Bachelor Degree Project

UNIVERSITY
OF SKÖVDE

1977

The Neural Correlates of Emotion and Reason
in Moral Cognition

Bachelor Degree Project in Cognitive Neuroscience
Basic level 22.5 ECTS
Spring term 2019

Ami Blomgren

Supervisor: Stefan Berglund
Examiner: Paavo Pylkkänen

Abstract

Humans are a social species. Automatic affective responses generated by neural systems wired into our brains create a moral intuition or "gut-feeling" of wrong and right that guides our moral judgments. Humans are also an intelligent, problem solving and planning species with neural structures that enable cognitive control and the ability to reason about the costs and benefits of decisions, and moral judgments, not the least. Previous research suggests that moral intuition and moral reasoning operates on different neural networks - a dual process of moral cognition, that sometimes gives rise to an inner conflict in moral judgments. Early lesion studies found correlations between damage to the ventromedial prefrontal cortex (VMPFC) and changes in moral behaviour. This has been further established through brain imaging studies and the suggestion is that VMPFC mediates affective signals from the amygdala in moral decision making and is highly involved in generating the gut-feeling of right and wrong. However, some moral issues are complex and demand higher level processing than intuition, and the dorsolateral prefrontal cortex (DLPFC) seems to be responsible for the rational, cost-benefit reasoning during moral judgments. Further, recent research suggests that during moral judgments, the brain employ neural systems that generates the representation of value, perspective and cognitive control as well as the representation of the mental and emotional states of others. The present thesis aims to investigate prominent and up to date research on the neural correlates of necessary components in moral cognition, and to examine the function of moral intuition versus reason in relation to current complex moral issues. Moral intuition is supposedly an adaption to favour "us" before "them", not to be concerned with large scale cooperation, which may explain why we treat many moral issues with ignorance. Understanding how the moral brain works involve understanding what sort of tasks the neural mechanisms in moral cognition evolved to handle, which may explain why some modern issues are so difficult to solve.

Keywords: Moral cognition, moral intuition, moral reasoning, dual process-theory, VMPFC, DLPFC

Table of Contents

1. Moral Cognition – Intuition and Rationality

Across the world, people live in societies in a broad variety of cultures, norms, religions and traditions that all include moral values. People need to make moral judgments in small or big issues in their everyday life. Morality often seems to be regarded as simple guidelines that a good member of a society should follow. But looking at it close-up, morality is complex, and moral issues tend to evoke emotional engagement because they commonly illustrate conflicts of interests between parties, or within oneself (e.g., "want" versus "should"). Topics such as death-penalty and abortion are constantly under debate, and our different interests are reflected in our political preferences just as well as in our everyday choices. We make moral judgments when deciding what to get for dinner: Should the meat be tasty but affordable? A bit more expensive but ethically produced? Or perhaps no meat at all? We judge others behaviour based on our morals as well. A small action such as tossing a piece of rubbish on the ground can sometimes cause an instinctive emotional reaction and we are quick to judge the action as morally wrong. Hence, our moral cognition covers issues from deeply disturbing actions such as murder and rape to egoistic actions of theft and fraud, and even just careless actions like littering or taking up too much space on the seat of the subway. Nevertheless, the same cognitive system is involved in these judgments, and also in guiding our own behaviour away from making such transgressions.

Morality and ethics have a long going history and have been discussed all over the world at various level of analysis. Great philosophical questions captivate and puzzle people today the same way they did to thinkers such as Plato, Aristotle and Buddha thousands of years ago. Normative and applied ethics are part of a variety of settings in modern days: Schools, hospitals, offices and sport-clubs to name a few. Most places that are dealing with the interacting with individuals have developed a code of ethics (that may or may not be official) that goes along with the moral norms and values of the actor. Morality seems to play a central

part in humanity and social interaction.

There is no global agreement on the definition of morality, but the word itself is derived from Latin and refers to the agreement of manners and behaviours in social groups (Moll, Zahn, de Oliveira-Souza, Krueger, & Grafman, 2005). Human behaviour that is adapted to suit social environments is considered pro-social behaviour, hence, positive for the survival of our species. From an evolutionary perspective, this would imply that the mechanisms that underpin the ability to interpret, understand and produce moral behaviour are inherited from our ancestors (Greene, 2014[a]). What can be traced back to the early development of moral behaviour may be the instinctive, intuitive feeling of right or wrong, commonly referred to as a "gut feeling" when making moral judgements. Whether this moral intuition is the determination in moral judgement has been tested by, for example, examining people's responses when confronted with a hypothetical offensive, yet harmless situation (e.g., eating one's dead pet dog). The responses found were typically emotion-intuitive: "it's just wrong…" rather than rational arguments. This indicates that the affective reaction is a good predictor for moral judgments (Greene, 2015; Haidt, 2001). However, Greene (2014[a]) means that there is a second mode to moral cognition: The ability of controlled reasoning and problem-solving. If the gut-feeling represent efficacy, the controlled cognition stands for flexibility (Greene, 2014[a]).

If humans were to make judgements based on the in-group cooperative function of moral intuition, our modern large-scale societies with its complex structures would not be possible. Our contemporary conditions demand the capability to integrate our controlled cognition and to make judgements based on reason. But what if there was only a controlled cognition? Is it possible to make moral judgements without the affective input, and if so, what function does the intuitive gut-feeling serve when our environmental conditions demand higher-level processing than in-group cooperation? Can there be a purely rational morality? The importance of this issue is more actual than ever, with growing societies, increased travelling and

globalization and development of advanced techniques that are even able to mimic human intelligence. Greene (2014ª) argues that we need to understand the difference between today's moral issues as compared to what our brains evolved to handle, as well as the neural structures of our moral cognition, in order to answer the stated questions.

In the present thesis, the term 'cognition' will be used to describe the mental processes, including emotions, that underpin certain behaviour. When necessary, higher-order processes will be referred to as "controlled cognition" and emotional processes as "affective cognition", "affect" or "emotions" to make a clear distinction between the two.

1.1 Aim, Structure and Method

The aim of this thesis is to investigate and present what the most recent research can reveal about the neural correlates of moral cognition. From the perspective of a dual process in moral judgment that includes emotion processing and controlled cognition, the aim is further to examine the function of both, as separate parts as well as the holistic process.

The theoretical background of moral cognitive neuroscience will be presented, along with the evolutionary perspective of morality and current theories. To reach the aim, a systematic literature search will be conducted where publications primarily from 2015 and up to present date will be considered. The selected material will be further restricted to publications mainly within the fields of cognitive neuroscience and psychology and consideration will be paid to the quality of the publication and journal. This will ensure that the most prominent and up-to-date research, both original studies and reviews in the field of moral cognitive neuroscience will be used as material and a discussion including critical evaluation, ecological validity and future directions will be provided.

2. Theoretical Background

2.1 Historical Background – From Philosophy to Biology

Morality and ethics have been a subject of interest for philosophers all over the world for thousands of years and is just as actual today. Although the word 'morality' includes many aspects and despite various cultural differences, moral theories generally share the same target; to define and describe universal rules to guide righteous behaviour. Still, not all people act according to what could be considered moral rules and human psychological factors would need to be considered to answer the question: Why act morally? (Doris, Stich, Phillips, & Walmsley, 2006).

Moral behaviour may be one of the most central parts in social behaviour and for the study of psychology, it has naturally been of interest for investigation: Do individuals construct the moral rules within oneself, or are they received from the outside? Do we get influenced by authorities or is it the social pressure of the society as a whole that affects us? These have been typical questions that well-known names within various subfields of psychology have tried to answer (Haidt, 2013).

