THE IMPETUOUS VOICE OF REASON

Emotion versus reason in moral decision-making

Bachelor Degree Project in Cognitive Neuroscience
Basic level 22.5 ECTS
Spring term 2018

Erik Svenning

Supervisor: Oskar MacGregor
Examiner: Stefan Berglund
Abstract

This is a review of what the currently dominant theories of moral decision-making are and where they derive from. While the introduction serves as a common ground to explain what moral decision-making is, the earlier parts of the thesis describe older traditionalist theories within the field, theories of emotional decision-making, in the form of the somatic marker hypothesis, as well as critique of the older traditionalist theories through the social intuitionist model. Both of these two theories are explained as the foundation of the current theories of moral decision-making and after establishing a clear basis on what the currently dominant theories of moral decision-making are built on, said theories are introduced in the form of the dual-processing theory and the event-feature-emotion complexes which are thoroughly reviewed, explained in detail and serves as the core of the text. This is afterward followed by criticism as well as arguments in favor of both theories as well as criticisms from other researchers who disagree with the methodology which the theories of moral decision-making are conducted on. The essay reviews the current state of the field of moral decision-making which has been split up into two different approaches, the locationist approach and the constructionist approach. The essay concludes that there are terms which needs to be clarified in order for the field to move forward and studies to be made regarding the social implications of gut reactions in moral decision-making.

Keywords: Moral decision-making, emotional decision-making, somatic marker hypothesis, social intuitionist model, the dual process theory, the event-feature-emotion complexes
# Table of Contents

1. Introduction .......................................................................................................................... 4
2. Moral reasoning in the traditional sense .................................................................................. 5
   2.1 The traditional and rationalistic theories of morality .................................................... 5
3. Somatic markers and social intuitions ..................................................................................... 7
   3.1 Somatic states and inducers .......................................................................................... 8
   3.2 The importance of emotions ......................................................................................... 9
   3.3 The social intuitionist model ......................................................................................... 12
4. A new and unusual perspective on moral decision-making ................................................... 14
   4.1 Two moral dilemmas .................................................................................................... 14
   4.2 A hypothesis forms ..................................................................................................... 16
   4.3 The scientific experiments of moral decision-making ................................................ 17
   4.4 The moral emotions .................................................................................................... 19
5. The neural correlates of moral emotions .............................................................................. 20
   5.1 Passive visual task ..................................................................................................... 21
   5.2 Visual sentence verification task ................................................................................ 22
   5.3 The event-feature-emotion complexes ....................................................................... 24
6. The dual-processing brain ..................................................................................................... 27
   6.1 Cognitive functions of the dual-processing brain ........................................................ 30
   6.2 Emotional processes of the dual-processing brain ..................................................... 31
   6.3 Utilitarian, or just careless? ......................................................................................... 33
7. Conflicting opinions .............................................................................................................. 35
   7.1 Two different approaches to decision-making ............................................................ 36
8. Discussion ................................................................................................................................ 38
9. Conclusion ............................................................................................................................. 40
References ................................................................................................................................... 42
1. Introduction

Decisions. Our lives contain millions of them and we make some of them in the blink of an eye. But what kind of process lies behind the curtain when making a decision? How are we able to make so many decisions so effortlessly, and how come that while some of the decisions we make are so easy and effortlessly made, others decisions seem impossible to decide? This essay reviews how the brain computes our moral decisions, especially the tougher decisions that concerns morals, values and have no right or wrong answers. We will also be taking a look at what the neural correlates of moral decision-making looks like when making difficult, moral decisions. In order to review this topic, we will be taking a closer look at older theories of moral decision-making. These older theories are not the main focus of this paper but are reviewed for an increased understanding of the current theories of moral decision-making. While the main concern is to review the current theories of moral decision-making, this essay will also bring some insight on newer perspectives regarding said theories and a critical perspective on how these current theories have evolved into the different approaches of moral decision-making that they represent today.

Current theories of moral decision-making suggest that we are driven by our emotions when making moral decisions (Bechara & Damasio, 2005; Haidt, 2001), but are emotions really the deciding factor when we decide what we want for ourselves? And more importantly, do we want emotions to be the deciding factor when we already have a set of values that we follow? Wouldn't you rather feel like you know what is right and wrong without the shadow of a doubt because it is logical rather than having a questioning feeling of wrongness even though you know that you are right?

Throughout our history, it has been commonly thought in scientific communities that moral decision-making is based on calculated reasoning and rationality, but more recent
research suggest that our emotions come to play more than we would like them to when making moral decisions and makes what rationally seem right feel wrong and unethical, even if we cognitively `know´ that it is right and fair. It seems like no matter how logical, reasonable and intellectual we become as we evolve with our society's law and order, all of us have that personal, egocentric and emotional `hunch´ that makes us question our own values and makes what seems logically right feel so wrong.

