objective data in order to avoid subjective mistakes by the forensic experts. Thereby, it would provide for more clarity and certainty regarding the choices made by the experts on the type of therapy and treatment, and also on risk assessment. In addition, neuroscience could help in making better and more effective adjustments of treatments according to the specific deviances and need of the delinquents. Neuroscientific technologies as tools to therapy could be moreover an alternative to preventive detention and lifetime surveillance, which would lastly result in avoiding an over-prediction of risk of sexual delinquents.

Besides these benefits, which rehabilitation systems could draw from neuroscience in the future, we have also seen that there are certain practical limitations to the application of neuroscience. These have to be addressed in the near future in order for neuroscience to be effectively applied. Moreover, ethical concerns are raised in this particular context, surrounding *inter alia* the concepts of privacy, autonomy and personality. As these rights must not be disregarded during the use of neuroscience, it is hence proposed to follow a scheme of informed consent to minder the constraints on the sexual offenders' rights as much as possible. At the same time, the interests of community must be safeguarded when applying neuroscience for the greater good of society. This different interest balancing will however lead to much discussion in the future.

Due to practical limitations, this paper did not analyse the psychological rehabilitative treatment in comparison with the neuroscientific approach. Therefore, it encourages to do so in future research. Furthermore, this study did not distinguish between different types of sexual offenders. It would be interesting and relevant to the underlying discussion if future research provides more insights into the heterogeneous group of sexual offenders. Additionally, it would be fruitful to conduct further research on more neuroscientific methods, in particular direct brain interventions.

Future debate on the underlying discussion might advance in several ways. Firstly, the debate could go into the question of whether neuroscience could be already applied to individuals who have not yet offended, but who have certain sexual deviances. However, the use of neuroscientific methods before any criminal activity entails a kind of self-incrimination. Secondly, a debate might ensue regarding the medical significance of sexual offences. If it were found that sexual offences are intrinsically tied to damaged brain structures or unusual brain activity, then these would constitute rather medical than legal problems. Would this then lead to a shift in the area of law? This question remains open for future discussion.

Bibliography

Arnow, B. A., Desmond, J.E, Banner, L.L., Glover, G. H., Solomon, A.; Polan, M. L., Lue, T.F, Atlas, S. W. (2002). Brain activation and sexual arousal in healthy, heterosexual males. *Brain*, Vol. 125, pp. 1014-1023.

Barrett, D. (2008). Sex offenders volunteer for 'chemical castration' drug treatment. *The Telegraph*. Retrieved from: *http://www.telegraph.co.uk/news/uknews/law-and-order/3966139/Sex-offenders-volunteer-for-chemical-castration-drug-treatment.html* .

Belcher, L. (2008). A Critical Evaluation of the Sex Offender Treatment Programmes used in Prisons. *Internet Journal of Criminology.*

Berlin, J. (1997). Chemical castration of sex offenders: "a shot in the arm" towards rehabilitation. *Whittier Law Review, Vol. 19*, No. 1, pp. 169-213.

Boyd, C. (2011). The impacts of sexual assault on women. *Australian Centre for the Study of Sexual Assault*. Retrieved from: *http://www.aifs.gov.au/acssa/pubs/sheets/rs2/rs2.pdf* .

Brooks, M. (2012). Neural Correlates of Hypersexuality in Parkison's Revealed. *Medscape*. Retrieved from: *http://www.medscape.com/viewarticle/766555* .

Buller, T. (2006). Brains, lies, and psychological explanations. In Illes, J. (ed.). *Neuroethics: Defining the issues in theory, practice, and policy.* Oxford University Press.

Canli, T. (2006). When genes and brains unite: ethical implications of genomic neuroimaging. In Illes, J. (ed.). *Neuroethics: Defining the issues in theory, practice, and policy.* Oxford University Press.

Center for Cognitive Liberty & Ethics (CCLE) (2003). *Frequently Asked Questions*. Retrieved from: *http://www.cognitiveliberty.org/faqs/faq_general.htm* .

Center for Sex Offender Management (CSOM) (2001). *Recidivism of Sex Offenders*. Retrieved from: *http://www.csom.org/pubs/recidsexof.html* .

Center for Sex Offender Management (CSOM) (2006). Understanding Treatment for Adults and Juveniles Who Have Committed Sex Offenses. Retrieved from: *http://www.csom.org/pubs/treatment_brief.pdf* .

Demitri, M. (2007). Types of Brain Imaging Techniques. *Psych Central*. Retrieved from: *http://psychcentral.com/lib/2007/types-of-brain-imaging-techniques/* .

Devlin, H. (2007). What is Functional Magnetic Resonance Imaging (fMRI)?. *Psych Central*. Retrieved from: *http://psychcentral.com/lib/2007/what-is-functional-magnetic-resonance-imaging-fmri/*

Dormont, D. et al. (2010). Neuroimaging and Deep Brain Stimulation. *American Journal of Neuroradiology, Vol. 31,* pp. 15-23.

Dreßing, H. et al. (2007). Welche Bedeutung hat die neurobiologische Forschung für die forensische Psychiatrie? *Forensische Psychiatrie, Psychologie, Kriminologie, Vol. 1,* pp. 241-248.

Dreßing, H. et al. (2008). Implications of fMRI and genetics for the law and the routine practice of forensic psychiatry. *Neurocase: The Neural Basis of Cognition, Vol. 14*, No. 1, pp. 7-14.

Easteal, P. (1994). *Voices of the survivors.* North Melbourne: Spinifex Press.

Easteal, P., McCormond-Plummer, L. (2006). *Real rape, real pain: Help for women sexually assaulted by male partners.* Melbourne: Hybrid Publishers.

Ford, P. J., Henderson, M. (2006). Functional neurosurgical intervention: neuroethics in the operating room. In Illes, J. (ed.). *Neuroethics: Defining the issues in theory, practice, and policy.* Oxford University Press.

Fortune, C.-A., Ward, T., and Willis, G.M. (2011). The Rehabilitation of Offenders: Reducing Risk and Promoting Better Lives. *Psychiatry, Psychology and Law, Vol. 19,* No. 5, 646-661.

Foster, K. R. (2006). Engineering the brain. In Illes, J. (ed.). *Neuroethics: Defining the issues in theory, practice, and policy*. Oxford University Press.

Fuchs, T. (2006). Ethical Issues in Neuroscience. *Current Opinion in Psychiatry, Vol.19,* pp. 600-607.

Greely, H. T. (2004). Prediction, Litigation, Privacy, and Property: Some Possible Legal and Social Implications of Advances in Neuroscience. *Neuroethics: The Neuroscience Revolution, Ethics, and the Law*. Retrieved from: *http://www.scu.edu/ethics/publications/submitted/greely/neuroscience_ethics_law.html* .

Greely, H. T. (2006). The social effects of advances in neuroscience: legal problems, legal perspectives. In Illes, J. (ed). *Neuroethics: Defining the issues in theory, practice, and policy.* Oxford University Press.

Gillespie, S. M. et al. (2012). Treating disturbed emotional regulation in sexual offenders: The potential applications of mindful self-regulation and controlled breathing techniques. *Aggression and Violent Behaviour, Vol. 17*, pp. 333-343.

Gordijn, B., Buys, A. M. (2010). Neural engineering. The ethical challenges ahead. In Giordano, J. J., Gordijn, B. (ed.). *Scientific and Philosophical Perspectives in Neuroethics.* Cambridge University Press.

Gordon, R. (2013). *Neuroethics Journal Club: Imaging Pedophilia.* Retrieved from: *http://www.theneuroethicsblog.com/2013/03/neuroethics-journal-club-imaging.html?spref=tw*

Gunkelman, J. D. , Johnstone, J. (2005). Neurofeedback and the Brain. *Journal of Adult Development, Vol. 12,* Nos. 2/3, pp. 93-98.

Grubin, D., Beech, A. (2010). Chemical castration for sex offenders. *BMJ, Vol. 340*, No. 74, pp. 433-434.

Hall, G. C. N. (1995). Sexual offender recidivism revisited: A meta-analysis of recent treatment studies. *Journal of Consulting and Clinical Psychology, Vol. 63,* pp. 802-809.

Hammond, D. C. (2008). What Is Neurofeedback? *Journal of Neurotherapy: Investigations in Neuromodulation, Neurofeedback and Applied Neuroscience, Vol. 10*, No. 4, pp.25-36.

Hanson, R. K., Bussière, M. T. (1998). Predicting Relapse: A Meta-Analysis of Sexual Offender Recidivism Studies. *Journal of Consulting and Clinical Psychology, Vol.66,* No.2, 348-362.

Hanson, R. K., Morton, K. E., & Harris, A. J. R. (2003). Sexual offender recidivism risk: What we know and what we need to know. In R. A. Prentky, E. S. Janus, & M. C. Seto (Eds.), *Annals of the New York Academy of Sciences: Vol. 989. Sexually coercive behavior: Under- standing and management* (pp. 154-166). New York: New York Academy of Sciences.

Hanson, R. K., Morton-Bourgon, K. E. (2005). The Characteristics of Persistent Sexual Offenders: A Meta-Analysis of Recidivism Studies. *Journal of Consulting and Clinical Psychology, Vol. 73,* No. 6, 1154-1163.

Hanson, R. K. et al. (2009). The Principles of Effective Correctional Treatment Also Apply To Sexual Offenders: A Meta-Analysis. *Criminal Justice and Behavior, Vol. 36,* No. 9, pp. 865-891.

Hornby, A. S. (1974). *Oxford Advances Learner's Dictionary of Current English*. Oxford University Press.

Illes, J., Racine, E, & Kirschen, M. P. (2006). A picture is worth 1000 words, but which 1000? In Illes, J. (ed.). *Neuroethics: Defining the issues in theory, practice, and policy.* Oxford University Press.

Joyal, C. C., Black, D. N., Dassylva, B. (2007). The Neuropsychology and Neurology of Sexual Deviance. A Review and Pilot Study. *Sex Abuse, Vol. 19*, pp. 115-173.

Klitzman, R. (2006). Clinicians, patients, and the brain. In Illes, J. (ed.). *Neuroethics: Defining the* issues *in theory, practice, and policy.* Oxford University Press.

Kobayashi, M, Pascual-Leone, A. (2003). Transcranial magnetic stimulation in neurology. *Lancet Neurology, Vol. 2*, pp. 145-166.

Kogel, C. H. de, Nagtegaal, M. H. (2006). Serious violent and sexual offenders with a mental disorder. *Boom Juridische uitgevers, WODC.*

Kutcher, M. R. (2010). The Chemical Castration of Recidivist Sex Offenders in Canada: A Matter of Faith. *Dalhousie L. K. Vol. 33,* pp. 193-216.

Levenson, J. S. (2009). Sex offence recidivism, risk assessment, and the Adam Walsh Act. *Sex Offender Law Report, Vol. 10*, No. 1, pp. 1-6.

Levenson, J. S., D'Amora, D. (2005). An Ethical Paradigm for Sex Offender Treatment: Response to Glaser. *Western Criminology Review, Vol. 6,* No. 1, pp. 145-153.

López-Ibor, J. J. et al. (2008). Transcranial magnetic stimulation. *Current Opinion in Psychiatry, Vol. 21*, pp. 640-644.

Lösel, F., Schmucker, M. (2005). The effectiveness of treatment for sexual offenders: A comprehensive meta-analysis. *Journal of Experimental Criminology, Vol. 1*, pp. 117-146.

Masterpasqua, F., Healey, K. N. (2003). Neurofeedback in Psychological Practice. *Professional Psychology: Research and Practice, Vol. 34*, No. 6, pp. 652-656.

McKenna, K. (1999). The brain is the master organ in sexual function: Central nervous system control of male and female sexual function. *International Journal of Impotence Research, Vol. 11,* No. 1, pp. 48-55.

Motiuk, L. L. (2000). The safe reintegration and risk management of violent, sex and repeat offenders in Canada. *Safe Reintegration of Sexual Offenders,* pp. 5-36.

Morse, S. J. (2011). Avoiding Irrational Neurolaw Exuberance: A Plea for Neuromodesty. *Law, Innovation & Technology, Vol. 3,* No. 2, p. 209.

Neurofeedback. (n.d.). *In NeuroRehabilitation & Neuropsychological Services, P.C.* Retrieved from: *http://thebrainlabs.com/neurofeedback.shtml* .

Pluta, R. M. et al. (2011). Deep Brain Stimulation. *The Journal of the American Medical Association. Vol. 305,* No 7, p. 732.

Raine, A. et al. (2000). Reduced prefrontal grey matter volume and reduced autonomic activity in antisocial personality disorder. *Archives of General Psychiatry, Vol. 57,* pp. 119-127.

Renaud, P. et al. (2011). Real-time functional magnetic imaging-brain-computer interface and virtual reality promising tools for the treatment of pedophilia. *Progress in Brain Research, Vol. 192,* pp. 263-272.

Rudolf Egg "Junge Sexualstraftäter. Rückfälligkeit und Gefährlichkeitsprognose" in Kammeier, H., Tondorf, G., Michalke, R. (2004). *Streben nach Gerechtigkeit: Festschrift für Prof. Dr. Günter Tondorf zum 70. Geburtstag am 8. Juli 2004*. LIT Verlag Münster.

Schiffer, B, Peschel, T., Paul, T., Gizewski, E, Forsting, M., Leygraf, N., Schedlowski, M., Krueger, T. H. C. (2007). Structural brain abnormalities in the frontostriatal system and cerebellum in pedophilia. *Journal of Psychiatric Research, Vol. 41,* No. 9, pp. 753-762.

Schiltz, K. et al. (2007). Brain pathology in pedophilic offenders: Evidence of volume reduction in the right amygdala and related diencephalic structures. *Arch Gen Psychiatry, Vol. 64,* No. 6, pp. 737-746.

Scott, C. L., Holmberg, T.(2003). Castration of Sex Offenders: Prisoners' Rights Versus Public Safety. *Journal of the American Academy of Psychiatry and the Law, Vol. 31*, pp. 502-509.

Seifert D., Jahn, K, Bolten, S., Wirtz, M. (2002). Prediction of dangerousness in mentally disordered offenders in Germany. *International Journal of Law and Psychiatry, Vol. 25,* pp. 51-66.

Seifert, D., Möller-Mussavi, S., Bolten, S. (2003). Aus dem Maßregelvollzug entlassene Sexualstraftäter. *Sexuologie, Vol. 10,* No. 1, pp. 14-20.

Sententia, W. (2004). Neuroethical considerations: cognitive liberty and converging technologies for improving human cognition. *Annals of the New York Academy of Sciences, Vol. 1013,* 221-228.

Seto, M. C. (2007). Pedophilia and sexual offending against children. *American Psychological Association.*

Shaw, E. (2012). Direct Brain Interventions and Responsibility Enhancement. *Criminal Law and Philosophy,* pp. 1-20.

Sifferd, K. (2013). *Chemical Castration as Punishment. Neuroethics & Law Blog*. Retrieved, from: *http://kolber.typepad.com/ethics_law_blog/2013/01/chemical-castration-as-punishment-by-katrina-sifferd.html*

Song, L., Lieb, R. (1994). Adult sex offender recidivism: A review of Studies. *Washington State Institute for Public Policy.*

Spalding, L. H. (1998). *Florida's 1997 Chemical Castration Law: A Return to the Dark Ages. Florida State University Law Review.* Retrieved from: *http://www.law.fsu.edu/journals/lawreview/issues/252/spalding.html* .

Stathopoulos, M. (2010). Measuring sexual offender recidivism. *Australian Centre for the Study of Sexual Assault Aware, Vol. 25.*

Steven, M. S., Pascual-Leone, A. (2006). Transcranial magnetic stimulation and the human brain: an ethical evalutation. In Illes, J. (ed.). *Neuroethics: Defining the issues in theory, practice, and policy*. Oxford University Press.

Stoléru S., Redouté, J., Costes, N., Lavenne, F., Bars, D. L., Dechaud, H., et al. (2003). Brain processing of visual sexual stimuli in men with hypoactive sexual desire disorder. *Psychiatry Research: Neuroimanging, Vol. 124*, pp. 67-86.

Thüringer Allgemeine (August 2011). *Länder beklagen Kosten der Sicherheitsverwahrung*. Retrieved from: *http://www.thueringer-allgemeine.de/web/zgt/leben/detail/-/specific/Laender-beklagen-Kosten-der-Sicherungsverwahrung-274610446* .

Vincent, N. (2012). Neurolaw and Direct Brain Interventions. Criminal Law and Philosophy.

Walsh, V., Cowey, A. (2000). Transcranial magnetic stimulation and cognitive neuroscience. *Nature Reviews Neuroscience, Vol. 1*, pp. 73-79.

Witzel, J. et al. (2008). Neurophilosophical Perspectives of Neuroimaging in Forensic Psychiatry – Giving Way to a Paradigm Shift? *Behavioural Sciences and the Law, Vol. 26,* pp. 113-130.

In how far are neurological rehabilitation methods for criminal offenders compatible with the concept of human dignity?

By Franziska Böhlke

Table of Contents

1. Introduction

In recent years neuroscience has experienced a drastic increase in popularity, driven by leaps of progress made in the field (Giordano, 2011, p. 412). By now, many interdisciplinary fields have emerged from it, two prominent examples being neurolaw and neuroethics (Shen, 2010, p. 352; Levy, 2008, p. 1). Neuroscience is nowadays also discussed in the context of the criminal justice system and may soon be used in the field of trial evidence, detecting biases in juries and judges, to make defendants competent for trial, and many other areas of the criminal law (Neurolaw: A video introduction). A very interesting field is rehabilitation of criminal offenders. The more neuroscience discovers about what is commonly termed the 'criminal mind', the more science attempts to find treatments that could help to correct deviant behaviour and reintegrate offenders into society (Greely, 2008, p. 1104). A number of direct as well as indirect methods of brain intervention are currently discussed in respect to their usefulness for this purpose. However, there are many caveats to such uses of brain intervention. This paper will deal with one of those caveats: the principle of human dignity. In this work, I wish to investigate in how far brain intervention for the purpose of rehabilitation of convicted criminal offenders is compatible with the notion of human dignity.

From the outset, it has to be clear that will not attempt a precise definition of human dignity. I, however, investigate the concept from a number of different angles in order to find possible clashes with brain intervention for the purpose of convict rehabilitation. I will first look into human dignity as a philosophical concept in historical and contemporary philosophy, and then go on to investigate human dignity as a legal claim in different national jurisprudences as well as in international law.

After that, I will offer a quick introduction to the different methods discussed for rehabilitation. Rehabilitative methods will be grouped into direct and indirect intervention. From this theoretical framework, I will discuss two issues vital to the concept of human dignity and in how far they are threatened by brain interventions, namely those of autonomy and individuality. The latter entails the concept of authenticity. This discussion prompts to look into the issue of consent. In that vein two further aspects of human dignity will be discussed, namely the subject-nature of human beings and equality.
A brief conclusion will round of the paper.

2. Human Dignity

Throughout history, the importance and understanding of the concept of Human Dignity has varied significantly (Spiegelberg, 2010, p. 42). Nowadays, there is an abundance of academic literature on this topic, the diversity of which does not allow for a uniform claim concerning its content and importance (Ibid., p. 62). This paper will not attempt to create such a claim, but rather investigate different notions and uses of human dignity in order to identify critical aspects where neurological rehabilitative methods may interfere with human dignity.

2.1. Human Dignity as a philosophical concept

2.1.1. Human Dignity – an ancient concept

The notion of human dignity relates back to ancient Rome. The original meaning of dignity (*dignitas*) referred to an attained social or political status. It was hence conceived in relation to society.[2] Cicero was the first author to mention the concept in relation to the special position human beings take in relation to the cosmos by virtue of his outstanding nature, meaning his rational capacity (Cicero and Miller, 1913, pt. 1 at 106, 109). However, the true shift from human dignity as a societal concept to an intrinsic feature of humanity can only be traced back to the Renaissance reaction to the pessimistic medieval vision of humanity. The gradual change was introduced by the famous poet and humanist Francesco Petrarca (1304-1374) (Englard, 2000, p. 1910). The notion of human dignity in the time of the Renaissance presented a struggle between religious connotations, which base the human being's position in the world on their divine creation in the image of God, and humanist ideas of writers such as Giannozzo Manetti (1396-1459), who identified a man's dignity as his creative powers (Ibid, p. 1910 – 1913).

The final transformation to dignity in its modern sense was a gradual process that took its final hold during the Enlightenment (Ibid, p. 1917). John Locke's and Samuel von Pufendorf's writings on the subject are illustrative of this shift. Locke (1632-1704) identified rational capacity, memory, consciousness, pursuit of happiness and responsibility before Divinity as the foundations of his individuality (Locke, 1690, chapter 27). Locke's contemporary, Pufendorf (1632-1694) starts his account of dignity as embodying his privileged position

2 This is of course a basic summary of the understanding of *dignitas*. For the detailed analysis see Viktor Pöschl; Der Begriff der Würde im Antiken Rom und Später. Germany: Universitaetsverlag Winter, 1989.

in a divinely created world (Raaflaub, 1974); but then refines this notion to base itself onto man's common rational nature (Englard, 2000, p. 1917). Both authors ultimately make a case for equality of all men out of their accounts for human dignity (Ibid). This connection has survived throughout the years and was taken up by the perhaps most important thinker on this subject, Immanuel Kant.

In his *Groundwork* Kant suggests the difference between relative worth on the one hand and inner worth, dignity, on the other. Everything that can be replaced by something else, hence on which a price can be put, has a relative worth. Such things serve the gratification of human needs, they exist as means to another end. Something of inherent worth, on the other hand, cannot be replaced. It exists as an end in itself. Kant goes on to saying that all rational creatures have an inner worth. From this wording alone it can be extracted that one critical criterion for a being to have a dignity, according to Kant, is rationality. He further states that rational beings have this dignity by virtue of their autonomy (Kant, 2011, p. 68f).

However, he puts forward a very complicated notion of rationality that is intensely intertwined with his teachings about morality. A good starting point is his differentiation of two realms: a sensible realm and a rational realm. Most creatures only inhabit the former. Within this, they act on their instincts in order to gratify their needs. They eat in order to satisfy their hunger. Human beings however inhabit the sensible as well as the rational realm, the latter being governed not by instincts and needs, but by pure reason, i.e. by our rationality. This is the realm in which we act according to the categorical imperative. This categorical imperative demands us to "[a]ct only according to that maxim whereby [we] can, at the same time, will that it should become a universal law." It is governed by the rationality common to all human beings and in absence of all personal inclinations, hence categorical. The categorical imperative always treats something as an end in itself, never as a means to an end and is Kant's universal moral law. It is also only in the realm of rationality that we can act autonomously. This is because when we operate in the sensible realm we act on inclinations we have not chosen in the first place. When we distance ourselves from all inclinations, however, we truly act autonomously. Still, at first glance it appears as though our freedom is very restricted within this rational realm, as we are constantly following a law, that being the categorical imperative. However, this is a law we give ourselves. When following a law of which we are the author we in fact act freely. Hence, it is clear that for Kant, the notions of rationality, morality, autonomy and freedom are virtually one and the same (Kant, 2011, p. 74 – 82).

When acting on the Categorical Imperative, we perform an act for the sake of the act itself; the act is an end in itself. According to Kant, this is also how we should treat everything that has a dignity. His second account of the categorical imperative states that a human being should never be used as a means to an end, but always as an end in itself.[3] This claim to be treated as a subject, rather than an object, is central to the principle of equality as described by Kant. Kant's conceptualisations on autonomy and self-legislation lead to a recognition of the same autonomy, freedom and respect-worthiness for all rational beings, the ultimate claim for equality (Kant, 2011 , p. 82f).

Very much in line with the general idea that human dignity is not only inherent in every individual person but also in humanity as such, Kant concedes that a person has a duty *vis-à-vis* humanity to treat him or herself with dignity. It is for this reason that he rejects the notion of suicide or consensual casual sex (Kant, 2011, p. 61).

2.1.2. Human Dignity in contemporary philosophy

Even after the substantial increase in importance the concept of human dignity experienced after World War II (Ploch, 2012, p. 897), the academic and philosophic debate is at no consensus about what it actually means (Spiegelberg, 2010, p. 53). In fact, many philosophers accept the concept as a societal good without explicit content, which makes it dependant on the cultural background and personal convictions of the beholder (Schachter, 1983, p. 849).

There are very few philosophers who have dared to attempt a qualitative definition of human dignity. It plays a central role in the writings about law and morals of the twentieth century Kantian Leonard Nelson. For him, it consists primarily in the 'capacity as a rational being to raise himself to a level of education where he can overcome practical error.'(Nelson, 1924, p. 115f) At a later point he declares that the dignity of man lies in his self-determination (Ibid, p. 358ff).

Two further noteworthy contemporary philosophers dealing with human dignity are Herbert Spiegelberg and Oscar Schachter. Spiegelberg (2010), after an impressive iteration of the historical accounts of human dignity, elaborates on two synonyms or connotational definitions of human dignity: dignity as worthiness of respect and dignity

3 It might be noteworthy at this point that Kant rejected the very notion of rehabilitation for exactly that reason, rehabilitation means using a human being for the good of society as a whole.

as intrinsic worth. He ultimately combines the two by first stating that respect implies certain non-interference with the subject worthy of such respect, i.e. the human being, and offering intrinsic worth as the characteristic of the human being which makes him worthy of it. He goes on to saying that '[h]uman dignity is the kind of intrinsic worth which attaches to a human being in his capacity of being a responsible person.'(p. 59f) He ultimately refrains from any further classification of human dignity and ultimately appeals to the philosophical debate to increase their dialogue in order to fill this seemingly empty concept (Ibid., p. 62). Schachter (1983), on the other hand, delivers a much more graspable concept of human dignity. His ultimate claim is that priority has to be given to personal beliefs, way of life, attitudes and conduct of public affairs (p. 849). His account is very interesting when considered in the context of brain interventions. He claims that 'the use of coercion, physical or psychological, to change personal beliefs is as striking an affront to the dignity of the person as physical abuse or mental torture.'(Ibid., p. 850) Schachter's very precise iteration of human dignity can be laid out to imply that respecting the dignity of a person means to respect his individuality, a connection that has already been drawn by John Locke and was later joined by Mahatma Gandhi when he said that 'it is beneath human dignity to lose one's individuality and become a mere cog in the machine.'(Attenborough, 2000, p. 23) Although Schachter fails to supply us with an analytical ground to his case of individuality, the connection between the two concepts, i.e. human dignity and individuality, does not require much technical and philosophical flexure. The Age of Enlightenment has initiated both of these concepts as a response to the depressing medieval notion of man and the human condition (Englard, 2000, p. 1917; Levy, 2007, p. 74). Further, our individuality and with it the struggle for authenticity is closely connected today's notion of individual autonomy, which in itself clearly is a pivotal point in the understanding of human dignity in both the historical as well as contemporary literature on the topic.

To summarize the discussions, it is obvious that the concept of human dignity is still highly debated. Many commentators have criticised it for being imprecise, not workable, and basically without content (Wetz, 2001, p. 311). I believe for this to be too pessimistic. There is clearly a lot of overlap among the thinkers from antiquity until today. The notion of rationality is already to be found in Cicero's writings and has survived throughout the ages, being recited by great thinkers such as Locke, Kant and Nelson. Autonomy and individuality is also mentioned throughout philosophical literature, be it directly, as for example autonomy in Kant or individuality in Schachter, or indirectly, as for example as self-determination with Nelson or non-interference with Spiegelberg. Equality is

mentioned directly by many authors, for example Locke, Pufendorf or Kant, but can also be derived from the basic fact that human dignity is something inherent in everyone merely by virtue of being human. Finally, Kant has added the subject-nature of human beings; a notion that was adopted by the academic society.

Nowadays the notion of human dignity has transcended from the philosophical sphere (Spiegelberg, 2010, p. 39). Especially in the area of criminal justice, of which rehabilitation is a part, it is important to consider its role as a legal claim. The question that has to be raised is thus whether the concept of human dignity as it is discussed in philosophy has been transposed into contemporary jurisprudence. This paper will go on to examining this by reviewing its use within the jurisprudences of the United States of America and Germany and then its use in international use, which in the same vein investigates its relation to human rights.

2.2 Human Dignity as a legal claim

2.2.1. USA

The United States of America is notable in that it is lacking any sort of concrete mention of dignity in any of its constitutional statutes. There are also very few cases in which dignity is mentioned at all, and the majority opinion on its importance has fluctuated over the centuries. However, the absence of the "intrinsic value of human beings" or human dignity should not be considered a conscious decision on the part of the writers of the American Constitution, considering the time in which it was written (Ploch, 2012, p. 923). Still, human dignity is not a widely-used legal concept in the United States, with very few rulings mentioning the concept at all (Ibid, p. 926). One of the most relevant cases regarding this topic is *Trop v. Dulles,* during which the Court cited that the Eighth Amendment had as its "basic (underlying) concept... nothing less than the dignity of a man". This was later affirmed in case law building on *Trop v. Dulles* such as *Gregg v. Georgia*.

Besides being of lesser use and impact in American jurisprudence, the legal claim of human dignity is also a lot less clear in regards to its content in American jurisprudence. In American case law we find very little evidence of such content. One case of note is *Laaman v. Helgemoe* (1977-2001), a class action civil rights claim regarding the New Hampshire State Prison's sub-adequate availability of rehabilitation programmes and sub-par living conditions for inmates. The district court referred to the Eight Amendment in its ruling, citing the importance of maintaining the physical, mental and emotional health and well-being of the inmates, and if prison conditions are inadequate and cause

cumulative damage to the inmates, their imprisonment "...does violence to our societal notions of the intrinsic worth and dignity of human beings..." and therefore violates the Eight Amendment and constitutes a 'cruel and unusual punishment'. It is interesting that on the one hand, American jurisprudence seems to adopt the Kantian notion of intrinsic worth, but on the other hand associates it with a 'societal notion'. Recently, also due to the increased concern about bioethics, there have been several efforts in the American legal and political debate to gain more clarity in respect to this concept. There are numerous essays written by several scholars who detail the concept of dignity from its ancient roots up to modern day interpretations and their application to science and justice, but since these have no strict legal value, the concept remains to be rather vague and comparatively underused. For example, Martha Nussbaum listed several criteria her essay "Human Dignity and Political Entitlements" to ensure human dignity is maintained, the most relevant to this essay being: life, bodily health, bodily integrity, and affiliation (Nussbaum, 2008).

2.2.2. Germany

In the German jurisprudence, on the other hand, human dignity plays a most noteworthy role. In article 1 of the *Grundgesetz*, the German Basic Law, it is stated that "Human Dignity shall be inviolable." Many central rights of the German Basic Law build on this strong central claim, a noteworthy example being the claim to equality of all human beings as held in the case *Life Imrpisonment* by the *Bundesverfassungsgericht* (The Federal Constitutional court, revered to as 'BVerfG' from now on). The respect for human dignity as enshrined in the Basic Law cannot be changed, it is protected by the so called *Ewigkeitsklausel,* enshrined in art. 79 of the Basic Law. The strong position of human dignity in German jurisprudence can be understood as a direct consequence of the atrocities of Nazi Germany (Gürber, 2009, p. 1).

In his influential commentary on art. 1 of the Basic Law in 1958, Günter Dürig highlights the subject nature of the human being to be the basis of human dignity as understood in German jurisprudence. His central claim is that human dignity is violated as soon as a person is objectified (Dürig, 1958, p.11):

> *'Die Menschenwuerde ist getroffen, wenn der konkrete Mensch zum Objekt, zu einem blossen Mittel, zur vertretbaren Groesse herabgewuerdigt wird.'*

The German Constitutional Court adopted and extended this in the *Life Imprisonment* case of 1977, where it stated that:

> *'It is the state's duty to respect and protect the dignity of man. This entails the notion of the human being as a moral creature with the right to self-determination and the liberty to develop its individuality freely. The individual has to accept the limits to his ability to act which the legislator has to draw in order to foster social coherence (...); still the autonomy of the person must be granted at all times (...). This means that every individual has to be considered an equal member of society with intrinsic worth. It hence contradicts human dignity to make the human being merely an object in the nation state (...). The sentence 'the human being has to be treated as an end in himself' takes unconditional effect in all fields of law; since the inalienable dignity of a human being consist out of his recognition as an autonomous person.'*

The influence of Kant is very obvious in German jurisprudence. It has hard copied several central claims of the philosophical accounts of human dignity: autonomy, individuality, equality and subject nature of the human being and the notion is known as *Objektformel* in German literature. It is interesting to note that German jurisprudence has taken over another important part of Kantian teachings about human dignity: the inalienability of one's own human dignity (Kant, 2011, p. 61). It is already hinted in the Life Imprisonment case, but was finally clarified in the case of a German stripper, who was not allowed to voluntarily strip if she could not engage her audience directly (Klug, 2003, p.143).

2.2.3. International Law and Human Rights

Although human dignity currently does not have a standing in the USA that is comparable to the one in Germany, the concept has become of utmost importance in the international legal discourse. The United Nation's use of the term "dignity" led to an adoption of its use by other bodies, catapulting its impact way beyond the boundaries of the human rights discussion (Ploch, 2012, p. 897). Nevertheless, it is exactly in that discussion though that the standing of human dignity in the international community becomes the clearest. It already appears in the Preamble of the Charter of the United Nations, where the UN states that it is 'determined' to reaffirm the faith in 'the dignity and worth of the human person'.[4] Dignity also features in Article 1 of the Universal Declaration of Human Rights,

4 There is an interesting debate on the exact difference between dignity and worth, see for example Spiegelberg 2010.

stating that '[a]ll human beings are born free and equal in dignity and rights.' Several subsequent instruments relate human dignity directly to concrete rights. A good example among them is the International Covenant on Economic, Social and Cultural Rights, which states in Article 13 that 'education shall be directed to the full development of the human personality and the sense of its dignity. Noteworthy are the Helsinki Accords. In Principle VII they sustain that participating states shall promote human rights and freedoms, 'all of which derive from the inherent dignity of the human person.' This is a very clear statement of international law concerning the relationship between human dignity and human rights. Of course, many problems with the claim of human dignity being the foundation for human rights can be found, first and foremost the fact that human dignity in itself is a concept in need of a foundation (Waldron, 2013). For the purposes of this paper, it suffices to point out the fact that this foundational claim exists and is not rejected by the international community.

The differences between the three discussed legal systems could not be greater. Where in the USA human dignity can hardly be said to play a significant role, the German jurisprudence has not only manifested it to be its perhaps most crucial constitutional value, but has also made it a workable concept. Workable parameters, like the ones found in German jurisprudence, might be exactly what internal law needs. Although it is clear that human dignity is of substantial significance there is no consensus about what exactly is meant when international law instruments speak of 'human dignity' (Waldron, 2013, p. 6). There are substantial confusions in semantics and conceptualisation of dignity (Spiegelberg, 2010, pp. 43 – 45). Generally, some scholars claim that it is merely used as a linguistic slogan (Macklin, 2003). Due to the many overlaps between different philosophers as well as between philosophers and jurisprudences, but as already mentioned, I believe that statement to be overly pessimistic. For the purposes of this paper, I will adopt workable parameters to assess the possible tensions between human dignity and neurological rehabilitative methods. An overview over the latter will be given next.

3. Neurological rehabilitative methods

With gasping leaps taken in the emerging field of neuroscience, the link between criminal behaviour and certain neurological particularities appears to become clearer (Greely, 2008, p. 1104). With a deeper understanding of the functioning of the human brain, experts speculate more and more about whether and which advances in the medical field can or should be used to rehabilitate criminal offenders in order to reintegrate them into society (Ibid., p.1103). A clear distinction has to be drawn between direct and indirect interventions. This distinction, however, is not self-evident. We are constantly moved or in fact altered due to the influence of outer circumstances, whether those circumstances are natural or another person makes, in the eyes of many scholars, no qualitative difference; either way they are out of our control (Bublitz & Merkel. 2009, p. 372). Levy identifies three characteristics by which direct interventions differ from indirect ones: direct interventions bypass rational capacities in ways indirect interventions do not, they implant an alien element that undermines authenticity, and they impose themselves over myself (Levy, 2007, p. 75). The concrete implications of these three differentiating aspects of direct and indirect intervention will be explored in part four of this paper.

As for brain interventions considered for the purposes of rehabilitation of criminal offenders I will confine myself to briefly describing three types of direct interventions used or considered for the use on criminal offenders, being neurosurgery, pharmaceuticals and Deep Brain Stimulation (DBS) and one central method classified as an indirect intervention, being Neurofeedback.

3.1. Direct Interventions

3.1.1. Neurosurgery

The most shocking account of invasive rehabilitation techniques poses neurosurgery. In 1949 Egas Moniz was awarded with the Nobel Prize for having invented the procedure known as prefrontal lobotomy (The Nobel Prize in Physiology or Medicine, 1949). This medical procedure aimed at severing the connections between the prefrontal cortex and deeper regions in the brain (Swayze, 1995, p. 507). The procedure was reported to have calmed its patients and made them more manageable; however it was also noted that the negative side effects ranged from putting the patients into a state of apathy to severe cognitive deterioration (Greely, 2008, p. 1111). Prefrontal lobotomy was eventually discredited in scientific as well as popular opinion after more than 30 000 Americans had already undergone this procedure (Greely, 2008, p. 1111). It is possibly for the reason of this

stigma that prefrontal lobotomy attached to neurosurgery as a rehabilitative method that there is very little research done in this field (Ibid p. 1112). Two notable exceptions is the research done on the destruction of the Amygdala in order to treat extreme aggressiveness and the research done in the field of addiction (Fountas & Smith, 2007, p. 710; Hall, 2006, p.1). However, both fields have experienced a drastic decline in research funding and have thus come to a halt (Greely, 2008, p. 1112).