The possibilities of a biological factor to determine the foundations of moral behaviour was suggested by biologist E.O Wilson in his work on sociobiology already in 1975, but by then, moral psychology was mainly the study of moral development and had more in common with philosophy than psychology as a field (Haidt, 2013). Wilson claimed that the underpinning mechanisms for morality were most likely shaped by natural selection: Affective responses created by what he referred to as the 'emotive centres' of the brain. This reconnaissance was put aside to advantage for theories like Kohlberg's approach to moral development where the cognitive component of reasoning was emphasized. However, Wilson's theory of the affective mechanism in moral cognition would be brought back to the spotlight, along with his work on sociobiology which now is known as evolutionary psychology. This was the result of a number

of trends that the study of psychology has gone through, where emotions have been re-considered as the central role in decision making due to their automatic, effortless and fast appearance. Moreover, research in the field of neuroscience made progress due to advancing imaging techniques and made it possible to look inside the neural structures and earlier studies of behavioural changes due to brain damage could be confirmed by brain imaging. (Haidt, 2013).

2.2 Morality in Evolution – an Innate Ability?

Based on the assumption of natural selection and survival of the fittest, the brain, just like any other organ in the body, is a product of evolution. Social behaviour and cooperation are vital in human development and the underlying functions that promote this kind of pro-social behaviour are regarded as evolutionary adaptions (Workman & Reader, 2014).

Moral behaviour does, with its typical properties of concern for others, classify as pro-social behaviour. Moral cognition with its underpinning mechanisms that enable this behaviour is biological adaptions wired into our brain. Altogether, moral cognition seem to be an innate ability in humans (de Oliveira-Souza, Zahn, & Moll, 2015; Yoder, Harenski, Kiehl, & Decety, 2015). This has been further investigated and established in studies on how children at a very young age can manage to determine moral behaviour in social interaction. In a study from 2007, Hamlin, Wynn & Bloom examined preverbal infants' ability to distinguish moral behaviour from other interaction by exposing them to very simple graphic videos of geometric figures with distinct characteristics (colour and shapes), moving around, towards and apart from each other in a landscape-setting. The video had two versions, and in one of the versions the figures were given human-like traits (big eyes). This made the movement of the figures seem like conscious behaviour, which in turn made it seem like one of the figures was helping the second one, while the third was sabotaging. After viewing the video the infants were presented to the same geometric figures in the form of toys which they could choose to pick up.

The group who viewed the version of the figures with eyes (i.e., watched behaviour rather that movement) rejected the unhelpful figure and favoured the helpful and the neutral one. Infants who watched the figures without eyes, however, showed no preference for any specific figure amongst the toys. This revealed that children as young as six months old show a significant preference for moral behaviour over other interaction. This discovery suggest that there is an innate, affective, automatic system in the human brain that, independent of reason (a criteria for reasoning is the cognitive ability of language, which the children in this experiment had not yet reached the age to gain) present to our awareness a feeling of like – dislike upon which we make judgments (Hamlin et al., 2007). However, what moral cognition actually refers to, how and why it evolved in the first place is still a topic of discussion.

The common agreement amongst researchers today is that morality evolved as psychological adaptions to facilitate cooperation (Greene, 2014[a]). These adaptions have been favoured by natural selection as an answer to selfish behaviour and exploitation of reciprocal altruism, not because it is nice, but because they are positive for human survival. Evolution has provided humanity with a reflex-like, automatic emotional response to actions, upon which judgments can be made. Quick, effortless and convincing, the only problem with this system is that it evolved to benefit ones' closest relatives and family, not to handle the complex moral issues of the large societies we live in today which demand higher level processing and flexibility (Greene, 2014[a]).

2.3 The Dual-Process Theory

Moral cognition seems to operate on two distinct and, from each-other, independent systems: The gut-feeling generating, automatic system, and the problem-solving, reasoning system. This is the general hypothesis about the dual-process moral brain, developed by Greene and colleagues (Greene, 2014[a]; Greene, 2014[b]). The automatic system is efficient and effortless while reasoning demands a higher level of processing but represent cognitive flexibility. These

two systems together provide a solution usually described as that of a dual-mode camera, where each setting (i.e., automatic and manual mode) serves a great function, given that the photographer knows when to use which mode (Greene, 2014[a]).

Humans have a variety of "automatic mode settings", information processing and sending signals going on without our conscious awareness: Regulation of the inner bodily states, planning and coordination of movement and control of breathing and heartbeat, to name a few. Also, the production of emotions are automatic processes – even if we can regulate them to some extent, it is difficult to choose to have an emotional experience. There still is no universal agreement of what emotions are, and emotional experiences can vary from very pleasant to very unpleasant with a range in-between. What seems to unify emotions is how they arise as a response to stimuli and create action tendencies in individuals (Helion & Ochsner, 2018).

2.4 The Implications of Cognitive Neuroscience

Greene (2014[b]) argues that contemporary sciences and techniques provide tools to look inside the "black box" housing the human mind. These possibilities enable us to trace down the inner workings that underpin moral behaviour. The field of neuroscience aims to investigate and enhance the understanding of how the neural systems in the brain work to solve problems that we define as moral and to understand the foundations of moral intuition, or gut-feeling not least. (Greene, 2014[b]). Haidt (2012), argues that one of the differences that distinguish solving moral problems from other, non-moral problems is the importance to provide reasons to justify moral judgments to others. This demands that the brain enables a representation of value and cognitive control. Cognitive neuroscience wants to understand the mind as part of the physical brain, and considering the mind as being most certainly involved in moral judgment, the aim of moral neuroscience is to map the neural activity that enables the process of making moral judgments (Greene, 2015).

2.5 Measuring Brain Activity and Detecting the Moral Brain

2.5.1 Lesion studies. One of the first steps in understanding how the brain enables the mind with its mental processes was to study the cognitive and behavioural changes in patients who suffered from brain damage. The famous example of Phineas Gage, who through traumatic brain injury that caused damage to his prefrontal cortex (PFC) and an impairment in his ability to reason and process emotion as a result (Damasio, Grabowski, Frank, Galaburda, & Damasio, 1994) has come to represent somewhat of a landmark for moral neuroscience.

Following the example of Gage, numerous studies have further investigated other patients with similar injuries. A study from the late 1990's examined the long term consequences of early life PFC lesions, and noted a deficiency in their ability to social and moral reasoning as well as impassibility to future consequences of decisions, similar characteristics as those known from antisocial personality disorder (ASPD) such as psychopathy (Anderson, Bechara, Damasio, Tranel, & Damasio, 1999). One of the theories that have come up from studying moral behaviour in psychopaths (a condition typically associated with serious moral deficits) is that the lack of grey matter in the VMPFC may cause a lack of somatic markers (emotions that are translated into unconscious bodily sensations), and in turn impair the individuals emotional guide in social or risky behaviour (Yang et al., 2005).

2.5.2 Technologies in cognitive neuroscience. To get access to the brain as "the black box", the technology of brain imaging is crucial to moral neuroscience. With a small variety of imaging techniques, both structural and functional, functional magnetic resonance imaging (fMRI) is most widely used in the research of the moral brain. By taking advantage of the magnetic properties of the red blood cells, fMRI detects the blood-oxygen dependent level (BOLD). This signal display what area of the brain is being most active or reactive to various stimuli in real-time (Verplaetse, DeSchrijver, & Braeckman, 2014). The use of fMRI has contributed with the possibilities to not only draw conclusions from changed behaviour due to

brain damage – but also to study what is going on within the brain of healthy individuals, performing certain mental tasks. What has come to the surface since these techniques were available is that there seems to be no moral centre in the brain, rather is it a complex set of systems that involve more or less the whole brain, but also that the activation of various system depends on the "value" of the moral dilemma (Abe, Greene, & Kiehl, 2018; Greene, 2015)

Other than techniques to measure brain activity, devices that can interfere with ongoing activity have also been used to detect morality in the brain. The usage of transcranial magnetic stimulation (TMS) can by creating a strong magnetic field interrupt the activity in chosen parts of the brain and act as a temporary lesion. (Young, Camprodon, Hauser, Pascual-Leone, & Saxe, 2010). Further, experiments where the hormone oxytocin, which is known to be involved in pro-social behaviour such as nursing, caring and social bonding. Studies where participants have been given intranasal injections of oxytocin have revealed effects on moral cognition (Verplaetse et al., 2014).