2. Moral reasoning in the traditional sense

If you were to ask a researcher 50 years ago how moral decision-making works the answer would be more tied to rationality and reasoning than one would think. Before we take a look at what recent research shows regarding moral decision-making it is important to have knowledge of these older more traditional theories of moral decision-making and how these theories has given rise to the field of moral decision-making as we know it today.

One could say that the field of moral decision-making as we know it today has risen because older theories of moral decision-making have reached a crossroads which questions the legitimacy of these older theories, not least based on more recent research regarding the neural correlates of moral decision-making. But what are these older theories? And what do they suggest?

2.1 The traditional and rationalistic theories of morality

As early as in ancient Greece, Plato devised a theory of a divided self where reason, which has its seat of power in the head, rules over the passions and desires which roam wildly in the chest and stomach (Haidt, 2003). A similar division between reason and emotion is shared between many different religions in the world. To have a voice of reason, to be moral and to be noble is regarded as higher motives and will be awarded by ascension to
heaven while emotions such as greed and lust are regarded as lower motives, in the forms of instincts which are to be considered animalistic (Haidt, 2003).

Kohlberg and Hersh (1977) and Piaget (1997) argued for an *a priori* rationalistic view of moral decision-making. Their view suggests that we can predict actions prior to the fact that they have been committed. Kohlberg and Hersh (1977) and Piaget (1997) undertook research regarding the moral development of children and argued that moral behavior is shown through rationality. By having parents teach their children morality, through moral constraints, the children learned how to act rationally and be moral individuals. As the children grew up and became adults they learned to follow their moral constraints because it felt right to do so, and not because their parents told them to. Kohlberg and Hersh (1977) and Piaget (1997) explained this feeling of following your moral constraints as a feeling of obligation, or a sense of duty, to fulfill your moral constraints. This made the children who were raised with moral constraints to become rational and moral thinkers. Kohlberg and Hersh (1977) and Piaget (1997) suggested that morality is of a rational nature that can be trained, and not something that is emotional and impulsive. Therefore, by logical deduction, making moral decision-making something that is rational and reasonable and therefore calculable and predictable. Kohlberg and Hersh (1977) argued that emotions do come into play in moral actions, but to a smaller extent than rationality since rationality is learned through a childhood of moral constraints.

The view that moral decisions are driven by a rational and calculated mind have been the dominant norm within the scientific community for a very long time. However, more recent theories suggest that humans aren't the rational, predictable agents they were originally thought to be (Damasio, 1994; Bechara, & Damasio, 2005). Haidt (2001) and Bechara and Damasio (2005) argue that emotions have a bigger impact in decision-making than Kohlberg and Hersh (1977), Piaget (1997) and older philosophers, such as Plato, once thought. One of
the most important hypotheses (within cognitive neuroscience) to suggest that emotions are a part of our decision-making is called the *somatic marker hypothesis*.

### 3. Somatic markers and social intuitions

The somatic marker hypothesis is a theory about how humans make decisions based on their emotions. The hypothesis was founded in 1994 when four researchers made a study about how damage to the human prefrontal cortex leads to a defect in decision-making (Bechara, Damasio, Damasio & Anderson, 1994). The word *somatic* originates from the Greek word *soma* (which means body) and a "somatic marker" could be explained as a body-related response to a stimulus that marks an emotion (Bechara & Damasio, 2005).

The hypothesis itself suggests that decision-making isn't as rational as explained by older theories, such as the *expected utility theory*, and that emotional mechanisms in the brain have a significant impact on the decisions we make as humans (Bechara & Damasio, 2005). The expected utility theory suggests that humans can be regarded as predictable, rational agents and that their decisions can be calculated (Von Neumann & Morgenstern, 2007). Bechara and Damasio (2005) argue that without the emotional mechanisms, we only have access to a reasoned cost-benefit analysis. This makes it impossible for us to make a decision based on what we want since our emotions need to indicate for us what is wanted and needed (Bechara & Damasio, 2005). Without our emotions, we would stand and weigh the pros and cons of different options until we were exhausted by all the options and wound up not choosing anything, or picking something out of frustration for being unable to understand what we want or need. (Bechara & Damasio, 2005).

But before we take a look at what evidence there is to support the hypothesis, and what implications it has we need to define what these emotional mechanisms are and how they enable us to feel emotions.
3.1 Somatic states and inducers

Bechara and Damasio (2005) suggest that we feel emotions as a reaction to stimuli that are emotionally-competent (capable of giving rise to an emotion) in our environment. These reactions could be explained through three different neural mechanisms of the brain: the central nervous system release of certain neurotransmitters, the brains reaction to a physical or physiological change in our body through the somatosensory cortex, and a noticeable change to the transmissions signals that the body sends to the somatosensory regions of the brain regarding its current state (damaged, tense, stretched, etc). These neural mechanisms together make the body and brain able to enact responses from stimuli in the form of emotions (Bechara & Damasio, 2005).