3.1.2. Pharmaceuticals

The use of pharmaceuticals in order to rehabilitate criminal offenders is actually nothing new to the criminal justice system. The most common example is chemical castration to treat sexual offenders (Ibid., p.1106). Several American as well as European states among which Denmark, the Netherlands, Norway, Germany and Switzerland allow for this procedure, albeit with differing legal hurdles. Despite its broad use, the technique is highly criticised by experts, mostly for its off-label use of the drug Depo-Provera, which in males results in difficulties having erections and ejaculations as well as a sharp decline in sexual impulses (Greely, 2008, p. 1107). Other research for the rehabilitation of criminal offenders by means of pharmaceuticals includes anti-addiction programs, the use of anti-psychotics and several advances in the field of drug vaccines (Greely, 2008, p. 1108 – 1110 and p. 1115).

3.1.3. Deep Brain Stimulation

A very interesting field that is currently on the rise is deep brain stimulation (DBS). Unlike neurosurgery, DBS works by strengthening certain brain regions, rather than having an ultimately destructive effect. One or more thin, insulated wires containing electrodes are surgically inserted into particular regions of the patient's brain. Those wires are connected to an "implanted pulse generator", which is then implanted under the shoulder or in the abdomen. The pulse generator then sends out electrical impulses through the leads at a particular voltage and frequency, which is regulated either by the physician and in some cases can be switched on and off by the patient himself (Ibid. p, 1113). DBS has been FDA approved for several medical conditions such as Parkinson's disease, essential tremor and dystonia (Kringelbach et al.. 2007, p. 623). Of particular interest for the field of rehabilitation of criminal offenders is the research done by Angelo Franzini in 2005. He and his team used DBS in the posteromedial hypothalamus of two patients with aggressive and disruptive behaviour. Both patients were mentally retarded and had not responded to any pharmaceutical treatment. The team of researchers reported consistent improvement in respect to the disruptive behaviour in both patients at the follow-up evaluation that took place one year later (Franzini et al., 2005, p. 63). There are several brain regions that

have been linked to criminal behaviour. Most prominently among these are different parts of the prefrontal cortex, an under-activation of which is associated with a lack of impulse control. Other areas that have been associated with criminal behaviour are the amygdala, the hypothalamic-pituitary-adrenal axis, the hippocampus and the corpus callosum (Greely, 2008, p. 1115). DBS is a plausible method to strengthen certain brain regions in which under-activation is associated with criminal behaviour, or also to inhibit other regions where an over-stimulation may lead to such unwanted behaviour. With its latter use it can serve the same function as neurosurgery, with the difference that this kind of treatment would be much more adjustable, intermittent, and most importantly, reversible (Ibid., p. 1114). On the other hand, it should be noted that at least the implantation of the electrodes bears all the risks of neurosurgery, and although DBS is shown to work in many cases, scientists still cannot explain exactly why it does so (Ibid., p. 1113). Of course the use of this method is still highly speculative, but the use of DBS is spreading rapidly. At the annual meeting of the Michigan Association of Neurological Surgeons, Dr Mark Hoeprich presented his proposal of the use of DBS for the rehabilitation of criminal psychopaths (An Analysis of the Proposal of Deep Brain Stimulation for the Rehabilitation of Criminal Psychopaths). It is very likely that the near future will bring more research of DBS use to rehabilitate criminal behaviour) (Greely 2008, p. 1113 – 1115).

3.2. Non-Invasive techniques

3.2.1. Neurofeedback

The most notable non-invasive technique that has been considered by experts to be useful in the criminal field is Neurofeedback, which is a type of biofeedback derived from electrical brain activity. Biofeedback is a conditioning procedure in which patients aim to gain self-control over physiological functions that usually are not consciously perceived or controlled (Moss & Kirk, 2004, p.1). Such functions are then converted into a visual or acoustic signal which is continuously fed back in real time (Heinrich et al., 2006, p. 4). For Neurofeedback in particular, the training tries to strengthen certain brain waves, whereas it tries to weaken others, as certain brainwave patterns have shown to correlate to very specific conditions such as ADHD or impulse control impairments (Hammond, 2011, p.2). In the course of the treatment, electrodes are connected to the patient's scalp, which measure the brain activity by means of a real-time fMRI or an EEG. The patient can see these measurements on a screen and by himself attempts to acquire the mental state that is meant to be strengthened. This method is particularly attractive for children, as the desired mental state can be portrayed, and positively enforced, in the setting of a computer game (Heinrich et al., 2006, p. 4.).

By now, there are actually several rehabilitation studies out especially on juvenile offenders. Quirk (1995) stated that "a subgroup of dangerous offenders can be identified, understood and successfully treated using this kind of biofeedback conditioning program" (p. 53). In 2005, two further studies were published. Smith and Sams noted an "improvement in cognitive performance as well as recidivism"(Smith & Sams, 2005, p. 88), and Martin and Johnson stated that five out of their seven subjects 'reflected gains in aspects of flexible problem-solving, improved regulation of emotional reactions and behavior, and inhibition of inappropriate responses.'(Martin & Johnson, 2005, p. 82) All of the above studies are concerned with incarcerated juvenile offenders suffering from a variety of disorders such as ADHD. The research in this field is certain to be continued.

The great advantage Neurofeedback offers is that the offender ultimately performs the action leading to the altering of the mind himself (Eagleman, Neurolaw: A video introduction). Invasive, direct interventions are often attacked on the basis that there is an "intervener", which is not the case with Neurofeedback. The patient can at any point in time decide to stop his endeavours to attain the desired brain frequency. The technique hence poses a lot less problems in respect to his dignity.

4. Neurological rehabilitation and Human Dignity

This part will aim to examine three aspects under which the relationship between the neurological rehabilitative methods described above and human dignity as a philosophical concept and legal claim can be questioned. What will not be elaborated on is whether they could pose a threat to the rational nature of a human being. However, as this principle is so commonly found in philosophical accounts of human dignity, I would like to take this opportunity to briefly explain why not. Rationality can very basically be understood as a normative notion of reason that stipulates that rational people should come to the same conclusion given the information at their disposal. It is the capacity to conform beliefs[5] with reasons to believe and to find optimal solutions by means of reason. Interfering with this basic capacity is not considered by anyone to be the aim of rehabilitative efforts; such a thing would be a humongous perversion of our criminal justice system. It cannot be denied that such effects have arisen in the past as negative and not anticipated side effects of certain treatments, most notably of pre-frontal lobotomy, but this paper

5 The term belief is used here in the broad sense of encompassing attitudes, opinions and the like, as opposed to the narrow use of the term that only has a spiritual connotation.

concerns itself with the normative question of the relationship between human dignity and brain intervention for the purpose of rehabilitation of convicted offenders, not its practical limitations.

I will hence elaborate on four other aspects that have proven to be vital to the concept of human dignity. First I will elaborate on the issues arising in respect to the autonomy and individuality of the person. Afterwards it will be examined whether several problems arising in their relationship to neurological rehabilitative methods can be remedied by informed and effective consent, which confronts us with difficulties in respect to equality and the subject-nature of human beings.

4.1. Rehabilitation, brain intervention and autonomy

There is no all-encompassing notion of autonomy, but rather many different theories about this concept that relate to different subject matters. For my purpose, it is enough to confine myself to autonomy as a condition for moral accountability (Bubitz & Merkel, 2009, p. 361).

An interesting account of autonomy is captured by Harry Frankfurt (1971) who is a very influential figure in the field of structural theories of autonomy makes a difference between first-order desires, which are desires to perform actions, and second-order desires, which have first-order desires as their objects. A second order desire consists out of approval or lack thereof in respect to a first order desire (p. 7). On Frankfurt's account, an agent is autonomous if her first and second-order desires are in harmony. Her effective first order desire is thus autonomous if she had a second-order desire to have the first-order desire and she also wanted that second-order desire to cause her to act, the latter being a second order volition (Ibid., p, 15f). In his later paper "Identification and Wholeheartedness" Frankfurt refined and simplified this concept by stating that the crucial point is the identification of the agent with her first-order desire (Frankfurt, 1988).

There is no obvious reason why neurological rehabilitative methods, be it direct or indirect techniques, should undermine this account of autonomy. As long as the intervention does not alter only the first-order desire, but secures a harmony between both, the agent would be deemed autonomous according to Frankfurt's approach (Shaw, 2012, p. 6). Frankfurt considers only internal conditions of agency. However, there is an important shortcoming of such structural theories.

Let us consider the situation famously described by Aldous Huxley in this novel *Brave New World*. Inhabitants of that world drug all of their sorrows and ambitions away with a drug called 'soma', which leaves them in a state of constant contentment. They have very few

desires, but they identify with those. According to structural theories of autonomy, such as Frankfurt's, these people are autonomous agents even though their content is derived only from their taking of a drug. Certainly this does not correlate with society's notion of autonomy (Bublitz & Merkel, 2009, p. 363). The same shortcoming of the structural approach is thinkable in manipulative two-person scenarios, which in the context of rehabilitation of criminal offenders could well take place if the offender is either not aware of the intervention or not aware of its full consequences (Ibid., p. 365). What these cases have in common is that the identification is brought about by heteronomous intervention. Historical theories of autonomy consider this aspect by adding that autonomous pro-attitudes have to have come about by an appropriate causal chain (Ibid., p. 363f). Fischer and Ravizza (1998) have identified autonomy as a historical concept, which at its core requires guidance control. This guidance control is defined to be present when the actions of an agent are the result of her own moderate reason-responsive mechanisms (pp. 34 – 51). In the case of brain interventions such as the ones considered for the treatment of criminal offenders, the problem is not that offenders may be stripped off their reason-responsive mechanisms, in fact one could imagine treatments that enhance such reason-responsiveness, but rather that the mechanisms do not appear to be her own mechanisms, but rather the mechanisms inherent in the intervention (Bubiltz & Merkel, 2009, p. 364).

As mentioned above, many scholars deny a qualitative difference between brain intervention and more traditional ways of altering minds and behaviour such as psychotherapy or simply argumentation (Greely, 2008, p. 1134). Levy (2007) however identifies three grounds for distinction: direct interventions bypass rational capacities in ways indirect interventions do not, they implant an alien element that undermines authenticity, and they impose themselves over myself (p. 75). The two latter points are more closely connected to authenticity, hence will be elaborated on at a later stage. The first element though is an interesting starting point for this discussion.

The question how direct interventions bypass rational capacities to a greater extent than indirect interventions do is difficult to answer when looking at it in the abstract. In the context of neuroenhancement, and the argument work for rehabilitation as well, Bublitz and Merkel (2009) state that '[d]irect interventions have an immediate impact on neuronal functioning, whereas traditional interventions change personality structures slowly and more holistically. Thus neuroenhancements may bypass the 'checks and balances' of an existing personality structure.' (p. 366) Bublitz and Merkel, however, go on to making the point that many actions not considered to be questionable have this effect. One example would be the increase of ephedrine production during physical activity. This clearly circumvents the 'checks and balances' functions of cognitively mediated brain

alterations, but it does not strike us to undermine the autonomy of the person pursuing the physical activity. They make the point that direct interventions do not undermine autonomy only by bypassing the 'checks and balances' rational capacities of the mind, but rather only have this effect in combination with a manipulation, an outer infringement of the right of the patients to self-determination (Ibid., p. 366f). Such infringement could, in the context of rehabilitation, take place either by leaving the offender ignorant of the treatment or leaving him ignorant of the full consequences, as already pointed out in the previous paragraph (Ibid., p. 365). Building on the structural as well as historical approach to autonomy, the researchers conclude that a person can be deemed autonomous if she (1) has the capacity to discerning right from wrong, (2) is reason-responsive, (3) has a minimal level of self-control, (4) has a minimal understanding of the world around her, (5) has not been manipulated in the above sense and (6) identifies with her traits.[6]

What does this mean for the neurological rehabilitative methods for criminal offenders? The easiest case is Neurofeedback. Bublitz and Merkel (2009) concede that self-induced alterations of the mind never infringe the principle of autonomy (p. 367). David Eagleman also stresses the point that the patient does the actual mind training, which is essentially all Neurofeedback is, by himself. He himself has to find the state of mind to achieve the desired brain frequency (Eagleman, Neurolaw: A video introduction). He can refuse to do so at any point in time, making it highly unlikely that the training could cause disharmony between his first- and second-order desires, and even less so do they bypass his rational capacities, as he is consciously performing the training. A more challenging case is presented by direct, invasive interventions. The requirements one to four as well as requirement six, i.e. the requirement that the agent identifies with her new traits, does not appear to be problematic in this context. The interesting part is the requirement of non-manipulation. Bublitz and Merkel (2009) propose, and I agree, that an agent is not manipulated if she has 'arranged for the intervention themselves and foreseen the result.'(p. 370) Such self-initiation at first sight appears paradoxical in the case of a direct brain intervention, but it appears that it could take place in the form of consent to the treatment. Whether such effective and valid consent, on which the compatibility of invasive rehabilitation techniques and the concept of autonomy seems to depend, can in fact take place in this field will be investigated in part 4.3.

It is important to note at this point that an alteration of the patient that does not leave him autonomous results in a shift of responsibility from the manipulatee to the manipulator

6 The points not discussed in this paper, i.e. points one to four are reflected in the Mc'Naghten Insanity Denfence. Ibid p. 361.

(Ibid., p. 371). The acknowledgement and restoration of exactly this responsibility is, however, one of the very vital elements that the concept of human dignity demands of us today. It is the flipside of its requirement to allow everyone to choose their individual way of life (Schachter, 1983, p. 850). In how far the requirement of individuality can be reconciled with neurological rehabilitative methods will be discussed next.

4.2. Individuality and authenticity

With the rejection of hierarchical social axiologies after the Middle Ages came an individualised and egalitarian understanding of the person. Suddenly a meaningful life was a life that suited us, a unique life we created, rather than just fulfilling the social role we were born into (Taylor, 1991, p. 25). The Romantic Movement culminated the slow growth of this individualistic conceptualisation of the self (Levy, 2011, p. 310). The possibility of choosing our path ourselves also brought an urge to *be ourselves,* a notion that is nowadays dubbed authenticity (Taylor, 1991, p. 29). In fact it appears that the entire notion of individuality rests on a presumption of authenticity. The very point of individualization is that we choose our own paths by virtue of being who we truly are, the latter being exactly the notion of authenticity. Individuality entails more than simply differing from others. Should individuality, in the sense of differing from others, be brought about by societal coercion it would defeat the very purpose of it being a derogation from a societal axiology towards a way of life that one's own choice. Of course the ethics of individualisation and hence authenticity are a lot more demanding that the social axiology they replaced (Taylor, 1991, p. 26). One suddenly could not find ones path, or even one's life's meaning, in an outward model embedded in society, but rather had to look inside oneself in order find or create such a life (Ibid.). Nowadays the notion of respecting the individuality of a person and hence their authenticity is one of the central claims made in relation to the human dignity of a person, as elaborated in part two of this paper. Living an individual, authentic life has moved from being a mere possibility, to being an ideal (Levy, 2007, p. 74).

A very elaborate academic debate is currently taking place concerning the threat neuroenhancements pose to authenticity (Levy, 2007; Elliot, 1998; Taylor, 1991; Bublitz & Merkel, 2009). Although they mostly deal with the implications of pharmaceuticals such as anti-depressants, most of the arguments hold true for prisoner rehabilitation as well. Authenticity is basically understood as being true to oneself (Levy, 2007, p. 73). However, there are competing notions as to how this is to be achieved. One camp belongs to the essentialists, who are very much in line with the German Romantics (Taylor, 1992, p. 26). According to Taylor, a strong proponent of the essentialist view, the authentic individual

looks within in order to find her own 'measure' (Ibid., p. 28). An authentic life hence builds on a pre-existing self with set standards. Elliot (1998) builds on this conception to make a strong case against neuroenhancements:

> *It would be worrying if Prozac altered my personality, even if it gave me a better personality, simply because it isn't my personality. This kind of personality change seems to defy an ethics of authenticity.*

Bublitz and Merkel (2009) clarify what is meant by 'personality' in this context. Meant is an alteration of personality traits through which we identify ourselves, and through which others identify us. The alteration appears inauthentic if it leads either party, ourselves or others, to call us 'not ourselves anymore.' (p. 360) It is easy to see how not only Prozac, but in fact every brain intervention threatens authenticity on this account. The claim here is that even if the personality of a person is altered in a manner that makes him more agreeable, or even better from a third person perspective, which rehabilitation would certainly aim to do, they would not be beneficial from a first person perspective (which is exactly the perspective from which authenticity has to be looked at) as it brings us further from who we really are. They introduce an alien element into the individual self, hence the alterations brought about are equally alien (Elliot, 1998, p. 182). As Levy (2011) eloquently states, '[i]t might *improve* me, but it would not improve *me'* (original emphasis) (p. 315). Direct interventions stand in strong antagonism to this conception of authenticity. Psychotherapy would in such case be preferable, as it explores our inner depths; hence solutions come from within us (Levy, 2007, p.75). Interesting is the role that Neurofeedback takes in this discussion. It again has to be emphasized that the training is actually a conscious, self-initiated process (Eagleman, Neurolaw: A video introduction). It is therefore hard to make a case in order to prove that it should not be in line with this conception of authenticity.

Let us turn to the other camp of authenticity: the existentialists. This view has gained more and more popularity, also through its convincing accounts made by one of the pioneers in the field of neuroethics, Neil Levy. Levy (2007) states that what is commonly forgotten by essentialists is that there is a whole other aspect to authenticity which is arguably even more closely connected to individuality: authenticity through self-creation (p. 104). Jean-Paul Sartre states that there is no pre-existing self as conceived by essentialists. To Sartre (1955), the authentic self realizes that there is absolutely nothing that binds it to a pre-existing essence of itself (p. 57). We are hence entirely free to be whoever we want to

be. Our future self does not need to correspond to our past self, and even if it does, that rests on an implicit choice we have made not to change (Ibid.). Of course the metaphysical foundation upon which this conceptualisation of authenticity rests is rather extravagant, but the enthusiasm with which Sartre's theory was embraced even beyond the academy proves of a certain resonance to the cultural roots of the self in our society (Levy, 2011, p. 312). Degrazia (2000) states that when considering authenticity as self-creation, brain intervention might be a means to achieving authenticity, rather than defeating it (p.43).
Today's society feels deeply attracted to both accounts (Parens, 2005, p. 34 – 41). Levy (2011), however, submits that the essentialist and existentialist approach can be reconciled. Self-discovery, the essentialist path to an individual and authentic life, may require us to change (p. 315). Imagine, as an analogy, the case of a patient with Gender Identity Disorder (GID). Sufferers complain about being born into the wrong body, i.e. into the wrong gender. A GID diagnosis qualifies a patient for a gender transformation, a very invasive surgery that changes a vital aspect of a person's personality (Ibid.). It is hard to imagine a brain intervention executed for the purpose of a prisoner's rehabilitation that could be as intrusive as a sex change. Despite this very intrusive measure that brings about such change, it is easy to believe that the patient would identify with himself better after the surgery. As Levy puts it, 'the possibility of radical alteration is understood as giving us the ability to conform ourselves to what we already, essentially, are.' (Ibid.) There is no reason why this should not be the case for criminal offenders. Many feel just as disconnected from their criminal urges as someone with GID feels from his or her biological gender (Eagleman, Neurolaw: A video introduction). This conceptualisation gives us the possibility of breaking the stalemate between the essentialist and existentialist accounts of authenticity (Levy, 2011, p. 315).
From the foregoing, brain interventions do not seem to pose a problem to authenticity, and hence individuality at all. They may enable us to become what we truly were in the first place. However, Bublitz and Merkel (2009) point out an important pre-requisite for authentic change: autonomy (p.370). In fact, both the essentialists and the existentialists have adopted this criterion into their theories (Ibid.). To the essentialists, autonomy consists of the conservation and unhindered development of a self by searching within, i.e. only through internal sources. Agents are thus only autonomous of all foregoing alterations can be traced back to preceding autonomous decisions. Personally I believe that this argument doesn't withstand criticism. It leads us to go back further and further in time, eventually ending up at the hour of birth of an agent, which certainly was not his autonomous decision. Existentialists, who believe in individuality and authenticity through self-creation, presume that an agent is autonomous only if he is in control of all transformations (Bublitz & Merkel, 2009, p. 370).

The exact criteria for a brain intervention not to be in contradiction to autonomy have already been set out in the previous section of this paper. However, it is important to point out again that for a change to be autonomous despite it bypassing rational capacities the way invasive methods do, the intervention has to be self-initiated and the complete result has to have been foreseeable (Ibid.). As already stated, whether the result is foreseeable is a matter of not having any manipulating factors in the equation and of the patient being fully informed. The much more interesting question is the self-initiation in the case of invasive rehabilitative methods. Could this take place in the form of consent to the procedure? Let us turn to that question.

4.3. Can consent serve as self-initiation?

Before diving too deep into the analysis, it should be noted that the need for consent as such is assumed in this paper. Not acquiring consent to treatment would attach a punitive character to such treatment, which would be in stark contrast to the point of rehabilitation and would in fact be a gross perversion of today's criminal justice system (Bomann-Larsen, 2011, p. 76). So, in theory, rehabilitation requires consent in order to be justified. Bomann-Larsen (2011) identifies two ways through which such consent may be rendered invalid. First, it is possible that consent is not *effective* (original emphasis) under constraining circumstances (p. 66). In the context of this paper, such restraining circumstances would be the threat of incarceration or other forms of punishment except for rehabilitation. Second, there may be constraints on what a person can consent to in the first place. Consent could be *normatively invalid* even if the formal conditions for effective consent are met (Ibid.). Let us consider the former case first.

Consent is effective when it is the expression of a free choice (Ibid. p. 68). Choice restricting influences such as coercions and manipulations are controlling factors that restrict choice and thus undermine consent. Beauchamp and Faden (1986) define coercion as being present when 'one party intentionally and successfully influences another by presenting a credible threat of unwanted and avoidable harm so severe that the person is unable to resist acting to avoid it.'(p. 339) The harm presented in the case of rehabilitation of offenders is of course the prospect of imprisonment or other punitive measures. However, for the purposes of this paper, I will only discuss the situation where rehabilitation leads to a full or partial remittal of an imposed sentence, not where a refusal of rehabilitation leads to prolonging of such. At first glance there does not seem to be a qualitative difference, in both cases consent to rehabilitation leads to freedom and refusal to incarceration, but the difference will become clear in what follows. The critical question that needs to

be asked is whether we are faced with an offer of rehabilitation or whether it actually amounts to a threat of incarceration. Imagine a cancer patient who will have to either undergo a very risky surgery or die a slow death caused by his cancer. We perceive him as effectively consenting to such surgery, even though the alternative he is faced with is arguably worse than incarceration. Intuitively, one might say that the difference consists in cancer being a natural occurrence; no other person has put him into such a situation. However, it has to be maintained that the offender has made himself liable to punitive measures like incarceration by committing a crime (Bomann-Larsen, 2011, p. 68). By stipulation, he deserves his punishment and would have had to execute it either way. This is where the distinction between offering leniency in respect to a sentence that is bestowed upon him rightfully and independently and threatening to increase a sentence is crucial. In the former situation he is in no qualitative different situation than the cancer patient, as the negative alternative would occur with or without the offer of rehabilitation. Rehabilitation is thus an offer, not a threat, and hence effective. The first requirement for consent is fulfilled. Let us now consider the second requirement, whether such consent is normatively valid.

Bomann-Larsen (2011) gives a very impressive outlook on the question of normative validity by focusing on the offer, rather than the acceptance. Some offers are in themselves wrongful, not because of the circumstances they were made in, but rather because they do not recognise the offeree as a moral equal, hence it is inappropriate (p.73). To Bomann-Larsen (2011), the inappropriateness is largely dependent on the relationship between offerer and offeree. Special relations create special permissions (p. 73). However, there are also some offers that are inappropriate *tout court* as no one is in a normative position to make them(Ibid. p. 74). In the case of rehabilitation of criminals, the offerer would be the state. What citizens are answerable for *vis-à-vis* the state would determine the scope of conditions the state can appropriately offer treatment for. Not all wrongs are also public wrongs and can as such be pursued by the criminal law (Ibid.). The appropriateness constraint thus demands of treatment to be as narrowly focused on the problem and should never go beyond what is needed to correct the behaviour for which the offeree was convicted.

More problematic is the second possibility put forward by Bomann-Larsen (2011), the fact that some offers are always inappropriate because no one is in the right position to make them. I would like to apply a notion put forward by Nicole Vincent to this matter.

4.3.1. Re-shaping Virtues

In the context of rehabilitation, Vincent (2009) distinguishes re-shaping an offender's capacities and re-shaping his virtue (p. 116ff). The former aims at improving certain trades that will help the offender execute his agency. Examples such as the treatment of a lack of volition control come to mind such as a neurofeedback study which is currently trying to execute such treatment for nicotine addicts. David Eagleman, who is part of the team of researchers, is planning on extending his research to criminal offenders (Eagleman, Neurolaw: A video introduction). The latter, on the other hand, aims at reshaping the offender's values (Vincent, 2009, p. 116ff). It would be treatment designed to prevent offenders from acting in a certain manner because they would, unlike before the treatment, perceive such behaviour to be the right thing to do (Ibid.). Increasing the offender's capacity for agency is not controversial in academic literature. Such treatment does not as such alter someone's way of thinking, but only allows that person to better control himself and act in accordance with his beliefs. Increasing his virtue responsibility, as Vincent terms the latter option of re-shaping values, is however strongly objected by many academics, notably in this context Elizabeth Shaw. Her strong opinion on this matter stems from the fact that such treatment would put the authorities, and arguably society as such, on a morally higher stance than the offenders (Shaw, 2012, p. 12). This has two important implications in respect to the offender's dignity. First, it would contradict the claim to equality of human beings, for the obvious reason that portraying the offender's virtues as worthy of correction makes him morally inferiour. Second, it amounts to an objectification of the offenders. Historically, society has always tried to single out deviant groups and contrast them to the rest of the population. The creation of a 'them' and an 'us' commonly occurs between society and criminal offenders, from which society tries to distance itself (Ibid.). This objectification becomes even more severe when considering that portraying criminals as morally inferior creatures strengthens their exclusion from society. The offenders are put into a light of being 'defect' objects whose convictions have to be straightened up before they can be integrated back into society. Re-shaping virtues conveys the image that something is wrong with the offender in principle, not that he has done a wrong on a particular occasion for which he is being punished (Ibid.).

The subject-nature of the human being is one of the most central elements of the notion of human dignity in its philosophical sense and as a legal claim. The principle that all human beings are equal is derived from this notion and constitutes one of the most central elements in many legal systems. Re-shaping an offender's virtue is exactly the type of treatment for which there cannot be a normatively valid consent, as they inherently amount to a violation of the human dignity of the offender by objectifying him and

portraying him as a lesser being than the rest of society. It is exactly for this reason that consent as self-initiation has to fail as well in the context of re-shaping virtues. So we finally come to the very core of the threat to human dignity posed by brain interventions for the purpose of rehabilitation of convicted criminals: The fact that it undermines the autonomy and individuality of the offender as he is in no position to validly strip himself off his own dignity by consenting to a re-shaping of his virtues and hence allowing himself to be objectified and portrayed as an inferior moral creature.

5. Conclusion

Evidently, a full analysis of all threats that the use of neurological rehabilitative methods poses to human dignity in its variety of understandings is very hard to make. In this paper, the implications they have on autonomy and personality have been thoroughly investigated. Subsequently, it was shown in how far and under what circumstances they violate the principle of equality and the subject nature of the person. All of these four are vital ingredients of human dignity in the classic, modern and legal understanding of the term.

Summarising all of the above, at first glance it appears as though indirect neurological interventions such as Neurofeedback do not violate the dignity of offenders. As the offender himself performs a conscious training, self-initiation is granted. The changes come about very holistically and consciously, even a change in virtue of the offender is more similar to changes brought about by psychotherapy than by direct brain intervention. Also, he is constantly aware of what he is doing and there is no alien element in the equation, so manipulation could not take place. This means that as long as the offender is fully informed about the training, has given his effective consent to it, and is in no way coerced to go through with it, the offender retains his autonomy and individuality.

Direct intervention, on the other hand, is quite obviously much more problematic. They bypass the rational capacities of the offender, introduce an alien element into the procedure, and impose themselves over the self of the offender. The only situation where this does not undermine the offender's autonomy and his individuality is when he has full knowledge of the medical consequences and the process is self-initiated. However, as we have seen, self-initiation by means of consent is very limited in the context of rehabilitation. Treatment has to be targeted as narrowly as possible, and there are situations in which the offer of treatment as such would be inappropriate because it is targeted at re-shaping the virtue responsibility of the offender.

Re-shaping citizens' values results in a violation of their dignity in every case. The method used does not play much of a role here, except perhaps on the practical level. As non-invasive methods do not bypass the neurological 'checks and balances' operates they might appear to be acceptable even for the notion of re-shaping the values of convicts. However, this is a bold misconception. It has to be highlighted that it is not only the act, but already the mere offer, the very idea, of re-shaping someone's virtues that results in objectification and moral inequality, and in the case of direct interventions bars the offender from retaining his autonomy and individuality.

With the new and exciting options science creates every day, the temptation of a utopian world, where science offers quick fixes for all kinds of disturbances seems almost in reach. Humanity has undergone a lot in the creation of the vital concepts that safeguard our very essence, perhaps the most important concept being human dignity. Due regard has to be given to it at all times, even when overriding them appears to bring about many benefits. Human dignity is not a matter of degree, it is an absolute.

Bibliography

Case law:

BVerfGE 45, 187, 227f – Lebenslange Freiheitsstrafe
Gregg v. Georgia, 428 U.S. 153 (1976)
Laaman v. Helgemoe , 269 (D.N.H. 1977)
Trop v. Dulles, 356 U.S. 86 (1958)

Books and articles:

An Analysis of the Proposal of Deep Brain Stimulation for the Rehabilitation of Criminal Psychopaths, Retrieved from: *http://www.destinationmi.com/documents/2011MANS presentation_ MarkHoeprich.pdf*

Attenborough, R. (2000). *The Words of Gandhi.* USA: Newmarket Press.

Bomann-Larsen, L. (2011). Voluntary Rehabilitation? On Neurotechnological Behavioural Treatment, Valid Consent and (In)approrpiate Offers. *Neuroethics 6,* pp. 65-77.

Bublitz, J.C. & Merkel, R. (2009). Autonomy and Authenticity of Enhanced Personality Traits. *Bioethics 23 (6),* pp. 360-374.

Cicero, M. T. & Miller, W. (1913). *Cicero. De Officiis. With an English Translation by Walter Miller.* London: Macmillan Company.

de Baets, A. (2007). A successful Utopia: The Doctrine of Human Dignity. *Historein Vol. 7: History and Utopia,* pp. 71-85.

Dürig, G. (1956). Der Grundrechtssatz von der Menschenwürde. S.L.: *Archiv fuer oeffentliches Recht 2*, pp. 117-157.

Englard, I. (2000). Human Dignity: From Antiquity to Modern Israel´s Consitutional Framework. *Cardozo Law Review,* Vol. 21:1903, pp. 1903-1927

Frankfurt, H. (1969). Alternate possibilities and moral responsibility. *Journal of Philosophy 66,* pp. 829-83.

Frankfurt, H. (1971). Freedom of the Will and the Concept of a Person. *Journal of Philosphy 68(1),* pp. 5-20.

Fischer, J. & Ravizza, M. (1998). *Responsibility and control: A theory of moral responsibility.* Cambridge: Cambridge University Press.

Fountas, K. & Smith, J. (2007). Historical Evolution of Sterotactic Amygdalotomy for the management of Severe Aggression. *106. J. Neurosurg,* 710, pp.710-713.

Giordano, J. (2012). Unpacking Neuroscience and Neurotechnology-Instructions not Included: Neuroethics Required. *Neuroethics 6,* pp. 441-414.

Greely, H. (2008). Neuroscience and Criminal Justice: Not Responsibility but Treatment´. *Kansas Law Review 56*, pp. 1103-1138.

Gürber, R. (2009). *Menschenwürde- Herzstück der Menschenrechte.* VKAS/ AGEAS Tagung vom 07./08.11.2009 in Quarten.

Hall, W. (2006). Stereotactic Neurosurgical Treatment of Addiction: Minimizing the Chances of Another "Great and Desprerate Cure.". *Addiction 101*, pp. 1-3.

Hammond, D., C. (2011). What is Neurofeedback: An Update. *Journal of Neurotherapy: Investigations in Neuromodulation, Neurofeedback and Applied Neuroscience,* 15:4, pp. 305-336.

Heinrich, H., et al. (2007). Annotation: Neurofeedback – Train your brain to train behaviour. *Journal of Child Psychology and Psychiatry 48 (1),* pp. 3-16.

Kant, I. (2011). *Grundlegung zur Methaphysik der Sitten.* In Werke, Band IV - Schriften zur Ethik und Religionsphilosophie [Ed.7] as compiled by Wilhelm Weischedel, Wiesbaden: Insel Verlag.

Heinz Klug, H. (2003). Symposium Article: The Dignity of the Montana Constitution: May Foreign Jurisprudence Lead the Way to an Expanded Interpretation. *Montana Law Review,* pp. 133-156.

Kringelbach, M., et al. (2007). Translational principles of deep brain stimulation. *Nature Reviews Neuroscience,* pp. 623-635.

Levy, N. (2007). *Neuroethics: Challenges for the 21st Century.* Cambridge: Cambridge University Press.

Levy, N. (2008). Introducing Neuroethics. *Neuroethics 1*, pp. 1-8.

Levy, N. (2011). Enhancing Authenticity. *Journal of Applied Philosophy 28 (3),* pp. 308-318.

Locke, J. (1690). *An Essay Concerning Human Understanding (Great Books of Philosophy. S.L.:* Prometheus Books.

Ruth Macklin, R. (2003). Dignity is a useless concept: It means nothing more than respect for persons or their autonomy. *BMJ*, pp. 1419-1420.

Martin, G. & Johnson, C. (2005). The Boys Totem Town Neurofeedback Project: A pilot study of EEG biofeedback with incarcerated juvenile felons. *Journal of Neurotherapy, 9(3),* pp. 71-86.

Martini, S. (s.d.) Die Formulierung der Menschenwürde bei Immanuel Kant´. Retrieved from: *http://akj.rewi.huberlin.de/projekte/seminararbeiten/marini2.pdf*

David Eagleman: The Brain and The Law: Neurolaw: A video introduction, Retrieved from: *http://www.neulaw.org/*

Nelson, L. (1924). *System der philosophischen Rechtslehre und Politik*. Gottingen: Offentliches Leben.

Nussbaum, M. (2008). Human Dignity and Political Entitlements. Washington DC.

„The Nobel Prize in Physiology or Medicine 1949". Nobelprize.org. Nobel Media AB 2013. Retrieved from: *http://www.nobelprize.org/nobel_prizes/medicine/laureates/1949/*

Ploch, P. (2012). Why Dignity Matters: Dignity and the right (or not) to rehabilitation from international and national perspectives. *International Law and Politics,* Vol. 44:887, pp. 887-948.

Douglas Quirk, D. (1995). Composite biofeedback conditioning and dangerous offenders: III. *Journal of Neurotherapy 1(2),* pp. 44-54.

Raaflaub, K. (1974). *Dignitatis contentio.* C. H. München: Beck´sche Verlagsbuchhandlung.

Schachter, O. (1983). Human Dignity as a Normative Concept. *The American Journal of International Law 77 (4),* pp. 848-854.

Shaw, E. (2012). Direct Brain Interventions and Responsibility Enhancement. *Criminal Law and Philosophy,* pp. 1-21.

Shen, F. (2010). The Law and Neuroscience Bibliography: Navigating the Emerging Field of Neuroscience. *International Journal of Legal Information 38,* pp. 352-434.

Smith, P. & Sams, M. (2005). Neurofeedback with juvenile offenders: A pilot study in the use of QEEG-based and analog-based remedial neurofeedback training. *Journal of Neurotherapy, 9(3)*, p. 87-99.

Spiegelberg, H. (2010). Human dignity: A challenge to contemporary philosophy. *World Futures: The Journal of Global Education 9 (1-2),* p. 39-64.

Swayze, V. (1995). Frontal Leukotomy and Related Psyhosurgical Procedures in the Era Before Antipsychotics (1935-1954): A Historical Overview'. *American Journal of Psychiatry 152*, pp. 505-515.

Vincent, N. (2009). Responsibility: Distinguishing Virtue from Capacity. *Polish Journal Philosophy 3(1),* pp. 111-126.

Waldron, J. (2013). Is Dignity the Foundation for Human Rights? [Working Paper]. *Public Law & Legal Theory Research Paper Series No. 12-73,* pp. 1-29.

Wetz, F., J. (2001). Die Würde des Menschen-Ein Phantom?. *ARSP*, Vol. 87, pp. 311-327.

Probation and effective rehabilitation – an alternative to incarceration?