3. Neural Correlates of Moral Cognition

The outcome of mapping the neural correlates of moral cognition depends first of all on how the term 'moral' is defined, and from that position, determine what abilities are necessary units in the production of moral behaviour.

The research in moral cognition and neuroscience has gone from looking for a specific neural circuit that would be active in moral judgment, to understanding just how complex the mechanisms of moral cognition is. Moral judgment demands the integration of both affective and cognitive components in various neural networks, depending on the characteristics of the moral issue (Greene, 2015). de Oliveira-Souza et al. (2015) asserts that moral behaviour is dependent on the understanding of socio-cultural norms and needs of others, along with the motivation to behave in accordance with that understanding. Further, they argue that acquiring knowledge of social norms may be used for selfish purposes and does thus not attribute as a

moral mechanism in itself. They suggest that the neural activity of selfish versus moral motivation should be investigated as a way to understand moral cognition. Behaviour in patients with VMPFC lesions has been described by carers to lack a sense of guilt and interpersonal warmth. They seem to have an intact knowledge of social rules when asked, but still fail to act morally in real life (de Oliveira-Souza et al., 2015).

Another possible cause of impaired ability to produce moral behaviour is the loss of (access) to social knowledge, which has been found in patients who suffer from frontotemporal lobar degeneration (dementia caused by a degree of atrophy to the frontal and anterior temporal lobes), specifically in cases when the right anterior temporal lobe was damaged. These patients were able to describe non-social behaviour (e.g., what it means to be diligent) without difficulty, but showed adversity to describe social behaviour (e.g., acting tactfully). The authors (de Oliveira-Souza et al., 2015) mean that findings as these may boil down to the hypothesis that motivation and knowledge are at least partly dissociable in moral cognition.

Studies using fMRI to research the neuroanatomical basis of moral cognition have detected certain structures which show increased activity to morally relevant compared to morally non-relevant stimuli. However, these structures which were fronto-temporo-subcortical networks, also show increased activity to other tasks within social cognition, such as theory of mind (ToM) (de Oliveira-Souza et al., 2015). This is similarly described by Yoder & Decety (2018) who argue that the same regions that underpin social-decision making play an important role in moral judgment as well. They mean that social-decision making and morality demand coordination of neurocognitive processes such as mental state understanding, stimulus-value association and response selection.

Greene (2015) means that to understand moral cognition, we need to understand the integration of the many neural systems that enable various representation: Value and motivation; the coordination of thought and action in accordance with internal goals; the ability

to imagine future events; the ability to theorize about other people's hidden mental states, to name a few. These systems underpin moral cognition as a whole but appear to have no specifically moral function of their own. He suggests that this can be done by studying the brain of people who continuously make basic moral transgressions (presumably caused by damage or deficit to certain neural structures, e.g., as in psychopathy) as well as the brains of healthy people and what reaction these transgressions cause and reflection of this in neural activity (Greene, 2015).

The early studies of patients with VMPFC damage revealed poor decision making caused by the disability to generate or access emotions to guide adaptive behaviour, which in turn highlights the importance of emotions in moral judgment. The impairment in moral behaviour seen in psychopaths is proposed to arise primarily from dysfunction of amygdala, along with abnormalities in the VMPFC. This is displayed in studies of individuals with psychopathic traits who reveal decreased activity of the amygdala than in healthy individuals, as a response to stimuli that indicate moral violations, harmful actions and emotional distress, as well as reduced VMPFC response to stimuli of the same valence (Greene, 2015).

3.1 Emotional Pressure and Internal Conflict: The Nature of the Problem

   3.1.1 Deontology and utilitarianism. Two branches within normative ethics that are commonly referred to in moral neuroscience is deontology and utilitarianism, two different views of the nature of morality. In the deontological view, the moral value of actions should be based on whether the action itself is good or bad, rather than what consequences it may have. This view differs from utilitarianism, where the moral value of actions should be determined by the outcome of the action. E.g., it is immoral to steal, because stealing is in itself a bad action (deontology). Or, if the action of stealing is "for the greater good" (that is, the benefits are greater than the cost), stealing is moral (utilitarianism) (Sinnott-Armstrong, 2003; Alexander & Moore, 2007). These two branches are important in the discussion of moral behaviour and

cognition because they often represent the interconnection – or gap between reason and emotion. This has been exemplified in numerous studies where the demand for cognitive control or moral gut-feeling varies, and sometimes compete with each other due to the complexity of the issue. (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001).

3.1.2 Personal involvement promotes deontological responding. A classical moral puzzle, known as "the trolley problem" and "the footbridge dilemma" that has been widely used to examine moral judgments and evaluations became the foundation of the dual-process theory. The hypothetical dilemma is to decide whether or not to kill one person to save five others. In the trolley problem, a train is approaching with five people stuck on the track. By pulling a switch, one can make the train change to a second track with one person stuck on it. In the footbridge dilemma, a train is approaching, and five people are stuck on the track. The only way to save them is to push a large person, a stranger, off a footbridge and stop the train. Previous research reveals that people prefer the utilitarian judgment in the trolley problem but deontological in the footbridge dilemma. The main difference between the two versions is the personal approach in the dilemma: Pulling a switch at a distance is in this case considered less personal than pushing someone off a bridge. Hence, personal involvement seems to elicit an internal conflict between reason to choose actions for the greater good and the instinctive resistance toward causing harm with direct physical contact.

In a study from 2001, Greene et al. designed a set of "personal" high conflict dilemmas along with a set of "impersonal" (less conflict) dilemmas (variations to the trolley and footbridge problems) and measured reaction time (RT) and brain activity when participants were confronted with the dilemmas and during their moral evaluation. The results from the study indicate that there are significant differences in neural activity in the process of moral decision making between the two (Greene et al., 2001). Responses to the "impersonal" dilemma signalled stronger activity in the frontoparietal control network (an area known to be

involved in cost-benefit reasoning), including the dorsolateral prefrontal cortex (DLPFC).

Responses to "personal" dilemmas elicited increased activity in areas associated with emotion

such as the amygdala and components in the default mode network (DMN), including the

medial prefrontal cortex (MPFC), medial parietal cortex and the temporoparietal junction

(TPJ). However, the study also revealed that the RT of participants who responded with

utilitarian judgments in the high conflict dilemmas was significantly increased compared to the

majority who gave deontological responses (Greene et al., 2001). This was interpreted as the

interference of the negative emotional response to cause harm, that delayed the cost-benefit

reasoning. These findings suggest that the inner conflict that arises with emotionally

incongruent moral evaluations operates on two separate systems, much like the influence of

automatic processes that influence responses in the Stroop-task (Greene et al., 2001; Greene,

2015). When examining responses from individuals with emotion-related deficits (e.g.,

frontotemporal dementia) to similar dilemmas the outcome was typically utilitarian approval

(Greene, 2015). Examining the responses of psychopaths and individuals with VMPFC damage

who were confronted with the trolley-problem indicated a five times higher chance of

utilitarian judgments as compared to healthy individuals, but with no significantly increased

RT, suggesting there is no inner conflict. They also show weak physiological response when

making such judgments, which can be compared to (healthy) individuals who are more

sensitive to physiological reactions and tend to make fewer utilitarian judgments (Greene,

2015). This gets further support from studies of skin conductance response (SCR) where

healthy controls signal strong SCR to utilitarian judgments compared to VMPFC patients

where no such response is seen when confronted with identical dilemmas (Rowley, Rogish,

Alexander, & Riggs, 2018). Greene (2015) argues that this may arise due to the loss of

information from the amygdala to VMPFC. The amygdala responds to harmful actions and it is

suggested that this information is signalled via the VMPFC, and from there further transferred

as physiological responses that promote deontological judgments. Thus, VMPFC patients do not get input from the amygdala, but the frontoparietal control network is still intact and due to this, utilitarian judgments are favoured. However, studies of psychopaths have shown that they respond normally to threat cues but weak to cues of distress (Greene, 2015).

