

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 revolutionize 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 necessarily 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 imaging 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 (VMPFC), 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 as an 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 (Enticott et 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 (Enticott et 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 VMPFC, 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, VPMC, 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*, VPMC 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., & Mamei, 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-interpeduncular 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., & Yehyaw, 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). Crazy 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 Addiction. *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). *Brevik nekter MR-undersøkelse av hjernen*. Retrieved from: <http://www.aftenposten.no/nyheter/iriks/Brevik-nekter-MR-undersokelse-av-hjernen-6676378.html#.UcMwgPn-Fcx>