Using neuroscience to facilitate rehabilitation methods

By Janika Bockmeyer

Table of Contents

1. Introduction

According to Jeffrey Rosen (2007), the breakthrough of neuroscience having an impact on law can be dated back the case of Weinstein from the early 1990s. The case dealt with a 65-year-old man who had brutally killed his wife and had thrown her out of the window afterwards to make it look like suicide. His lawyers then suggested not to hold him guilty due to a mental disorder namely an abnormal cyst, which encompassed the brain "like a spider web" (Rosen, 2007). In the end, Weinstein was held guilty but his charge was reduced to manslaughter only. This case therefore implies what value neuroscience might add to law in the future.

In recent years critiques of collective sentencing and imprisonment have gained importance. Alarming numbers of overcrowded prisons and extraordinary high rates of recidivism have drawn attention towards legal proceedings and the imposition of sanctions and sentences. Moreover, assessments of forensic psychologists appeared to be of rather less accuracy in terms of predicting the propensity of a perpetrator to re-offend. At the same time, the field of neuroscience has experienced significant progress in exploring our brains and the connection to our minds. More precisely, the research on correlations between specific brain functioning and appertaining human behaviour has remarkably advanced in recent years. Certain methods have been developed allowing for brain imagining and lie detection to a certain extent. For this reason, the field of 'neurolaw' has emerged with emphasis on the impact of neuroscience on law. Proponents of the latter suggest that neuroscience may serve as evidence to support solving questions of guilt and punishment and help to advance the forecast of future criminal behaviour. Especially in the light of emerging neuroscientific findings both legal and neuroscientific scholars have argued for a reform of the justice systems towards more individualized litigation and a greater focus on rehabilitation instead of incarceration.

Conventional rehabilitation methods such as occupational and psychological therapies, however, largely experience scepticism and reluctance among the public and policy makers due to rather unsatisfactory results (Chen & Shapiro, 2007). Nevertheless, a growing number of scholars believe that neuroscience may indeed add value to rehabilitation methods and thus improve general results. Although imprisonment brings along some problematic issues, according to the German Ministry of Justice (2013) probation as an alternative also shows its limitations namely high rates of recidivism of those who were out of prison based on probationary sanctions. Nevertheless, linked to effective

neuroscientific rehabilitation methods probation could gain more popularity. Therefore, the question arises in how far could probation be an alternative to imprisonment when linking it to effective neuroscientific rehabilitation methods? To which degree would this be legally and ethically acceptable?

To answer those questions, this paper briefly assesses the debate of imprisonment and recidivism first in relation to retributivism of society. Then brain disorders and different methods of rehabilitation are discussed in particular real-time fMRI neurofeedback. The following section elaborates the concept of probation and its potential as alternative to incarceration if linked to neuroscientific rehabilitation. The paper concludes with explaining the problematic nature of certain rehabilitation measures and an assessment of legal and ethical issues.

2. Punishment: The Problem of Incarceration

When a person commits an offence of unlawful character he or she will be punished for doing so. This concept of punishment is utterly accepted by society and every child is raised and educated with regard to it. Nevertheless, due to the fact that offences may be of different severity punishment has to be appropriate as well as proportional. At present times, however, overcrowded prisons imply potential overcriminalization of offences. Additionally, incarceration bears high costs for both prisoners and society.

2.1. Criminal Law Theory: Retributivism and Society

Criminal justice practice involves three major components: criminalization, enforcement and punishment. Criminal law theory addresses all three components. This paper, however, focuses only on punishment as component. There used to be two main approaches to justify punishment that are commonly recognized: retributivism and utilitarianism, which may be considered as the best-known version of the consequentialist theory (Brown, 2012). Retributivist adherents believe that punishment needs to be imposed on the offender because he or she deserves it. Retributivism may also be called the 'agent-relative' doctrine because it requires that the perpetrator's culpability alone determines the degree of punishment. As Moore (2010) states: "[f]or a retributivist, the moral responsibility of an offender also gives society the duty to punish" (p.90). Moreover, supporters believe in a broader authority of the state meaning a wider range of obligations than it is believed

in other approaches. In this context, only the punishment by the state for wrong actions meets offender and victim with respect. In fact, these norms rather than the aim to improve public order and safety are the state's main goals for enforcing criminal law.
Opposed to this, adherents of the utilitarian theory believe that punishment needs to have the goal of improving the safety and well being of society as a whole (Gruber, 2010). Utilitarianism as an approach focuses only on the consequences of an action and does not consider the intrinsic character of a plot itself. In the context of this theory, it is assessed whether criminal punishment in fact has a net benefit for the society. To be more concrete, punishment is regarded as tool used "to deter, rehabilitate, or incapacitate, so its form should be designed to serve those goals" (Brown, 2012, p.74). In other words, utilitarians perceive punishment as an instrument designed particularly to foster a specific aim. Jeremy Bentham was the first to clearly formulate utilitarianism as a theory with the purpose that society should maximize its utility meaning the minimization of aggregate pain and maximization of aggregate pleasure (Bentham, 1970). Moreover, he was the first to unambiguously state that prevention of criminal acts shall be given priority: "general prevention ought to be the chief end of punishment, as it is its real justification" (Bentham, 1970, p.396).

In the beginning of the twentieth century retributivism lost importance and was neglected as a legitimate goal of the society. However, towards the end of the century the retributivist theory gained attention again and finally established itself as the dominant theory of criminal prosecution. Jean Hampton (1991-1992) stated that "[t]here has been a steady rise in the popularity of retributivism over the last decade, which is surpassing given its near death in the 1950 and 1960s" (p.1659). Several courts incorporated retributivism in their criminal jurisdictions and several states of the US even adopted retributivist features in their penal codes.

Generally speaking, criminal law with distinction to civil law has a rather harsh character as it defines offences by the fault that has been done and the blameworthiness. Nevertheless, many observers argue that especially the Anglo-American codes "over-criminalize, meaning that statutes label conduct as criminal that should not be so label[l]ed because the conduct is not sufficiently harmful and wrongful, and committing it does not manifest culpability" (Brown, 2012, p.29). Although it spread in recent years, the Problem of overcriminalization as presented in the United States appears differently and more limited in Europe. However, "European countries are increasingly creating crimes that prohibit conduct well before it causes harm" (Molina, 2011, p.127). Despite that, in

continental Europe the principle of culpability is largely respected meaning a person may only be punished if he or she is culpable. In this context, Europe does not have those strict liability offenses and therefore less overcriminalization. Moreover, continental Europe's law system prevents offenders from being punished twice for the same perpetration (Molina, 2011). Neuroscience with its new findings will certainly have an impact on law meaning in particular "people's moral institutions about free will and responsibility" (Green & Cohen, 2004, p.1775). In this context, "a shift away from punishment aimed at retribution in favo[u]r of a more progressive, consequentialist approach to the criminal law" may be expected (p.1775).

In a general sense, the criminal justice systems around the world show retributivist features. Moreover, retributivist principles have a powerful moral and political appeal (Tonry, 2004). This is because the society largely belongs to the libertarian notion of free will, which is perceived as partly integral of human dignity (Green & Cohen, 2004). Libertarian free will implies that human beings are free agents being alone responsible for their actions and free from any determination or constraints. Moreover, all "free will theists" believe that libertarian freedom is crucial for moral responsibility. Therefore, Libertarian freedom is the freedom to act on own accounts with sufficient control of one's nature, predisposition and desires, such as pride and jealousy. Responsibility, in this regard means that one had a free choice and could have acted differently (Clarke, 2003). As Kant (2002) already put it in his work, punishment must be adequate and proportionate to a person's internal 'wickedness' rather than only serving future social welfare (p.179). Nevertheless, it has to be stated that retributivism as it was 20 years ago is hardly any existent. The emphasis shifted also towards safety and well being of society instead of just punishment because the offender deserves it. At present times, the common legal approach towards criminal law does indeed show also utilitarian elements whereas a retributivist tenancy and the aim compensate the victims of an offense with 'adequate' punishment prevail (Green & Cohen, 2004). In that context, the persistent criminal law theory could be described as utilitarian retributivism. Certainly, there is always the tension between sanctions that work and sanctions that hurt (Cullen & Gendreau, 2001). Nevertheless, prosecution has rather departed from the purely punitive approach towards a more societal approach.

2.2. Incarceration and Recidivism: A Complex Relation

The term rehabilitation is rather broadly used meaning any form of treatment after criminal offences. Incarceration in this sense is the most common method used to treat

criminal offenders (Eagleman & Flores, 2012). There is a common presumption that one more offender in prison is one less criminal person in the streets (Cullen, 2006). For a long time now, it is generally accepted and believed that incarceration as such is a deterrent tool against potential criminal offenders. More specifically, deterrence theories assume that harsh punishment is likely to reduce future individual criminal behaviour. In this context, the public as well as policy makers are largely convinced that serving a prison sentence has "powerful deterrent effects" (Gendreau et al., 2000).

In recent years there has been a substantial increase in the number of prisoners, which suggests that imprisonment is the most frequently used form of sanction to incapacitate offenders (Drago et al., 2011). Certainly, this is also linked to overcriminalization as mentioned earlier. For example between 2001 and 2010 the number of prisoners increased by 41,54 percent in the France, by 21,9 percent in Italy, by almost 4 percent in the Netherlands and by 26,7 percent in the United Kingdom. In the United States between 2001 and 2009 the numbers of prisoners increase by 17,2 percent (Eurostat, 2009). However, imprisonment as treatment bears non-neglectable physical costs and social consequences for the respective perpetrator and society. For instance, serving a prison term largely implies the destruction of social contacts and personal relationships of the perpetrator. Moreover, it may damage employment opportunities and thus takes away the basis for an independent life after prison. On top of this, certain studies suggest that serving a prison term implies the production of more criminality. In this context, the data of specific research clearly suggests that prison, as deterrent tool, is not supported. Rather the opposite is assumed meaning that prisons can be perceived as "schools for crime" (Gendreau et al., 2000, p.4). This is mainly because being in jail means being introduced to other criminal offenders with diverse criminal backgrounds. It is not unusual that criminal circles are being established in prison, which may lead to potential drug use and membership of gangs both during the time of sentence and afterwards (Eagleman & Flores, 2012). For this reason, Cullen (2006) argues that all policies supporting this common presumption that offenders in prison do not pose a grave risk by recklessly distributing short and long prison sentences are "destined for failure" (p.667).

Sutherland (1939) already identified that "the essential reason why a person becomes criminal is that they have been isolated from the culture of the law-abiding group" (Sutherland, 1939, p.595). In that sense, criminality is the by-product of isolating a person from culture and therefore, imprisonment as it implies even greater isolation is not likely to enhance this problem. Early criminologists understood that any treatment could only

be successful and effective if the target is the cause of recidivism (Cullen & Gendreau, 2001). Already Sutherland (1939) recognized that "a large proportion of the offenders under the care of any agency are recidivists" (p.585). Moreover, he suggested that harsh punishment of the offender is ineffective because it does not address the cause that once produced the criminal act. In their study of 2007 on harsh prison conditions and recidivism Chen & Shapiro (2007) made use of individual-level data and found evidence that being incarcerated in higher security prisons indicates that after release the person shows a higher tendency to re-offend. On top of this, Drago, Galbiati & Vertova (2011) argue that "[a]lthough it is to some extent a popular view that being tough on inmates can "rehabilitate" them, we do not find evidence supporting the idea that harsher prison conditions reduce recidivism" (p.107). The empirical analysis based on the Italian Department of Prison Administration (DAP) rather showed that harsh prison conditions in general increased rates of recidivism. In particular the growth in numbers of deaths among Italian prisoners revealed a significant increase in re-offences of fellow inmates (Drago, Galbiati & Vertova, 2011).

The Office of Justice Programs (OJP) belongs the United States Department of Justice. It is an agency with emphasis on crime prevention through research and development. The graph below withdrawn from their database illustrates the rates of re-arrests of offenders released in 1994 from prisons in 15 States of the United States.

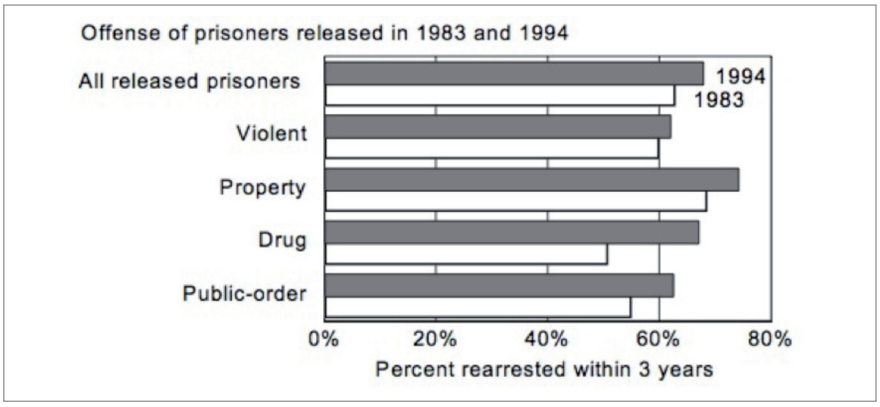


Figure 1: Re-arrests of prisoners released in 1983 and 1994 in the US; Source: Langan & Levin (2002), Special Report of the Bureau of Justice Statistics, June 2002

In 1994 almost 300 000 prisoners were release in 15 States of the US. Of those a percentage of 67.5 was rearrested within three years after the first criminal act. In comparison, a similar

study of 1983 estimated only 62.5 percent of rearrests among the released. The highest rates of rearrests as illustrated in the graph can be found among property offences with more than 70 percent. Nevertheless, drug and public order offences as well as violent acts show also high numbers above 60 percent. The record makes clear that a considerable amount more than half of the released offenders in those 15 States repeatedly committed a criminal offence within three years after release. Generally speaking, this implies how serious the problem of recidivism in fact is. Those findings are supported by the study of Spohn & Holleran (2002) where they found no evidence that incarceration lowers the rates of recidivism of drug offenders in any kind. Instead the opposite had to be noted stating that offenders who had to serve a prison term generally showed higher rates of re-offending (Spohn & Holleran, 2002).

In Europe on the contrary, general rates of recidivism depend on the effectiveness of the different systems for criminal justice. So far there is little harmonization and standardization achieved between the European member states. Nevertheless, there are several common features that can be discovered as identified by the European Sourcebook of Crime and Criminal Justice Statistics. The most important predictors of re-offending are past criminal offences meaning "the highest rates being from offenders with the longest criminal history" (Aebi et al., 2010, p.294). However, no relationship could be identified between the degree of seriousness of an act and rates of recidivism. In Germany, for instance, a study from 2004 until 2007 of the Ministry of Justice showed that of all incarcerated people in that time frame only 33,7 percent re-offended within three years after release (German Ministry of Justice, 2013). According to the Ministry of Justice of the United Kingdom (2013), England and Wales have recidivism rates of only 26,9 percent in terms of an offence committed within one year after release with a court conviction. These numbers do not appear particularly high, however, they include all offenders regardless whether they were imprisoned in the first place or convicted differently. In other words, these numbers incorporate any offences of civil law, public law and criminal law and therefore the propensity of being incarcerated is usually rather low in particular with the former two kinds of offences. The Dutch Ministry of Justice (2013) recently published a study on recidivism rates in the Netherlands with distinction to whether were incarcerated or punished otherwise. These numbers appear to be much more expressive as almost 50 percent of former inmates re-offended within two years after release.

Already Sutherland (1939) favoured correctional policies that would facilitate social contacts of offenders. For that reason, he suggested probation as tool of 'punishment' in

order to allow and support perpetrators to establish social relations within the society rather then isolating them. Furthermore, he argued to make use of parole, which is similar to probation but only begins after a certain period of serving a prison term. In this context, the most beneficial option would be individualized 'punishment' serving as an effective tool due to the fact that imprisonment alone does not necessarily protect society in the future (Sutherland, 1939). This "leads to the inescapable conclusion that, when it comes to reducing individual offender recidivism, the [successful method] is appropriate cognitive-behavioural treatments which embody known principles of effective intervention" (Gendreau et al., 2000, p.4).

3. Brain Disorders and Different Methods of Rehabilitation

Any kind of obstacle that hinders a person to think and decide freely and independently may be considered as mental or brain disorder. Such disorders clearly affect a person's ability to choose and thus it has also an effect on the process of making decisions (Buchanan, 2000). Consequently, "this could indeed explain why [it] undermine[s] our capacity – at least in some instances – to conform our conduct to the requirements of the law" (Meyen, 2013, p.94). Kalis et al. (2008) identified three different stages in the decision-making process of a person. First, options are generated, second, options are selected and third, the action is initiated. Each of those different phases of decision-making can potentially be affected by mental disorders.

Recent studies showed evidence that people with psychopathic tendencies have some kind of attention-deficit disorders. It appears that those people have difficulties to identify reasons against performing a certain action (Shaw, 2012). Furthermore, Breiter et al. (2001) explain a relation between a dysfunction of neural mechanisms and impulsive behaviour. The former can potentially lead to different impulse disorders, such as abuse of drugs and gambling issues. ADHD patients[7], as another example, showed correlations with taking higher risks when making decisions (Meyen, 2013). Those capacity deficits may be potential targets for neuroscientific rehabilitation methods. More precisely, neuroscientific methods, such as fMRI Neurofeedback, might facilitate a "better distinction between those cases in which a mental disorder is related to an increased risk of recidivism and those in which

7 ADHD stands for 'adult attention deficit hyperactivity disorder'

this is not at all the case" (Meyen, 2013, p.97). The following sections outline the challenges for neuroscientific rehabilitation measures and specific methods, which may be used.

3.1. The Problem of Awareness

A brain injury through whatever reasons may cause impaired self-awareness and harm rational decision-making. Nevertheless, the absence of awareness of brain disorders does have consequences for patients participating in rehabilitation programs and thus their well-being. This is in particular because the main goal of rehabilitation methods is to address and eliminate disorders in the brain of the patient. More precisely, "[a]wareness is an important issue to address in brain injury rehabilitation" (Fleming & Ownsworth, 2006, p.475). Therefore, it is essential to improve self-awareness among offenders suffering brain abnormalities. Furthermore, self-awareness has to be achieved prior to cognitive rehabilitation measures in order to make the latter effective and successful (Simmond & Fleming, 2003). In this context, the terms awareness and disorders in the brain have to be defined. Although, there is no unilateral definition of it, Prigatano and Schacter (1991) put emphasis on self-assessment describing awareness as "the capacity to perceive 'self' in relatively 'objective' terms while maintaining a sense of subjectivity" (p.13). The term disorder on the other hand was clarified by Barco et al. (1991) as "inability to recognize deficits or problem circumstances caused by neurological injury" (p.129). As mentioned earlier, there is a need for awareness interventions as part of the rehabilitation programmes, which are individually tailored for the clients. The difficulty, however, lies in the fact that awareness as such cannot be measured in numbers but rather has to be inferred (Simmond & Fleming, 2003). On top of this, there has been almost no critical analysis of the issue, although it is generally understood that neuroscientific rehabilitation methods are profitably for the clients and society.

Prigatano (1991) states that a difference could be detected between patients who successfully completed the rehabilitation program and patients who abandoned the treatment. The former showed a good self-awareness, whereas the latter showed discrepancies between the ratings of members of staff and their self-awareness. These results shows two things: first, that self-awareness plays a significant role in the results of effective rehabilitation and second, that if there is no self-awareness effective rehabilitation may not have the desired impact on the client. In other words, the general problem of self-awareness lies in its character of being a "substantial barrier to successful rehabilitation outcome" (Prigatano, 1999, p.146).

3.2. Rehabilitation and Correctional Measures

As mentioned earlier, incarceration as 'rehabilitation method' portrays a rather problematic issue bearing decisive costs and consequences. Opposed to this common presumption that offenders are harmless when being in prison there is also the common wish that imprisoned perpetrators would leave prison less criminal than when they were incarcerated in the first place (Cullen, 2006). Rehabilitation in terms of enhancing criminal offenders still holds a rather unpopular position because it potentially brings benefits to the perpetrators and thus is considered as 'welfarist' (Gruber, 2010). This was brought to a point where voices even called for a permanent incapacitation and thus incarceration to ensure the safety of society. This is why utilitarian adherents used deterrence as their crucial justification for punishment.

Many rehabilitation methods can be considered as progressive concept in line with the "culture of control" by Garland (2001), which describes current legal systems as more adaptive towards social control. Moreover they can be said to compete with the system of 'mass incarceration' (Nagin et al., 2006). The challenge therefore is to bridge the gap between simple incarceration and policies for correctional practice. In this context, however, a clear vision of rehabilitation and correctional practices is needed, in particular because the range of possible corrections differs immensely in terms of severity of the intervention (Cullen, 2007). For this reason, it's the task of criminologists to establish clear guidelines for rehabilitative and correctional measures.

Most incarcerated people show severe difficulties in controlling their impulses. However, it would be wrong to assume that those people are actually oblivious to what is best for them. It is rather the frontal lobe that is responsible for long-term considerations and that is sometimes powerless against short-term desires and urges treated in the amygdala (Eagleman & Flores, 2012). Hence, it is essential to bear in mind that "the brain operates like a team of rivals" (Eagleman & Flores, 2012, p.165). In that sense, it is important to include the latest scientific findings and incorporate them into the justice system in order to better understand what actually is taking place in the minds of criminal offenders. Cullen & Gendreau (2001) even suggest reducing punishment and harming while giving offenders certain rights in order to keep them in the community by any means.

The task therefore is to create a system providing for individual risk assessment, such as a 'neurocompatible criminal justice system'. While no system should treat similarly

situated individuals differently fair systems should not fail to account for the differences" (Eagleman & Flores, 2012, p.166). Up to the present individual risk assessment has indeed taken place already. For instance differences are made between adults and adolescence due to their different stages in the development of controlling impulses. Moreover, there is made a distinction between crimes of passion and premeditated offences. Nevertheless, the problem of recidivism is not yet solved as mentioned earlier and therefore, individual risk assessment has to be extended. In particular because reports on criminal offenders often differ dramatically from how they actually behave once released. For instance, when assessing the condition of sex offenders before being released "psychiatrists and parole board members had the same predictive accuracy as coin-flipping" in determining whether a person would be likely to re-offend (Eagleman & Flores, 2012, p.167).

Recently, evidence has grown that rehabilitative treatments for offenders in statistical terms do in fact reduce the recidivism rates. In addition it has been found that punishment-oriented treatments for offenders are rather ineffective and do not improve the numbers of re-offenders (Cullen, 2006). Furthermore, the expense factor of rehabilitation treatments is decisive due to the fact that – according to a growing body of evidence – the latter is much more cost effective than conventional punishments, such as incarceration. Nagin et al. (2006) conclude that especially with emphasis on juvenile offenders the threshold to an approach of rehabilitation, which is in particular public reluctance, is more presumed than fact.

Neuroscientific rehabilitation methods do include occupational therapies, psychological therapies and brain interventions. Occupational therapies in that sense imply the use of treatments geared to further develop, recover, or maintain daily routine and working skills of patients suffering from mental disorder. Psychotherapy implies therapeutic treatment aiming at an increase in the sense of well-being of the patient. Finally, brain interventions can be performed through different methods, such as electric stimulation, pharmaceuticals and surgery. This paper, however, focuses on real-time fMRI neurofeedback only and its potential contributions to the criminal justice systems. Some scholars describe real-time fMRI neurofeedback as non-invasive (Caria et al., 2012), whereas others such as Greely (2012) define all sorts of behavioural treatments as brain interventions. This paper assumes real-time fMRI to be among non-invasive behavioural treatments and thus no brain intervention as such.

3.3. Real-Time fMRI Neurofeedback

Real-time functional magnetic resonance imaging (fMRI) is a method, which detects changes in hemodynamic properties of the brain meaning changes in the blood-oxygen level- dependent (BOLD) in relationship with engagement in different mental tasks. Real-time fMRI or fMRI Neurofeedback allows the observer to measure brain activity while the person being tested is addressing different tasks. This subsequently allows the experimenter to non-invasively study the effects of brain activity on behavioural characteristics by choosing specific tasks for the person to be tested (Caria et al., 2012). Prof. Dr. Goebel, professor for cognitive neuroscience at Maastricht University, described real-time fMRI as close to meditation in which they found which parts of the brain visualize different states of mind.[8] It therefore can be described as visualization of task-related brain activity or as 'frontal lobe workout' according to Eagleman and Flores (2012). Presumably, the most striking application of neurofeedback is the possibility to take possession of "volitional control of localized brain activity using real-time fMRI [...] protocols" (Caria et al., 2012, p.487).

Certainly, an individual study for each patient has to be designed in advance defining the physiological target and response and thus the study depends on the behavioural effects that are desired. The procedure works as follows: a patient lies in the scanner and receives online information through a screen on how active a particular part of his or her brain is at the moment. The delay of the protocol as such is only a few seconds and thus there is almost real time transfer of data from the person's brain, namely from the artifact detection to the estimations of activation. The neurofeedback then is projected on the screen and can be directly observed by the patient. The latter then attempts to control the activation in the targeted brain area by using different mental strategies. These can include "anything from simple finger tapping to mental imagery or complex cognitive tasks" (Sulzer et al., 2012). Neurofeedback as such is usually presented as 'thermometer display' or scrolling curve showing the activation of the brain. Each run may take up to 15 minutes and may be repeated up to five times within the session. Usually patient enjoying this treatment do have around ten sessions. The overall aim is to get patients to practice volitional control activation in specific parts of their brains. In other words, when a patient performs a task the neurofeedback tells him or her directly afterwards how active certain

8 Personal Meeting with Prof. Dr. Goebel on April 23, 2013 at the Faculty of Psychology

brain areas were during the specific performance of the task. The patient's assignment then is to work on this brain activity and gain control over it using the given feedback. After the patient has successfully completed the sessions he or she will be tested whether the acquired skills can be applied without feedback while performing different tasks in different settings.

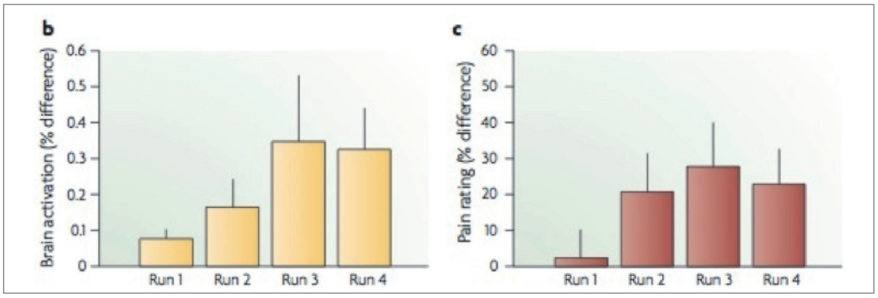


Figure 2: Impact of training on brain control and pain perception; Source: deCharms (2008)

Figure two illustrates in how far training can lead to spatial control over activation of areas in the brain and a change in for instance pain perception thereof. Part b) of the figure shows an increase in the average difference in functional MRI (fMRI) signals received from a patient's brain. Each run consisted of five cycles, in which the patient was asked to switch between increasing and decreasing brain activity. As can be detected from the bars, there is a learning process in which patients showed increasing control over their brain activation. Part c) of the figure shows the perception of pain of the same patient. In this experiment it can be clearly detected that there is a correspondence between the degree of brain activity and the perception of pain. More precisely, due to more runs the patient was able to increasingly control his or her subjective pain experience (deCharms, 2008).

Generally speaking, the more we learn about brain activities and their impact in behaviour, the more rational approaches one can make towards solving potential disorders. Up to present times, neurofeedback deals with disorders in the brain namely stroke, addiction and autism, which are decisive challenges of public health for the method (Caria et al., 2012). Nevertheless, the method has high potential for more development in particular towards patients with neurological or psychological disorders, as less is currently known about the latter. Therefore, it is particularly crucial to distinguish between training to improve circuits, which show deficits and training of 'compensatory' circuits, which are to replace lost functions. In recent years, real-time fMRI studies have shown that learned

control over brain activity in certain areas responsible for motor, sensory, cognitive and emotional processing can be acquired in relatively few neurofeedback sessions. The process of learning includes "updating expectations of the outcome proportionally to prediction error, in a way that across trials the expected outcome converges to the actual outcome" (Caria et al., 2012, p.492). In other words, real-time fMRI neurofeedback contributes to changes and potentially to an optimisation in the patient's behaviour.

4. Probation as Alternative: Effective Rehabilitation in lieu of Jail

As mentioned earlier, rehabilitation methods enjoy controversial debates among the public and policy makers. Nevertheless, certain studies have suggested that the public opinion on such measures is not as negative as often presumed in particular with juvenile offenders. Previously, public surveys often revealed that more severe methods of punishment were demanded also towards juvenile offenders. Admittedly, this was mostly in the context of extremely violent offences, such as school shooting and other rampages. When surveyed in a neutral context, however, citizens were much more willing to support rehabilitative programs instead of incarceration for juvenile offenders (Nagin et al., 2006). In a specific study on treatments of juvenile offenders, Nagin et al. (2006) found that "respondents on average expressed somewhat greater willingness to pay for rehabilitation [...] than for longer incarceration [...] of youths charged with serious crimes—and even greater willingness to pay for an early childhood prevention program" (p.642). These findings serve as evidence that citizens are generally willing to pay for rehabilitative treatments that assure to diminish crime among juveniles and for preventive programs instead of longer periods of imprisonment. Moreover, the results are of importance for policy makers as the latter "often justify expenditures for punitive juvenile justice reforms on the basis of popular demand for tougher policies" (Nagin et al., 2006, p.627). Furthermore, Farrington and Welsh (2007) found that there is also growing proof that early programs for interventions are effective and successful. These findings in turn might imply the possibility of a coherent framework of neuroscientific interventions from the moment of committing the crime until adulthood. In that sense, policy makers should move away from quickly supporting punitive treatments and rather consider rehabilitative measures as response to public opinion.

This leads to the question of probation and its possible benefit when linking it to rehabilitative measures. Probation is a form of sentencing that allows the perpetrator to stay out of prison with judicial supervision and under certain conditions determined by the court. It is a legal tool to both deter and punish criminal offenders. Generally speaking, the measure is considered as a "rehabilitative process intended to give an offender the opportunity to develop skills and moral principles necessary to forestall future criminal activity" (Stickels, 2007, p.33). In practical terms, however, perpetrators mostly fail to satisfy their probationary guidelines usually resulting in a recantation of probation and thus imprisonment. In 1991, nearly half of all prisoners in the United States committed their latest offense while being out of prison on probation (DiIulio, 1997). This is supported by the study of the Germany Ministry of Justice (2013) on recidivism where perpetrators with probationary sentences show higher rates of recidivism. The main reason for this phenomenon is that offenders after conviction are left without support for reintegration in society away from the criminal environment. As DiIulio (1997) puts it "we spend next to nothing on the systems, and get about what we pay for" (p.41). In order to make probation a more effective and thus successful toll, investments have to be made. In addition, an improvement in probationary statistics would serve the general interest due to the fact that probation is by far more cost effective than incarceration.

Linking it to neuroscientific methods, such as fMRI Neurofeedback, the process of probation could be supported. Unfortunately, statutes and laws often require incarceration of offenders having committed certain acts. At the same time, however, prosecutors often face cases in which they have to impose incarceration despite the fact that the risk of recidivism is low and rehabilitation rather likely. Generally speaking, a low propensity of re-offending and good prospects of rehabilitation could cause incapacitation of prisoners to be useless (Seave, 1993). In those cases lawmakers should be encouraged to foster probationary sentences based on correctional and rehabilitative measures. This approach should be pursued in particular, when there is a low risk of recidivism given. Certainly, there might be the possibility of disparities between offenders having committed similar crimes in terms of punishment. Nevertheless, the states of mind and different motivations should be taken into consideration and thus individualized sentencing in order to ensure effectiveness and long-term societal benefits. Therefore, even if the law required imprisonment prosecutors could downward punitive measures towards probationary penalties in correlation with rehabilitative measures (Seave, 1993).

Generally speaking, it is recommendable to actively engage an offender in his or her own rehabilitation process. With real-time fMRI direct feedback is given to the patient and therefore success can be directly measured. In that sense the patient is part of a 'game' against his or her own brain and thus actively engaged in the process. As mentioned earlier the feedback is usually indicated in a thermometer display or scrolling curve but may also be shown as virtual reality in terms of reaching for a coffee mug or computer games (Sulzer et al., 2013). By designing real-time fMRI as a sort of 'computer game' and thus making it increasingly visual the method is more accessible for the patient. On top of this, the method might arouse the patient's ambition to achieve high scores in the 'game' and hence the willingness to participate in more sessions in order to continuously increases the scores. In particular juvenile offenders could show high interest in participating in fMRI neurofeedback due to its appealing character and resemblance to computer games. Nevertheless, the method might also be attractive for adults due to its simple handling and room for self-determination.

If, however, there is a high risk of re-offending in the beginning, parole sentences could portray another alternative. Parole refers to the early release of a convicted offender before the actual term of prison ends. If the criminal commits again an unlawful act within a certain period of time set by the court the remaining time from the first conviction will be added to the second one (Fabel & Meier, 1999). In other words, if probation is too risky at the time of the trial then imprisonment combined with effective rehabilitation methods, such as real-time fMRI, should be an option. If the offender then shows good conduct and progress in his or her neurofeedback results then the parole boards could be consulted and decide upon a possible early release. Members of parole boards generally estimate the propensity of an offender towards new criminal behaviour. As Eagleman & Flores (2012) argue, the reductive accuracy in this context is rather low and flawed. Therefore, real-time fMRI could contribute to more accurate parole decisions and thus simultaneously help to decrease recidivism rates.

5. Legal and Ethical Issues

"As neuroscience learns more about the causes of human behavio[u]rs, it will give us new ways to change those behavio[u]rs" (Greely, 2012, p.163). Nevertheless, neuroscience as a tool to determine a person's culpability remains a controversial issue. This is mainly due to the complexity of the brain and the difficulty to assess whether certain behaviours are

to be labelled as disease or not. The distinction between a normal brain and an abnormal brain is not made naturally but by norms determined by humans, which are changed and adjusted every once in a while. For instance, it is rather commonly accepted that rehabilitation methods are used to cure depression. However, using such methods to treat 'wrong' sexual orientations would most probably cause a controversial public debate. As mentioned earlier, imprisonment of perpetrators is a controversial issue especially with regard to re-integration after having served a prison term. Nevertheless, the limitations of incarceration are well known and can be predicted whereas for instance neuroscientific rehabilitation methods may cause negative side effects that are not predictable at the moment.

Generally speaking, ethical issues do have legal implications in any case. Neuroscience poses a controversial topic and certainly has legal issues in particular with regard to the fundamental principle of autonomy. Despite the fact that autonomy as principle appeals universally it experiences different applications in different legal systems. The principle as such is rather simple allowing for self-determination and self-rule for states, associations and individuals. In other words, one is allowed to do what he or she wants, however, with some moral restraints not to overexcite the own autonomy to a degree where it affects another individual's autonomy. It implies that the state is not allowed to interfere with a person's life "except to the extent that this interference is warranted by the common good of society as a whole" (Sellers, 2008, p.2). Autonomy as such is one of the most important justifications in law as the latter protects liberty and autonomy of an individual. For this reason, autonomy can be regarded as a product of law. The principle of autonomy is also closely linked to the concept of privacy, which prevents unwanted intervention into private lives of individuals. Sellers (2008) defines privacy as "the negative expression of the positive value expressed by autonomy" (p.2). Despite all differences in legal systems, the common denominator is protection of liberty. In this context, privacy and autonomy can be regarded as fundamental elements of liberty and therefore of law as such.

The right to privacy is a human right and understood as describing the area in which individuals can act autonomously. It implies the control over personal space, flow of information and relationships (Post, 2010). In this context, the question arises whether neuroscience in particular brain imagining (fMRI) poses a risk to the right of privacy. The most commonly known fear of neuroscientific methods is that it may force people to reveal private thoughts against their own will possibly even without knowing it. Neuroscientific methods are currently on an ascending branch and thus its possibilities

in the future appear to be great and ground-breaking. Another important aspect is the concept of bodily integrity referring to the physical inviolability of human bodies. In this context, again the principle of autonomy is highlighted together with self-determination over the own body (Post, 2003). Many neuroscientific methods among others drugs pose a risk to this principle as they 'invade' the human body. Real-time fMRI Neurofeedback as such is generally seen as non-invasive. Nevertheless, the scanning of brain activity during task performance still means 'looking into a person's body' and in that sense even neurofeedback could arouse debate. In other words, even this method may be regarded as infringing bodily integrity and the right to privacy.

Sellers (2008) stated that any form of 'invading' the brain infringe the right to privacy and thus bodily integrity and are therefore prohibited unless they serve the society. The question arising in this context is where the line between unlawful invasion and the benefit for society is drawn. In order to facilitate an appropriate use of neuroscientific methods certain rules should be determined allowing it to be a rather undisputed tool.

First of all, neuroscientific interventions of any kind need to be a voluntary act. However, even if an adult offender is given the choice of receiving such treatments it is difficult to assess whether the choice was fully freely made. Pressure may come from different directions meaning family members, colleagues and society in general (Greely, 2012). Therefore, the offer of a treatment as such must be genuine and not a threat. On top of this, the intervention must only be aimed at addressing the specific behaviour for which the offender was convicted (Shaw, 2012). Secondly, the aim should by no means be to perform a fundamental personality change. Any efforts made to convert the offender's opinion should be based on rational dialogue. Neuroscientific methods, however, could be used to facilitate this intended moral dialogue by enhancing the offender's capacities. For instance, neuroscience could help to improve attention capacities of perpetrators and thus foster their ability to consider different options before performing an act. Moreover, it could support to improve an offender's warning system in order to help him or her to resist self-defeating behaviour (Shaw, 2012). Nevertheless, it has to be stated that the line between enhancing capacities and re-shaping the offender's values is rather thin.