Some researchers argue that a problem arises when using the trolley-problem as a measurement because the moral dilemma is quite extreme: The utilitarian response demands the violation of highly regarded deontological rules (do not kill), and this may cause a misunderstanding of the connection between executive reasoning and utilitarian judgments. In studies of moral judgments where the dilemma demands more plausible transgressions (e.g., lying to avoid hurting someone's feelings), a majority of the responses are utilitarian (Rowley et al., 2018). However, in this example, also the intuitive feeling was to lie. Further, the perception of morally righteous actions can change, depending on the social context in which they are presented. This is displayed in a study where the trolley-problem is re-designed into a situation where one has to decide whether a doctor should perform surgery on a healthy person, taking his organs for transplant and causing his death, to the benefit of five others who will be saved from organ-failure that would otherwise cause their death. In this situation, most participants rejected the utilitarian option, even though it did not imply performing the violation of the healthy person by own hands (Rowley et al., 2018; Van Bavel, FeldmanHall, & Mende-Siedlecki, 2015).

Another study examined the responses to the trolley-problem and footbridge-dilemma from participants who were under the influence of alcohol, measured by blood alcohol concentration (BAC) in relation to utilitarian judgments. The results of the study revealed a significant correlation between BAC and utilitarian responses to the high conflict dilemma (i.e., acute effects of alcohol led to increased willingness to push someone off a bridge to stop the train). As alcohol intoxication is known to impair higher-order cognitive reasoning simultaneously as

it increases emotional reactivity, the utilitarian responses in this study may contradict Greene's dual-process model. However, acute effects of alcohol are also strongly associated with a reduced aversion towards harming others, and self-reports of disinhibition were positively correlated with utilitarian judgments. Consistent with VMPFC patients, chronic drug/alcohol dependence seems to cause an increased preference for utilitarian judgments (Duke & Bègue, 2015).

3.1.3 In lack of emotion. Several neuroimaging studies reveal greater activity in the lateral regions of PFC in psychopaths than healthy controls as a response to affective stimuli. This may be regarded as either compensatory processes for the loss of paralimbic input or a propensity for a top-down controlled response to salient stimuli. In studies of criminal psychopaths, higher scores on the Psychopathy Checklist-Revised (PCL-R) are predictive for utilitarian judgments (Yoder et al., 2015) as well as reduced activity in the anterior cingulate cortex (ACC) (Abe et al., 2018). The ACC is commonly involved in high conflict cognitive tasks (e.g., the Stroop task), impulse-control and responding to moral dilemmas. Abe et al. (2018) describe the relationship between the ACC and DLPFC as the "bottom-up conflict" being detected in the ACC, and the "top-down" controlled task-response managed by the DLPFC. A study on dishonest behaviour in individuals with psychopathic traits found that those who scored high on the PLC-R revealed a significant correlation between reduced left-ACC activity and reduced reaction time, when engaging in dishonest/immoral behaviour, suggesting that the ACC fail to signal distress cues in the bottom-up process.

Yoder et al. (2015) mean that recent research indicates that psychopaths show differences in the allocation of attention to socially or morally relevant information compared to healthy individuals, which in turn may be regarded as disruptions in both the bottom-up and top-down processing. They examined how individuals with psychopathic traits (range from low to high) responded to morally explicit versus morally implicit (visual) information. They found that

high psychopathy was predictive for decreased connectivity in TPJ and anterior insula (aINS), regions that are associated with emotional awareness. Further, reduced connections between the right amygdala and VMPFC was displayed in the participants with high PCL-R scores in the implicit task, which brings further support to the argument that disruptions in the connectivity between amygdala and prefrontal areas are part of the neurobiology that underpins psychopathy. Psychopathy was also correlated with reduced activity in several areas of the salience network in the explicit task. This network has its core nodes in the dorsal aspects of ACC and aINS and is responsible for the allocation of attention to salient information, but also for the dynamic shifting between cognitive control and DMN (Yoder et al., 2015; Chiong et al., 2013). DMN is a large scale network with core nodes including the posterior cingulate cortex (PCC), MPFC and the angular gyrus, and is suggested to manage a self-referential cognitive system and is known to be highly active when the brain is at wakeful rest, that is, when not engaging in any external tasks (Davey, Pujol, & Harrison, 2016). Studies in moral cognition, however, suggest that the DMN is involved also in moral judgment.

3.1.4 In lack of social knowledge or missing out on social cues. When examining the responses to personal versus impersonal dilemmas, Greene (2015) and colleagues found that personal dilemmas elicited increased activity in the DMN in healthy individuals. This was interpreted as a response to the self-referential properties of DMN, rather than an active reflection of emotional engagement of this system. The responses from psychopathic individuals, on the other hand, reported hypoactivity in DMN when engaging in moral dilemmas. The neural connectivity of DMN includes mnemonic, semantic and limbic structures and is suggested to enable the integration of internal or external salient stimuli with one's current social context (Spreng & Andrews-Hanna, 2015). The core nodes of DMN integrate information from areas of two sub-systems that include dorsomedial prefrontal cortex (DMPFC), TPJ, lateral temporal cortex (LTC), temporal pole, hippocampus (HC),

parahippocampus (PHC), retrosplenial cortex (RSC), posterior inferior parietal lobe (PIPL) and VMPFC (Andrews-Hanna, Smallwood, & Spreng, 2014).

Amongst the areas included in the DMN and in relation to moral cognition, the TPJ is involved in functions that are of special interest. Research indicates that the right TPJ manages the process of morally relevant information through e.g., redirecting responses, involvement in motion perception and, the ability of theorizing about other people's actions, thoughts and intentions, known as theory of mind (ToM) (Yoder et al., 2015). The ability of ToM does not necessarily demand an internal representation of others emotional state; indeed, psychopaths are known to be very successful at mentalizing without empathic input. In moral intuition, however, studies show that affective-sharing and empathy are highly involved and rely on affective components, hence, known as affective-ToM (Rowley et al., 2018).

When evaluating the morality of an action, the perceived intention of the action is essential for the outcome. Healthy individuals mainly judge an incident less harsh knowing it was accidental, rather than intentional (e.g., accidentally hitting another person's car when trying to avoid hitting a cat on the road is less likely to be judged as a moral transgression, than hitting someone's car for fun), and in this process, ToM is highly useful. Studies of brain activity in ToM mentalizing indicate that the neural correlates are, to a high extent, overlapping with those of moral cognition (Bzdok et al., 2012). Greene (2015) means that the TPJ is particularly sensitive to intended harmful actions and that a disruption of the TPJ activity results in outcome-focused judgments (e.g., if no harm is done than the action was not morally wrong, regardless of the intention). Similar tendencies have been noted in patients with damage to VMPFC, an area that is necessary for affective-ToM (Rowley et al., 2018). Moreover, individuals with high-functioning autism, a condition associated with impaired ability for ToM, generally judge accidental harm as wrong, with little consideration of the intention. Individuals with autistic spectrum disorders (ASD) typically express a reduced ability of selective attention

in social contexts as well as a lack of interest in others emotions (Barak & Feng, 2016).

Mentalizing about other people's minds is in real-life a fast and intuitive process that requires attention to meaningful details in social contexts. Further, to understand and be able to produce context-appropriate behaviour, recognition and knowledge of social behaviour is necessary. The storage of context-independent knowledge of social behaviour has been identified as located in the anterior temporal lobe. This has been studied in patients with semantic dementia and further established in fMRI and TMS studies. Further research suggests that the storage of knowledge of social concepts overlap in the left middle anterior temporal area that is active in mentalizing tasks. This could be explained by the access of social conceptual knowledge to interpret the intentions of other people. To produce adaptive behaviour, however, the ventral frontal cortex is suggested to link to associative social knowledge such as the sequence of actions. This could also explain why individuals who suffer from ventral frontal lesions can behave in ways that may be in accordance with their character, but not context appropriate (de Oliveira-Souza et al., 2015).

Research that aims to explore any potential differences in moral cognition between genders has not been able to present any consistent results so far. Some suggest that females are higher than males in levels of empathy, which can find support by comparing the parental investment between sexes (Decety & Yoder, 2016). Other studies suggest that men and women process information differently and that this leads to disparity in moral reasoning. Neither one of these claims have yet been confirmed by additional data and the discussion is contentious (Decety & Yoder, 2016). Nevertheless, research suggests that differences in the neurophysiology that underpins social behaviour do imply differences between gender. The hormone oxytocin, associated with pair-bonding and sensitive parenting, is suggested to have a main target in the amygdala and is involved in the formation of responses to possible threats. Studies where intranasal injections of oxytocin have been distributed to participants have revealed reduced

activity in the amygdala as a response to fearful faces and other harm-inducing stimuli (Bernhard et al., 2016). In a study from 2019, blood samples from 45 male and 45 female participants of the same age were examined and revealed significantly higher levels of oxytocin in the female sample (Marazziti et al., 2019). Further, recent research has found correlations between differences in the oxytocin receptor gene and the probability of utilitarian versus deontological responding to moral dilemmas. The same gene variations have been suggested to correlate to the probability of psychopathy versus ASD (autism spectrum disorder) (Bernhard et al., 2016). Males are overrepresented in both psychopathy and ASD, and a study on acquired sociopathy indicates that men are more likely to develop ASPD (anti-social personality disorder) from brain lesions than females. The male sample revealed correlations between injuries in the right or both cerebral hemispheres and acquired sociopathy, whereas most women with acquired sociopathy suffered from bilateral lesions (de Oliveira-Souza, Paranhos, Moll, & Grafman, 2019).

As mentioned above, many of the neural regions involved in moral cognition are overlapping, and some researchers mean that even though it may be possible to determine structures to be involved mainly as affective or cognitive components, a strong distinction between the two may not be a realistic view of the holistic process (Yoder et al., 2015). Moreover, it is also suggested that the study of responses to hypothetical dilemmas fails to take consideration to the influence of socio-emotional factors and that humans are unable to accurately imagine emotional and motivational states that they are not currently experiencing (Helion & Ochsner, 2018). Individuals confronted with a hypothetical dilemma tend to report that they would behave honestly, not selfish, and in other ways morally appropriate. The same dilemma in more realistic settings, however, usually elicit tendencies to prefer momentarily self-benefits over previously stated intentions to behave altruistic, hence, traditional research in

moral cognition may not capture the psychological and neural processes that underpin authentic moral behaviour (Van Bavel et al., 2015).

4. Motivation to Be Good

Greene (2017) proposes that examining responses to the trolley-problem and footbridge dilemma constitutes a specifically interesting case due to the relatively high consistency across cultures. And indeed, in various cultures and geographic regions, different variations of norms and rules are applied to societies, but still, there are some core norms that seem to be consistent such as 'do not kill' or 'do not steal' (Yoder & Decety, 2018). There is also a strong tendency across various cultures to be willing to punish violations of moral norms; a demand for justice. Recent research suggests that justice motivation, preference for fairness and avoidance of inequity can be seen in early development, but that there also are individual differences in justice sensitivity which in turn predicts altruistic behaviour (Decety & Yoder, 2016). Justice sensitivity is suggested to reflect an individual's perception of injustice. However, if the injustice is perceived from self-interest or interest for others has a significant influence on the production of moral behaviour. Research indicates that self-focused justice sensitivity may lead to reduced pro-social behaviour and distrust of others intentions. Individual's ability to empathize, which includes components of affective sharing, perspective taking and empathic concern, has for long been regarded as a determinant factor in dispositional justice sensitivity with concern for others. Recent research, however, proposes that individual differences in justice sensitivity and moral decision making do not support emotional reactivity as the central component. A study in neurodevelopment examined pro-social sharing behaviour in pre-school children. The results reveal that emotional processes were involved, but only differences in controlled cognitive activity could predict actual sharing behaviour. Nevertheless, the investigation of moral cognition in psychopaths, who lack in empathy and concern for others supports the claim that empathy may play a critical role in justice motivation and moral

behaviour (Decety & Yoder, 2016). Beyond empathy, other emotions that are highly associated with moral motivation are the feelings of guilt, indignation and disgust. Guilt is widely assumed to serve a pro-social function in moral cognition by preventing us from harming others and by engaging us in reparatory behaviour when harm has been done. Moral indignation (anger), on the other hand, typically arises when disapproving of the actions of others, and aids the motivation to "punish wrong-doers". Moral indignation is proposed to arise from the violation of rules that regard autonomy, such as harm or unfairness. Disgust closely resembles feelings of anger but is suggested to answer stronger to actions that violate "purity", e.g., taboo sexual acts (Giner-Sorolla & Chapman, 2017). Studies with fMRI have shown increased activity in the dorsal aspects of the orbitofrontal cortex (OFC) associated with feelings of moral indignation and disgust, as well as when making moral decisions to punish disliked actors (de Oliveira-Souza et al., 2015). Greene (2014[a]) suggests that the motivation to punish noncooperators is a pro-social strategy; our righteous indignation drives us to punish wrong-doers, even when we have nothing in it to gain and even when the cost is higher than the gain. Further, he argues, emotions such as shame and guilt, can be regarded as self-punishing and encourage us to behave honourably. However, in many cases when transgressions are made, we are also willing to regard those actions as mistakes and be forgiving, given that the actor expresses remorse, guilt or embarrassment. Moreover, we tend to express gratitude, as an encouragement to cooperative behaviour. Greene (2014[a]) proposes that being equipped with emotions that seem to specifically aid cooperative behaviour, cooperation is typically intuitive. This has been further established in experiments where participants were confronted with a dilemma and their decision could be either cooperative or selfish. The results of cooperative decisions revealed a correlation to decision time: The faster people decide, the more they decided to cooperate (Greene, 2014[a]).

Beyond the consciously experienced emotions, another highly influential factor for the

motivation of moral adaptive behaviour is assumed to be the unconscious neural activity that generates the response of somatic markers. These responses are suggested to be emotion-related signals, translated to physiological arousal that guide individuals towards safe decision options and prevent risky behaviour and are generated by the VMPFC (Yip, Stein, Côté, & Carney, 2019). Also, the lateral aspects of the ventral PFC (VLPFC) seem to be involved in moral judgments and are suggested to be responsible for regulating amygdala activity and resolving the cognitive dissonance that arises when accepting unfair offers. This is usually implemented by a re-evaluation of the emotional stimulus to make it seem less unsettling, which have revealed increased activity in the VLPFC (Helion & Ochsner, 2018). The brain not only uses distress cues to avoid bad decisions but also reward signals when indulging in cooperative behaviour (Greene, 2015).