6. Conclusion

This paper has outlined the pressing problems judicial systems face at present times. Recidivism rates are very high and do not show a tendency to decrease in the near future. Moreover, prison populations show likewise extraordinary high numbers and with regard to recent years the number of incarcerated offenders is about to even increase in the future. At the same time, however, national societies in particular in the United States and in Europe show features of retributivism believing as mentioned before that one more offender in prison is one less criminal in streets (Cullen, 2006). Concluding from that there is a general need for reform due to the fact that the current system appears flawed and ineffective. The growing field of neuroscience holds great potential for new options, which could help to improve current judicial systems. Despite the controversial debates about neuroscience and its legal and ethical implications, it is a field of research with new technologies and methods with great capabilities. In that sense, we doubtlessly should make use of it in order to extract the best and benefit to improve the current systems. Albeit, it has to be stated that at present times individual risk assessments are done already to a certain degree and prosecution shows a more proactive behaviour in particular with juveniles.

This paper offers an approach on how neuroscientific methods in particular real-time fMRI or Neurofeedback could be used in a beneficiary way to improve the system and thus enlarge the benefit of the society. Certainly, this approach is opposed to conventional retributivism, as it is understood that retributivist thinking brings no benefits to the society as a whole but rather attempts to compensate the victims alone. In this context, offenders should not only receive imposed punishment but rather be actively engaged in it. Conventional incarceration does not live up to these expectations as it implies passive behaviour of the offender. Rehabilitative measures on the contrary provide for active engagement of the perpetrator and thus appear to be much more useful. In particular fMRI neurofeedback with its 'game-like' method allows the offender to actively participate in his or her rehabilitation process and thus implies potential higher success. Consequently, this paper argues that probation should more often be considered as alternative to incarceration when being linked to fMRI neurofeedback. As mentioned earlier, certain studies predict positive changes towards recidivism rates when making use of such neuroscientific methods. Nevertheless, it is important to mention that in reducing recidivism rates the focus should not only be on rehabilitation methods and brain disorders due to the fact that biological facts are not the sole cause of misconduct. Not all people

with mental deficits engage in offences, whereas not all offenders do in fact have brain disorders. Certainly, there are many other factors involved, such as the environment of the offender and most importantly parental education. To be more precise, there is a mutual influence between mental capacities and other influencing factors. This paper, however, emphasises only the neuroscientific elements and their impacts.

To conclude with, societies would benefit more from effective rehabilitation methods instead of simple incapacitation, which may cause potential re-offences. The approach does not necessarily substitute contemporary legal systems. It should be rather seen as an added value and a possible alternative to present methods. If the use of neuroscientific rehabilitation methods appears to be not successful there is still the opportunity to return to the old model of incapacitation. If, however, real-time fMRI Neurofeedback unfolds great potential in the upcoming years the method could also be utilized as preventive measure for 'danger groups'. In particular juveniles with problematic backgrounds who are predestined to get into trouble could be made familiar with the 'game' of Neurofeedback. Among those could be children from deprived areas, where parents do not have the capabilities to care for their children. A project as such could be conducted in cooperation with child protective services.

Bibliography

Aebi, M. F., Aubusson de Cavarlay, B., Barclay, G., BaGruszczynska, B., Heiskanen, M., Harrendorf, S., Hysi, V., Jaquier, V., Jehle, J. M., Kilians, M., Shostko, O., Smit, P. & Þórisdóttir, R. Wetenschappelijk Onderzoek- en Documentatiecentrum, (2010). *European Sourcebook of Crime and Criminal Justice Statistics* (Fourth Edition) Ministerie van Justitie.

Barco, P. P., Crosson, B., Bolesta, M. M., Werts, D., & Stout, R. (1991). Training Awareness and Compensation in Postacute Head Injury Rehabilitation. In J. S. Kreutzer & P. H. Wehman (Eds.), *Cognitive rehabilitation for person's with traumatic brain injury: A functional approach* (pp. 129-146). Baltimore: Paul H. Brookes.

Bentham, J. (1970). J. Burns & H. Hart (Eds.), *An Introduction to the Principles of Morals and Legislation*. London, England: The Athlone Press.

Breiter, H. C., Aharon, I., Kahnemann, D., Dale, A., & Shizgal, P. (2001). Functional Imaging of Neural Responses to Expectancy and Experience of Monetary Gains and Losses. *Neuron*, 30(2), 619-639.

Brown, D. K. (2012). Criminal Law Theory And Criminal Justice Practice. *The American Criminal Law Review*, 49(1), 73-103.

Buchanan, A. (2000). *Psychiatric Aspects of Justification, Excuse and Mitigation*. London: Jessica Kingsley.

Caria, A., Sitaram, R., & Birbaumer, N. (2012). Real-time fMRI: A Tool For Local Brain Regulation. *The Neuroscientist, 18*(5), 487-501.

Chen, K. M., & Shapiro, J. M. (2007). Do Harsher Prison Conditions Reduce Recidivism? A Discontinuity-based Approach. *American Law and Economics Review* 9, 1-29.

Clarke, R. (2003). *Libertarian Accounts of Free Will*. New York, United States: Oxford University Press.

Cullen, F. T. (2006). It's Time To Reaffirm Rehabilitation. *Criminology & Public Policy*, 5(4), 665-672.

Cullen, F. T. (2007). Make Rehabilitation Corrections' Guiding Paradigm. *Criminology & Public Policy*, 6(4), 717-727.

Cullen, F. T., & Gendreau, P. (2001). From Nothing Works To What Works: Changing Professional Ideology In The 21st Century. *The Prison Journal*, 81(3), 313-338.

deCharms, R. C. (2008). Applications Of Real-Time fMRI. *Natural Reviews Neuroscience*, 9, 720-729.

DiIulio, J. J. (1997). Reinventing Parole and Probation: A Lock-'Em-Up Hard-Liner Makes The Case For Probation. *The Brookings Review*, 15(2), 40-42.

Drago, F., Galbiati, R., & Vertova, P. (2011). Prison Conditions and Recidivism. *American Law and Economics Review*, 13(1), 103-130.

Eagleman, D. M., & Flores, S. I. (2012). Defining a Neurocompatibility Index for Criminal Justice Systems: A Framework to Align Social Policy With Modern Brain Science. *The Law of the Future and the Future of the Law*, 2 (1), 161-172.

Eurostat, 2009; retrieved from: *http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=crim_pris&lang=en*

Fabel, O., & Meier, V. (1999). Optimal Parole Decisions. *International Review of Law and Economics*, 19, 159-166.

Farrington, D. P. & Welsh, B. C. (2007). *Saving Children from a Life of Crime: Early Risk Factors and Effective Interventions*. New York: Oxford University Press.

Fleming, D. J. M., & Ownsworth, T. (2006). A Review of Awareness Interventions in Brain Injury Rehabilitation. *Neuropsychological Rehabilitation: An International Journal*, 16(4), 474-500.

Garland, D. (2001). *The culture of control: Crime and social order in contemporary society*. New York, United States: Oxford University Press.

Gendreau, P., Goggin, C., Cullen, F. T., & Andrews, D. A. (2000). The effects of community sanctions and incarceration on recidivism. *Forum on Corrections Research*, 12, 10-13.

Greely, H. T. (2012). Direct Brain Interventions to "Treat" Disfavored Human Behaviors: Ethical and Social Issues. *Clinical Pharmacology and Therapeutics*, 91(2), 163-165.

Green, J., & Cohen, J. (2004). For the law, neuroscience changes nothing and everything. *The Royal Society,* 359(1451), 1775-1785.

Gruber, A. (2010). A Distributive Theory Of Criminal Law. *William and Mary Law Review,* 52(1), 1-73.

Hampton, J. (1991-1992). Correcting harms versus righting wrongs: The Goal of retribution. *UCLA Law Review*, 39, 1659-1702.

Kalis, A., Mojzisch, A., Schweizer, T. S., & Kaiser, S. (2008). Weakness Of Will, Akrasia, And The Neuropsychiatry Of Decision Making: An Interdisciplinary Perspective. *Cognitive, Affective, & Behavioral Neuroscience*, 8(4), 402-417.

Kant, I. (2002). *The Philosophy of Law.* Translated by Hastie, W. New Jersey: The Lawbook Exchange, LTD.

Langan, P. A., & Levin, D. J. U.S. Department of Justice, Bureau of Justice Statistics. (2002). *Recidivism of Prisoners Released in 1994* (NCJ 193427). Washington, DC.

Meyen, G. (2013). A Neurolaw Perspective on Psychiatric Assessments of Criminal Responsibility: Decision-making, Mental Disorder, and the Brain. *International Journal of Law and Psychiatry*, 36, 93-99.

Ministry of Justice of Germany (Bundesministerium der Justiz, BMJ); retrieved from: *http://www.bmj.de/DE/Recht/Strafrecht/KriminologieKriminalpraevention/_doc/Rueckfallstatistik_doc.html;jsessionid=4E79B250E34A07322106133397E7CB7C0.1_cid289?nn=1470118*

Ministry of Justice of the Netherlands (Ministerie van Veiligheid en Justitie); retrieved from: *http://www.wodc.nl/onderzoeksdatabase/actualisering-recidivemeting-sancties-2010.aspx#paragraph_meerinfo*

Ministry of Justice of the United Kingdom; retrieved from: *https://www.gov.uk/government/publications/proven-re-offending--2*

Molina, F. (2011). A Comparison Between Continental European And Anglo-American Approaches To Overcriminalization & Some Remarks On How To Deal With It. *New Criminal Law Review*, 14(1), 123-138.

Moore, M. S. (2010). *Placing Blame: A Theory of the Criminal Law.* Oxford Scholarship Online.

Nagin, D. S., Piquero, A. R., Scott, E. S., & Steinberg, L. (2006). Public Preferences For Rehabilitation Versus Incarceration Of Juvenile Offenders: Evidence From A Contingent Valuation Survey. *Criminology & Public Policy,* 5(4), 627-652.

Post, S. G. (2003). Bodily Integrity. A Plausible Argument. *Journal of Disability Policy Studies*, 13(4), 261.

Prigatano, G. P. (1991). Disturbances of self-awareness of deficit after traumatic brain injury. In G. P. Prigatano & D. L. Schacter (Eds.), *Awareness of Deficit after Brain Injury: Clinical and Theoretical Issues* (pp. 111-126). New York: Oxford University Press.

Prigatano, G. P., & Schacter, D. L. (1991). Introduction. In G. P. Prigatano & D. L. Schacter (Eds.), *Awareness of Deficit After Brain Injury: Clinical and Theoretical Issues* (pp. 3-15). New York: Oxford University Press.

Rosen, J. (2007, March 11). The Brain On The Stand. *The New York Times*. Retrieved from: http://www.nytimes.com/2007/03/11/magazine/11Neurolaw.t.html?pagewanted=all&_r=0

Seave, P. L. (1993). Rehabilitation, Non-recidivism, and Probation: The Sentencing Commission's Unwanted Stepchildren. *Federal Sentencing Reporter*, 5(4), 223-224.

Sellers, M. (2008). *Autonomy in the Law*. Dordrecht: Springer.

Shaw, E. (2012). Direct Brain Interventions And Responsibility Enhancement. *Criminal Law and Philosophy,* 1-20.

Simmond, M., & Fleming, J. M. (2003). Occupational Therapy Assessment of Self-Awareness Following Traumatic Brain Injury. *British Journal of Occupational Therapy*, 66(10), 447-453.

Spohn, C., & Holleran, D. (2002). The Effect Of Imprisonment On Recidivism Rates Of Felony Offenders: A Focus On Drug Offenders. *Criminology*, 40(2), 329-358.

Stickels, J. W. (2007). The Game of Probation. *Journal of Police and Criminal Psychology,* 22, 33-43.

Sulzer, J., Haller, S., Scharnowski, F., Weiskopf, N., Birbaumer, N., Blefari, M. L., Bruehl, A. B., Cohen, L. G., deCharms, R. C., Gassert, R., Goebel, R., Herwig, U., LaConte, S., Linden, D., Luft, A., Seifritz, E. & Sitaram, R. (2013). Real-time fMRI Neurofeedback: Progress and Challenges. *NeuroImage, 76*, 386–399.

Sutherland, E. H. (1939). *Principles of criminology* (3rd ed.). Chicago: J. B. Lippincott.

Tonry, M. 2004 *Thinking About Crime: Sense and Sensibility in American Penal Culture*. New York: Oxford University Press.

A neuroscientific perspective on cognitive and volitional impairment in criminal irresponsibility assessments: a case for a capacity-based approach

By Nina Koivula

Table of Contents

1. Introduction

"If a doctor were to bleed his patients with leeches today, or if a psychiatrist were to attribute insanity to the moon, the hue and cry would be tremendous. And yet instance after instance may be pointed out wherein the law has remained, sometimes for hundreds of years, curiously rigid, despite the changes in scientific opinion upon which that law was based. Many rules in the criminal law are still affected by early views concerning psychology, which views are now outmoded or repudiated by newer discoveries through experimentation" (Woodbridge, 1939, p. 822). Even though these words date back over 70 years, they still hold weight today.

The law remains reluctant of joining forces with neuroscience to better understand human behaviour despite the fact that "preliminary biological explanations" exist for a number of relevant phenomena (Garland & Glimcher, 2006, p. 131). Contemporary brain imaging techniques have enabled the study of law-related notions such as consciousness, morality and intent, to name a few (Gazzaniga, 2008, p. 412). Others argue that neuroscience is not advanced enough to uncover mental content that is pertinent to the law (Morse, 2011, pp. 849-850). It has also been suggested that even if it were possible to prove a precise correlation between the requirements for criminal responsibility and certain neural patterns, these patterns could only amount to "evidentiary support for the assertion that the criterion in question was in fact satisfied at the time of the crime" (Morse, 2006, p. 399). Plausibly, these claims do not warrant ignorance towards existing neuroscientific research, which is not insignificant in volume. Moreover, it is hard to see why the society as a whole would not benefit from legal determinations which are as rigorous and precise as possible. It could be argued that every insight – whether neuroscientific, sociological or evolutionary, for example – provides a "reality" of human behaviour from a distinct, but complementary point of view (Bufkin & Luttrell, 2005, p. 185).

1.1. Law and neuroscience

Neurolaw can take at least three different forms. In academia, these have been labelled as the *law of neuroscience, neuroscience of the law* and finally, *neuroscience in law*, which is of most relevance to this paper as it covers the neuroscientific research into legally relevant aspects of human behaviour and the cognitive processes that underlie it. Findings in the fields of moral and legal cognition as well as impulse and behavioural control have the potential to enhance assessments of criminal (ir)responsibility (Klaming & Koops, 2012, pp. 228-229). Nonetheless, the law has yet to accommodate neuroscience in a systematic

manner. Despite the procedural challenges which it inevitably presents, neuroscience does not create new problems for the law as such. The legal questions at the core of (ir) responsibility determinations must be answered in spite of any scientific contribution to the process. Of course, it is not certain that the use of neuroscientific models would simplify rather than complicate such determinations (Aharoni & Funk & Sinnott-Armstrong & Gazzaniga, 2008, p. 146).

The law is pragmatic in nature and aims to resolve these complex questions using its own framework and conceptions. Despite the fact that its questions are ultimately legal in nature and must be answered as such, the law should not and does not turn a blind eye to science in general. Whereas it is evident that law and science fulfill distinct roles in society, with the former maintaining the philosophical notion of justice and the latter aspiring to illustrate and interpret concrete phenomena, the law must to a certain extent rely on science, primarily in its investigative stages (Eastman & Campbell, 2006, p. 312). Importantly, science operates using a much higher threshold of certainty than the law – commonly, 95% compared to 51% – even if legal certainty concerns causation and scientific certainty can only show a relationship (Garland & Glimcher, 2006, p. 131).

Criminal law perceives humans and behaviour through a "folk-psychological" lens (Morse, 2011, p. 839). The same is clearly not true for the field of neuroscience, which takes a reductionist approach. Furthermore, the law – albeit perhaps more the common law – stresses the importance of precedent in legal decision-making, whereas neuroscience looks predominantly into the future. The hypothetical conflict between the ideas of free will and determinism should also be mentioned (Martell, 2009, p. 124). However, this as well as considerations of the "fundamental psycho-legal error" can be dismissed in so far as the aim of this thesis is to merely suggest ways in which the discovery of evidence used in legal responsibility determinations could be improved through the use of neuroscience. Those who reject the use of neurological data on the premise that it will somehow distort the facts or decide the case on its own should bear in mind that to explain is not to excuse (Aharoni et al., 2008, p. 146). It must be noted, of course, that the proper role of neuroscience in the realm of the law is not a question for science to answer, but rather a legal verdict. Fundamentally, criminal law and neuroscience are difficult to reconcile. This is, however, precisely what neurolaw aims to *do*.

At the moment, it seems that neuroscience only holds the potential to yield a so-called internal contribution to the law – that is, something that strives to change or explain a collection of "legal doctrines, practices or institutions" whilst assenting to them in general (Morse, 2011, p. 843). It may thus be premature or even imprudent to claim that neurolaw will somehow revolutionarize law. However, even some of the most sceptic

academics acknowledge the possible value of neuroscience in illuminating matters such as legal insanity (Morse, 2011, p. 845). Progress is obstructed by the fact that the amount of empirical neuroimaging studies which address legal questions is currently very low, as most of the advancement has taken place in the 21st century. This is because 'neurolegal' research is notably difficult to design and carry out (Meynen, 2013, p. 96).

Nonetheless, I argue that neuroscience has something to contribute to the law – more specifically, evaluations of criminal irresponsibility. Criminal irresponsibility, or 'legal insanity', involves the assessment of a defendant in order to determine – as crudely generalized – whether they will be confined in a prison or a mental institution. Such an evaluation is most often the responsibility of a psychologist or a psychiatrist, and involves the subjective appraisal of whether the individual was acting under such a mental disturbance at the time of the offense as to exculpate him from responsibility, either completely or partly.

Structural neuroimaging can provide an insight into lesions or any other pathologies, such as tumours which may be present in the defendant's brain. Functional neuroimaging, on the other hand, has the potential to additionally reveal abnormalities in neural metabolism and, to some extent, disclose whether a defendant's cognitive functions have been compromised. Neuroscience, as opposed to traditional assessment methods such as interviews, is not as vulnerable to malingering nor subjective or biased appraisals of third parties concerning the defendant's personality and behaviour (Vincent, 2011, pp. 38-39).

The aim of this thesis is therefore to analyze research on neural correlates of legally relevant behaviour in the context of criminal irresponsibility with a focus on both cognitive and volitional impairment in an attempt to map regions of interest (ROIs). The thesis will conclude with suggestions on how the criminal law can benefit from neuroscience in this regard.

1.2. Criminal responsibility

Notwithstanding of the opinion of the law, there is a general presumption that individuals are responsible for the acts they undertake. In cases where these acts are contrary to the law, the determination of criminal responsibility becomes necessary. No uniformly accepted theory exists for explaining the assignment of criminal responsibility to a particular person, but several proposals have been put forward over the years (Wilson, 2009, p. 473). The generic agency theory suggests that responsibility is a consequence of only those actions that "reflect on them as agents" (Husak, 2013, p. 57). There are three more specific categories of theories which are most commonly discussed – namely,

character, choice and capacity theories. Character theories hold that perpetrators should be criminally responsible for their actions only to the extent that such actions are representative of their character. Choice theories purport that responsibility follows only if the perpetrators made the choice to undertake the act and they had a satisfactory variety of choices to make. Lastly, capacity theories suggest that the crucial factor in determining criminal responsibility is whether the perpetrators possessed the capacity or opportunity in order to have acted in a different manner. This stresses the capacity of the individual to comprehend their actions as well as their volitional abilities (Tadros, 2007, p. 22). On the other hand, the social theory purports that the determination of criminal responsibility is a purely social appraisal without a personal mental component. In practice, the two elements which are most often cited to constitute at least a part of the criteria for criminal responsibility are the *actus reus* and the *mens rea* (Wilson, 2009, pp. 473-474). Of the above, the capacity theory will serve as the theoretical foundation for this thesis.

It has been suggested that criminal responsibility is above all a normative determination, owing partly to the fact that a "brain correlate of responsibility" does not exist (Klaming & Koops, 2012, p.2). It can be inferred that the same holds true for criminal irresponsibility.

1.3. Criminal irresponsibility

For one reason or another, Western legal practice and jurisprudence has generally accepted the intuitive idea that there are circumstances in which an individual should not be held responsible for their illicit actions (Wilson, 2009, p. 473). Examples of this are acts committed by the immature and, more importantly, the mentally disordered. Such an approach has been justified on a number of grounds, including the argument that it would be against the *nulla poena sine culpa* principle, and public morals to punish an individual who cannot act rationally or control their behaviour due to a mental disturbance. It has also been postulated that punishing a defendant who is deemed irresponsible would be meaningless in the context of both specific and general deterrence (KKO:2000:126).

The desire to not punish unaccountable individuals is manifested in various legal systems in the form of something akin to an 'irresponsibility clause' or 'insanity defense'. The composition of the test varies, with some legal systems opting for purely cognitive and others for both cognitive and volitional criteria. The cognitive prong typically measures the defendant's ability to understand their act and/or appreciate its wrongfulness (Corrado, 2010, p. 508) Volitional standards, on the other hand, relate to the ability to control one's behaviour so that it stays in conformity with the law (Carrido, 2012, p. 310).

To name a few examples of irresponsibility clauses, Section 16 of the Canadian Criminal Code excludes from criminal responsibility those who commit an act or omission whilst suffering from a mental disorder which made the perpetrator "incapable of appreciating the nature and quality of the act [or omission] or of knowing that it was wrong". It thus relies only on the cognitive test.

On the contrary, whereas it is not a legal system per se, the Rome Statute of the International Criminal Court states in Article 31 that a person is not criminally responsible in the case where at the time of their conduct, the actor suffers from a mental disease or defect which devastates the "capacity to appreciate the unlawfulness or nature of his or her conduct, or capacity to control his or her conduct to conform to the requirements of law". The consideration of both cognitive and conative impairment is evident.

The Criminal Code of the Netherlands provides in Article 39 that a person is exempt from punishment if the act they have committed cannot be attributed to them as a result of "poorly developed or pathologically disturbed mental capacities". It does not suffice to prove that the mental disorder or deficiency was present at the time of the act, but it must also be shown that there is a link between the condition and the criminal act. Dutch law focuses on two separate elements in determining responsibility: the ability of the perpetrators to tell right from wrong and their perceived liberty to decide whether to act or refrain from acting (Klaming & Koops, 2012, p.8). Again, the cognitive and volitional limbs are both present.

A more comprehensive statutory clause of irresponsibility can be found in Finnish law, where Chapter 3 Section 4 subsection 2 of the Criminal Code states:

> *"the perpetrator is not criminally responsible if at the time of the act, due to mental illness, severe mental deficiency or a serious mental disturbance or a serious disturbance of consciousness, he or she is not able to understand the factual nature or unlawfulness of his or her act or his or her ability to control his or her behaviour is decisively weakened due to such a reason".*

Essentially, therefore, there is a requirement of either a diagnosed mental illness or a severe form of another mental impairment at the time of the act which also must have caused the unlawful behaviour. The law distinguishes between what is a clearly two-branch cognitive limb as well as a volitional limb. The cognitive limb prima facie accommodates both the incomprehension of circumstances pertaining to reality as well as the 'unlawfulness' of the act. The 'unlawfulness' could be seen to be two-dimensional and include both 'wrongfulness' in the moral sense as well as 'illegality'. This is supported

by both the Finnish word used in place of 'unlawfulness', namely, oikeudenvastaisuus – which translates roughly to 'contrary to justice'– as well as a judgment of the Finnish Supreme Court. The judgment, discussing criminal irresponsibility, refers to a person who cannot understand the "legal and moral meaning" of their act (KKO:2005:48).

Some legal systems take only the standard of either moral or legal wrongfulness while others maintain that knowing that the actions were contrary to the law does not suffice to exclude irresponsibility (Sinnott-Armstrong & Levy, 2011, p. 304). Whereas it is obvious that legal and moral values do not always coincide, it may be beneficial to include both in the definition for irresponsibility, for example in cases where the person was not aware of the legislation in force. Both the narrow legal wrongfulness test and the broader moral wrongfulness should be acknowledged (Yeo, 2008, p. 252).

What is notable about the volitional limb in Finnish law is that it does not explicitly demand a complete loss of control on behalf of the perpetrator, but uses the standard of "decisively weakened". Due to its comprehensiveness, relative clarity and functional nature, the Finnish model will be used as a point of departure for the further analysis of the legal criteria for criminal irresponsibility in this thesis.

1.3.1. Insanity of irresponsibility? Preliminary considerations.

Labelling defendants 'insane' is problematic as well as stigmatizing since it echoes "an ill-informed and insensitive public's perception of people who suffer from mental disorders" (Reider, 1998, p. 341). The title of this thesis also reflects this consideration. Using the term 'irresponsible' or even 'unaccountable' rather than 'insane' is arguably a more sophisticated and objective representation of mentally disordered criminal defendants. This is in conformity with the norms of 21st century civilization. Such terminology also better reflects the fact that the criteria for the test are ultimately legal. Sanity and insanity are remarkably obscure and intangible concepts for psychiatrists, judges and laymen alike, and should thus be discarded from legal use.

1.3.2. Case Breivik

A recent high-profile criminal case highlighting the volatile nature of neuropsychiatric examinations in determining criminal irresponsibility was the trial of mass murderer Anders Behring Breivik in Norway. The national laws state that a defendant is deemed irresponsible if he is "psychotic, unconscious or severely mentally retarded" at the time of the offense. In establishing whether the defendant is responsible or not, the Norwegian practice is to appoint two forensic psychiatrists who undertake an assessment of the accused individual. The court can either comply with or reject the conclusions of the experts.

The first evaluation consisted of analysing the interrogations carried out by the police, 36 hours of personal interviews with the defendant as well as discussions with his mother. This stage was concluded in November 2011, four months after the crimes were committed. The experts submitted Breivik was psychotic during the massacre as well as the psychiatric assessment (Melle, 2013, p. 16-17). They diagnosed him with paranoid schizophrenia. Nevertheless, in January of the following year, the court called for a new assessment. The re-evaluation included an additional inpatient observation, but was otherwise methodologically congruent with the first (Melle, 2013, p. 18). Regardless, the experts came to a different conclusion after concluding their work in March. They found that Behring Breivik was not psychotic during his acts nor after they took place, but suffered from a severe form of narcissistic personality disorder (NPD) accompanied by compulsive lying. As regards the prior diagnosis of schizophrenia, the second pair of psychiatrists saw that the defendant did not satisfy the criteria found in the International Classification of Diseases (ICD-10) system. On the contrary, it has been argued that his symptomatology would have fulfilled the requirements of the Statistical Manual of Mental Disorders (DSM-IV) – the other near-universal diagnostic instrument – for schizophrenia.

Nevertheless, he was declared criminally responsible first by the experts and subsequently by the court, which has the final say in the matter. The court saw that since Behring Breivik was capable of controlling his impulses during the questioning and since several mental health professionals were of the opinion that his symptoms pointed towards a personality disorder, he was not psychotic. The court also relied heavily on the results of the three-week inpatient observation (Melle, 2013, p. 19). The fact that the court acts as the ultimate authority on what seems to be the *clinical* determination whether a person was 'psychotic' is problematic. Furthermore, the court, through the interpretation of the experts, treated diagnostic criteria *de facto* as legal rules. It is also notable that the court dismissed the contradictory nature of the two reports as mere "differing interpretations of similar observations", without having regard to the fact that the examinations were initiated six months apart (Melle, 2013, p. 20).

First and foremost, the Behring Breivik trial served to demonstrate how thin the line between the verdict of criminally responsible and criminally irresponsible can be, and how inconsistently similar symptoms can be interpreted by different forensic psychiatrists. Unfortunately, in some legal systems, an irresponsibility determination may be a question of life and death. Interestingly, whereas the first pair of psychiatrists seemed to view his extreme ideas and thoughts as signs of delusion, the second pair interpreted them as compulsive lies. As regards the procedural aspects of the responsibility determination, it is not insignificant that the second assessment period ended eight months after the

mass murders took place, keeping in mind that the determination of responsibility must necessary reflect the defendant's mental state at the time of the acts. It is also notable that no functional or structural neuroimaging was used in the evaluation, but this was solely due to the fact that Behring Breivik refused to undergo a scan (Aftenposten, 2011). Thus any considerations of how neuroscience could have contributed to the assessment are unfortunately reduced to mere speculation.

1.3.3. Psychiatric diagnoses in the context of irresponsibility

It has been argued that in the context of irresponsibility considerations the law ought to avoid using the classic categories of mental diagnoses and instead undertake "a more individualised evaluation" of the specific capacities of the defendant (Reider, 1998, p. 333). This is precisely what the author wishes to advocate with the present thesis – an approach outside of the diagnostic labels, focusing on the effect of and not the cause behind the pathological or functional disturbance. Arguably, this would also make the framework more easily adaptable to the law. It is nonsensical to propose that psychiatric evaluations be abolished in this context, but they could be at the very least corroborated by neuroscientific evidence. This coincides with the opinion that despite scientific advances, the irresponsibility determinations of the future will still depend on interpersonal assessments and interpretations (Melle, 2013, p. 20).

2. Neuropathology

The human brain is arguably one of the biggest puzzles in the human body that has yet to be thoroughly decoded, much owing to its tremendous complexity. Neuroscientific evidence supports the argument that at least some distinct patterns of behaviour have a neural correlate in the form of a specific region in the brain, and that damage to these areas can account for behavioural abnormalities (Batts, 2009, p. 265). It has however been demonstrated that a person may act in extremely corrupt and violent ways even if their brain scan results are not atypical. The reverse is also true – some persons do not engage in criminal activity despite suffering from considerable brain damage in regions linked with abnormal behaviour (Sapolsky, 2004, p. 1794). Ultimately, current neuroscience can only deal with correlations, not causation. In summary, one could say that each mental event is embodied within – but not equal to – corresponding neural events (Martens, 2002, p. 175). From the law's point of view, the most informative technology seems to be functional magnetic resonance imaging, commonly known as fMRI (Morse, 2011, p. 849).

It is studies using this technology which form the core of the research presented below.
It is also important to consider the temporal element in the context of irresponsibility. There is, to date, no means by which it is possible to measure neural activity in an individual's brain *in flagrante delicto*. This necessarily means that neuroscience cannot show how the brain of the offender was functioning at the time of the act. However, structural neuroimaging can reveal "temporally stable" abnormalities which may have been present at the time of the act and which can compromise one's mental capacities in a manner that is relevant to the law (Morse, 2011, p. 850). Nonetheless, neuroscience has shown that the traditional division of psychiatric and neurological disorders into ones with either 'organic' or 'nonorganic' etiology is largely outdated. This is especially relevant to considerations involving disorders which were traditionally seen as nonorganic, such as psychoses. These have now been linked to alterations in neurological structures (Manor & Tyano 1999, p. 415-419).

2.1. Neuroanatomy

The human prefrontal cortex (PFC) is responsible for executive functions and complex mental processing, such as abstract thought and problem-solving. It is commonly divided into five regions: the orbitofrontal, dorsolateral, dorsomedial, ventrolateral and the ventromedial cortex (Forbes & Grafman, 2010, p. 306). Because the isolation of neural areas is not unequivocal, some areas can only be distinguished by their functioning and not anatomy. Nevertheless, the OFC-DLPFC-DMPFC-VLPFC-VPMFC distinction will be applied throughout this thesis.

2.2.1. Neuroplasticity

The idea of neuroplasticity lies at the very heart of behavioural neuroscience. In this context, research in both animals and humans has shown that the brain can undergo remodeling on several levels outside the typical stages of development, and that this is linked to behavioural differences (Kolb & Whishaw, 1998, p. 44). Remodeling is demonstrated by for example neurogenesis, the creation of new nerve cells, and pruning, the elimination of redundant neural connections. Neuroplasticity is the hypernym used to denote such changes (Kays & Hurley & Taber 2012, p. 119). Brain plasticity has implications towards a number of fields relevant to neurolaw, such as neurorehabilitation. The fact that the brain is – to some extent – a flexible organ is also relevant to irresponsibility considerations due to the underlying assumption that neural correlates to behaviour exist. The criminal law, of course, is only interested in neuroplasticity insofar as it triggers

behavioural changes. What one should take away from all this is that the human brain is not a static entity, and this poses additional challenges to neuroimaging in *ex post facto* irresponsibility determinations.

2.2. Mental illness

The Finnish Criminal Code's prerequisite of "*mental illness, severe mental deficiency or a serious mental disturbance or a serious disturbance of consciousness*" ought to be interpreted as a catch-all phrase. In this context, it must be stressed that what amounts to a mental illness, disability or other similar concept in the eyes of a specific criminal justice system is primarily determined by factors such as the public policy of the legal order in question (Slovenko, 1999, p. 180). Accordingly, this element of irresponsibility determinations will not be discussed further.

2.3. Cognitive impairment: preliminary considerations

Before any substantial discussion can take place, it is important to consider the limitations of neuroscience in this regard. It is practically impossible to determine through neuroimaging whether an individual truly understood the factual nature of their act or made a moral or a legal judgment at the time of the crime and also understood this specific judgment. What neuroscience can and should do is to shed light on the capacities an individual may possess or lack to understand what they are doing or make moral or legal judgments in a normal manner.

2.4. Cognitive impairment: understanding factual nature

One must turn to case law for guidance on the interpretation of the 'factual nature' criterion. The Finnish Supreme Court, on appeal, discusses a ruling by the Kajaani District Court which found a person criminally irresponsible due to delusional disturbances which had an adverse effect on his capability to normally understand the factual nature and unlawfulness of his act (KKO:2009:56). By analogy from other irresponsibility tests, it can be opined that understanding the factual nature pertains to the "appreciation of the physical characteristics of the act done and of the material circumstances in which it occurs". In this determination, the standard of whether the defendant's interpretation of physical reality is deemed to be sufficiently deviant from that of an average person can be used (Gotlieb, 1956, pp. 272-273).

When examining whether an individual was able to understand the factual nature of their act, the fundamental issue is that of perception. Ways in which human perception can be altered to the extent that the person "loses touch with reality" is for example through vivid hallucinations or delusions, commonly those related to psychosis (Redding, 2006, p. 81). Owing to space constraints, this section will only discuss the neuroscientific substrates of delusions and hallucinations. This is due to the fact that both pathologies are firmly related to violent behaviour (Nordström et al., 2006, pp. 192-193).

In layman's terms, hallucinations can be referred to as "crazy perceptions" and delusions as "crazy beliefs" (Morse, 1999, p. 155). Neuroscience has shown interest in specifically the perceptual type of consciousness, which is equated with awareness (Bennett, 2008, p. 916). Perception can be detected in the brain as activation in cortical sensory areas. It is, however, unlikely that there is one distinct cortical sub-region which houses the capacity to "be aware or conscious of that which is perceived" (Bennett, 2008, p. 917). Schizophrenic hallucinations have been associated with lesions in the respective visual and acoustic pathways in the brain (Bennett, 2008, pp. 922-923).

A study using PET scans to analyse a patient group presenting with enduring and fixed delusions and hallucinations found a relationship between such symptoms and increased or decreased activity in several neural regions. Hallucino-delusional manifestations were associated with an increase in regional cerebral blood flow in the left medial temporal lobe (LMTL), left ventral striatum (LVS) as well as Broca's area, most commonly implicated in production of speech. These symptoms were also linked to reduced blood flow in the left lateral temporoparietal cortex as well as the right posterior cingulate (RPC). The evidence that reality distortion entails the dysfunction of Broca's area alongside specific regions of the temporal lobe which are frequently associated with 'monitoring of self-generated mental activity' has been corroborated in other research (Liddle, 1997, pp. 334-335).

Researchers using convergent functional genomics have extracted, from a cluster of 40 000 genes and expressed sequence tags, 7 blood biomarkers as indicators of hallucinations and 31 blood biomarkers as indicators of delusions (Kurian et al., 2011, pp. 44-46). It can be opined that the discovery of such biomarkers may pave the way for a neurobiological diagnostic test, something that could in theory be administered to the defendant very soon after their arrest. This could – at least hypothetically – provide a more accurate assessment of the person's capacities at the time of the offense. Of course, it is far from obvious that hallucino-delusional experiences necessarily negate the capacity to understand the factual nature of the *particular act* that the individual is prosecuted for (Broome et al., 2010, p. 184).

Nevertheless, the importance of undistorted perception to the more complex cognitive appraisal of the act is self-evident. *"An agent cannot appreciate the wrongfulness or criminality of an act if she cannot appreciate the nature and quality of the act, especially if the relevant parts of the nature and quality of the act are those that matter to criminality and wrongfulness"* (Sinnott-Armstrong & Levy, 2011, p. 312). In other words, if it is evident that the defendant could not understand the factual nature of what they were doing, the moral and legal implications of their actions are consequently also lost on them.