5. A Rational Morality

If motivation to produce moral behaviour lies in automatic processes, emotions and cues of distress or signals of reward, do we even need reason? According to Greene (2017), moral intuition is reasonably reliable in everyday social life and guides us towards behaviour which encourages our "in-group" thinking (cooperation is favoured over selfishness). However, the cooperative nature of moral intuition seems to be limited to the in-group. When it comes to the interests of out-groups the gut-feeling is typically not other-focused and caring, but rather selfish and set to out-conquer. Considering the societies we live in today, it seems obvious that we need the ability to reason about moral problems. One might even ask if the moral intuition may be too inflexible and unsophisticated to deal with the complexity of the issues that we need to take a stand in today. Is it possible then, to have a purely rational morality? Greene (2016) describes a modern problem that exemplifies the trolley – and footbridge dilemma that arises with advancing technologies and artificial intelligence (AI): When a driverless car needs to choose, should the decision be for the sake of the many at the cost of the few, even if the few

are the passengers in the car (i.e., if a car with one passenger is headed towards five people on the road and the only option is to either hit the five to save the passenger, or to make a turn straight into a concrete wall and save the five for the cost of the passenger? Moreover, who is suitable to make this decision? This moral concern is highly current, driverless cars are expected to be part of future transportation and the technology is readily available in many new cars on the roads. A study from 2016 investigated the response to such dilemmas amongst the public. The results reveal an overall preference of utilitarian AV's, even when the illustrated dilemma involved themselves and a member of their family as passengers (Bonnefon, Shariff, & Rahwan, 2016). However, the willingness to buy such a vehicle was low, even in the self-protective case scenario (where the car would always favour the safety of the passenger, even if it meant the sacrifice of many), but particularly low in the imagined utilitarian scenario including themselves and a family member (Bonnefon et al., 2016). Greene (2016) argues that whatever ethical principles will be approved in this issue, it will meet critique: Deciding on software programmed to privilege the passenger at all cost will be unacceptable due to the willingness to cause multiple harm, and the utilitarian option and acceptance to sacrifice the passenger will be challenged in the question of who would want to buy a vehicle that is willing to kill you. The purely rational AI can offer promising expectations of AV's ability to reduce traffic-related harm and pollution and overall increased traffic efficacy and safety. Still, it is not possible to program software to be seeking justice or be morally motivated, and this seems to cause discomfort and aversion amongst the public (Greene, 2016).

6. Discussion

The aim of this thesis is to explore research on neural mechanisms that underpin moral cognition, and further to investigate the - from each other independent, but in moral cognition intertwined - components of emotions and higher-level process of reasoning. To achieve the

aim, a systematic literature search that includes recent and relevant research within various fields of psychology and cognitive neuroscience have provided the material for this thesis. With consideration to our contemporary societies with their complex moral issues and advancing techniques that to some extent seem to require an emotion-less (purely rational) morality, special interest was paid to examine the functions of moral intuition.

From the take off for the dual-process theory which provided support for two distinct neural systems, much research has come to focus on the absence of components that underpin each system. This includes studies of various forms of brain damage such as lesions studies, neural atrophy and congenital injuries where the damaged neural structures may explain, or associate to expressions of abnormal moral behaviour (de Oliveira-Souza et al., 2015). Naturally, examining psychopaths in moral neuroscience can reveal a great deal about the absence of morally relevant emotions, how this affects moral judgments and the neuroanatomical structures that enable this effect. The VMPFC has been identified as a vital component in the production of pro-social behaviour since the early lesion studies and this has been consistently confirmed (Greene, 2015; Rowley et al., 2018). The dysfunction of connectivity between the amygdala and VMPFC is commonly regarded as the core neural deviation that causes the callousness towards others observed in psychopaths. However, the amygdala and VMPFC display normal activity as a response to threat cues, indicating the information that gets through is selective. The ACC, involved in cognitive error-detection and core node in the salience network did reflect reduced activity in both Abe and colleagues' study on dishonesty in psychopaths (Abe et al., 2018) and Yoder and colleagues' (Yoder et al., 2015) study on psychopaths' disability to pick up on socially relevant stimuli. The salience network is responsible for the allocation of attention to salient information, and the results from these studies may be interpreted as the reduced activity in the ACC reflects a loss of attention to distress cues, and, at least partly due to this fails to detect the cognitive error that healthy

people display. It seems plausible that not picking up on distress cues may result in impaired moral behaviour since it is typically the salient information that generates an emotional response (e.g., empathic response) which prevent us from behaving in any way that may cause others harm.

As it seems, there are distinguishable differences in neural structure and biological factors that are assumed to underpin moral behaviour (Bernhard et al., 2016; de Oliveira-Souza et al., 2019; Marazziti et al., 2019). Differences in oxytocin-levels between individuals have been observed to correlate to differences in moral behaviour (Bernhard et al., 2016), but still, no studies have confirmed any moral differences between males and females, even with results that support a significant difference in the level of oxytocin (Marazziti et al., 2019). However, as other studies also discuss the possibilities of higher empathy in women to look after and care for their children, the higher levels of oxytocin could possibly reflect a capacity for extended empathic concern in the maternal aspect. How and why this does not affect moral decision-making, or at least not been observed as to do so (Decety & Yoder, 2016), however, is still intriguing, particularly as empathy has been named one of the more essential emotions in moral behaviour. Nevertheless, the role of empathy as a predictor for altruistic behaviour has also been questioned recently after revealing how only controlled cognitive activity would predict actual sharing behaviour in children (Decety & Yoder, 2016). Whether this finding can be applied to larger contexts is still unclear, and it seems reasonable to question whether it reflects a certain stage in the moral development in children or if the same pattern can be observed in a wider age range.

Another neural network that frequently shows up in the moral neuroscience literature is the DMN, and the fact that there are notable differences in the activity of the DMN in healthy people compared to psychopaths when engaging in moral reasoning does suggest that the network may have an influence on moral judgments (Greene, 2015). The hypoactivity of the

DMN observed in psychopaths when actively responding to high-conflict moral dilemma may reflect how they treat such issues as any form of problem-solving. Increased activity in the cost-benefit reasoning components of DLPFC supports that the engagement rather lies in controlled processes, similar to solving mathematical problems (which in general do not require any self-referential input from the DMN), than feelings of personal involvement and reflection. The DMN is generally most active when not engaging in any external tasks, and the case of healthy individuals who exhibit increased activity of the DMN in high-conflict moral judgement (Greene, 2015) may support the idea that high-conflict dilemmas cause feelings of personal involvement which in turn activate the self-referential DMN, rather than employ solely the control network. Treating high conflict moral dilemmas such as the footbridge-case similar to a mathematical problem where five is better than one indicates a high sense of rationality, but imaging to push someone off a bridge without doubt or remorse is not compatible with moral behaviour. Healthy people generally do not consider human lives as bricks in an abacus, the moral intuition signals strongly against causing anyone harm, even if it is the rationally appropriate decision. Personal, high-conflict dilemmas may demand a sort of "what does this mean to me" reasoning, rather than "what is the appropriate thing to do", hence increased activity in the DMN.