2.5. Cognitive impairment: understanding unlawfulness: morality

The first caveat in the context of moral cognition is the nature of morality itself. By its very nature, morality, "a code of values and customs that guide social conduct", is subject to varying content (Mendez, 2009, p. 609). Nevertheless, moral cognitive neuroscience aims to discover moral correlates in the brain and discern their function.

The second caveat has to do with the different ways that morality can be tested in a research setting – for example by using vignettes to describe hypothetical scenarios or presenting the participant with images. Research has also demonstrated that neural response patterns vary for different kinds of moral problems (Funk & Gazzaniga, 2009, p. 679). Moreover, variables such as the predictability of the experimental patterns and the duration of the exposure to stimuli have evoked different neural reactions (Forbes & Grafman, 2010, pp. 307-308). An additional problem is that the whole situation is often, especially in a laboratory environment, lacking ecological validity. Socio-cultural differences as regards the participants' conception of morals also constitute a variable which is difficult to control. Presumably in an attempt to attain a level of coherence, it appears that contemporary research on moral cognition has largely relied on the Greene et al. test battery from 2001 (Lotto et al., 2013, n.a.). The battery, which contains an extensive set of non-moral, impersonal moral as well as personal moral dilemmas, is readily available online.

Philosophers maintain that morality can be of either the "descriptive" or "normative" type, that is, moral values which are tied to and upheld by a certain social group, or a moral code that is common to all reasonable persons irrespective of the specific moral norms of their surroundings (Mendez, 2009, p. 609). It has also been argued that the moral aspect of wrongfulness should be divided into "objective" and "subjective" components, the former describing a mentally disordered individual deprived of the capacity to appreciate that the society condemns his actions, and the latter to an individual who knows that his actions are against the law but who considers them to be "personally morally justified" (Knoll & Resnick, 2008, pp. 93-94). This paper aims to predominantly address normative morality together with the objective component of moral wrongfulness.

2.5.1. Neural correlates of moral cognition

The understanding of the moral aspect of unlawfulness implies that the applicant should have made or attempted a moral judgment of their actions at the time of the offense. The word 'understand' suggests that the standard of 'knowledge' is insufficient. For the purposes of this thesis, the word 'moral judgment' is defined as an evaluation established through assessments of the appropriateness of one's own conduct in accordance with socially constructed notions of right and wrong (Moll et al., 2005, p. 807). A successful moral judgment necessarily implies underlying moral knowledge.

It has been argued that morality originated as a result of evolutionary development in the prefrontal cortex (Knabb et al., 2009, p. 222). Corroborative scientific evidence drawn from several disciplines appears to confirm that morality is indeed innate and bound to the brain (Shoemaker, 2012, p. 816). A recent study found that transcranial direct-current stimulation could alter moral conduct and judgment in healthy individuals (Fumagalli & Priori, 2012, p. 2008). The neural regions recruited for moral cognition in normals have also been shown to be approximately the same as the ones which are related to both developmental and acquired sociopathy (de Oliveira-Souza & Moll, 2009, p. 267). The same overlap is true for the regions connected to antisocial behaviour in general. Interestingly, research implies that instead of problems regarding moral knowledge, antisocial groups principally exhibit deficiencies in "feeling" what is moral (Raine & Yang, 2006, pp. 209-210). Research has pinpointed what is referred to as a "neuromoral network", allowing humans to react to moral problems which they are faced with (Mendez, 2009, p. 608). This network is fairly complex and has so far been shown to include several regions of the brain. In this context, it has been put forward that any moral judgment and behaviour demands the combination of several neural processes: *"the decoding of signals perceived by the sensitive organs (thalamus), the activation of basic emotions (anteromedial temporal lobe, brain stem, and the nuclei of visceromotor centres), the awareness of the relevance and importance of the stimuli (VMPFC and OFC), and the implementation and control of potentially related forms of behaviours (frontal lobes)"* (Marazziti et al., 2013, p. 7).

However, what has received most attention in this regard is the ventromedial prefrontal cortex (VMPFC) (Moll & de Oliveira-Souza, 2007, p. 319). Functional magnetic resonance imagining can show the activation of the VMPFC during both exercises which necessitate explicit moral judgments, as well as passive observing of morally significant images. Research suggests that the VMPFC is especially involved in personal moral dilemmas – in which a direct action by the individual could lead to another person being severely harmed (Mendez, 2009, p. 610). This is relevant to the consideration of crimes where the accused is suspected of directly inflicting morally inappropriate harm on the victim.

An experiment conducted on 12 normals and 7 individuals with lesions located in the VMPFC tested for deviations in moral decision-making by presenting the participants personal and impersonal moral dilemmas as well as non-moral dilemmas to control for general executive defects. Owing to the role of the VMPFC in moral decision-making as presented in neuroimaging literature, it was hypothesized that the VMPFC-damaged participants should be more ready to accept moral wrongs in personal dilemmas (Ciaramelli et al., 2007, p. 85). The findings of the study are as follows: the patient group did not show abnormalities when solving non-moral problems or impersonal moral dilemmas when compared to the control group, but approved personal moral wrongs more often and faster than the normals. In other words, the patients, unlike the normals, did not authorize fewer personal moral wrongs in comparison to impersonal moral wrongs (Ciaramelli et al., 2007, pp. 87-88).

Moreover, the time that patients required for rejecting moral wrongs corresponded to the time that the normals took, indicating that the patients' moral knowledge was intact and that they could utilize it. The proposition that VMPFC lesions do not deprive individuals of their moral knowledge is corroborated by other studies. All in all, the findings of the experiment attest that VMPFC damage can cause a very specific type of deficiency related to personal moral judgment, despite the fact that the patients have retained moral knowledge and the capacity to judge impersonal moral problems in a normal manner. Interestingly, individuals with frontotemporal dementia have demonstrated comparable behaviour (Ciaramelli et al., 2007, p. 89). The results of this study should however be interpreted with caution, not least because of the small sample size.

Lesions or disorders affecting the right VMPFC have been also shown to impair moral emotions (Mendez, 2009, p. 616). Moreover, the VMPFC has been associated with enabling persons to conform to social norms (Fumagalli & Priori, 2012, p. 2007). On the other hand, studies conducted on healthy subjects indicate, in addition to VMPFC activity for a variety of tasks requiring moral reasoning, the involvement of the frontopolar cortex (FPC in this context, also referred to as the frontopolar prefrontal cortex or FPPFC). This activity in the VMPFC-FPC was accompanied by activations in the anterior temporal cortex (ATC), superior temporal sulcus (STS) in the temporal lobe as well as the limbic structures (Moll & de Oliveira-Souza, 2007, p. 319). The limbic system incorporates the amygdalae, hippocampus, hypothalamus as well as the cingulate cortex and the basal forebrain. It has been argued that the comprehensive functioning of the limbic system, which feeds input to the prefrontal cortex, is critical for proper moral judgment (Casebeer, 2003, pp. 843-844). Systematic VMPFC-FPC activation has been linked to both passive responses to morally salient stimuli and explicit individual moral judgments (Moll & de Oliveira-Souza, 2007, p. 321).

Research on normals has further shown the engagement of the amygdalae, orbitofrontal and ventrolateral cortex (OFC/VL) as well as the dorsolateral prefrontal cortex (DLPFC) in moral tasks (Mendez, 2009, p. 609). Generally, the DLPFC has been associated with rule-based executive functions such as decision-making (Schleim et al., 2011, p. 49). Activity in the DLPFC has been correlated with responding to specifically impersonal moral tasks (Mendez, 2009, p.610). In contrast, it has also been claimed that demanding personal moral dilemmas, as opposed to simple ones, elicit "control-related" activity in the anterior DLPFC after initial "conflict-related" activity in the anterior cingulate cortex (ACC) (Funk & Gazzaniga, 2009, p. 679). Corroborating evidence exists indicating the engagement of the ACC in moral judgments. The DLPFC and the ACC are, to a great extent, connected, meaning that they also have "functional similarities" (Knabb et al., 2009, p. 224). Repetitive transcranial magnetic stimulation as applied into the right DLPFC region was shown to modify moral judgment in 24 healthy participants. More specifically, high-conflict personal (subjective) moral judgments were affected (Tassy et al., 2012, pp. 283-287).

As regards the OFC, located in the ventral prefrontal cortex (VPFC), case studies suggest that adults who suffered OFC trauma in their early childhood before the age of 16 months have disturbances in their moral reasoning. Furthermore, the evidence indicates that such lesions interfere with both moral decision-making as well as actual knowledge of right and wrong (Redding, 2006, p. 74). Persons with OFC trauma sustained as an adult, whilst demonstrating equivalent behavioural abnormalities, achieved normal results in standardized moral reasoning examinations in comparison to those with childhood OFC trauma, who exhibited puerile reasoning (Casebeer, 2003, p. 843).

It should be noted that there is a lack of clear consensus concerning the precise anatomical boundaries of the VMPFC and its relation to the orbitofrontal cortex (OFC) (de Oliveira-Souza & Moll, 2009, p. 261). It is argued that the VMPFC overlaps, in part, with the medial sections of the OFC. This makes it problematic to separate these regions in literature. Trauma to these areas may include degenerative disorders of the brain, strokes, tumours, and excisions due to surgery as well as various head injuries (Zald & Andreotti, 2010, p. 3378). A 2012 meta-analysis of existing neurological research about moral cognition found concurrent activity in the ventromedial (VMPFC), frontopolar (FPPFC), and dorsomedial (DMPFC) prefrontal cortices, the temporoparietal junction (TPJ), the precuneus, posterior cingulate cortex (PCC), the left amygdala, the right temporal pole as well as the right middle temporal gyrus (RMTG) (Bzdok et al., 2012, p. 787). On the other hand, a 2013 review article encompassing three decades of research points at a primary role for the VMPFC, DLPFC, VLPFC and OFC as well as the amygdala in human morality. In addition to the four primary neural regions, some research in normals has indicated the involvement of the inferior parietal lobes, the TPJ,

the anterior insula, the anterior cingulate gyrus (ACG), posterior cingulate cortex (PCC), the posterior superior temporal sulcus (PSTS), the precuneus as well as the mesolimbic pathway (MLP) and the ventral striatum (Marazziti et al., 2013, p. 4)

Current evidence from healthy brains thus points towards the fact that the frontal lobe is largely responsible for moral behaviour. The temporal and parietal lobes as well as subcortical structures such as the amygdalae have also been implicated in moral judgments (Fumagalli & Priori, 2012, pp. 2008-2011). Some even go as far as to say there is "remarkable agreement between functional imaging and clinico-anatomical evidence" concerning the neural regions associated with moral cognition – most prominently, the FPPFC together with Brodmann's area 9 (BA9), the OFC, the posterior superior temporal sulcus, insula, precuneus, ACC, the anterior temporal lobes and the limbic regions (Moll et al., 2005, p. 800). It should be noted here that BA9 is a part of the DLPFC (Martins-de-Souza et al., 2011, p. 2347).

2.6. Cognitive impairment: understanding unlawfulness: legality

Legal wrongfulness has been described, inter alia, as the individual's "concrete understanding at the time of the offense that his act was against the law" (Knoll & Resnick, 2008, p. 93). Unlike the neuromoral network pointing to the localisation of normative morality, the legal element of unlawfulness is more subjective, at least in content. This does not mean, however, that it is impossible to shed light on the neurological correlates of legal judgment.

2.6.1. Neural correlates of legal cognition

A recent fMRI study on 40 healthy individuals, of which half were lawyers and the remaining half other academics, found that performing legal and moral judgments employ the same regions of the brain, namely the DLPFC, the posterior cingulate gyrus and precuneus as well as the left temporoparietal junction (TPJ). The fact that the two types of judgments produced comparable neural responses points towards a 'considerable overlap in cognitive processing' for moral and legal dilemmas.

As already stated above, the DLPFC is associated with deliberating on explicit rules. In this context, it was found that legal judgments, when compared to moral judgments, were correlated with substantially greater activation of the left DLPFC, implying that legal decisions relied less on intuition and more on clear-cut rules. In the legal condition greater activation was also observed in the left middle temporal gyrus (LMTG) (Schleim et al., 2011, p. 55). The methodology of the study involved short stories which were open

to interpretation from both a legal and moral perspective and the participants had to determine whether the conduct was right or wrong in the condition (legal or moral) which they were assigned to (Schleim et al., 2011, p. 49).

2.7. Volitional impairment: decisively weakened ability to control behaviour

The Criminal Code's reference to "decisively weakened ability to control behaviour" in the context of volitional impairment is manifestly vague. The sparse case law of the Finnish Supreme Court on irresponsibility has considered, in this instance, neuroscientific evidence attesting to brain damage which has weakened an individual's impulse control (KKO:2008:79). Another case presents considerations of similar nature, namely the capability to regulate one's actions and to refrain from or to discontinue their act (KKO:1987:130). In layman's terms, one might want to refer to the apparent standard as that of willpower.

It can be opined that the behavioural control criterion of the criminal law is not so much about being able to produce voluntary acts but rather being able to prevent acts which do not necessarily reflect the individual's will: the penal law is concerned about one's ability to conform their behaviour to the law, not whether one is capable of instigating voluntary acts. Some legal systems abide by what is known an as irresistible impulse doctrine. The irresistible impulse standard considers whether the accused had the capacity to choose their course of action or the ability to regulate their behaviour (Carrido, 2012, p. 316).

As mentioned above, the volitional prong of the irresponsibility test has been a matter of considerable controversy. Whereas some maintain that it should be excluded from irresponsibility considerations altogether, some argue that it should in fact be the only constituent element of the test (Corrado, 2009, pp. 482-483). The middle ground asserts that both cognitive and volitional impairment can and should be accommodated by the law. Those who oppose the volitional aspect commonly justify their point of view by the fact that it is problematic to evaluate in comparison to the cognitive prong (Penney, 2012, p. 101).

Modern neuroscience, with its growing body of evidence attesting to an organic basis for volitional impairment, makes a strong case for the justification of a volitional prong in irresponsibility evaluations (Sapolsky, 2004, p. 1790). As a result of such findings, calls have been made for a "neurojurisprudence" to emerge and for the introduction of control tests (Redding, 2006, p. 53). However, a troublesome feature of the existing literature is that it approaches behavioural control (or the lack of it) from a variety of perspectives. Owing to

space constraints, this thesis will focus on two of these; namely, behavioural inhibitory control and impulsive action. The relevant literature appears to disagree on the precise relationship between the two concepts, but for the purposes of this thesis it suffices to note that they are interconnected.

Behavioural inhibitory control (BIC) is regarded as a necessary tool for persons within any societal order (Yuan, Meng, Jang, Yao, Hu & Yuan, 2012, p. 240). Its function is to repress either an implicit or explicit response to stimuli, which has been attributed primarily to the OFC region. Research has suggested the intercommunication of a number of neural structures, both in the cerebral cortex as well as the subcortical region in relation to inhibitory control. A model of inhibitory control, *executive inhibition*, posits that inhibition consists of cognitive inhibition, behavioural inhibition as well as interference control. Difficulties in executive inhibitory control have been linked to impulsivity. It is typically reasoned that impulsive action results from the "inhibitory dyscontrol" of an improper response, and that adequate inhibitory control allows for an interval, thus enabling normal executive functions to take place (Enticott et al., 2006, p. 286). A study on 31 healthy adults using cognitive tests and self-reports found that particular types of inhibitory dyscontrol could be, to a certain extent, responsible for impulsive conduct (Enticottet al., 2006, p. 292). Impulsivity can be defined as a "predisposition towards rapid, unplanned reactions to internal or external stimuli with a lack of regard for the negative consequences of these reactions to the impulsive individual or to others" (Muresanuet al., 2012, p. 16). Some academics have criticized impulsivity research for "imprecise understanding of the underlying cognitive cause" behind the phenomenon (Enticottet al., 2006, p. 286). It is also important to note that as is the case with measuring moral cognition, there are several possible techniques of testing and measuring impulsiveness. This may result in a mass of literature that consists of results attained with divergent methodology. On the contrary, it has been claimed that neuroimaging research on impulsiveness "has used a fairly narrow set of task paradigms [which] share the characteristic that successful performance requires the inhibition of a prepotent response" (Congdon & Canli, 2008, p. 1454). Whereas impulsivity *per se* is usually measured using self-reports, prepotent response inhibition can be measured using a variety of experimental tasks (Aichert et al., 2012, p. 1017).

An example of this is the standard Go/NoGo test, which analyses an individual's ability to inhibit a response during NoGo trials in spite of a predominant Go response which results from a comparatively much higher incidence of Go trials within the test (Zald & Andreotti, 2010, p. 3381). The utility of this method will be elaborated further below. A second commonly used test is the Stop-Signal task, which similarly requires the suppression of a response. The task consists of instructing the subjects to react to a specific set of stimuli but to suppress

their response when they perceive a stop signal (Congdon et al., 2008, p. 27).
It appears, however, that these two tasks are not interchangeable. A study on 504 normals found that self-reported impulsivity using the Barratt Impulsivity Scale was associated with the Go/NoGo but not the Stop-Signal test or the two supplementary test patterns. A recent meta-analysis of neuroscientific research also found that activation in the right inferior parietal lobule (RIPL) and the right middle frontal gyrus (RMFG) during Go/NoGo tasks was considerably higher when compared to performance in the Stop-Signal task. Furthermore, additional research seems to back the hypothesis that it is indeed the Go/NoGo paradigm which is more successful in predicting trait impulsivity (Aichert et al., 2012, p. 1026).
It should be noted that these tasks, much like other standardized laboratory experiments, suffer from a lack of ecological validity. This is demonstrated in the context of impulsivity research by the fact that they fail to elaborate on important moderating factors such as autonomic physiological arousal (Enticott et al., 2006, p. 286). In practice, this means that the results cannot be easily transferred outside of the laboratory environment because of their limited applicability to real-life situations. On an entirely different note, it should be noted that the metaphysical conflict of determinism and free will is excluded, to the extent that it is possible, from the following analysis.

2.7.1. Neural correlates of impulsivity/behavioural control

The overwhelming majority of neuroscientific evidence points towards the significance of the frontal lobe in behavioural control. Impulsivity is often related with aggression, and a large proportion of research has investigated impulsive aggression in particular. In this context, it must be noted that criminal irresponsibility is not theoretically restricted to crimes which involve aggression per se.
Early studies using the Go/NoGo test in primates demonstrated a link between OFC lesions and poor inhibitory performance. Subsequent research on humans supports these initial findings, linking lesions in the prefrontal region to difficulties in the NoGo trials (Zald & Andreotti, 2010, p. 3381). More specifically, activity during Go/NoGo tasks indicates the predominant involvement of a neural system consisting of the VPFC, DLPFC, parietal cortex as well as the striatum and ACC (Liu et al., 2012, p. 2).
It has been suggested that atypical impulsivity in humans follows from an imbalance of the neural circuit connecting the limbic system to the frontal lobe region. The two elements can be referred to as the impulsive system, which includes the amygdala producing an instantaneous signal of comfort or discomfort, and the reflective system, which employs the VMPFC and analyses the feedback and reflects on the long-term effects of different

behavioural responses, acting as a buffer. In other words, impulsivity is allegedly caused by the reflective system's inability to moderate the commands relayed from the impulsive system (Penney, 2012, p. 100).

There are at least two variants of frontal lobe pathology – one affecting the VMPFC and the other the DLPFC, of which the former seems to predispose individuals to impulsive behaviour and committing impulsive acts of violence, and the latter corrupts "judgment and moral reasoning" (Redding, 2006, p. 68). Nonetheless, as presented above, the VMPFC has also been constantly associated with moral decision-making.

Research has also shown that persons with VMPFC lesions show more impulsivity in comparison to those with lesions elsewhere in the frontal cortex outside of the OFC (Matsuo et al., 2009, p. 1189). In comparison to normals, impulsive individuals have also exhibited decreased gray matter volume in the hippocampus and DLPFC. More specifically, subjects with impulse control disorders displayed less activity in the DLPFC than normals when undertaking aggression control tasks (Penney, 2012, p.100).

study using voxel-based morphometry (VBM) found that normal subjects who scored high on the Beckett Impulsivity Scale presented with lesser right and left OFC gray matter volumes in comparison to those whose scores were low. An association between a low volume of gray matter in the ACC and high impulsivity was also implied by the results, supporting the hypothesis that the VMPFC is relevant for impulse control at least through these two regions (Matsuo et al., 2009, pp. 1191-1194).

As already mentioned, the ACC is associated with the volitional control of behaviour (Peoples, 2002, p. 1623). The supplementary motor area (SMA) has also been implicated in response inhibition. A meta-analysis of 11 studies which utilized either simple or complex Go/NoGo tasks – the latter demanding additional working memory capacity – found concurrent activation of the pre-SMA in both categories of studies. The fact that the pre-SMA was the only region to feature independently of the task implies that it is crucial for the inhibition of responses (Simmonds et al., 2008, p. 230). A review on both transcranial magnetic stimulation and transcranial direct current stimulation concluded that results achieved using these methods implicate – *inter alia* – the pre-SMA, inferior frontal gyrus (IFG) as well as the frontal eye fields in effective response inhibition (Juan & Muggleton, 2012, p. 67).

As demonstrated above, not all evidence linked to behavioural inhibition points solely towards the frontal lobe. A comprehensive new study on mice implicated the medial habenula-interpeduncular nucleus (mHb-IPN) pathway in inhibitory control, regarding both impulsiveness and compulsiveness (Kobayashi et al., 2013, pp. 17-18). The results as such, of course, cannot be generalised to human biology, but may provide an interesting framework for future studies in humans.

Psychopharmacology may have something to contribute to the discussion as well. Dopamine, a neurotransmitter, has been associated with impulsivity in both human and animal studies (Congdon & Canli, 2008, p. 1459). Low levels of serotonin have also been correlated with impulsivity, especially in the cases of criminal adults with violent tendencies (Reider, 1998, p.325). The implications on serotonin levels are also anatomically relevant owing to the fact that the PFC has a high concentration of serotonin receptors. Furthermore, the PFC is connected to serotonergic nerves, and thus it is reasonable to predict that prefrontal disturbances have an effect on neural serotonin activity (Bufkin & Luttrell, 2005, p. 185).

In adolescents, it has been observed that the incomplete development of the frontal lobe presents itself as impulsiveness (Redding, 2006, p. 65). From the cognitive point of view, children demonstrate very limited prefrontal cortex function, including but not limited to, moral reasoning and proper impulse control. This is hardly surprising considering the well-established fact that the myelination of the prefrontal cortex is not complete until early adulthood (Sapolsky, 2004, p. 1792). It may be interesting to note that research points towards the heritability of impulsiveness. Three separate twin studies proposed a genetic influence of an estimated 45%, 44% and 45%, respectively (Congdon & Canli, 2008, p. 1458). Interestingly, there also appears to be a link between genetics and dopamine. Using an endophenotype-based approach, aiming to pinpoint such 'intervening variables' between genetic information and a specific behaviour that are susceptible to variation in alleles, researchers found a connection between two dopamine-related polymorphisms and behavioural inhibition on a Stop-Signal test. More specifically, those individuals out of a sample of 119 normals who possessed both the dopamine receptor D4 (DRD4) 7-repeat allele and the 10/10 genotype of the dopamine active transporter (DAT), displayed the longest stop signal response times indicating poor inhibitory control. In other words, the trait for inhibition was correlated with the dopamine-related genetic makeup of persons. The study also found that the Stop-Signal test was a more accurate measure of behavioural inhibitory control than the Barratt Impulsivity Scale (Congdon et al., 2008, pp. 27-31).

2.8. Summary of findings

The elementary review of neuroscientific research as presented above perhaps raises more questions than it can provide answers. Nonetheless, a number of legally relevant neural correlates of behaviour could be identified. These consist of regions of interest (ROIs), overarching neural networks as well as neurochemical and genetic factors. The findings will be summarized in the framework and order of Chapter 3 Section 4 subsection 2 of the Finnish Criminal Code.

The examination of the first relevant criterion, *understanding the factual nature of the act,* was undertaken from the perspective of perception and reality distortion, with a specific focus on the putative neural basis of hallucinations and delusions. Perception was associated with activity in the corresponding cortical sensory regions, but most likely cannot be localized into a single cortical sub-region. A PET scan experiment – the results of which are supported by additional research – found indicators of a connection between constant hallucino-delusional symptoms and malfunction in temporal regions as well as Broca's area. Lesions in auditory and visual pathways have been implicated in hallucinations. Most interestingly, a groundbreaking study identified blood biomarkers pertaining to hallucinations and delusions. All in all, research suggests an organic basis for reality distortion in this context.

The second criterion, *understanding unlawfulness* was divided into two components – *moral unlawfulness* and *legal unlawfulness*. Moral unlawfulness was investigated through the much-researched concept of moral cognition – undertaking moral judgments – including aspects of moral knowledge. Evidence indicates that the brain regions necessary for moral cognition are correspondent in healthy and dysfunctional populations. In fact, what has emerged is a so-called neuromoral network.

Much of the literature has focused on the VMPFC region. VMPFC damage is particularly associated with abnormalities in judging personal moral dilemmas despite the fact that moral knowledge is intact and no deviance is present when evaluating impersonal moral dilemmas. Activation in the VMPFC when undertaking moral tasks is associated with concurrent activity in the FPC (alternatively FPPFC) as well as the ATC and the STS, along with limbic structures. It has been postulated that successful moral judgment requires normal functioning of the limbic system.

Another neural area which has been the source of considerable scrutiny is the DLPFC, a part of the prefrontal cortex connected to rule-based executive functions. There is some disagreement about the precise involvement of the DLPFC in moral cognition, however. The region has been implicated in both impersonal as well as personal moral judgment, together with activity in the ACC. It has been postulated that damage to the OFC also leads to impairment in not only moral cognition but also moral knowledge.

Moreover, the combined output of two extensive review articles, from 2012 and 2013, respectively, implicates the VMPFC, VLPFC, FPPFC, DMPFC, DLPFC, OFC, ACG, TPJ, the left amygdala, precuneus, anterior insula, inferior parietal lobes, PSTS, PCC, right temporal pole, RMTG, mesolimbic pathway as well as the ventral striatum in moral cognition.

It is immediately apparent just how significant the prefrontal cortex is for human morality. In comparison, research on legal cognition is very scarce. The existing fMRI literature has specifically compared legal judgments with moral judgments, and the findings suggest that there are notable similarities in the regions employed for both types of judgments. Examples of this are the DLPFC and the TPJ. Nevertheless, in the legal condition, activation in the left DLPFC and LMTG was increased when compared to activity during moral judgments.

The third criterion was that of *behavioural control*. This was considered from the point of view of both impulsivity and response inhibition, as the precise nature of the relationship between the two is under dispute. Research suggests that prefrontal pathologies have an adverse effect on response inhibition. Functional imaging has linked the DLPFC, VPFC, ACC, parietal cortex and the striatum with inhibition during experimental conditions. Further regions of interest seem to be the pre-SMA as well as the IFC and frontal eye fields. The mHb-IPN pathway has also been associated with inhibitory control in rodents. A study on endophenotypes discovered a relationship between response inhibition and two dopaminergic gene polymorphisms, namely DRD4 and the DAT, implying a link between genetics and dopamine in this regard.

As regards impulsivity *per se*, VMPFC pathologies have been associated with proneness towards committing impulsive actions, including violent ones. Studies have also indicated that hippocampal and DLPFC gray matter is reduced in persons with impulsive symptoms but not in healthy individuals. Similarly, an experiment utilizing VBM demonstrated that normals with high impulsivity scores had reduced OFC gray matter volume when compared to those participants with low scores. The same study also implicated reduced ACC gray matter volume in high impulsivity scores. Of the most common neurotransmitters, dopamine and serotonin are said to play a role in impulsive behaviour. Lastly, three twin studies came to the conclusion that impulsiveness may well be an inherited trait.

It must be acknowledged that this review does not offer a sufficient analysis of procedural problems that the interpretation and application of neuroscientific information raises in the legal realm. Its approach is also notably reductionist. Moreover, an additional caveat is that of the modal fallacy – the incorrect assumption that poor performance in tasks involving neuroimaging necessarily means that the subject *lacks* the capacity to perform well. It may very well be that the individual possesses the capacity but does not utilize it (Vincent, 2011, p. 45).

3. Suggestions

3.1. General suggestions

The reviewed literature makes a strong case for the existence of neurological correlates of reality distortion, moral knowledge and cognition, legal judgment as well as behavioural inhibitory control and impulsivity. The primary role of the prefrontal cortex in cognitive and volitional impairment is indisputable. What is the law make of these findings? The studies examined above have been conducted for purely experimental purposes and serve the interests of science. Moreover, their ecological validity is low, a concern that has already been addressed above. The division of human behaviour into components that are deemed relevant to criminal responsibility is also inevitably artificial. This separation of the three main elements of moral and legal cognition and behavioural control in no way implies that they are unrelated to each other – simply put; they had to be examined separately for the purposes of coherence.

Perhaps the logical next step would be to begin to study the aforementioned phenomena in a forensic setting. This would involve the composition of neurolaw test paradigms especially adapted for forensic use (Meynen, 2013, p. 96). A commendable example of the use of neuroscience in this context is provided by the Rigoni et al. 2010 case study *'How neuroscience and behavioral genetics improve psychiatric assessment: report on a violent murder case'*. Arguably, it is also becoming a necessity for legal professionals to acquaint themselves with neuroscience. The process of familiarization could even be initiated on the level of legal education. However, one must be careful to not place too much emphasis on the role of neuroscience in the law. After all, "[i]f the brain findings and behavior are inconsistent, the behavior must be our guide" (Morse, 2007, p. 13).

3.2. Specific suggestions

On a more specific note, the new findings on blood biomarkers of hallucinations and delusions warrant additional research due to their possible utility to the law. As regards the neuroscientific study of moral cognition, it could be beneficial for forensic experts to adopt a version of the Greene et al. 2001 standardized test battery, or at least use it as a starting point in the development of fMRI 'neuroforensic' tests. It has been successfully repeated in several studies since its publication and it appears to produce consistent results. Of special relevance to irresponsibility determinations are its personal moral

dilemmas. Legal cognition, on the other hand, ought to be studied more extensively in order to better understand judgments pertaining to explicit rather than implicit (moral) rules. Finally, it is suggested that the Go/NoGo paradigm of response inhibition executed in combination with fMRI could serve as one preliminary measure of volitional impairment, *mutatis mutandis*.

4. Conclusion

To conclude, the law should not become fixated on diagnostic labels, but rather focus on the effect of the disturbance on the defendant's behaviour. A mental disorder or neural dysfunction on its own can never suffice to exculpate. Nevertheless, the law ought not to ignore neuroscientific advancements but rather attempt to harness these tools in order to facilitate and improve the determinations that it must undertake.

The findings of this thesis point towards the usefulness of structural and functional neuroimaging, and to a certain extent, genetics and neurochemistry in corroborating forensic assessments of criminal irresponsibility. Of course, moral and legal cognition and behavioural inhibition are rather crude representations of cognitive and volitional impairment. The law prosecutes humans, not brains, and its determinations are and will remain normative.

Nevertheless, a capacitarian neurolaw approach could assist in the objective evaluation of the capacities of an individual to understand the factual nature and unlawfulness of their act and to conform their behaviour to the law. No capacity implies no responsibility, but contemporary neuroscience is not fit to make these assessments on its own. This does not mean, however, that it has nothing to offer. As long as the limitations of neuroscience are understood, it poses no threat to the administration of justice. Perhaps it is time to stop calling neurolaw the law of the future, and call it the law of the present instead. This is not to suggest that neuroscience should suddenly be absorbed into the legal realm, but rather that it be treated as fact not science-fiction. As long as the aim of neuroimaging is to explain human behaviour and the aim of the law is to control it, there should be no reason for law to reject the assistance of neuroscience in irresponsibility assessments.

Bibliography

Case law of the Finnish Supreme Court:

KKO:1987:130
KKO:2000:126
KKO:2005:48
KKO:2008:79
KKO:2009:56

Books and articles:

Husak, D. (2013). Reply: The Importance of Asking the Right Question: What is Punishment Imposed For? In R.L. Christopher (Ed.), *George Fletcher's Essays on Criminal Law* (pp. 53-57). Oxford: Oxford University Press.

Klaming, L. & Koops, B.J. (2012). Neuroscientific Evidence and Criminal Responsibility in the Netherlands. In T.M. Spranger (Ed.), *International Neurolaw. A Comparative Analysis* (pp. 227-256). Heidelberg: Springer.

Sinnott-Armstrong, W. & Levy, K. (2011, p. 312). Insanity defenses. In J. Deigh & D. Dolinko (Eds.), *The Oxford handbook of philosophy of criminal law* (pp. 299-334). Oxford: Oxford University Press.

Tadros, V. (2007). *Criminal Responsibility.* Oxford: Oxford University Press.

Aharoni, E., & Funk, C., & Sinnott-Armstrong, A., & Gazzaniga, M. (2008). Can Neurological Evidence Help Courts Assess Criminal Responsibility? Lessons from Law and Neuroscience. *Annals of the New York Academy of Sciences 1124*, pp. 145-160.

Aichert, D.S., & Wöstmann, N.M., & Costa, A., & Macare, C., & Weinig, J.R., & Möller, H-J., & Rubia, K., & Ettinger, U. (2012). Associations between trait impulsivity and prepotent response inhibition. *Journal of Clinical and Experimental Neuropsychology 34*(10), pp. 1016-1032.

Batts, S. (2009). Brain lesions and their implications in criminal responsibility. *Behavioral Sciences and the Law 27*(2), pp. 261-272.

Bennett, M.R. (2008). Consciousness and hallucinations in schizophrenia: the role of synapse regression. *Australian and New Zealand Journal of Psychiatry 42*, pp. 915-931.

Broome, M.R., & Bortolotti, L., & Mameli, M. (2010). Moral Responsibility and Mental Illness: A Case Study. *Cambridge Quarterly of Healthcare Ethics 19*(2), pp. 179-187.

Bufkin, L., & Luttrell, V.R. (2005). Neuroimaging Studies of Aggressive and Violent Behavior : Current Findings and Implications for Criminology and Criminal Justice. *Trauma, Violence, & Abuse 6*(2), pp. 176-191.

Bzdok, D., & Schilbach, L., & Vogeley, K., & Schneider, K., & Laird, A.R., & Langner, R., & Eickhoff, S.B. (2012). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Struct. Funct 217*, pp. 783-796.

Carrido, M. (2012). Revisiting the Insanity Defense: A Case for Resurrecting the Volitional Prong of the Insanity Defense in the Light of Neuroscientific Advances. *Southwestern Law Review 41*, pp. 309-330.

Casebeer, W.D. (2003). Moral cognition and its neural constituents. *Nature Reviews Neuroscience 4*, pp. 841-846.

Ciaramelli, E., & Muccioli, M., & La'davas, E., & di Pellegrino, G. (2007). Selective deficit in personal moral judgment following damage to ventromedial prefrontal cortex. *SCAN 2*, pp. 84-92.

Congdon, E., & Canli, T. (2008). *A Neurogenetic Approach to Impulsivity. Journal of Personality 76(6) pp. 1447-1484.*

Congdon, E., & Lesch, K.P., & Canli, T. (2008). Analysis of DRD4 and DAT Polymorphisms and Behavioral Inhibition in Healthy Adults: Implications for Impulsivity. *American Journal of Medicinal Genetics Part B: Neuropsychiatric Genetics 147 B,* pp. 27-32.

Corrado, M. (2010). The Case for a Purely Volitional Insanity Defense. *Texas Tech Law Review 42,* pp. 481-511.

de Oliveira-Souza, R., & Moll, J. (2009). The Neural Bases of Normal and Deviant Moral Cognition and Behavior. *Top Magn. Reson. Imaging 20*, pp. 261-270.

Eastman, N., & Campbell, C. (2006). Neuroscience and legal determination of criminal responsibility. *Nature Reviews Neuroscience 7,* pp. 311-318.

Enticott, P.G., & Ogloff, J.R.P., & Bradshaw, J.L. (2006). Associations between laboratory measures of executive inhibitory control and self-reported impulsivity. *Personality and Individual Differences 41*, pp. 285-294.

Forbes, C.E., & Grafman, J. (2010). The Role of the Human Prefrontal Cortex in Social Cognition and Moral Judgment. *Annu. Rev. Neurosci. 33,* pp. 299-324.

Fumagalli, M. & Priori, A. (2012). Functional and clinical neuroanatomy of morality. *Brain 135*, pp. 2006-2021.

Funk, C.M. & Gazzaniga, M.S. (2009). The functional brain architecture of human morality. *Current Opinion in Neurobiology 19*, pp. 678-681.

Garland, B., & Glimcher, P.W. (2006). Cognitive Neuroscience and Law. *Current Opinion in Neurobiology 16,* pp. 130-134.

Gazzaniga, M.S. (2008). The Law and Neuroscience. *Neuron 60,* pp. 412-415.

Gotlieb, A.E. (1956). Intention, and Knowing the Nature and Quality of an Act. *The Modern Law Review 19*(3), pp. 270-275.

Juan, C-H., & Muggleton, N.G. (2012). Brain stimulation and inhibitory control. *Brain Stimulation 5*, pp. 63-69.

Juan, Y., & Meng, X., & Yang, J., & Yao, G., & Hu, L., & Yuan, H. (2012). The valence strength of unpleasant emotion modulates brain processing of behavioral inhibitory control: Neural correlates. *Biological Psychology 89*, pp. 240-251.

Kays, J.L., & Hurley, R.A., & Taber, K.H. (2012). The Dynamic Brain: Neuroplasticity and Mental Health. *J Neuropsychiatry Clin. Neurosci. 24*(2), pp. 118-124.

Knabb, J., & Welsh, R.K., & Ziebell, J.G., & Reimer, K.S. (2009). Neuroscience, Moral Reasoning, and the Law. *Behavioral Sciences and the Law 27,* pp. 219-236.

Knoll, J.L., & Resnick, P.J. (2008). Insanity Defense Evaluations: Toward a Model for Evidence-Based Practice. *Brief Treatment and Crisis Intervention 8*(1), pp. 92-110.

Kobayashi, Y., & Sano, Y., & Vannoni, E., & Goto, H., & Suzuki, H., & Oba, A., & Kawasaki, H., & Kanba, S., & Lipp, H-P., & Murphy, N.P., & Wolver, D.P., & Itohara, S. (2013). Genetic dissection of medial habenula-interprenducular nucleus pathway function in mice. *Frontiers in Behavioral Neuroscience 7*, pp. 1-20.

Kolb, B. & Whishaw, I.Q. (1998). Brain Plasticity and Behavior. *Annu. Rev. Psychol. 49*, pp. 43-64.

Kurian, S.M. & Le-Niculescu, H., & Patel, S.D., & Bertram, D., & Davis, J., & Dike, C., & Yehyawi, N., & Lysaker, P., & Dustin, J., & Caligiuri, M., & Lohr, J., & Lahiri, D.K., & Nurnberger Jr, J.I., & Faraone, S.V., & Geyer, M.A., & Tsuang, M.T., & Schork, N.J., & Salomon, D.R., & Niculescu, A.B. (2011). Identification of blood biomarkers for psychosis using convergent functional genomics. *Molecular Psychiatry 16*, pp. 37-58.

Liddle, P.F. (1997). Dynamic neuroimaging with PET, SPET or fMRI. *International Review of Psychiatry 9*, pp. 331-337.

Liu, J., & Zubieta, J-K., & Heitzeg, M. (2012). Sex differences in anterior cingulate cortex activation during impulse inhibition and behavioural correlates. *Psychiatry Res. 201*(1), pp. 54-62. [Author manuscript].

Lotto, L., & Manfrinati, A., & Sarlo, M. (2013). A New Set of Moral Dilemmas: Norms for Moral Acceptability, Decision Times, and Emotional Salience. *J. Behav. Decis. Making*, pp. n.a. doi: 10.1002/bdm.1782

Manor, I. & Tyano, S. (1999). Organic disorders and psychosis. *Current Opinion in Psychiatry 12*(4), pp. 415-419

Marazziti, D., & Baroni, S., & Landi, P., & Ceresoli, D., & Dell'Osso, L. (2013). The neurobiology of moral sense : facts of hypotheses ? *Annals of General Psychiatry 12*(6), pp. 1-12.

Martell, D.A. (2009). Neuroscience and the Law: Philosophical Differences and Practical Constraints. *Behavioural Sciences and the Law 27,* pp. 123-136.

Martens, W.H.J. (2002). Criminality and Moral Dysfunctions: Neurological, Biochemical, and Genetic Dimensions. *International Journal of Offender Therapy and Comparative Criminology 46*(2), pp. 170-182.

Martins-de-Souza, D., & Guest, P.C., & Steeb, H., & Pietsch, H., & Rahmoune, H., & Harris, L.W., & Bahn, S. (2011). Characterizing the proteome of the human dorsolateral prefrontal cortex by shotgun mass spectrometry. *Proteomics 11*, pp. 2347-2353.

Matsuo, K., & Nicoletti, M., & Nemoto, K., & Hatch, J.P., & Peluso, M.A.M., & Nery, F.B., & Soares, J.C. (2008). A Voxel-Based Morphometry Study of Frontal Gray Matter Correlates of Impulsivity. *Human Brain Mapping 30*, pp. 1188-1195.

Melle, I. (2013). The Breivik case and what psychiatrists can learn from it. *World Psychiatry 12*, pp. 16-21.

Mendez, M.F. (2009). The Neurobiology of Moral Behavior: Review and Neuropsychiatric Implications. *CNS Spectr. 14*(11), pp. 608-620.

Meynen, G. (2013). A neurolaw perspective on psychiatric assessments of criminal responsibility: Decision-making, mental disorder, and the brain. *International Journal of Law and Psychiatry 36*, pp. 93-99.

Moll, J. & de Oliveira-Souza, R. (2007). Moral judgments, emotions and the utilitarian brain. *TRENDS in Cognitive Sciences 11*(8), pp. 319-321.

Moll, J., & Zahn, R., & de Oliveira-Souza, R., & Krueger, F., & Grafman, J. (2005). The neural basis of human moral cognition. *Nature Reviews Neuroscience 6*, pp. 799-809.

Morse, S.J. (1999). Craziness and Criminal Responsibility. *Behavioral Sciences and the Law 17*, pp. 147-164.

Morse, S.J. (2006). Brain Overclaim Syndrome and Criminal Responsibility: A Diagnostic Note. *Ohio State Journal of Criminal Law 3*, pp. 317-412.

Morse, S.J. (2007). Voluntary Control of Behavior and Responsibility. *The American Journal of Bioethics 7*(1), pp. 12-13.

Morse, S.J. (2011). Avoiding Irrational NeuroLaw Exuberance: A Plea for Neuromodesty. *Mercer Law Review 62*, pp. 837-859.
Muresanu, D.F., & Stan, A., Buzoianu, A. (2012). Neuroplasticity and impulse control disorders. *Journal of the Neurological Sciences 316*, pp. 15-20.

Nordström, A., & Dahlgren, L., & Kullgren, G. (2006). Victim relations and factors triggering homicides committed by offenders with schizophrenia. *The Journal of Forensic Psychiatry & Psychology 17*(2), pp. 192-203.

Penney, S. (2012). Impulse control and criminal responsibility: Lessons from neuroscience. *International Journal of Law and Psychiatry 35*, pp. 99-103.

Peoples, L.L. (2002). Will, Anterior Cingulate Cortex, and Addition. *Science, New Series 296*(5573), pp. 1623-1624.

Raine, A. & Yang, Y. (2006). Neural foundations to moral reasoning and antisocial behavior. *SCAN 1*, pp. 203-213.

Redding, R.E. (2006). The Brain-Disordered Defendant: Neuroscience and Legal Insanity In the Twenty-First Century. *Villanova University School of Law Working Paper Series Paper 61,* pp. 51-127.

Reider, L. (1998). Toward a New Test for the Insanity Defense: Incorporating the Discoveries of Neuroscience Into Moral and Legal Theories. *UCLA Law Review 46,* pp. 289-342.

Sapolsky, R.M. (2004). The frontal cortex and the criminal justice system. *Phil. Trans. R. Soc. Lond. B 359*, pp. 1787-1796.

Schleim, S., & Spranger, T.M., & Erk, S., & Walter, H. (2011). From moral to legal judgment: the influence of normative context in lawyers and other academics. SCAN 6, pp. 48-57.

Shoemaker, W.J. (2012). The social brain network and human moral behavior. *Zygon 47*(4), pp. 806-820.

Simmonds, D.J., & Pekar, J.J., & Mostofsky, S.H. (2008). Meta-analysis of Go/No-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. *Neuropsychologia 46*, p. 224-232.

Slovenko, R. (1999). The Mental Disability Requirement in the Insanity Defense. *Behavioural Sciences and the Law 17*, pp. 165-180.

Tassy, S., & Oullier, O., & Duclos, Y., & Coulon, O., & Mancini, J., & Deruelle, C., & Attarian, S., & Felician, O., & Wicker, B. (2012). Disrupting the right prefrontal cortex alters moral judgement. *SCAN 7,* pp. 282-288.

Vincent, N.A. (2011). Neuroimaging and Responsibility Assessments. *Neuroethics 4,* pp. 35-49.

Wilson, S. (2009). Criminal responsibility. *Psychiatry 8*(12), pp. 473-475.

Woodbridge, F. (1939). Some Unusual Aspects of Mental Irresponsibility in the Criminal Law. *Journal of Criminal Law and Criminology 29*(6), pp. 822-847.

Yeo, S. (2008). The Insanity Defense in the Criminal Laws of the Commonwealth of Nations. *Singapore Journal of Legal Studies* pp. 241-263.

Zald, D.H., & Andreotti, C. (2010). Neuropsychological assessment of the orbital and ventromedial prefrontal cortex. *Neuropsychologia 48*, pp. 3377-3391.

Newspaper articles:

Aftenposten. (2011). *Breivik nekter MR-undersøkelse av hjernen.* Retrieved from: *http://www.aftenposten.no/nyheter/iriks/Breivik-nekter-MR-undersokelse-av-hjernen-6676378.html#.UcMwgPn-Fcx*

To what extent is the taking and use of neuroscientific evidence compatible with the rights enshrined in the European Convention of Human Rights?

By Petar Lozev

Table of contents

1. Introduction

Criminal law is arguably among the most important parts of any system of law as its purpose is to counter serious forms of socially undesired behavior such as assault on one's property, physical integrity and life. Therefore in order to fulfill the high expectations society has of it, it should be equipped with the best tools to find out the truth, determine who is guilty and either punish or send them to rehabilitation. To achieve these goals it has the power to gather all kinds of evidence and invade people's liberties and private lives.

These extensive prerogatives are controlled to a large extent, but not only, by the protection from the state's intrusion into private life enshrined in human rights documents such as the right to fair trial and the right to privacy in the European Convention on Human Rights (ECHR). The protection granted to individuals in this way, however, is far from absolute and it is often reduced in the process of solving particularly important cases. Thus the answer to whether human rights are infringed can often be ambiguous, policy-driven and depend on the balancing of interests in the particular case.

On the other hand, in the past twenty years society has seen the rise of a neuroscience, which has made numerous discoveries relating to the structure and functioning of the human brain. On the basis of the results obtained in these studies, claims have been made that neuroscience will be able to drastically change the legal systems and criminal law in particular. However, such a drastic challenge to the criminal justice system seems to hold little promise at present (Greene & Cohen, 2004; Morse, 2008; Morse, 2011).

What seems to be a realistic goal for neuroscience is to redefine some of law's concepts by shedding light on the functioning of the brain and the thought process (Greene & Cohen, 2004; Morse, 2008). For example, being able to 'read' directly the defendant's brain and thought processes is expected to be a powerful tool in the hands of the prosecution that would allow them to gather evidence more efficiently, with less mistakes and therefore will lead to more just verdicts. However, the maxima 'the end justifies the means' is not necessarily valid in criminal law and the means have to be evaluated by themselves.

This paper endeavors to discuss whether the *compulsory* taking and use of neuroscientific evidence in the form of fMRI lie-detection and Guilty Knowledge Tests ('GKT' from now on) as well as Brain Fingerprinting ('BF' from now on) detection of existing knowledge from the defendant in criminal proceedings complies with the right to fair trial and particularly the right to silence which is part of it; and the right to privacy as enshrined in the European Convention on Human Rights. The first part of this paper will discuss the three types of neuroscientific evidence envisaged as they are in their present form and it will provide some further information on the current use of neurosciences in the courtroom. The

second part will deal with the issue of the right of fair trial laid down in Art 6 ECHR. Before providing an answer to the question, a comparison will be made with the federal law of the United States ('US law' from now on) in the context of the Fifth Amendment of the Constitution of the United States ('US Constitution' from now on) in search for possible solutions. At the end a conclusion on the compatibility of neuroscientific evidence with the right to a fair trial and more particularly the right to silence will be drawn. The third part will be a discussion on the right of privacy (Art 8 ECHR) and will follow the same general scheme used in the previous one. The fourth part will provide a conclusion to the paper and an overview of the results reached.

2. Neuroscientific evidence

Neuroscientific evidence in the form of Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET) and Single-photon emission computed tomography (SCEPT) scans has been used for quite a while in US courts to prove brain damage as result of accidents such as lesions (Moriarty, 2008, pp. 3, 12-13, 19; Appelbaum, 2009; Grafton, 2010). Particularly in criminal cases such evidence has been used to try to prove diminished responsibility or insanity defense by arguing that the brain damage indicated that the person had diminished cognitive capabilities. However, this evidence has been admitted primarily in the penalty phrase as a mitigating circumstance to reduce sentence, because admissibility standards are not so strict at that stage (Moriarty, 2008, p. 17; Grafton, 2010, p. 62).

Considering the fact that for the greater part these tests collect images of one's tissues and are not aimed at obtaining information about the actual functioning of the person's brain, they can be equated with medical tests to see if someone is wounded (Raichle, 2010 p. 12). However, the problem of interpretation as to what that structural damage means or whether it existed at the time of the commission of the crime is encountered which acts as a major impediment to their wider use in the stage where the verdict is given (Moriarty, 2008, pp. 14-16; Aharoni et al., 2008, pp. 5-6, 9).

2.1. fMRI testing

On the contrary, functional Magnetic Resonance Imaging (fMRI), which is already familiar to the general public thanks to popular media, measures the so-called blood level oxygenation dependent (BOLD) signals. This means that it detects the oxygenation of

the blood and thus where it flows to. Based on that, scientists infer which regions of the brain are activated and work harder (Moriarty, 2008, pp. 4-5; Raichle, 2010, p. 13). The key difference from older MRI method is that it measures brain activity even though indirectly, rather than just structure.

Possibly more importantly for criminal law, scientists claim to be able to tell when someone is lying by measuring which regions are activated (Davatzikos et al., 2005; Ganis et al., 2003; Kozel et al., 2004; Kozel et al., 2004a; Langleben et al., 2002) and even possibly establish experimental knowledge of something (Gamer et al., 2012, pp. 7-9). To add to that, it is in theory possible that fMRI BOLD test will be used to prove that someone is actively suffering from pain (Fields, 2010). However, people would probably undergo this type of tests voluntarily since it would usually be evidence to prove that they have suffered damage and therefore in their favor.

Considering the quickly growing body of academic work on the topic of fMRI lie-detection, this seems to be among the most eagerly expected new developments neurosciences may bring in the foreseeable future. Therefore, it is logical to analyze its compatibility with human rights as well as the use of fMRI for establishing experimental knowledge of something (Bles & Haynes, 2008, pp. 3-4).

While older types of Guilty Knowledge Tests and lie detection, with which crime-related information was sought, relied on reading the physiological responses of the person as a whole, these new technologies promise to allow tester to detect the signs of the actual recognition of the information (MacLaren, 2001, p. 674). Thus the chain between the existence of the incriminating memory and the prosecution becomes much shorter and the chances for drawing mistaken inferences are reduced. Furthermore, this would also allow drawing much more reliable conclusion as all the factors of the environment as well as other stress-related physiological responses are excluded as possible causes for the strong reaction to the actual information (MacLaren, 2001, pp. 675-676). This additional reliability seems to give an edge to neuroscientific methods of gathering evidence which is a valid ground to assume that the prosecution would try to utilize them as soon as they are admissible.

However, the admission of GKT for detecting experimental knowledge by means of an fMRI is very difficult at the very least, because at the time the blood flow starts to alter and can be measured, one is conscious what they are thinking about. Thus, one can change their thoughts and from a purely practical perspective there is no way to detect the fleeting reminiscence of the person who is being subjected to this test (Goebel, 2013). Up to now, the evidence gathered by fMRI scans and the results thereof have not made their way into the courtroom because they do not cover admissibility standards for expert

evidence. Among the main issues is that the connection between the increased blood flow and the corresponding type of the mental activity attributed is not causal. Instead scientists work with correlates (Moriarty, 2008, p. 5) and problems of interpretation may arise especially in the case when two experts conflict (Aggarwal, 2009, p. 3). What is more, the regions which are usually associated with one activity may be responsible for a whole range of similar activities, thus aggravating the problem of interpretation (Bles & Haynes, 2008, p. 6).

Moreover, neither the motivation of the subjects in the tests was similar to what the motivation of a defendant in a criminal case would be, nor the complexity of the situation was comparable (Bles & Haynes, 2008, pp. 5-6). This problem may seriously undermine the results of fMRI tests as it relates to their greatest weakness - BOLD signal in cortical areas is susceptible to deliberate distortion by subjects (Bles & Haynes, 2008, p. 9; Goebel, 2013). However, the real issue with these current weaknesses is that they are not imputable only to technology, but to brain functioning itself. Therefore remedying them may be quite a challenge for neuroscience

2.2. EEG/Brain Fingerprinting

Electroencephalography (EEG) detects the electrical signals that the brain emits as a result of its activity by means of electrodes attached to the scalp. This technology has been used to show the defendant images or ask questions which only someone present at the crime scene could have known – GKT (Bles & Haynes, 2008, p. 7; Rosenfeld, 2001; Farwell et al, 2012, p. 117). The strong signal received from the so-called P300 wave has been interpreted to indicate that the person has knowledge of the thing shown (Farwell et al., 2012, pp. 115, 117-118). The strength of this signal does not depend on the person's active response, but rather on indirect physiological markers whether the interviewee has knowledge of the information (Bles & Haynes, 2008, p. 7). What is more, the electric signals EEG detects are the results of neuronal activity and therefore very quick, fleeting mental processes can be detected.

Again the reliability of the evidence is questioned as the meaning given to the electrical signals from the brain is dependent on interpretation. (Rosenfeld et al., 2004) Furthermore, claims have been made that it is highly resistant to countermeasures, even though they have been contested, but have not been proven to be unfounded (Farwell et al., 2012, p. 141; Rosenfield, 2005; Rosenfield & Labkovsky, 2007).

Both the ability to prove that the person tested possesses knowledge which only the perpetrator may have or to detect lies could prove to be a powerful tool in the hands of

the prosecution. That is why it is so vital to see if their use complies with human rights.
Despite the current problems of admissibility of neuroscientific evidence due to its low reliability, the discussion on this first line of protection against the threat of irrelevant or unreliable evidence is not central to this work and therefore it will be left aside for the reasons that the standards of admissibility are set by national law and not by the European Court of Human Rights (ECtHR or 'the Court' from now on) (*Jalloh v. Germany,* pars. 94-95). Therefore, the discussion on the main issue will be from a futuristic viewpoint from the premise that such evidence is admissible, trustworthy and scientists can indeed use neuroscientific methods as a reliable tool for lie-detection and administering GKT.

3. Right to fair trial – Art 6 ECHR

The right to fair trial enshrined in Article 6 of the European Convention of Human Rights (ECHR) secures the defendant two main guarantees: equality of arms and *le droit á une procedure contradictoire* which can roughly be translated in English as an 'adversarial proceedings' though it would not convey the precise meaning. Equality of arms entails that both the prosecution and the defense should be entitled to be heard on equal terms while the 'adversarial proceedings' aspect entails that the defense should be notified of the prosecution's materials against them and given a chance to answer it (Marty & Spencer, 2002, p. 45).
Thus a literal reading of Art 6 does not provide what is termed 'right to remain silent and privilege against self-incrimination' (the right to remain silent from now on, if I refer to either aspect of the right specifically, I will note it) which might be threatened when one is subjected to the taking of evidence via fMRI or Brain fingerprinting. However, the ECtHR has read the right to remain silent and the privilege against self-incrimination as one aspect of fair trial relying on a broader interpretation of the article in *Funke v France* (see also Wu, 2011, pp. 38-39, 44-45).
One has to remember that the right to remain silent is not a specific autonomous right, but falls within the general notion of 'fair hearing'. Therefore, if a violation against the right to remain silent alone would have to amount also to a breach of Art 6 ECHR, it would have to deprive the defendant of the essence of their right (*Jalloh v Germany*, pars. 96-97).
The first part of the analysis of the applicability of the right to remain silent is that it applies only to people charged with a criminal offence. The notion 'charged with a criminal offence' has been given an autonomous and wide interpretation by the ECtHR (Berger, 2005, pp. 346-349). However, as the starting assumption of this paper is that the person

from whom neuroscientific evidence is taken is a defendant in a criminal case, this is not problematic.

To add to this, the Court has held evidence which was not prima facie incriminating, but was later deployed in a criminal trial is also under the protection of the right to remain silent (*Saunders v United Kingdom*). Therefore, if neuroscientific evidence is taken from a witnesses and the testimony is in no way incriminating for that person, it will most probably be deemed admissible. However, if it was later used against that witness, it would be considered to be a protected kind of evidence.

Second, in judging if the essence of the right is extinguished the Court has held in *Allan v United Kingdom* (par. 44) that it considers first the nature and degree of compulsion used to obtain the evidence, second whether there are safeguards in place against state misuse of its powers and third how the evidence has been utilized (see also Wu, 2011, p. 46). The degree of compulsion has been defined as 'coercion such as to render his right not to incriminate ineffective' in *Serves v France* (par. 46; see also Berger, 2005, pp. 355-356). However, as the paradigm that is discussed in this paper – when the defendant is forced to take the fMRI or Brain Fingerprinting test, is in itself a situation of compulsion by the state, this requirement for the applicability of the right to remain silent is fulfilled.

The final requirement for the application to the right to remain silent is that the particular type of evidence is protected or as it was termed by the Court it has to be 'dependent of the will of the defendant'. In this regard the ECtHR has connected the right to remain silent with the will of the defendant not to disclose evidence (*Saunders v. United Kingdom*, par. 69). What is more, the Court has also granted protection from compulsion by the drawing of adverse references against the defendant when they have made use of the right to remain silent (*John Murray v United Kingdom*; see also Wu, 2011, pp. 45-46).

It seems that the key distinction that the Court makes is that the right to remain silent doesn't apply to evidence which exists independently of the defendant's will such as bodily samples including blood, urine, hair, voice samples or documents obtained pursuant to a warrant (*Saunders v United Kingdom*, par. 69; *Jalloh v Germany*, par. 102).

Applying the three-fold analysis of *Alan v United Kingdom*, however, the Court seems to give also particular importance to the compulsive force used to obtain even 'real' evidence (*Jalloh v Germany*, pars. 114, 116, 118). If the compulsion used is much more than the required and necessary interference with one's integrity to obtain the aforementioned 'real evidence', then even that may be a violation of the right to remain silent.

The main problem in front of the present analysis is that the Court does not give much guidance as to where the dividing line between evidence that exists dependent on the person's will and one that exists independently of it lies in regard physiological processes and it is difficult to tell on which side neuroscientific evidence will fall.

3.1. Fifth Amendment and the right not to be witness against oneself

Therefore this paper turns for arguments as to how to deal with the issue at hand to US federal law since it has more experience with this matter because the right not to be witness against oneself has been dealt by federal courts on the basis of the Fifth Amendment of the US Constitution and is clearly comparable to the right to remain silent in Art 6 ECHR. Even though the theoretical and philosophical foundations of the Fifth Amendment are not clear, its practical application has been discussed and defined at length (Allen & Mace, 2004, pp. 243-246).

The US Supreme Court has developed a three-fold test as to whether a particular piece of evidence is protected (*Fisher v United States*), which is very familiar to the ones used by the ECtHR. The first two inquiries are whether there was compulsion and whether the evidence gathered was incriminating (for a more thorough discussion on the two see Allen & Mace, 2004, pp. 250-259). In the present analysis both are not included, because a priori the evidence has been taken compulsory – either by a judicial order or by force; and the assumption is that the evidence is taken from the defendant in the course of criminal proceedings with the object of incriminating them.

The question therefore rest on the third question - whether the three kinds of neuroscientific evidence will be considered 'testimonial' evidence. In *Schmerber v. California* the Supreme Court of the United States created the distinction between 'testimonial and communicative' and 'real and physical' evidence which seems similar to the approach later developed by the ECtHR in *Saunders v United Kingdom* and *Jalloh v Germany*.

However, in the same case Justice Brennan created an exception to the general distinction for polygraph testing. The wording of his argument is based on the fact that determining one's guilt on the basis of physiological responses would go against the spirit of the Fifth Amendment. This exception, however, has proved difficult to conceptualize for legal scholars (Allen & Mace, 2004, pp. 260-265) and some have argued that it is a non-binding dicta (Holley, 2009, p. 19). As basically both fMRI and EEG GKT measure the physiological responses of the person – be it either blood flow to the brain or electric impulses, one could argue that this exception can be used precisely in this case.

This distinction has been further refined and developed by scholars. On the one hand, some have drawn the dividing line at whether the act of communication itself can be considered testimonial and whether the information itself is of such nature (Stroller & Wolpe, 2007, pp. 367-368). Thus, the degree of control the defendant has over the transmission of the evidence becomes paramount.

On the other hand, some analysts argue that 'testimony' should be understood as the results of cognition that allow holding a proposition true or not true. Thus the acquisition, storage, retrieval and use of knowledge which has been caused by the state would be protected 'testimony' (Allen & Mace, 2004). Following this reasoning seems to lead to the conclusion that the 3 types of neuroscientific evidence in question are protected 'testimony' as they include cognition and in particular the retrieval of knowledge stored in one's brain.

Further scholarly arguments are based on the connection between the Fifth Amendment and the protection of privacy; however, this discussion will be left for later (Stroller & Wolpe, 2007, pp. 370-372).

3.2. Analysis of compatibility of fMRI lie-detection / fMRI guilty knowledge tests and Brain fingerprinting with the right to fair trial

Considering the case law of the ECtHR the key question in determining whether subjecting a suspect in a criminal trial to forced neuroscientific testing will be in violation of their right to remain silent will depend on whether the evidence produced will be considered to exist independently of the person's will or not (*Jalloh v Germany*, par. 102).

To begin with, the question of the compatibility of fMRI lie-detection turns out to be quite straightforward. Considering that lie-detection requires the active cooperation of a person and their actual responses, it clearly depends on the person's will as the actual cooperation of the person is required at least to answer the questions posed in order to determine whether these responses are truthful. And in such cases both Courts have made it explicit that this is a case where the right to remain silent applies in full power. Thus, without much problem one can conclude that administering fMRI lie-detection and forcing one to answer the questions is in violation of the right of fair trial in Article 6 ECHR.

The argument which asserts the compatibility of the neuroscientific evidence gathered by GKTs with the right to remain silent is based on the idea that what is measured during the factual process of taking of the evidence are the defendant's bodily processes - blood flow and electrical impulses. This argument is supported by the materialistic neuroscientific stance which views consciousness, memory and cognition as a multitude of physiological processes in the brain. Therefore according to this line of reasoning, evidence obtained via fMRI and EEG should indeed be considered to be real evidence.

This argument finds conceptual support in the notion that the right to silence itself is created to assure testimony are reliable and there are no forced false confessions, thus

resulting in verdicts that actually serve justice. It follows that once the defendant becomes a reliable source of evidence, the prosecution should not doubt to use that evidence to serve justice in the most efficient way possible and therefore satisfy the societal desire for risk reduction by removing dangerous persons from society.

The fact that both the blood flow within the brain and the electrical impulses are controlled by processes which are not part of our conscious thought process supports this line of reasoning – we do not generally have the power to direct our blood flow to sub-cortical areas (Goebel, 2013), nor do we control the electrical signals within our brain that occur in such a quick fashion as measured by BF. Therefore in the light of this argument, neuroscientific tests are be compared to blood content or body temperature tests aimed at determining the state of the person at a particular time (such as the a test to determine whether one has high levels of alcohol in their blood).

Furthermore, this argument is supported in the fMRI and Brain Fingerprinting admissions of GKTs by the non-existent control the person has over the transmission of the evidence. Since the dependence of the evidence on the person's will grants them control over its transmission (as the person can then stop other people's access to it at will), the opposite is therefore also necessarily correct. In view of the fact that the brain's regional increases in blood flow and electric activity in the cases of GKT are mostly independent on our will in the short term, this line of arguments asserts that the defendant will not enjoy the right to silence on the basis of the ECtHR's case law in regard to the three types of neuroscientific evidence which are a discussed.

Contrary to that stands the argument centered on the notion that what is protected by the right to remain silent is actually the result of human cognition and the fact that so much emphasis is put on the defendants' testimony is that this used to be the only way to have access to their mental processes which are the essence of what is to be protected by the right to remain silent

This argument draws strong philosophical support from a retributivist view of the justice system which views the defendant as a subject and not merely as an object and a source of evidence. When elevated to the status of subject, the defendant deserves to be treated as such and their most sacred inner self – cognition, should remain inviolable.

In support of this argument legal scholars and neuroscientists assert that cognitive processes do not exist independent of people's will and memory as one among them should be viewed as part of person's will (Farell, 2010, p. 94). In essence, the fact that in the long run people can will what they want to think about and remember, even if they are given conflicting outside stimuli (Goebel, 2013), adds credibility to the assertion that memory is indeed part of their will (and therefore dependent on it). This ability to exert

some measure of conscious control over some of their mental states is why techniques for brain enhancement such as neurofeedback have been reported to have great success in the context of treatment of psychological issues such as depression (Goebel, 2013; Hammond, 2005; Won Choi et al., 2011). Moreover, considering that the use of fMRI for GKT requires the compliance of the subject in holding the thought for several seconds while blood flow alters accordingly (Goebel, 2013) provides a further proof that people's thoughts in that instance are their own conscious cognition and gives an indication that the 'control argument' is not fully applicable to either kind of fMRI testing.

On the other contrary, if it would be possible to make the defendant think for a longer period of time about a particular event, fact or place by means of specific outside stimuli and to measure the results of this process, thus giving insight into the content of that person's thoughts, one would have a result very similar to forcing one to talk about these things.

Furthermore, the language used by the Court in its case law seems to place the stress on the defendant's will to communicate the evidence which seems to support an argument centered on the protection of the defendant's cognitive processes rather than control as naturally one's will to communicate refers mainly to the results of their cognitive processes.

To add to this, the ECtHR has not made any exception to the distinction of evidence similar to the US Supreme Court's with polygraph. However, the underlying logic employed by Justice Brennan that seems to be the protection of privacy may prove in the future persuasive if privacy is also threatened and thus leading the ECtHR to embrace it. This exception provides further support to the view that neuroscientific evidence would violate the right to fair trial.

Another quite different peril to the right of fair trial in general poses the danger of assigning too much importance to the neurological evidence. As neither judges, nor juries are unerring in their assessment of evidence, the colorful results of fMRI scans which are backed by a relatively new and revolutionary science may induce such mistakes. This problem is aggravated by the relatively low knowledge of judges and juries in the area of neurosciences at present. However, this situation of lack of experience in dealing with such matters may (and most surely will if courts are often confronted with neuroscientific evidence) change considerably in the future and therefore this argument is not really persuasive ground for considering neuroscientific evidence a threat to the right to fair trial.

A situation which is an interesting exception from the two argument outlined above is the situation whereby even neuroscientific evidence which is considered to exist independently of the person's will, may fall within the exception provided for in *Jalloh v Germany*: the Court there allowed for the protection of real evidence when the intrusion

into one's personal integrity to obtain the evidence was too great. In this instance, the defendant has to then rely on the violation of other rights such as the right to privacy and even more importantly - the right to be free of inhuman and degrading treatment to invoke this defense as the Court has held that violation of other rights could lead to a violation of the right to fair trial as well.

At present, it would be wishful thinking on my part to believe that the Court will embrace any of the two arguments that have been constructed by this paper when faced with such a case. This is so especially considering that a lot of time may pass and many factual and legal circumstances may change until the judges are faced with a case dealing with this issue.

However, considering the practical necessity for active response by the person tested in the case of fMRI lie-detection it is almost certain that this will be an instance when the defendant enjoys the right to remain silent irrespective of the viewpoint taken.

In the case of fMRI GKT the decision is not as clear-cut. However, due to the necessity that the person has to think about the relevant information for a couple of seconds before blood flow alterations occur, one can infer that the person has a significant amount of control over the results of the test. Having that in mind, it seems that in this instance both arguments would support a ruling that the defendant's right to remain silent will be endangered in this case.

The third type of neuroscientific evidence – Brain Fingerprinting, seems to be the one that will raise the most controversy. On the one hand side, the person tested has no effective control over the information transmitted as the P300 wave appears only 300 ms after the stimulus is given and is physically impossible to suppress. Furthermore, as the assertion that it is resistant to manipulation has not been refuted convincingly, the conclusion that one is not in control of the result of the tests cannot be challenged effectively thus lending further credibility to this line of reasoning. These two give a strong indication that a control-based reasoning may actually find the subjects of Brain Fingerprinting tests not to enjoy the right to silence.

On the other hand, memory and the retrieval of information from it is indeed a cognitive process and therefore should deserve the status of protected evidence from the viewpoint of the cognition argument. Opponents of the cognition line of argument may, however, claim that what BF tests (and not only) is not cognition itself, but physical changes which are interpreted to have a specific meaning. Thus the results from BF are equated with the tests for other 'similar' physical changes such as facial expression or body movements which are also used for drawing inferences about one's state of mind and therefore BF is just a way to get more accurate measurements (Holley, 2009, pp. 20-21). However, there is a strong intuitive feeling stemming from Cartesian philosophy that draws a distinction

between physical changes in the brain that affect our mental processes and in the rest of the body, which leads us to think that the processes that occur within our brain and are therefore linked to our cognition deserve a higher level of protection, thus making this counterargument less persuasive. What is more, if one takes the US Supreme Court's exception for polygraph GKT further, it is not difficult to see how it can be extended by analogy to cover brain fingerprinting as well.

Overall, if I am to make an educated guess as to the line of reasoning that the Court might follow should such a case arise, my prediction would be that the 'cognition' argument or a variation thereof would form the basis of the ECtHR's reasoning. This should not come as a surprise considering that the Court, in fulfilling its task as an enforcer of the Convention, has more often than not taken strong stances on the protection of human rights and has had little tolerance to their limitation by states. Hence this judicial activism could play an important role as it is not difficult to envisage how the use of neuroscientific evidence may intrude into areas the protection of which is among the core values of a democratic society such freedom of thought (Brems, 2005, pp. 303-305). Furthermore, curtailing the prosecution's right to use neuroscientific evidence obtained from the defendant does not prevent them from uncovering the truth, but makes that task more difficult. While on the contrary, the use of neuroscientific evidence from the viewpoint of the 'cognition' argument threatens to extinguish the essence of both the right to remain silent and the right to privacy. Applying this analysis by analogy from cases of conflicts of human rights, it seems important to provide protection for the defendant, rather than take it away and give less freedom to the prosecution by reasoning in the line of the 'control' argument (Brems, 2005, pp. 303-304).

4. Right to privacy

In the case law of the European Court of Human Rights the analysis as to whether there has been a violation of Art 8 ECHR is usually divided into five parts. The first question that the Court asks is whether the relevant information obtained or stored is covered in the concept of 'privacy'. The second inquiry is whether there has been an actual interference with the right.

Thirdly the Courts looks at whether it has been done in accordance with the law. If a particular state activity is to fulfill this criterion, there will have to be thorough regulation on the issue in order to prevent the arbitrary interference with people's lives. Therefore, the old Acts on criminal procedure will have to be amended to accommodate the new

developments, or probably the old rules will be applied to neuroscientific evidence with the help of some interpretation of the concepts laid down therein. The question of legality, however, is not really the focus of the discussion, so it will be assumed that there is relevant domestic law which regulates the matter.

The fourth requirement that has to be met to avoid infringing the right of privacy is that the acts must have been undertaken in pursuit of one of the interests laid down in art 8(2) ECHR. This is also not really a problem as the evidence from the defendant in this paper is taken for the public safety and the prevention of crime.

The fifth question the Court asks is whether the gathering of the information is necessary in a democratic society. This is a proportionality requirement aimed at restricting governmental interference to the minimum level necessary (Trechsel, 2005, pp. 535-540).

The present discussion on whether the information obtained by fMRI and EEG scans is protected by the right to privacy will be focused on the first two issues that are dealt by the ECtHR, because they appear to be most problematic in the case discussed.

The Court has used the notion 'legitimate expectation' a person has for the respect of their private life to determine which information and what use of that information fall within the scope of the right (*Von Hannover v Germany*, pars. 51, 69). It is, however, difficult to set down concrete formula the Court uses for establishing whether one harbors such 'legitimate expectations', because States parties to the Convention enjoy a margin of appreciation due to cultural differences. Therefore, each state can engage in a relatively independent evaluation of the right to privacy and the person's legitimate expectations (Bignami, 2008, pp. 212-214, 239, *Von Hannover v Germany,* par. 57). However, one should not forget that the Court itself can (and will) often engage in an evaluation of its own and in such cases it has attached importance to the question whether the information gathered related to private or public matters (*Peck v United Kingdom*, par. 61). To add to this, considering the fact the defendant is protected from having to provide the prosecution with incriminating evidence by the right to remain silent, it can be argued that such 'legitimate expectations' indeed exist in regards to this information.

Therefore, in view of these indications that the person may have 'legitimate expectation', what the Court and States can do is essentially evaluate whether the public interest in obtaining a particular piece of information outweighs the person's interest in keeping it secret. The question is not an easy one to solve and factors such as the nature of the information and the degree of its intrusiveness are very important.

To begin with, the question whether there has been interference with the right hinges to a great extent on the duration and seriousness of the interference as well as the duration for keeping the information. In *Jalloh v Germany* (pars. 70-71, 79) the Court clearly

demonstrated that that the mode of taking evidence is also to be considered – the more intrusive the scientific method of taking of evidence, the more serious justifications should be given even for physical evidence. Taking this rationale to its logical end it was held that evidence taken in violation of Art 3 ECHR and thus obtained by means of inhuman and degrading treatment can never be used to prove the defendant's guilt, because this would render the trial unfair and also in violation of Art 6 per se (*Jalloh v Germany*, pars. 105-106). On the other hand, the present discussion will not turn to the problem of storing of the data obtained with the neuroscientific evidence and its later use, because this is another aspect of the right and questions raised are quite different.