Examining impaired social and moral cognition and behaviour that correlates to brain damage help us understand the function of each specific neural component that is necessary to produce adaptive behaviour. The solid research on psychopaths' neural structure and poor moral behaviour do explain a lot about the function of integrating emotions in decision making. Still, not being a psychopath does not equal being able to produce moral behaviour. As observed in research on neural atrophy in patients with various forms of dementia that affects the storage of social knowledge (de Oliveira-Souza et al., 2015), it is not possible to produce appropriate behaviour without understanding one's surroundings. Likewise, as displayed in

individuals with ASD, impaired ability to understand others intentions also causes deficits to moral cognition (Greene, 2015). It seems plausible to assume that difficulties of mentalizing may cause additional distress to these individuals, as they are not able to distinguish accidental harm from intended harm, hence, emotions cut from the context do not serve a specific moral function. The observation of dysfunctional moral behaviour as a result of impaired knowledge or impaired emotional responding gives an insight to the nature of our moral brains; remove one of the components and what you get is not morality minus that specific part. E.g., without the morally relevant emotions, a person does not produce a purely rational moral behaviour, but rather there is a loss of motivation to cooperate, as observed in psychopaths. This explains how the systems are distinguishable from each other, but also intertwined. The distinction between intuition and reason observed in the experiment using the trolley-dilemma (Greene et al., 2001) was criticised by Rowley et al. (2018) for using such high conflict dilemma that would not give a fair representation of reasoning and utilitarian judgment. The results from their experiment on moral judgment revealed that most people are willing to lie to a friend to avoid hurting her feelings. This does display that intuition can be utilitarian and that people are willing to break deontological rules (such as "do not lie") for the greater good even with personal involvement. However, in this case, the moral transgression of lying is to avoid hurting someone, and the moral intuition was congruent with the utilitarian reasoning, hence, there was no inner conflict. The results from this may be interpreted as additional support to the idea that moral intuition is signalling strongly against any type of behaviour that involves hurting someone, and particularly someone close. Duke & Bègue (2015) study also questioned the accuracy of the dual-process theory by claiming that the acute effects of alcohol make people more prone to make utilitarian judgements. While it is true that alcohol increase emotional reactivity (which according to the dual-process should cause a deontological preference), it is not necessary the emotional elements that work inhibitory towards harming others that increase. The authors do,

however, also discuss other effects of alcohol that may affect the responses, and it may seem more plausible that bluntness towards harming others explain the willingness to push someone off a bridge, rather than a nonspecific increased emotional reactivity.

A major challenge in moral neuroscience research is to provide high ecological validity in the experiments. As described by Helion & Ochsner (2018) and Van Bavel et al. (2015), experiments that include hypothetical dilemmas where participants are asked to take a stand does not always reflect behavioural reality. When mapping neural correlates to moral judgment in laboratory settings, using brain imaging techniques, the possibilities to study actual behaviour is very restricted. Further, to set up a real experiment of e.g., the trolley problem would meet further resistance due to the quest which is (somewhat ironically in the context) highly immoral to put participants through. The tremendous pressure that the trolley problem would imply in real life would most likely reveal more about the moral brain than is achievable today. Put together, this to some extent may indicate that much of the research on the topic today measures brain activity that reflects people's imaginary moral behaviour. Nevertheless, repeatedly conducted studies have provided consistent results regarding the correlation between short reaction time and preference for decisions that aid cooperation (Greene, 2014[a]), hence, support the assumption of an automatic intuitive system. This same system is supposedly what can be observed in the young children in Hamlin et al., (2007) experiment, a preference for altruistic behaviour and justice. A variety of emotions seem to fuel moral intuition and create a strong motivation to behave accordingly. Nevertheless, people commonly make questionable moral choices. Greene (2014[a]) suggests that this ingenious intuitive system developed, as one side of the coin, to ensure cooperation within groups back in history when our societies were significantly smaller than they are now. The other side of the coin is the self-focused, competing intuitive reaction towards out group-members (Greene, 2014[a]), which still is highly actual and may explain much about various conflicts of interests worldwide, and to some extent

help us understand how they arose and why it is so difficult to solve them. Moral intuition is proposed to be reliable in everyday social life (Greene, 2016), but the neural circuits that are engaged in the process where tailormade to handle issues far ago in history, not of the kind that we meet today. This seems apparent in the choices we meet frequently which imply consequences for various parts: The clothes we buy, the food we consume, how we travel (to name a few) and more generally, where we spend our money. These choices, to name a few, all include a moral evaluation even if we are not always aware of it. In the situation where our cost-benefit reasoning does not get any distress cues from the possible harm of others, it is plausible that it dominates our decision. It seems reasonable to assume that the problem that arises with this, is comparable to what was observed in Greene et al. (2001) study on emotional engagement in moral decision-making: even with knowledge of the circumstances that underly our decisions, we need to feel personally involved to get the cues of distress and somatic markers that keep us from making choices that cause others harm. This explains how it is possible to, for example, make the decision to buy a piece of clothing that is so cheap that it seems inevitable that someone along the line of producing the piece did not get paid (knowledge of circumstances). But not seeing anyone suffer and not appreciating oneself as being directly responsible for this situation of harm (not feeling personally involved) means an absence of inner conflict, hence, nothing stops us from buying the piece. To speculate, the same absence of inner conflict may also to some extent explain the brutal nature of online bullying; it seems that online digital social interaction may not activate an affective response because no salient information that we are programmed to react to is available behind a screen. Further on, the inherent endeavour to avoid the unpleasant feeling of cognitive dissonance has been noticed in the regulation of signals from the amygdala, suggested to be handled by the VLPFC (Helion & Ochsner, 2018). Cognitive dissonance typically occurs when our behaviour and values are incongruent and make us look for reasons that justify or support our actions, rather than

information that would make us more uncomfortable and possibly motivate us to change behaviour.

The limitations of our tribalistic brain certainly meet many challenges in our contemporary societies, and the study on attitudes towards autonomous vehicles (Bonnefon et al., 2016) illustrates an example of how a moral puzzle commonly used as a hypothetical issue in studies on moral psychology, has become a present problem. And judging from the results, the public seems highly ambivalent towards the question. It may be reasonable to assume that part of the issue lies in whether humans can relate to artificial intelligence as a moral agent. From what is suggested about moral cognition, it seems that for us to appreciate an action as morally righteous or unjustifiable, we need to understand the intention behind it and be able to relate to that intention (as in ToM and affective-ToM). If the action is not appreciated as morally righteous, we can still be willing to forgive the wrong-doer if he shows remorse, guilt or embarrassment. What does this mean in relation to the AV's then? How do we understand the intentions behind an action where the agent has no motivation of its own to behave morally? The study by Bonnefon et al. (2016) indicates that the public can relate to the rationality of utilitarian decisions, but still feel uncomfortable to let a vehicle be responsible for the safety of oneself and family members without any emotional attachment. This suggests that our moral brain which seems to consist of a complex system of affective signals and controlled reasoning have difficulties to relate to emotionless actions, which further raise the question if we would accept an exclusively rational morality as moral.

7. Conclusion

The neural correlates and networks involved in moral cognition seem to cover both affective and controlled information-processing. This reflects the complexity of moral behaviour that has been a large part of human success and development. The neural correlates of moral cognition are widely overlapping, which makes it difficult to determine whether they

are components of an affective or controlled network. The networks are also highly intertwined and work both as triggers and inhibitors on each other.

The key role of moral intuition seems to be its function as motivation to produce pro-social, unselfish behaviour. Supposedly this system developed as a pro-social adaption to aid cooperation within group and its original function also imply certain limitations when it comes to modern moral issues which demand reasoning. To manage large-scale cooperation, share resources, solve and avoid conflicts the controlled processes involved in problem-solving, planning and rationality are necessary. However, even with every tool in the problem-solving tool-box, we also need to allocate the motivation to use them. Understanding the neural structures and functions that underpin moral behaviour helps us to understand the capacities and limitations of our moral cognition.