It seems that the Court would evaluate the conflicting interests in each case with view of the particular facts of the case rather than some general principles. Therefore, in search for further arguments and guidance this paper turns once more to US law.

4.1. Fourth Amendment and the right not to be subject to unreasonable searches and seizures

First and foremost, the US Constitution does not contain a general right of privacy as in the ECHR. Instead the interpretation of the Fourth Amendment used by the Supreme Court in the case *Katz v United States* where the two-fold formula for assessing whether the guarantee against warrantless unreasonable searches and seizures was breached will be used. The test consisted of the questions whether the person has specific expectation for privacy and whether those were reasonable (*Katz v United States*, 361). Yet, this is not equal to a general right to privacy as the Supreme Court stated itself. The protection granted to individual privacy by the Fourth Amendment applies only in some instances of government intrusion into private life (*Katz v United States*, 350; Wimberly, 2007, pp. 295-296). As the present analysis focuses an intrusion into the functioning of one's brain by the government with the purpose of using the results as evidence at a criminal trial, the conclusion is that the Fourth Amendment serves the same purpose as the right to privacy in the ECHR and a meaningful comparison can be made.

However, if the US police have a valid search warrant issued by a judge, this protection is not applicable (Halliburton, 2007, 340-341; Wimberly, 2007, p. 294). Therefore the analysis turns to unwarranted searches and the permissibility of the evidence gathered thereof.

To begin with, in the cases of unwarranted searches courts would balance legitimate expectations for privacy with the governmental interest in the information very much like the ECtHR (Wimberly, 2007, pp. 294-295). In determining whether one has reasonable expectations of privacy worthy of protection in unwarranted searches the US courts

consider several factors. The main one is that a person retains no 'legitimate expectations' in regard behavior that is exposed to the public or even to a limited circle of people. This is also extended to cover information which can be gathered only thanks of the capabilities of technologically sophisticated methods of observation. This principle is qualified in cases of novel technologies - if the technology used is not of general public use, then the person will be considered to have retained such interest even when it relates to evidence-gathering procedure which is less intrusive than a classical search (*Kyllo v United States,* 34-35; see also Kerr, 2004, pp. 831-837). Thus the more commonplace brain scanning of all types becomes, the more peoples' legitimate expectations for privacy will diminish (Halliburton, 2007, pp. 348).

What is important in this line of reasoning is that there is no determination whether it is morally appropriate to use such technologies (Halliburton, 2007, p. 346). The lack of specific boundaries means that there is a real possibility that at some point there are no boundaries and no inviolable private sphere is left and in my opinion this is a highly undesirable situation.

Last, but not least, some scholars have been argued that even today the analysis of the Fourth Amendment revolves around concepts, notions and expectations from property law (Kerr, 2004, pp. 827-831). However, such line of reasoning has little utility when applied to the right of privacy, because the ECtHR has never relied on such arguments. This way of thinking becomes even more inappropriate when one considers the fact that human bodies and the working of their brains are not property, as the opposite would violate human dignity.

4.2. Analysis of compatibility of neuroscientific evidence with the right to privacy

The starting point of the analysis is whether the information at stake is protected. As the taking of neuroscientific evidence includes a scan of human brains, the information retrieved arguably forms part of one's private life and thus one has 'legitimate expectation' of privacy in that regard. Furthermore, inferring certain mental states and feelings by measuring brain activity allows scientists access to thoughts and feelings one has consciously chosen to keep for oneself and to exclude the outside world from. Such a decision to exclude others from the information has been held to be an indication for the strong expectations one has for it remaining private and therefore the possible existence of 'legitimate expectations' (*Niemietz v. Germany*, par. 29).

Moreover, Cartesian philosophy and Lockean natural-law, which have shaped the perception of what constitutes the modern person, have deeply embedded into society the notion that the essence of humanity is the freedom of thought and conscience. Since even simple observation of our behavior often limits it to a great degree and behavior is result of human will it is possible to conclude that observation will also limit cognitive processes (Halliburton, 2007, pp. 331-340). This intuitive notion is used to support the claim that people should retain legitimate expectations of privacy in regard to all their cognitive processes even if they are defendants at criminal trial.

Contrary to that, as fMRI testing requires the complete stillness of a person, complying with this may be interpreted as a functional consent to the taking of the test even in the face of the fact that the person does their best to frustrate the test results. Such consent can then be used to argue that one has exposed the information and should therefore not retain their legitimate expectations. However, one can argue that there has been no real consent, because the defendant has been forced to take the tests and their will has been overborne and therefore this line of reasoning does not appear very persuasive.

Therefore the conclusion is that both fMRI and BF tests have the potential to violate the right to privacy, because the defendant would have retained their 'legitimate expectations' of privacy in regard the information that has been gathered.

Hence the second part of the analysis, which seeks to answer whether there will be an actual violation of the right when neuroscientific evidence is used, becomes paramount.

The first aspect of this analysis is a balancing exercise between the prosecution's interest in determining whether the defendant is lying in the testimony given or has knowledge of some details of the crime and the person's interest in privacy.

This balancing naturally hinges on the type of information that may be gathered, but the threshold is different from the one used to determine the existence of 'legitimate interests' (*Peck v United Kingdom*). Considering that most recent developments in the area of fMRI scans allow determining one's most general thoughts, but not their specific content (Goebel, 2013), it seems that the threat to any information of personal character is quite limited. This gives preference to the prosecution's interests and therefore to the conclusion that there is no violation. What is more, as the person has great degree of control over their persistent thoughts, it can be argued that they retain control over the collateral information that the prosecution may uncover. Therefore the argument that fMRI scans can discover irrelevant information which is very personal (McMonagle, 2007) is not very persuasive at the moment.

It is imperative, however, that this balancing be undertaken each time neurosciences advance and the more precise information can be inferred, because then greater the

protection has to be afforded to the defendant and their non-crime related knowledge.
Even though the spectrum of what fMRI lie-detection and GKT can uncover in the person's cognitive processes at the moment is not very broad, it will increase in the future. Therefore it seems that the balancing exercise with regard fMRI methods of obtaining evidence will have to be undertaken with increasing care and intensity each time.
In the case of BF this balancing does not seem to raise many issues considering that the capabilities of EEG at detecting thought processes are quite limited and the information that can be gathered is relates only to whether one is familiar with some information (Farwell, 2012, pp. 115-116). Therefore, on application of the reasoning from *Peck v United Kingdom* the person's interest in privacy may be held to be diminished compared to those of the government. Hence, it seems that the interests of the prosecution outweigh the interests of the defendant relatively easily in the case of BF and therefore the defendant's right to privacy would probably have to be curtailed to allow the taking and use of BF evidence.
The next step in the present analysis is to determine whether there is actual interference with the right to privacy in view of the degree of intrusiveness of the measures taken.
In the case of scanning one's brain to see if the person is in possession of particular information it can be argued that there is indeed interference on the basis of the fact that this allows to extract information which the person has decided not to share.
On the one hand, the nature of the interference in the case of fMRI scans whereby one has to remain still in a huge magnet, be subjected to strong noise and magnetic fields does not seem to be very problematic considering the ECtHR's adjudication based on the ways evidence is taken. Possibly in the case when the test subject suffers from aggravating conditions such as claustrophobia the test would be so stressful for the person tested that it will amount to inhuman and degrading treatment and thus a violation Art 3 ECHR, therefore rendering the evidence unusable. However, in normal circumstances this method does not seem invasive enough because there is no intrusion into one's physical integrity, nor is there any interference into the psychological functioning of the person.
On the other hand, EEG testing with electrodes on the scalp does not seem to raise any issues in that regard because of the relatively low level of discomfort for the person who is being tested (Farwell et al., 2012, p. 122).
The last point that should be raised in the present analysis focuses on the last requirement – whether the invasion into privacy is necessary in a democratic society. This condition is a safeguard that could have either a very negligible role or a very prominent one, depending on how the evidence is used. This clause provides protection against the misuse of previously gathered evidence by the state as well as against the ever more invasive types of surveillance that are used.

Therefore if neuroscientific evidence's existence is limited to the case only and is destroyed immediately afterwards, then an assessment of the first two criteria will probably be sufficient and this will not be a ground for asserting unlawful violation of privacy. On the contrary - cases of creation of large databases with the results of neuroscientific testing seem bound to be indefensible in the light of modern social philosophy.

In my opinion, this part of Art 8(2) ECHR will play an important role in future as its task is to prevent the executive power from taking the road to becoming the all-knowing and all-powerful totalitarian government depicted in George Orwell's *1984*. In the foreseeable future such a scenario seems highly unlikely and only the fMRI technology seems to pose serious foreseeable threat to the defendant's right to privacy at a criminal trial, but it is nevertheless important that this safeguard is in place.

5. Conclusion

This paper has taken three state-of-the-art types of neuroscientific evidence and analyzed their compatibility with human rights as enshrined in the European Convention of Human Rights. The approach used was futuristic – the reliability of the evidence and the results of its analysis are considered to be unerring so that the discussion can be focused on the compatibility between their forceful taking and use and the right to silence and the right to privacy.

Having looked into the case law of both the European Court of Human Rights and the United States Supreme Court on broadly similar matters, general arguments supporting both the thesis that neuroscientific evidence is in accordance with human rights and the claim that it violates them have been put forward and discussed.

In the case of the right to silence two opposing lines of reasoning were developed. According to the first one, only the testimony over which the defendant has conscious control is protected with the goal to ensure proper administration of justice and to exclude forced false confessions. The opposite argument is that the results of one's cognitive processes lay at the heart of the right to remain silent and they should be protected.

When applying the case-law of the ECtHR it is almost certain that fMRI lie-detection will be deemed to be a case when the defendant enjoys the right to remain silent, because their actual cooperation and answers to the questions posed are required. However, its use on witnesses is restricted to a much lesser degree which makes the distinction between defendant and witness of paramount importance.

The compatibility of fMRI based Guilty Knowledge Tests with human rights is not so clear, because there may be problems of interpreting what is 'evidence existing independent of the defendant's will'. However, considering both arguments, there is a strong indication that it will also prove to be a kind of protected evidence whereby the defendant enjoys the right to remain silent.

On the other side of the clarity scale is the case of Brain Fingerprinting. As it is the least intrusive test which measures something that people have no conscious control of and none of the two arguments seems to be prevailing and applicable to a greater degree, the decision may hinge on the facts of the case or the sentiment of the court.

During the analysis of the protection afforded by the right of privacy it became clear that in both jurisdictions the 'legitimate expectations' one has are balanced against the interests of the prosecution in uncovering the particular type of information.

On the one hand, in view of the fact that fMRI is capable of unveiling more that only crime-related information, the conclusion is that if not at present, then in the future serious issues with the right of privacy may arise.

On the other hand, as the results of BF only indicate whether one is familiar with a piece of information, it seems that the relative importance of this information is not so high as to overcome the interests of investigation.

In conclusion, it appears that the current possibilities of neuroscience to delve into human brains are much more limited that one might think based on popular media publications. In fact, at present there is actually little that threatens cognitive liberty and the right to privacy in normal circumstances. But this also seems to be the moment to think exactly how much does society value these rights and how they should be balanced against the interests that exist in crime prevention and swift administration of justice. For if society would start thinking on those questions when it is already faced with the issue and difficult decisions have to be made under pressure, there is a great chance that grave mistakes will be made.

Bibliography:

Case law:

Allan v United Kingdom ECHR (2002)
Fisher v United States, 425 U.S. 391 (1976)
Funke v France ECHR (1993)
Jalloh v Germany ECHR (2006)
John Murray v United Kingdom ECHR (1996)
Katz v United States, 389 U.S. 347 (1967)
Kyllo v United States, 533 U.S. 27 (2001)
Niemietz v Germany ECHR (1992)
Peck v United Kingdom ECHR (2003)
Schmerber v. California, 384 U.S. 757 (1966)
Saunders v United Kingdom ECHR (1996)
Serves v France ECHR (1997)
Von Hannover v Germany ECHR (2004)

Books and articles:

Aggarwal, N. (2009). Neuroimaging, Culture, and Forensic Psychiatry. *Journal of the American Academy of Psychiatry and the Law 37(2),* pp. 239-244.

Aharoni, E, Funk, C, Sinnott-Armstrong, W and Gazzaniga, M. (2008). Can Neurological Evidence Help Courts Assess Criminal Responsibility? Lessons from Law and Neuroscience. *The Year in Cognitive Neuroscience 1124*, pp. 145-160, retrieved from: *http://onlinelibrary.wiley.com.ezproxy.ub.unimaas.nl/doi/10.1196/annals.1440.007/abstract*

Allen, R & Mace, K. (2004). The Self-Incrimination Clause explained and its Future Predicted. *The Journal of Criminal Law and Criminology 94(2)*, pp. 243-294.

Appelbaum, P. (2009). Through a Glass Darkly: Functional Neuroimaging Evidence Enters the Courtroom. *Law & Psychiatry 60(1)*, pp. 21-23, retrieved from: *http://ps.psychiatryonline.org/article.aspx?articleID=100059*

Berger, M. (2005). Europeanizing Self-Incrimination: The Right to Remain Silent in the European Court of Human Rights. *Columbia Journal of European Law 12*, pp. 339-382.

Bignami, F. (2008). Case for Tolerant Constitutional Patriotism: The Right to Privacy before the European Courts. *Cornell International Law Journal 41*, pp. 211-250.

Bles M & Haynes J, D. (2008). Detecting Concealed Information Using Brain-Imaging Technology. *Neurocase 14(1),* pp. 82-92, retrieved from: *http://www.tandfonline.com.ezproxy.ub.unimaas.nl/doi/abs/10.1080/13554790801992784*

Davatzikos, C., K. Ruparel, Y. Fan, D.G. Shen, M. Acharyya, J.W. Loughead, R.C. Gur, and D.D. Langleben. (2005). Classifying spatial patterns of brain activity with machine learning methods: Application to lie detection, *NeuroImage 28*, pp. 663-668, retrieved from: *http://www.sciencedirect.com.ezproxy.ub.unimaas.nl/science/article/pii/S1053811905005914*

Brems, E. (2005). Conflicting human rights: an exploration in the context of the right to a fair trial in the European convention on human rights. Human Rights Quarterly 27(1), pp. 294-326.

Farell, B. (2010). Can't Get You Out of My Head: The Human Rights Implications of Using Brain Scans as Criminal Evidence. *Interdisciplinary Journal of Human Rights Law 4(1),* pp. 89-95.

Farwell, L. (2012). Brain fingerprinting: a comprehensive tutorial review of detection of concealed information with event-related brain potentials. *Cognitive Neurodynamics 6(2),* pp. 115-154, retrieved from: *http://link.springer.com.ezproxy.ub.unimaas.nl/article/10.1007%2Fs11571-012-9192-2*

Fields, H. (2010). Can Neuroscience Identify Pain?. In Mansfield, A., Gazzaniga M. [eds.]; A Judge's Guide to Neuroscience, Santa Barbara, University of California Santa Barbara, retrieved from: *http://www.sagecenter.ucsb.edu/sites/staging.sagecenter.ucsb.edu/files/file-and-multimedia/A_Judges_Guide_to_Neuroscience%5Bsample%5D.pdf*

Gamer, M, Klimecki, O, Bauermann, T, Stoeter, P, Vossel, G. (2012). fMRI-activation patterns in the detection of concealed information rely on memory-related effects. *Social Cognitive and Affective Neuroscience 7*, pp. 506-515, retrieved from: *http://scan.oxfordjournals.org.ezproxy.ub.unimaas.nl/content/7/5/506*

Ganis, G., Kosslyn, S.M., Stose, S., Thompson, W.L., Yurgelun-Todd, D.A. (2003). 'Neural correlates of different types of deception An fMRI investigation'. *Cerebral Cortex 13(8),* pp. 830-836, retrieved from: *http://cercor.oxfordjournals.org/content/13/8/830*

Grafton, S. (2010). Has neuroscience already appeared in the courtroom?. In Mansfield, A., Gazzaniga M. [eds.]; A Judge's Guide to Neuroscience, Santa Barbara, University of California Santa Barbara, retrieved from: *http://www.sagecenter.ucsb.edu/sites/staging.sagecenter.ucsb.edu/files/file-and-multimedia/A_Judges_Guide_to_Neuroscience%5Bsample%5D.pdf*

Greene, J. & Cohen, J. (2004). For the law, neuroscience changes nothing and everything. *Philosophical Transactions of the Royal Society B: Biological Sciences 359*, pp. 1774-1785, retrieved from: *http://www.ncbi.nlm.nih.gov.ezproxy.ub.unimaas.nl/pmc/articles/PMC1693457/?tool=pmcentrez&rendertype=abstract*

Halliburton, C. (2007). Letting Katz out of the Bag: Cognitive Freedom and Fourth Amendment fidelity. *Hastings Law Journal 59,* pp. 309-368.

Halliburton, C. (2009). How Privacy Killed Katz: a tale of cognitive freedom and the property of personhood as Fourth Amendment norm. Akron Law Review 42(3), pp. 803-885.

Hammond, C. (2005). Neurofeedback Treatment of Depression and Anxiety. *Journal of Adult Development 12*, pp. 131-137, retrieved from: *http://link.springer.com.ezproxy.ub.unimaas.nl/article/10.1007%2Fs10804-005-7029-5*

Holley, B. (2009). It's All in Your Head: Neurotechnological Lie Detection and the Fourth and Fifth Amendments. *Developments in Mental Health Law 28(1),* pp. 1-24.

Kerr, O. (2004). The Fourth Amendment and New Technologies: Constitutional Myths and the Case for Caution. *Michigan Law Review 102(5)*, pp. 801-888.

Kozel, F., Revell, L., Lorberbaum, J., Shastri, A., Elhai, J., Horner M. et al. (2004). A pilot study of functional magnetic resonance imaging brain correlates of deception in healthy young men. *The Journal of Neuropsychiatry and Clinical Neurosciences 16*, pp. 295-305, retrieved from: *http://neuro.psychiatryonline.org/article.aspx?articleID=101888*

Kozel, F., Padgett, T., George, M. (2004). A replication study of the neural correlates of deception. *Behavioral Neuroscience 118 (4)*, pp. 852– 856, retrieved from: *http://web.ebscohost.com.ezproxy.ub.unimaas.nl/ehost/pdfviewer/pdfviewer?sid=614090d3-459a-4f21-927c-c2c60e3c3de5%40sessionmgr15&vid=1&hid=9*

Ronald Kulich, R., Macyewicz, R. & Scrivani, S. (2009). Functional Magnetic Resonance Imaging (fMRI) and Expert Testimony. *Pain Medicine* 10(2), retrieved from: *http://onlinelibrary.wiley.com.ezproxy.ub.unimaas.nl/doi/10.1111/j.1526-4637.2009.00567.x/abstract*

Langleben, D. D., Schroeder, 1 L., Maldjian, J. A., Gur, R. C., McDonald, S., Ragland, J. D. (2002). Brain Activity during Simulated Deception: An Event-Related Functional Magnetic Resonance Study. *NeuroImage 15,* pp.727-732, retrieved from: *http://www.sciencedirect.com.ezproxy.ub.unimaas.nl/science/article/pii/S1053811901910031*

Delmas-Marty, M., Spencer, J.R.. (2002). European Criminal Proceedings, Cambridge, Cambridge University Press.

MacLaren, V. (2001). A Quantitative Review of the Guilty Knowledge Test. Journal of Applied Psychology 86(4), pp. 674-683.

McMonagle, E. (2007). Functional Neuroimaging and the Law: A Canadian Perspective. *The American Journal of Bioethics 7(9),* pp. 69-70, retrieved from: *http://www.tandfonline.com.ezproxy.ub.unimaas.nl/doi/abs/10.1080/15265160701518854*

Moriary, J. C. (2008). Flickering Admissibility: Neuroimaging Evidence in the US Courts. *Behavioral Sciences and the Law 26,* pp. 29-49, retrieved from: *http://onlinelibrary.wiley.com.ezproxy.ub.unimaas.nl/doi/10.1002/bsl.795/abstract*

Morse, S. (2008). Determinism and the Death of Folk Psychology: Two Challenges To Responsibility from Neuroscience. *Minnesota Journal of Law, Science & Technology 9,* pp. 1-36.

Morse, S. (2011). Avoiding Irrational NeuroLaw Exuberance: A Plea for Neuromodesty. *Mercer Law Review 62,* pp. 837-860.

Raichle, R. (2010). What is an fMRI?. In Mansfield, A., Gazzaniga M. [eds.]; A Judge's Guide to Neuroscience, Santa Barbara, University of California Santa Barbara, retrieved from: *http://www.sagecenter.ucsb.edu/sites/staging.sagecenter.ucsb.edu/files/file-and-multimedia/A_Judges_Guide_to_Neuroscience%5Bsample%5D.pdf*

Rosenfeld, P., Soskins, M., Bosh, G., Ryan, A. (2004). Simple, effective countermeasures to P300-based tests of detection of concealed information. *Psychophysiology 41(2),* pp. 205-219, retrieved from: *http://onlinelibrary.wiley.com.ezproxy.ub.unimaas.nl/doi/10.1111/j.1469-8986.2004.00158.x/abstract;jsessionid=8ACBDD82A825EA63FE3BB10FED30380C.d04t01*

Rosenfield, P. (2005). "Brain Fingerprinting": A Critical Analysis. *The Scientific Review of Mental Health Practice 4(1),* pp. 20-37, retrieved from: *http://cns.bu.edu/~gsc/Articles/Brain_FingerPrinting.pdf*

Rosenfield, P. & Labkovsky, E. (2007). New P300-based protocol to detect concealed information: Resistance to mental countermeasures against only half the irrelevant stimuli and a possible ERP indicator of countermeasures. *Psychophysiology 47(6),* pp. 1002-1010, retrieved from: *http://onlinelibrary.wiley.com.ezproxy.ub.unimaas.nl/doi/10.1111/j.1469-8986.2010.01024.x/abstract*

Stoller, S. & Wolpe, P. (2007). Emerging Neurotechnologies for Lie Detection and the Fifth Amendment. *American Journal of Law and Medicine 33,* pp. 359-376, retrieved from: *http://repository.upenn.edu/cgi/viewcontent.cgi?article=1032&context=neuroethics_pubs*

Won Choi, S. et al. (2011). 'Is Alpha Wave Neurofeedback Effective with Randomized Clinical Trials in Depression? A Pilot Study'. *Neuropsychobiology 63*, pp. 43-51, retrieved from: *http://www.karger.com.ezproxy.ub.unimaas.nl/Article/Pdf/322290*

Trechsel, S. (2005). Human Rights in Criminal Proceedings, Oxford, Oxford University Press.

Wimberly, M. (2007). Rethinking the Substantive Due Process Right to Privacy: Grounding Privacy in the Fourth Amendment. *Vanderbilt Law Review 60,* pp. 283-324.

Wu, W. (2011). Interrogational fairness under the European Convention on Human Rights. *International Journal of Law, Crime and Justice 39,* pp. 37-59, retrieved from: *http://www.sciencedirect.com.ezproxy.ub.unimaas.nl/science/article/pii/S1756061611000231*

Meetings:

Meeting with Professor Reiner Goebel, 23 April 2013, Maastricht University Brain Imaging Center

All Internet sources have been last visited on 17 June 2013

Violation of Article 19 of the UN Convention on the Rights of the Child: An analysis from a neurobiological point of view with regard to criminal behaviour

By Birgit Thun

Table of Contents

1. Introduction

Jan Willems writes in his article on the Rights of the Child that every fifth child suffers from maltreatment, even every third is not securely attached to its parents (2012, p. 5). What *attachment* means has been explained in the Encyclopedia of Child Development as the relationship between parents and their children. Parents need to be attentive and supportive towards their children. Maltreatment and neglect lead to insecure attachment resulting in, according to this article, "problem behaviours and psychopathologies" (Attachment Synthesis, 2012, p. i). Children, especially infants, seem to be particularly vulnerable to such treatment because their brains are still in the process of development. Such a kind of treatment has been explicitly forbidden by Article 19 of the United Nations (UN) Convention on the Rights of the Child (CRC). It demands of states parties to take all appropriate measures guaranteeing an environment in which children can grow up holistically. The Committee on the Rights of the Child, however, indicates that measures taken by states for this purpose are in many cases neither sufficient nor adequate (2011, p. 6). This may be the source of high costs on the criminal justice system as the Committee argues that children growing up in a violent environment are more likely to develop criminal behaviour (p. 7).

Criminal behaviour, in the case of persons below the age of 18 called delinquency, is a concept stemming from the legal field, according to which a person shows "conduct that does not conform to the legal or moral standards of society" (Deliquency, n.d.). The notion of antisocial behaviour in psychopathology features similar characteristics described as "behaviour usually marked by aggression but representing transgressions against societal norms" (Smith & Stern, 1997, p. 383). Chakraborty et al. confirm that criminal behaviour is defined by law, and that it is thus not usable in biology. This leads to the use of the notion of antisocial behaviour instead and the investigation thereof by researchers. They identify three ways of defining antisocial behaviour: the first one would be an equation with criminal behaviour and delinquency, which were explained above. This, however, seems to be problematic because a definition thereof changes over time and from one legal system to another. A second possibility is to investigate Antisocial Personality Disorder (ASPD) because it often involves the commission of criminal acts resulting from neglect of rights of fellow human beings. The third possible way of how to define antisocial behaviour is a focus on characteristics that are at risk for developing criminal behaviour, such as aggressiveness and impulsivity (Chakraborty et al., 2011, pp. 37-38).

I am seeking to find out what is meant by holistic development from a neurobiological perspective and which consequences child maltreatment has on this development. In

particular, I am focussing on criminal behaviour as a possible consequence. Hence, my research question states as follows: In how far does a violation of Article 19 of the UN Convention on the Rights of the Child lead to criminal behaviour?
My paper is structured as follows: First, I am describing the methodology used in my research in order to obtain meaningful and reliable results. A short explanation of the theoretical framework is following, which shall lay down the relation between processes occurring in the brain, its influence on behaviour and environmental impact on these processes. This framework will be used throughout the paper and will be turned back to in the end. Thereafter, insecure attachment will be further elaborated on. The section afterwards deals with the content of Article 19 of the CRC and with problems that experts see in a case of a violation thereof. The consequences of such a violation will then be considered from a neurobiological perspective, which means that the focus will be purely on the effects of maltreatment on the neurobiology of children. This will be followed by an analysis of the neurobiology of criminal personality traits. The findings are discussed afterwards followed by a conclusion consisting of a small summary and some remarks.

2. Methodology

In order to answer my research question, I am analysing articles of experts in the field of neurobiology. As I am not a neurobiological expert, I will have to rely on their descriptions of studies, their results, interpretations and opinions on this issue. However, in order to guarantee reliability, I am analysing several articles on the same issue so that a control mechanism is given. When writing from a human rights perspective, I am using primary literature wherever this is possible rather than repeating what has been said in secondary literature on the respective issue. This guarantees an own approach to the subject and more reliability and objectivity.
I have chosen literature in form of books which deal with the development of the brain during childhood in general and how it may develop under atypical circumstances. Furthermore, I have used the search option in the data bases "BioMedCentral" and "Psychology and Behavioral Sciences Collection" using the words *child maltreatment, consequences*, *criminal behaviour, (early) stress, neglect, aggression* and *neurobiology* in different combinations for rather general and/or explanatory articles, and also more specific notions, such as *norepinephrine, dopamine* and *serotonin* in combination with *violence* or *aggression*.

I am seeking to find an answer to my research question by analysing texts which give me an insight into the neurobiology of maltreated children. Thereby, it must be explained that I do not look at each specific kind of maltreatment which is listed in, for instance, Article 19 of the CRC. My aim is to draw a conclusion on that provision as a whole and not only on one kind of maltreatment in particular, which in itself can, as will be seen below, subdivided into even more categories. Hence, I chose instead to analyse the effects of stress as a result of maltreatment. This is followed by an investigation of the neurobiology of criminal behaviour. For this purpose, criminal behaviour has to be defined. As established in the introduction, this is problematic as the term is defined by legal systems and not according to neuropsychological criteria. Moreover, it differs from one system to another. Thus, the term "antisocial behaviour" is chosen, which is usable in neuropsychology. It can be analysed in three ways, of which I choose the third one. Hence, my analysis of criminal behaviour is based on three points that may be characteristic of criminal behaviour, namely aggression or violence, missing empathy, and immorality. Methodological limitations will be mentioned in the conclusion.

3. Theoretical framework: an integrated model of neuropsychological and transactional paradigms

Technologic advances have helped to explain more and more disorders by pointing to processes occurring in the brain rather than by considering behavioural or environmental influences as the sole explanation possibilities (Teeter Ellison & Semrud-Clikeman, 2007, pp. 1-2). The field of science examining the brain and its relation to a person's behaviour is called neuropsychology. It is founded on the notion that "[a]ll behaviour – including cognitive processes, which are essentially psychological – is mediated by the brain and central nervous system and their integrated and supporting physiological systems" (p. 3). However, neuropsychology does not suffice to explain the relation between different factors, such as environmental influence, behaviour and psychological aspects. Therefore, other theories should be used as well if one wants to have the full picture of the subject matter to be examined (ibid.).

In a transactional model, it is explained how environmental factors influence the child's development. First of all, the development of the central nervous system is dependent on both biogenetic factors and environmental impact, such as complications during birth or pre- and postnatal toxins or insult. The child's intellect, perception and cognitive

capacity are shaped by biological factors only, namely various brain regions, which in turn are affected by cortical and subcortical structures. These various brain regions also have an influence on the academic, behavioural and psychosocial development of the child and may be the source of disorders, just as environmental impacts. Such can be school or family, the latter of which includes the standard of living, but also the way parents treat their children (Teeter Ellison & Semrud-Clikeman, 2007, pp. 6-7). For instance, childhood stress can lead to psychopathologies. In this way, the child's development is shaped by experiences it makes. Thereby, it adapts to the circumstances in which it lives, which enables it to survive under these circumstances, but which are not adequate in a different environment (p. 7). This assumption is also the basis of the alternate developmental pathway approach, which will be dealt with in a later section.

4. The sources of insecure attachment

Attachment has been defined in the introduction as parent-child relationship that is important for the child's development in many respects, as for example its personality. Thereby, the provision of security and affection plays an especially important role (Gervai, 2009, p. 1). There are different types of attachment, which can be observed upon the performance of a so-called Strange Situation Procedure (SSP), during which the child is separated from its parent in order to observe its reaction upon reunion. Insecure attachment is shown when the child avoids its parents or when it cannot calm down; disorganised attachment, the most extreme form of insecure attachment, manifests itself in a mixture of attitudes which contradict each other. Insecure attachment is problematic for the further development of the child as it may lead to behavioural disorders (Attachment Synthesis, 2012, p. i). The kind of attachment that applies in a particular case is said to depend mainly on the sensitivity that parents provide their child with (p. ii). This would imply that children who are treated with sensitivity always develop secure attachment to their parents, while those confronted with neglect, abuse or violence are always insecurely attached. However, of those children that are maltreated, there is still a number of up to 30% that are securely attached (Gervai, 2009, p. 2). This leads to the conclusion that there are other factors that play a role as well in the formation of a certain kind of attachment. Indeed, it has been found that "anomalies of caregiver's mental state and behaviour had only low explanatory power in accounting for attachment disorganization" (p. 3). In addition to the environmental factor, there also seems to be a biological component accounting for the type of parent-child relationship (p. 4), namely

the 48 base pair tandem repeat (48 bp VNTR) of the D4 dopamine receptor (DRD4) gene which has been associated with the child's temperament. Indeed, the 7-repeat variant of the 48 bp VNTR polymorphism has been found overrepresented in children insecurely attached to their parents. It thus seems that it has protective effect if the 7-repeat allele is not transmitted facilitating the development of secure attachment (p. 5). There are other genes that modulate the effect of the kind of rearing on the child's development. For instance, the monoamine oxidase A (MAOA) gene has been found to regulate the effect of maltreatment on the development of antisocial behaviour (p. 6). These findings must be kept in mind when the rest of the paper is read: the kind of rearing, which also implies maltreatment, does not seem to have an effect on the behavioural development of the child if its genetic predisposition grants protective effect. Hence, the observations made below are valid only for those children without such protective effect.

5. Article 19 of the UN Convention on the Rights of the Child: a human rights perspective

The purpose of this section is the introduction of Article 19 CRC, which forms the basis of the present paper. The CRC was adopted on 20 November 1989. It is one of several UN Human Rights Conventions that have been adopted on the basis of the provisions of the Universal Declaration of Human Rights making these legally binding. It was negotiated and adopted by the UN Commission on Human Rights. A number of 193 states are a party to it, 140 so far have signed (United Nations Treaty Collection, n.d.; Schmidt, 2010, p. 404). Article 19 (1) of the CRC states as follows:

> States Parties shall take all appropriate legislative, administrative, social and educational measures to protect the child from all forms of physical or mental violence, injury or abuse, neglect or negligent treatment, maltreatment or exploitation, including sexual abuse, while in the care of parent(s), legal guardian(s) or any other person who has the care of the child.

The Committee on the Rights of the Child, established under Article 43 CRC, has made a general comment on this article. Such comments by the treaty bodies serve to interpret and to specify convention provisions (Schmidt, 2010, p. 409). General Comment 13 is used to specify the content of Article 19 CRC and to interpret it in the context of the whole

convention. Therefore, it refers to the different components of subparagraph 1 and defines these further. For instance, neglect implies physical, psychological, and educational neglect and neglect of the child's health. These different forms of neglect are themselves further specified. If a parent does not pay attention to the child's well-being by allowing it to be injured or by denying it basic needs such as clothes or food, this is physical neglect. Neglect of health means that the parents do not provide their child with the necessary medical care. Psychological or emotional neglect refers to a state in which the child does not receive affection and attention from its parents and does thus not feel loved and cared for (Committee on the Rights of the Child, 2011, pp. 8-9). In this way, all the different components of Article 19(1) CRC are subdivided and then further specified. The Committee seems keen on not leaving forgotten any possible form of maltreatment. Therefore, it underlines that "all forms of violence against children, however light, are unacceptable" (p. 8).

A child rights approach shall guarantee all the rights to children that are enshrined in the Convention. Children shall be considered as individual persons, having their own rights with an entitlement to a fulfilment of those. This implies that the child must have the right to express its opinion which shall be heard, laid down in Article 12. Article 3 provides that everything that concerns the child shall happen in its best interest. The best interest principle shall be interpreted in line with the whole Convention, excluding the justification for violence in the child's best interest (Committee on the Rights of the Child, 2011, pp. 23-24). The notion of development in Article 6 CRC shall be understood "as a holistic concept, embracing the child's physical, mental, spiritual, moral, psychological and social development" (p. 24). Only if all the components implied in this term are fulfilled, holistic development is ensured (Willems, 2012, pp. 11-12). Willems refers to holistic, or healthy, child development as the principal aim of the CRC. In his view, *physical* implies mainly "early brain" (2012, p. 11) development. This early brain development is the focus of the next section

6. Article 19 of the UN Convention on the Rights of the Child: from a neurobiological point of view

As indicated before, the duty to guarantee holistic child development includes the development of the child's brain (Willems, 2012, p. 11). Teicher et al. write that consistent stress during childhood greatly affects brain development regarding composition as well

as operation (Teicher et al., 2002, p. 397). In order to understand the impact of early stress on this development, it must first be set out how the brain generally develops during childhood.

6.1. The developing brain

According to the transactional paradigm, both biological and environmental factors shape the brain in its development. This can be confirmed from a neurodevelopmental point of view: the basic composition of the brain is determined by genetic information. It is then further developed by experiences we make, meaning that the neuronal connections in the brain are subject to influences from outside and are affected by those. The first stage in brain development is taken already before birth: approximately 250,000 brain cells are created every minute leading to the fastest growth phase of the brain of the whole development (Ellison & Semrud-Clikeman, 2007, p. 19). Neurons are guided to change position arriving at their final, genetically determined location and forming branches with a view to building connections. What follows is called cell death: this is a process whereby more than half of these neurons disappear because they are not needed (Teicher et al., 2002, p. 398; Ellison & Semrud-Clikeman, 2007, pp. 20-21).