Multicultural societies, severely unbalanced distribution of resources worldwide, pollution, climate change and global conflicts of interest are a few of the major challenges that we face today. Research in moral cognitive neuroscience entails implications at a societal level: Understanding the neural structures that underpin our cognitive conditions may help us understand why it is so difficult to find solutions in some moral issues, and how we should best relate to these issues to prevent biased judgments.

References

Abe, N., Greene, J. D., & Kiehl, K. A. (2018). Reduced engagement of the anterior cingulate cortex in the dishonest decision-making of incarcerated psychopaths. *Social cognitive and affective neuroscience*, *13*(8), 797-807.

Alexander, L., & Moore, M. (2007). Deontological ethics. Retrieved 2019 April 30 from https://stanford.library.sydney.edu.au/archives/win2015/entries/ethics-deontological/

Anderson, S. W., Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1999). Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nature neuroscience*, *2*(11), 1032-1037

Andrews-Hanna, J. R., Smallwood, J., & Spreng, R. N. (2014). The default network and self-generated thought: component processes, dynamic control, and clinical relevance. *Annals of the New York Academy of Sciences*, *1316*(1), 29-52.

Barak, B., & Feng, G. (2016). Neurobiology of social behavior abnormalities in autism and Williams syndrome. *Nature neuroscience*, *19*(5), 647.

Bernhard, R. M., Chaponis, J., Siburian, R., Gallagher, P., Ransohoff, K., Wikler, D., ... & Greene, J. D. (2016). Variation in the oxytocin receptor gene (OXTR) is associated with differences in moral judgment. *Social cognitive and affective neuroscience*, *11*(12), 1872-1881.

Bonnefon, J. F., Shariff, A., & Rahwan, I. (2016). The social dilemma of autonomous vehicles. *Science*, *352*(6293), 1573-1576.

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 Structure and Function*, *217*(4), 783-796.

Chiong, W., Wilson, S. M., D'Esposito, M., Kayser, A. S., Grossman, S. N., Poorzand, P., ... & Rankin, K. P. (2013). The salience network causally influences default mode network activity during moral reasoning. *Brain*, *136*(6), 1929-1941.

Damasio, H., Grabowski, T., Frank, R., Galaburda, A. M., & Damasio, A. R. (1994). The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science*, *264*(5162), 1102-1105.

Davey, C. G., Pujol, J., & Harrison, B. J. (2016). Mapping the self in the brain's default mode network. *Neuroimage*, *132*, 390-397.

Decety, J., & Yoder, K. J. (2016). Empathy and motivation for justice: Cognitive empathy and concern, but not emotional empathy, predict sensitivity to injustice for others. *Social neuroscience*, *11*(1), 1-14.

de Oliveira-Souza, R., Zahn, R., & Moll, J. (2015). 11 Neural Correlates of Human Morality: An Overview. In Wheatley., T & Decety. J (Eds), *The moral brain: a multidisciplinary perspective*, (p.183-195). Cambridge: MIT press

de Oliveira-Souza, R., Paranhos, T., Moll, J., & Grafman, J. (2019). Gender and Hemispheric Asymmetries in Acquired Sociopathy. *Frontiers in psychology*, *10*.

Doris, J., Stich, S., Phillips, J., & Walmsley, L. (2006). Moral psychology: Empirical approaches. Retrieved 2019 April 30 from https://plato.stanford.edu/entries/moral-psych-emp/

Duke, A. A., & Bègue, L. (2015). The drunk utilitarian: Blood alcohol concentration predicts utilitarian responses in moral dilemmas. *Cognition*, *134*, 121-12

Giner-Sorolla, R., & Chapman, H. A. (2017). Beyond purity: Moral disgust toward bad character. *Psychological Science*, *28*(1), 80-91.

Greene, J. (2014[a]). *Moral tribes: Emotion, reason, and the gap between us and them*. New York: Penguin Books.

Greene, J. D. (2014[b]). Beyond point-and-shoot morality: Why cognitive (neuro) science matters for ethics. *The Law & Ethics of Human Rights*, *9*(2), 141-172.

Greene, J. D. (2015). 12 The Cognitive Neuroscience of Moral Judgment and Decision Making. In Wheatley, T., & Decety, J (Eds.), *The moral brain: A multidisciplinary perspective* (p. 1013-1023) Cambridge: MIT press

Greene, J. D. (2016). Our driverless dilemma. *Science*, *352*(6293), 1514-1515.

Greene, J. D. (2017). The rat-a-gorical imperative: Moral intuition and the limits of affective learning. *Cognition*, *167*, 66-77.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*(5537), 2105-2108.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological review*, *108*(4), 814-834

Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. London: Penguin Books

Haidt, J. (2013). Moral psychology for the twenty-first century. *Journal of Moral Education*, *42*(3), 281-297.

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*(7169), 557-560

Helion, C., & Ochsner, K. N. (2018). The role of emotion regulation in moral judgment. *Neuroethics*, *11*(3), 297-308.

Marazziti, D., Baroni, S., Mucci, F., Piccinni, A., Moroni, I., Giannaccini, G., ... & Dell'Osso, L. (2019). Sex-Related Differences in Plasma Oxytocin Levels in Humans. *Clinical Practice and Epidemiology in Mental Health: CP & EMH*, *15*, 58-63.

Moll, J., Zahn, R., de Oliveira-Souza, R., Krueger, F., & Grafman, J. (2005). The neural basis of human moral cognition. *Nature reviews neuroscience*, *6*(10), 799-809.

Rowley, D. A., Rogish, M., Alexander, T., & Riggs, K. J. (2018). Counter-intuitive moral judgement following traumatic brain injury. *Journal of neuropsychology*, *12*(2), 200-215.

Sinnott-Armstrong, W. (2003). Consequentialism. Retrieved 2019 April 30 from https://stanford.library.sydney.edu.au/archives/win2015/entries/consequentialism/

Spreng, R. N., & Andrews-Hanna, J. R. (2015). The default network and social cognition. *Brain mapping: An encyclopedic reference*, *1316*, 165-169.

Van Bavel, J. J., FeldmanHall, O., & Mende-Siedlecki, P. (2015). The neuroscience of moral cognition: From dual processes to dynamic systems. *Current Opinion in Psychology*, *6*, 167-172.

Verplaetse, J., DeSchrijver, J., & Braeckman, J. (2014). Introduction. In Verplaetse, J., DeSchrijver, J., Vanneste, S., & Braeckman, J (Eds.), *Moral Brain*. Dordrecht: Springer.

Workman, L., & Reader, W. (2014). *Evolutionary psychology* (3. ed.). New York: Cambridge University Press.

Yang, Y., Raine, A., Lencz, T., Bihrle, S., LaCasse, L., & Colletti, P. (2005). Volume reduction in prefrontal gray matter in unsuccessful criminal psychopaths. *Biological psychiatry*, *57*(10), 1103-1108.

Yip, J. A., Stein, D. H., Côté, S., & Carney, D. R. (2019). Follow your gut? Emotional intelligence moderates the association between physiologically measured somatic markers and risk-taking. *Emotion*.

Yoder, K. J., Harenski, C., Kiehl, K. A., & Decety, J. (2015). Neural networks underlying implicit and explicit moral evaluations in psychopathy. *Translational Psychiatry*, *5*(8), e625.

Yoder, K. J., & Decety, J. (2018). The neuroscience of morality and social decision-making. *Psychology, Crime & Law*, *24*(3), 279-295.

Young, L., Camprodon, J. A., Hauser, M., Pascual-Leone, A., & Saxe, R. (2010). Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proceedings of the National Academy of Sciences*, *107*(15), 6753-6758