New neurons are produced before birth only by the time of which the brain has adult characteristics, although it is still immature (Ellison & Semrud-Clikeman, 2007, p. 20). Afterwards it is only in the hippocampal gyrus that neurons are created. Until the age of fife, the brain becomes bigger, which is for a big part due to myelination of fiber tracts, which is a process whereby exchange of information is accelerated. In the prefrontal cortex (PFC), this process takes place at a later stage (Teicher et al., 2002, p. 398). Myelination happens in the context of development of other capabilities, such as social and cognitive behaviour. This implies that outside events can affect myelination, which again affects the child's learning process. Just as before birth, an overproduction takes place, this time however not of neurons but of synapses. In contrast to the process taking place before birth, the synapses are pruned back without neurons being deleted anew. This process is different from one brain region to the other. The frontal and parietal cortices, for instance, are those gray matter regions with the greatest increase, while the basal ganglia decrease. Other brain areas that increase are the hippocampus, amygdala, and corpus callosum (Teicher et al., 2002, pp. 399-400; Bremner, 2008, pp. 12-13). During the period of overproduction, new information is absorbed in a process which takes a lot of time and energy. At the end of this process, connections not needed are pruned back. It is this process of overproduction and subsequent pruning, which makes children especially vulnerable to stress, such

as maltreatment by their caregivers, because it can have consequences on the brain's development that are not reversible (Teicher et al., 2002, pp. 399-400).
Ellison and Semrud-Clikeman conclude that "[a]lthough genetic factors certainly map the nature and course of neuronal development, environmental factors have a significant influence on the developing nervous system" (2007, p. 22), meaning that genetics lay the foundation for our abilities, but environmental factors influence their maturation (p. 50). To guarantee well-proceeding brain development, stimulation must be appropriate (p. 22). This refers to the way parents interact with their children. Parents should be attentive towards their children and show them that they are cared for to evoke a positive effect in them. In other words: children need to be securely attached to their parents (p. 50). This is in line with the human rights perspective, according to which children only grow up holistically if the parent-child relationship is oriented towards the child's best interest. If children do not live under such circumstances, but for instance are neglected, and are thus insecurely attached to their parents, this affects their development in a negative way (p. 50). This issue is going to be dealt with in detail under the following subtitle.

6.2. Consequences of child maltreatment on the developing brain

It might seem necessary to define child maltreatment before I can go on further in my research. As was made clear in Article 19 of the CRC, child maltreatment can have various facets, and it might seem useful to focus on one in particular in order to have meaningful results in the research. However, researchers in this area have mainly focused on the consequences of stress resulting from maltreatment such as Navalta, Tomoda and Teicher, who claim that "the stress that results from CA [child abuse] has an unfavorable effect on neurodevelopment and, consequently, behavioural development" (2008, p. 50). This focus on stress rather than on one kind of maltreatment also enables me to draw a more general conclusion in the end concerning all kinds of maltreatment listed in Article 19 CRC and not merely one. Therefore, I am first defining stress in order to investigate in the second subsection its effects on brain development. The behavioural aspect will be dealt with in a later section.

6.2.1. Defining child maltreatment as stressor

According to the handbook of stress, "[s]tress can be defined as any challenge to the homeostasis of an individuum that requires an adaptive response of that individuum" (Steckler, 2005, p. 25). It is constituted by three elements, namely by a stressor, the assessment thereof, and a reaction. A stressor is an aversive stimulus that is possibly

harmful to the one experiencing it as an alteration of his environment (ibid.). Two types of stressors can be distinguished: firstly, there are those that intrude in a disagreeable way into the field surrounding the person concerned. Secondly, stressors can manifest themselves as a removal from that familiar field. Both can happen once or several times, shortly or for a long period, or they can happen continuously. The field in which the person finds himself and that is intruded by the stressor can be either external or internal. If the external field is affected, this most probably causes psychological stress, which refers to those stressors that are assessed as being stressful themselves. It is moreover considered "as an asymmetry between the motivational systems of reward and punishment" (p. 26). If, by contrast, the internal field is concerned, this results in physical stress, whereby the stressor itself is mostly not assessed but the stress is rather an automatic reaction to it. Physical and psychological stress can furthermore be distinguished by the brain regions that it activates: the latter has an impact on the higher brain regions, the former on the lower ones. However, the actual distinction between physical and psychological stressors is not always clear, and it can often be argued in both ways (pp. 26-27).

Wallick analysed among others child maltreatment as one type of stressor. She states that the exact definition of what maltreatment means differs from one legal system to another. For the purpose of the present paper, the list of different maltreatment forms laid down in Article 19 CRC serves as definition. However, as the stress resulting from this maltreatment is analysed rather than the different forms themselves, an exact definition is not of importance for the purposes of this paper. It merely needs to be emphasised that all the forms listed in that provision belong to the category of maltreatment and are therefore able to evoke the stress effect that is analysed here. This enables a general conclusion valid for the whole provision. For a confirmation of the categorisation as maltreatment of a certain kind of acts of some parents against their children, one might have a look at the Maltreatment Classification System developed by the University of North Carolina within the framework of their longitudinal studies of child abuse and neglect. Similar to Article 19 CRC, different forms of physical abuse, including violence against the child, sexual abuse, physical neglect, emotional maltreatment, and educational maltreatment are distinguished (English & LONGSCAN Investigators, 1997).

The child's affected environment can be defined as its external environment. Maltreatment would thus be classified as psychological stress, also because of the asymmetry of punishment and reward argument. According to Wallick, maltreated children are "motivated more by the necessity to avoid pain than by pleasurable reward, thus throwing out of balance the normal regulatory function of the pleasure principle" (1990, p. 205). The probably most meaningful criterion in a classification of child maltreatment as stressor

is the brain regions that are affected by it. I am going to deal with this in the second subsection.

6.2.2. Effects on the developing brain

According to Teicher et al., maltreatment during childhood, and the stress resulting from it, has effects on the developing brain because at that time it is not fully developed yet so that brain development can still be easily influenced by outer circumstances (2002, p. 397). In particular, they refer to those processes that have been claimed to be vulnerable to outside events above, namely the formation of neurons, overproduction of synapses, pruning, and myelination leading to different consequences in the various parts of the brain (2003, p. 33). There have been differing hypotheses about the concrete consequences of childhood stress on brain development. Teicher et al. suggest that the "alterations in neurodevelopment represent an adaptive, alternative developmental pathway" (p. 39), allowing the child to adapt to its stressful environment, rendering it however unable to adapt to a friendly environment. Ayoub and Rappolt-Schlichtmann confirm this hypothesis and claim that there seems to be reliable proof for it. They write that the theory of alternative developmental pathways form part of developmental traumatology, which consists of research on trauma, developmental psychology and developmental psychopathology. Thereby, emphasis is set on the effect of the environment, in particular a stressful environment due to maltreatment, on the development of children (2007, p. 306). The developmental pathway approach is founded on the notion that "skills are constructed gradually over the course of development through practice in real activities during interaction with others and independently" (p. 308). Thus, everyone's development is shaped and influenced by experiences one makes and by the spheres in which one moves. Thereby, a variation of abilities evolves which is different for every child dependent on the environment by which it is surrounded. Thus, if the child is maltreated, it develops in a different way than children growing up in a friendly and affective environment. To be more specific, maltreated children separate events and the fears resulting from these to protect themselves from being overwhelmed by those fears. If the events causing their fears continuously occur, such a separation, together with the way of thinking and feeling connected to it, becomes habit. This is then called traumatogenic alternative developmental pathway (p. 309).

As established before, child maltreatment may be considered as stressor through which stress hormones, in particular corticosteroids, are released. Under ordinary circumstances, the brain is protected from a high level of such a release, but certain stressors can lead to a high level of corticosteroid release (Teicher, 2002, p. 402). Corticosteroids have great effect

on brain development, namely glucocorticoids in children are responsible for permanent decrease in brain weight. Moreover, through an "N-methyl-D-aspartate (NMDA) receptor-dependent glutamate excitatory pathway" (ibid.), granule cell mitosis is prevented after birth in the cerebellum and dentate gyrus, which in turn, disturbs myelination. This is enabled by the hormones norepinephrine and vasopressin. Apart from that, neural morphogenesis is changed in some brain regions (p. 401).

Stress not only affects myelination, neural mitosis, and morphogenesis, but it also the stress-response system, which consists of three main components: firstly, the hippocampus together with the hypothalamic-pituitary-adrenal (HPA) axis regulating the feedback of cortisol. Secondly, a response mechanism to norepinephrine and adrenaline, responsible for among others fight or flight reactions, is needed, which is constituted by the amygdala, locus coeruleus (LC), adrenal gland, and sympathetic nervous system. Thirdly, vasopressin and oxytocin activate the release of adrenocorticotropin hormone (ACTH). Through experiences, changes in this stress-response system can be developed. Stress leads to permanent changes. For instance, neglect changes the consistence of gamma-aminobutyric acid (GABA)-benzodiazepine so that less GABA-A receptors are to be found in the amygdala and the LC. Neglect also leads to a higher level of the corticotropin-releasing hormone (CRF) mRNA, while it leads to a lower level of $\alpha 2$ noradrenergic receptors in the LC so that the feedback inhibition of noradrenergic neurons is suppressed. Neglect as well as extended stress reduces the amount of glucocorticoid receptors in the hippocampus. Furthermore, negative feedback inhibition of cortisol is decreased, while the level of vasopressin mRNA in the hypothalamus is increased so that more ACTH and corticosterone are created. To summarise the molecular effects of early stress, it leads to increased levels of norepinephrine, vasopressin and corticosteroid (Teicher, 2002, pp. 400-401).

If one investigates the structural consequences of early stress on the brain, the focus should be on those parts that have a large quantity of glucocorticoid receptors, on those parts whose neural development is not completed on birth, and on those with an extended development after birth. These parts are especially sensitive to cortisol neurotoxicity rendering them vulnerable to continuous stress during childhood. A part of the brain with these characteristics is the hippocampus (Teicher et al., 2003, p. 37). At different steps during development, there is a different amount of synapses in this area. After birth, there is an especially high amount due to overproduction as already explained before. As a result of stress, however, this overproduction does not take place. Pruning nevertheless does take place leading to a lasting state of too low a concentration of synapses (p. 34). Moreover, pyramidal cells can be changed or even destroyed as a result of stress, and the creation of new granule cells is prevented (Teicher et al., 2002, p. 402).

A region with even more glucocorticoid receptors, and also fulfilling the other two criteria for stress-sensitive brain regions, is the cerebellar vermis making it thus especially sensitive to stress (Teicher et al., 2003, p. 36). It is responsible for regulating the level of glucocorticoid through a connection to the hypothalamus. This may be important in controlling the effects of stress and neglect on the brain and behaviour (Teicher et al., 2002, p. 410). For persons who were abused or neglected during childhood, there is less perfusion of the cerebellar vermis so that it is not as active as with people without such a history, which implies that it is functionally affected so that it cannot exercise its controlling function (Teicher et al., 2003, p. 35; Ayoub & Rappolt-Schlichtmann, 2007, p. 315; Teicher et al., 2002, p. 406).

The amygdala is also affected by childhood stress. It is an important region for the occurrence of kindling, which may lead to neuronal excitability through sporadic stimulation possibly resulting in seizures. GABA is responsible for diminishing such excitability. However, as established above, the amount of GABA receptors is reduced as a result of stress. This should mean, argumentum e contrario, that neuronal excitability is not reduced. Apart from that, the serotonin and dopamine innervations are permanently affected with the level of dopamine being increased and the level of serotonin being decreased in the amygdala. Moreover, the size of the left amygdala has been found to decrease. Together with abnormal development of the hippocampus, these findings may result in temporal limbic seizures (Teicher et al., 2002, pp. 403-404).

The corpus callosum is also vulnerable to early stress because glial cell division may be hindered which is needed for myelination (Teicher et al., 2003, p. 35). Moreover, it has been found out that the volume of the corpus callosum becomes smaller in the middle parts, which may lead to less communication between the hemispheres of the cerebrum. This effect has been observed especially with boys that are neglected by their parents and with girls who have been sexually abused. Moreover, child maltreatment leads to a more dominant right hemisphere, which is responsible for emotions, in particular the negative ones (Teicher et al., 2002, p. 405; Teicher et al., 2003, p. 35; Ayoub & Rappolt-Schlichtmann, 2007, pp. 319-320; p. 321).

The cerebral cortex, in particular the PFC, also has many glucocorticoid receptors rendering it vulnerable to glucocorticoid toxicity resulting from stress (Teicher et al., 2003, p. 36). Arnsten confirms this and states that "even mild uncontrollable stress can rapidly impair PFC functions" (2009, p. 410). The cerebral cortex is a brain region developing slowly, especially the PFC with myelination extending until the mid-twenties (Teicher et al., 2003, p. 36). The ventromedial PFC (VMPFC) is connected to subcortical brain regions which are responsible for emotions and which its task is to control. The dorsolateral PFC (DLPFC),

in contrast, is connected to sensory and motor cortices so that its responsibility is the control of thought, action, and attention. Inappropriate behaviour shall be prevented by the right inferior PFC (rIPFC). If stress does not disturb the functioning of the PFC, these different parts work together to "orchestrate the brain's activity for intelligent regulation of behaviour, thought and emotion" (Arnsten, 2009, p. 411). The PFC is responsible for functioning as inhibitor to the stress response of subcortical areas as well of the HPA axis (Teicher et al., 2003, p. 36). Stress leads to a release of dopamine and norepinephrine by the amygdala to the PFC. The stress regulation by the PFC is consequently affected, while the amygdala is strengthened in its activity so that behaviour regulation as done by the amygdala instead of by the PFC meaning that it is directed by emotions and reflexes rather than by deliberation. If this happens chronically, the result is thus a weakened stress response by the PFC and instead promotion of the stress response by the amygdala (Ayoub & Rappolt-Schlichtmann, 2007, p. 320; Arnsten, 2009, p. 411; p. 418). Stress has a great impact on the structure of the PFC: it requires only one week of stress for the dendrites of the PFC to alter. Sometimes, even a one-time impact of stress is sufficient to affect the PFC (Arnsten, 2009, p. 418). Hence, adults who were maltreated while they were been children, showed a smaller volume of the left DLPFC and of the right medial PFC (McCrory et al., 2010, p. 1085). Ayoub and Rappolt-Schlichtmann state that a further effect of early stress on the PFC is retarded development thereof (2007, p. 320).
Teicher et al. conclude that the brain shapes according to experiences one makes. Thus, also the brains of children who are under continuous and severe stress shapes according to this negative experience. This creation of an alternate developmental pathway allows those children to cope with the stress of abuse or neglect they have to live with (Teicher et al., 2002, p. 414).

7. Neurobiology of criminal personality traits

As the antisocial – rather than the criminal – person's behaviour has been investigated in the field of psychology, I use these sources for my research on the neurobiology of the criminal person, focussing thereby on the character traits and behaviour that are typical for criminal persons as defined before, such as immoral behaviour, lack of empathy, and impulsive aggression, "which often culminates in physical violence" (Davidson et al., 2000, p. 591). First, it will be described which regions of the brain are responsible for moral, empathetic and non-violent behaviour in order to, as a second step, examine what may be different for people showing the negation of all these characteristics, which may be a sign for criminal behaviour as it is defined as such a negation.

7.1. Brain regions responsible for moral, empathetic and non-violent behaviour

The question is if morality, which is defined as "that code of values and customs which informs social conduct" (Marazziti et al., 2013, p. 3), really is regulated by processes in the brain or if it is rather a social construct that our brain has no influence on. The famous case of Phineas Gage, who had his frontal lobe smashed by a metal bar and whose behaviour completely changed afterwards into immoral and antisocial attitudes, leads Marazziti et al. to the conclusion that ethical and social behaviour is determined by processes that occur in the brain. Brain damage can lead to a reversal of such personality traits (2013, p. 2). The brain regions involved in morality are three parts of the frontal cortex: the VMPFC, the DLPFC, and the orbitofrontal and the ventrolateral cortex (OFC/VL) as one region. More specifically, the VMPFC is involved in the innate moral sense, which refers to the moral sense that is active if we have to take a decision in a moral problem involving ourselves and the possibility that someone else might be harmed. The DLPFC is responsible for moral problems of a more general nature, but it is also active in situations in which the innate moral sense is needed. The OFC/VL area regulates aversive reactions, it alters reactions depending on how another person reacted, and it controls impulsiveness created by the amygdala (Marazziti et al., 2013, p. 4), which will be explained below.

The VMPFC and the OFC/VL are also involved in the creation and activation of the second characteristic that is investigated here, namely empathy: the VMPFC is responsible for the ability to understand someone else's perspective on a certain issue, the OFC/VL for the emotional kind of empathy, which refers to the ability to feel with someone else and to share this person's feelings (Marazziti et al., 2013, p. 5; Rodrigo et al., 2010, p. 2). There are more areas in the brain, which are active in the creation and expression of emotions, such as empathy. Especially in the case of negative emotions, the insula is activated. The same process occurs when such emotions of fellow human beings are observed, in particular those that are dear to us. The relation between the insula and another part of the cerebral cortex, the anterior cingulated cortex (ACC), is significant as well for the experience of emotions and the ability to feel with others. The ACC relates to the autonomic nervous system, which is responsible for automatic reactions in situations when they are needed, and it also reacts in situations when something does not seems to be right (Rodrigo et al., 2010, p. 2). Thus, the ACC is not only important for empathy, but also for morality. Together with the OFC/VL, it belongs to the limbic system, which is said to be important for empathy. A further part of it is constituted by the amygdala. It is equally involved in learning and experiencing emotions, and it is responsible for reactions to the fear of others

(ibid.; Marazziti et al., 2013, p. 4). How empathy is created may be explained with the aid of so-called mirror neurons, which are active when someone pursues an action himself and also when he observes someone else pursuing this action. Marazziti et al. explain that the "neurons of the observer 'mirror' what is taking place in the mind of the observed subject, as if it were the observer that was carrying out the action" (2013, p. 5; Rodrigo et al., 2010, p. 2). Thereby, other brain areas become active, namely the inferior area of the anterior central gyrus, the posterior area of the inferior central gyrus, and the anterior rostral part of the inferior parietal lobe. Marazziti et al. emphasise, however, that it is only a hypothesis that mirror neurons are the involved in the creation of empathy (2013, p. 5).

Emotions also play a role in violence and aggression, more specifically negative emotions, such as impulsive aggression, negative affect and anger, and the inability to control them. This leads to the conclusion that non-violent people are able to control such negative emotions (Davidson et al., 2000, p. 593; p. 591). This is normally done via a mechanism, whereby an area in the PFC, which probably is the OFC, works as an inhibitor to the amygdala (p. 592). The amygdala responds to signs of threat, such as a fearful face, whereby it is important to note that the response to a fearful face is stronger than to other facial expressions, as for example an angry face. Anger is rather associated with a more active OFC and ACC, which is a mechanism that shall regulate the strength of the anger expressed (p. 591). The OFC is involved in the process of reversal learning which means that someone having reacted negatively in one case learns from this in another and restrains this negative feeling (p. 592). Based on the knowledge gained from this subsection, the next one lays down what happens in the respective brain regions if one's behaviour is immoral, unsympathetic and violent.

7.2. Neurobiology of immoral, unsympathetic and violent behaviour

Marazziti et al. indicate that studies looking at the neurobiology of criminals are restricted, but that "some criminals often display unspecific alterations at temporal level, or disturbance of other brain areas" (2013, p. 5). For instance, people behaving in an indifferent and unemotional way often show abnormal functioning of the VMPFC, the OFC/VL and the amygdala and those who are violent often have defective frontal lobes (Rodrigo et al., 2010, p. 3; Silver et al., 2003, p. 757). Teicher et al. add that irritability of the amygdala can be the reason impulsive violence and for loss of self-control (Teicher et al., 2002, p. 408).

Davidson et al. claim that persons with a defective OFC and those vulnerable to impulsive aggression are deficient at reversal learning, meaning that they are not able to suppress

their negative emotions (2000, p. 592). This is because if the OFC is defect, it cannot inhibit the reaction to aversive stimuli which is regulated by the amygdala (ibid.). This leads to an overreaction that would normally be regulated by the frontal cortex. More specifically, stimuli that would normally have no great effect now lead to exaggeratedly angry and/or aggressive reactions. Hence, if the neocortex, especially the OFC, is defective, the consequence can be anger that is not inhibited and needs not a lot of provocation so that the respective person has no regard for the effect on others that his or her anger has and the behaviour related to it. Equally, a defective amygdala can lead to violent behaviour (Silver et al., 2003, p. 757). Moreover, aggressiveness results from kindling in the amygdala (p. 758; Davidson et al., 2000, p. 593). Marazziti et al. add that early abnormal, or reduced, development of the amygdala may result in abnormal development of the VMPFC and the OFC/VL. They specify that this "would lead to an erroneous association between actions that are harmful to others and negative reinforcement of the discomfort of the victim" (2013, p. 6). Abnormal functioning of the ACC and insula, however, is rather associated with psychopathic behaviour. If the limbic and paralimbic circuitry are less active, this makes it more difficult to be empathetic, especially with the fear of someone else. It is also associated with egoistic and antisocial behaviour, such as the refusal to help, and it hinders a person from recognising when his actions render others uncomfortable. Moreover, mirror neurons, which are involved in the activation of empathy, are suggested to be abnormal in callous persons (Rodrigo et al., 2010, pp. 2-3).

More observations that were made in the brains of criminals are put forward by several authors, such as abnormalities of the temporal lobe. For instance, the so-called "dyscontrol syndrome" (Silver et al., 2003, p. 757) results from epilepsy of the temporal lobe, for which suspiciousness and no control over impulsiveness are characteristic (ibid.). Such seizures of the temporal lobe can moreover be responsible for aggressive behaviour (Teicher et al., 2002, p. 408). The hypothalamus is responsible for "fight or flight" responses linked to the autonomic reaction mechanism. If it is impaired, this may result in aggressiveness, even without being provoked (Silver et al., 2003, p. 757). A smaller volume of gray matter in the prefrontal cortex may be connected to a decreasing ability to automatically respond in a given situation, which might have been responsible for violent criminal acts in the past. Marazziti et al. add that the smaller the prefrontal cortex, the greater is the risk for antisocial behaviour. They claim that the cerebral cortex of children showing no affection seems to mature later than for other children. Davidson et al. affirm this with a finding of "hypoactivation in prefrontal territories including lateral and medial zones of the PFC" in a study of 41 murderers (2000, p. 593; Marazziti et al., 2013, p. 6). Violence is not only characterised by neurofunctional but also by also by neurochemical processes.

Norepinephrine, for instance, may be responsible for aggressive behaviour, such as sham rage, affective aggression or fighting as a result of shock (Silver et al., 2003, pp. 757). Release of norepinephrine can be the consequence of amygdaloid stimulation. However, it is not entirely clear to which extent there really is such a correlation as studies are pointing into different directions. As Siegel states, "[t]he most likely possibility is that various noradrenergic receptor subtypes mediate different effects on aggressive processes" (2005, p. 170). Serotonin, in contrast, has been studied more in this context. Less activity of serotonin, in particular if that is the case in the PFC, has been found out in various studies to lead to aggressiveness, in particular impulsive aggressiveness because serotonin is an inhibitor of such a kind of aggressiveness. In part, this low activity of serotonin can be made responsible for hyporesponsive cortisol secretion which, together with a low basal cortisol level, may permanently alter cortical and subcortical connections. Such a phenomenon has been found with people behaving aggressively (Silver et al, 2003, pp. 757-758; Davidson et al., 2000, pp. 592-593; Rodrigo et al., 2010, p. 3). By contrast, the level of dopamine has to be raised to effect aggressive behaviour (Silver et al., 2003, p. 757; Siegel, 2005, p. 175). GABA increase, in turn, is associated with less aggressiveness (Silver et al., 2003, p. 758).

These findings are compared to the neurobiology of maltreated children in the next section, which elaborates on the behavioural consequences of maltreatment, having regard especially to those characteristics that are at risk for developing criminal behaviour.

8. Discussion of findings

According to the transactional model, neurobiology is impacted by environmental occurrences. Or, to say it differently – in terms of the alternate developmental pathway approach – neurobiology adapts to the environment and the experiences that we make therein. This was observed in section 6.2 of the present paper. The child that is permanently subject to maltreatment by its parents adapts to these circumstances: the brain develops in a way that is different from that of children raised in an affective environment and that allows the child concerned to survive under the negative impact that it has to experience. Neuropsychology and the transaction model both concede furthermore an influence of neurobiology on the behaviour of the respective person. Teicher et al. state that this alternate developmental pathway leads to, among others, a hostile attitude (2002, p. 415). At another point, they speak of an alteration in social behaviour, which they explain by the release of glucocorticoids as a result of stress (p. 401). In this way, behavioural

consequences can be found for the development of each of the brain regions described in section 6.2. The hippocampus is supposed to be an important inhibitor of environmentally inappropriate actions (Teicher et al., 2003, p. 37) which leads to the conclusion that an affected hippocampus cannot exercise its inhibiting function. Hence, there seems to be a likelihood that a person with an affected hippocampus may behave inappropriately to his surroundings (Ayoub & Rappolt-Schlichtmann, 2007, p. 311). This might be an indicator of antisocial behaviour, but it does not necessarily have to lead to criminality.

The PFC is said to be responsible for making us "rational, intellectual, and moral entities" (Mobbs et al., 2007, p. 693). This function, however, is impacted by chronic stress which leads to an emotional and impulsive reaction instead of a deliberative one, which would normally be regulated by the VMPFC together with the DLPFC and the rIPFC. Instead, response to stress is controlled by the amygdala. It seems thus that the inhibitor function of the amygdala by the PFC is affected so that stressful situations lead to anger that is not inhibited leading to impulsive and aggressive reactions. Moreover, the volume of the left DLPFC becomes smaller, which might impair the function it has in the context of morality. The volume of the right medial PFC also becomes smaller as a result of stress. As said before, a smaller volume of grey matter in the PFC may have been the source of criminal acts in the past, and less activity in the PFC has been found in a study involving murderers. The VMPFC is responsible for remembering experiences and for modulating emotions and behaviour according to these memories. It has moreover the task to inhibit behaviour that is not appropriate to one's environment (Arnsten, 2009, pp. 410-411). It also directs the innate moral sense and the ability to understand the point of view of someone else. The OFC has the task of controlling impulsiveness and the emotional kind of empathy. These functions do not seem to be directly affected by stress, but indirectly they might as the PFC matures later as a result of stress. Such a late development has been found in children that do not show affection. It might also account for acts of delinquency if the regions responsible for empathy and morality develop later than normally.

As stated above, the right hemisphere of the corpus callosum is supposed to be responsible for negative emotions. If it is more dominant than the left one, this may hence lead to a prevalence of negative emotions. In an experiment, Ayoub and Rappolt-Schlichtmann observed that maltreated infants had problems telling positive stories dealing with an interaction of themselves and another person. Instead they preferred to tell negative ones, whereby they even turned positive stories into negative ones. Negative stories are also told in a more complex and active way than positive stories; they are often violent or aggressive. These infants see themselves as the bad protagonist, while non-maltreated infants consider themselves to be good (2007, pp. 321-322). This, together with the findings

of a smaller corpus callosum and reduced communication between the hemispheres, may have the consequence that the person with such a brain development becomes angry and aggressive if confronted with situations of danger or loss (Teicher et al., 2002, p. 414).
Changes in the amygdala and the limbic system can be responsible for a fight-flight response and aggressive reactions as well as impulsive violence (Teicher et al., 2002, p. 414; Teicher et al., 2003, p. 37). More concretely, irritability has been associated with the reduced size of the left amygdala side (Teicher et al., 2002, p. 404). As has been explained above, the reduction of GABA leads to kindling. At another point, it has been explained that irritability of the amygdala may be the source of violence, aggressiveness, and missing self-control. The effects on the amygdala and the hippocampus have also been found to lead to seizures of the temporal lobe, which has later been called "dyscontrol syndrome", which was said to be responsible for suspiciousness, impulsiveness and aggressiveness. Moreover, stress-induced changes in the amygdala, but also in the cerebellar vermis, may lead to violent behaviour. In this regard, Ayoub and Rappolt-Schlichtmann state that severe maltreatment may have the effect of "frequent impulsive violence, increased fear and negativity" (2007, p. 316). Apart from that, a reduced development of the cerebellar vermis plays a significant role in maintaining the stress effects flowing from the corpus callosum and the hemispheres described above (Teicher et al., 2002, p. 415).
Comparisons can also be made between the neurochemistry of stress and that of criminal characteristics: an increase of GABA has been stated to be negatively correlated with aggressiveness, meaning that a reduced level of GABA resulting from stress may result in more aggressiveness. Stress raises the level of norepinephrine. Although findings have pointed into different directions, it may be a source of different forms of aggressive behaviour. The level of dopamine has been found to increase, while that of serotonin decreases. These processes correspond exactly to what has been said about what has to occur to these neurotransmitters in order to lead to aggressiveness, in particular impulsive aggressiveness in the case of serotonin.
The cerebellar vermis is particularly vulnerable to stress because of its high density of glucocorticoid receptors. This factor renders the hippocampus and cerebral cortex vulnerable areas as well. The effect of stress on the cerebellar vermis may be violent behaviour, and the maintenance of angry and aggressive reactions resulting from stress effects on the corpus callosum and the hemispheres. Anger, aggressiveness, and impulsiveness are also consequences of stress and the effect it has on the amygdala and the PFC and the changes of dopamine, serotonin, norepinephrine and GABA levels. A striking fact is that a smaller volume and hypoactivity of the PFC have been found in criminals, in particular murderers. The PFC's responsibility in morality and empathy

may also account for these findings. Violence, aggressiveness and impulsiveness are characteristics that have been named in the beginning to define criminal behaviour. There seems thus to be a link between maltreatment of children and criminality: maltreatment is a stressor evoking changes in the brain bearing the risk to develop criminal behaviour. Maltreatment can be considered to be a psychological stressor because in addition to the external environment and reward/punishment arguments the affected brain regions seem to confirm this hypothesis. The child adapts to its malevolent environment. Not only neurostructural and neurochemical changes occur, but these changes also have an impact on the child's behaviour: it reacts more aggressively which enables it to survive in its malevolent surroundings. It makes it, however, difficult to adapt its behaviour to a friendly environment. Hence, the brain changes lead to that behaviour established before: violence, impulsiveness and aggressiveness. These factors do not necessarily lead to criminality, but they represent at least a risk for such behaviour.

Hence, the thesis may be confirmed that a violation of Article 19 CRC, which prohibits child maltreatment, has social costs in the form of high criminality rates. From a human rights perspective, it was argued that a parent-child relationship should be based on the child's best interest and its holistic development, which included, among others, brain development. If the relationship is in fact formed according to the child's best interest, this means that maltreatment does not occur. The child's opinion and its healthy development are of major interest to the parents. Moreover, the parents treat their child with affection, which in a different part of the paper has been stated to be the basis for a normal, thus holistic, brain development. This implies that there would be no reason for the brain to develop according to the alternate developmental pathway approach. There would be no reason to develop aggressive behaviour. The compliance with Article 19 CRC is thus of utmost importance not only for the protection of the child and its holistic development but also for society as a whole. The compliance with Article 19 CRC might reduce the amount of people behaving impulsively, aggressively and violently. The consequence may be a reduction of criminal acts.

9. Conclusion

The present paper sought to answer the question of in how far a violation of Article 19 of the CRC leads to criminal behaviour. Through an analysis of the impact of early stress on the brain and the neurobiology of people exhibiting character traits that might be at risk for developing criminal attitudes I arrived at the conclusion that there is a connection between the neurobiology of maltreated children and people showing impulsive, violent and aggressive behaviour. There is a connection in so far as the child's brain develops according to the alternate developmental pathway approach in such a way that the brain adapts to the unfriendly surroundings of the child. These alterations, however, continue to exist in a friendly environment and are alterations of a kind that is often found in people exhibiting criminal behaviour.

This paper focused on maltreatment as a possible source of criminality. Of course, there may be many different factors that may be a reason for becoming criminal. It shall also not be read as a generalisation. The paper is not supposed to accuse people with a history of maltreatment of becoming automatically criminal. It rather states that there might be higher risk that people with such a negative background behave aggressively and impulsively, which may in the end be the reason for criminal acts. My findings are moreover restricted by the fact that genetic predisposition may award protective effect meaning that maltreatment has no influence on the behavioural development of the person concerned.

A limitation of this paper may be that it analysed maltreatment as stressor and the effects of stressors in general on the brain. However, it could be seen that different kinds of maltreatment may have different consequences as well, also depending of the sex of the person concerned. Hence, it would be interesting to investigate in detail which kind of maltreatment listed in the CRC accounts for which kind of changes in the different brain areas. Moreover, it could be further researched how either maltreatment or its effect can be prevented by neuroscience and in how far such prevention mechanisms were ethically justifiable or in how far they would be compatible with the principles of the CRC or the UN human rights Conventions in general.

Bibliography

Arnsten, A. F. T. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nature Reviews Neuroscience, 10,* 410-422.

Attachment Synthesis (2012). In *Encyclopedia on Early Child Development*. Retrieved from: *http://www.child-encyclopedia.com/pages/PDF/synthesis-attachment.pdf*

Ayoub, C.C. & Rappolt-Schlichtmann, G. (2007). Child Maltreatment and the Development of Alternative Pathways in Biology and Behaviour. In D. Coch, G. Dawson & K. W. Fischer (Eds.), *Human Behaviour, Learning, and the Developing Brain* (pp. 305- 330). New York: The Guilford Press.

Bremner, D. (2008). The Neurobiology of Trauma and Memory in Children. In M. L. Howe, G. S. Goodman, & D. Cicchetti (Eds.), *Stress, Trauma, and Children's Memory Development* (pp. 11-49). Oxford: Oxford University Press.

Chakraborty, Dr., N. K., Upreti, P. N., & Mishra, A. (2011). Genetic and Environmental Influences on Criminal Behaviour. *Calcutta Law Times, 1,* 36-48.

Committee on the Rights of the Child (2011). *General comment No. 13 – The right of the child to freedom from all forms of violence*. New York: United Nations.

Davidson, R. J., Putnam, K. M., & Larson, C. L. (2000). Dysfunction in the Neural Circuitry of Emotion Regulation – A Possible Prelude to Violence. *Science 289,* pp. 591-594.

Delinquency (n.d.). In *Encyclopaedia Britannica*. Retrieved from: *http://www.britannica.com/EBchecked/topic/156597/delinquency*

English, D. J. & the LONGSCAN Investigators (1997). Modified Classification System (MMCS). Retrieved from: *http://www.iprc.unc.edu/longscan/pages/maltx/index.htm*

Gervai, J. (2009). Environmental and genetic influences on early attachment. *Child and Adolescent Psychiatry and Mental Health, 3.*

Marazziti, D., Baroni, S., Landi, P., Ceresoli, D., & Dell'Osso, L. (2013). The neurobiology of moral sense: facts or hypotheses? *Annals of General Psychiatry 12*.

McCrory, E., De Brito, S. A., & Viding, E. (2010). Research Review: The neurobiology and genetics of maltreatment and adversity. *Journal of Child Psychology and Psychiatry, 51,* 1079-1095.

Mobbs, D., Lau, H. C., Jones, L. O., & Frith, C. D. (2007). *Law, Responsibility, and the Brain*. Retrieved from: *http://papers.ssrn.com/sol3/papers.cfm?abstract_id=982487*

Navalta, C. P., Tomoda, A., & Teicher, M. H. (2008). Trajectories of Neurobehavioral Development. In M. L. Howe, G. S. Goodman, & D. Cicchetti (Eds.), *Stress, Trauma, and Children's Memory Development* (pp. 11-49). Oxford: Oxford University Press.

Rodrigo, C., Rajapakse, S. & Jayanada, G. (2010). The "antisocial" person: an insight in to biology, classification and current evidence on treatment. *Annals of General Psychiatry, 9*.

Schmidt, M. (2010). United Nations. In Daniel Moeckli (Ed.), *International Human Rights Law* (pp. 391-431). Oxford: Oxford University Press.

Siegel, A. (2005). *The Neurobiology of Aggression and Rage*. Boca Raton: CRC Press.

Silver, J. M., Anderson, K. E., & Yudorfsky, S. C. (2003). Violence and the Brain. In E. Todd & Farah, M. J. (Eds.), *Behavioural Neurology & Neuropsychiology* (pp. 755-762). New York: McGraw-Hill.

Smith, C. A. & Stern, S. B. (1997). Deliquency and Antisocial Behaviour: A Review of Family Processes and Intervention Research. *Chicago Journals, 71,* 382-420.

Steckler, T. (2005). The Neuropsychology of Stress. In T. Steckle, N. H. Kalin, & J. M. H. M. Reul (Eds.), *Handbook of Stress and the Brain* (pp. 25-42). Amsterdam: Elsevier.

Teeter Ellison, P. A. & Semrud-Clikeman, M. (2007). *Child Neuropsychology – Assessment and Interventions for Neurodevelopmental Disorders.* New York: Springer Science + Business Media.

Teicher, M. H., Andersen, S. L., Polcari, A., Anderson, C. M., Navalta, C. P., Kim, D. M. (2002). Developmental neurobiology of childhood stress and trauma. *Psychiatric Clinics of North America, 25*, 397-426.

Teicher, M. H., Andersen, S. L., Polcari, A., Anderson, C. M., Navalta, C. P., Kim, D. M. (2003). The neurobiological consequences of early stress and childhood maltreatment. *Neuroscience and Behavioural Reviews, 27*, 33-44.

United Nations (1989). *Convention on the Rights of the Child.* New York: United Nations.

United Nations Treaty Collection (n.d.). Retrieved from: *http://treaties.un.org/Pages/ViewDetails.aspx?src=TREATY&mtdsg_no=IV-11&chapter=4&lang=en*

Wallick, M. M. (1990). Developmental Sources of Stress. In J. D. Noshpitz & R. D. Coddington (Eds.), *Stressors and the Adjustment Disorders* (pp. 189-216). New York: Wiley & Sons.

Willems, J. (2012). It takes a SMECC to raise a child - Meeting Basic Developmental Needs of Newborn Persons. Retrieved from: *http://arno.unimaas.nl/show.cgi?fid=25622*