Now that we have reached an understanding of what gives rise to an emotion neurologically and physically, how are these emotions triggered in our everyday lives and how do they affect our decision-making? Bechara and Damasio (2005) argues that the emotionally-competent stimulus is the exclusive trigger of emotional responses and they call these emotionally-competent stimuli primary and secondary inducers.

Primary inducers are stimuli which are encountered in our immediate environment. They induce somatic states which could be seen as either pleasant and full of joy or unpleasant and fearfull (like seeing a dog across the street while being afraid of dogs). Primary inducers can also be experienced through learned knowledge of concepts, hearing about a personal accomplishment, or hearing about a personal failure. An example would be hearing that a relative has been in an accident. You might not be at the place of the accident as you hear of it but it still induces the same feelings as if you would have been there (Bechara & Damasio, 2005).
Secondary inducers are thoughts and memories of primary inducers that evoke the same emotion as the primary inducer did when it was encountered. However, the emotion is only felt when the primary inducer is actively thought of (Bechara & Damasio, 2005). An example would be thinking about the event where you encountered the dog across the street or when thinking about the time you heard of the relative who had been in an accident. Bechara and Damasio (2005) also suggest that secondary inducers could induce an emotion by imagining primary inducers happening to you, even though they haven't actually happened.

Evidence shows that the amygdala is a vital neural system for inducing emotions from primary inducers while the ventromedial prefrontal cortex (VMPFC) is a neural system that is necessary for inducing emotions from secondary inducers (Bechara & Damasio, 2005). It is important to note that since primary and secondary inducers rely on different neural systems, they can also be active at the same time. This simultaneous activation of the neural systems can create some problems in interpreting the results of experiments since it might be hard to know which of the inducers caused said results if they were both active at the same time.

3.2 The importance of emotions

The most common way of showing evidence for the function of emotions in decision-making is with a test called the Iowa Gambling Task (ITG).

In the ITG, participants are first given a finite amount of money to play with. The participants are then faced with a choice of drawing cards from one out of four different decks (A, B, C, and D) with the option of switching between which deck to draw from at any time (Bechara et al., 1994; Bechara, Tranel, & Damasio, 2000; Damasio, 1996). Drawing cards from decks A and B result in a net loss in the end, even though drawing from decks A and B proves to be advantageous in the beginning because of the high payout. Drawing cards
from deck C and D result in a net gain, since the payout increases over time, but is disadvantageous in the beginning because of the low initial payout (Bechara et al., 2000).

Results from studies using the IGT show that both participants with bilateral damage to the VMPFC and participants with bilateral damage to the amygdala tends to pick cards from decks A and B without considering the consequences of the future net loss as the penalty of deck A and B increases, while controls tend to draw cards exclusively from decks C and D after realizing the increase in penalty from decks A and B (Bechara & Damasio, 2005).

Results from other studies using IGT shows that we make decisions on which deck to draw from before we are even conscious of it. This is usually measured with an anticipatory skin conductance response (SCR; Bechara, Damasio, Tranel, & Damasio, 1997). The anticipatory SCR measures autonomic nervous system control which indicates emotional responses from the body as control participants are pondering on which card to select during the IGT (Bechara & Damasio, 2005). The anticipatory SCR shows that control participants unconsciously receive emotional signals from their bodies that bias decisions in an advantageous direction, which in the IGT is to pick a more profitable deck when drawing from an unprofitable deck. This emotional signal is split into three different periods: pre-hunch period, hunch period and conceptual period (Bechara et al., 1997). The pre-hunch period shows a significant increase in the anticipatory SCR compared to the normal baseline state (or pre-punishment period as Bechara et al., 1997 refer to it as). This pre-hunch period was however only discovered in control participants during IGT, suggesting that the VMPFC is involved in the creation of the pre-hunch. Bechara et al. (1997) could also see that the pre-hunch period started after about 10 cards (on average) had been picked by the control participants. After the control participants had been picking about 50 cards, the hunch period began (Bechara et al., 1997). The important distinctions between the pre-hunch period and the
hunch period is that in the hunch period the SCR seemed to subside when picking cards from good decks (C and D) compared to bad decks (A and B) and that the control participants started to report a feeling of "liking or "disliking" certain decks and could start to guess which decks were safe or risky, but was unable to back that feeling up with anything more than a hunch (Bechara et al., 1997). Just like with the pre-hunch, patients with VMPFC damage was unable to generate a hunch period. The last stage, the conceptual period, was seen after picking about 80 cards where the control participants were able to accurately articulate the nature of the task as well as being able to tell the good decks from the bad decks with an explanation of why they were good or bad (Bechara et al., 1997).

Bechara and Damasio (2005) argue, with the evidence from the studies made with IGT, that their somatic marker model explains the importance of emotional inputs in decision-making and that the unconscious pre-hunch and hunch periods are evidence which shows that these emotional inputs needs to be accounted for. Bechara and Damasio (2005) also argue in favor of the prospect theory made by Tversky and Kahneman. The prospect theory suggests that individuals are irrational decision makers and when faced with situations where they have high risks of losing their money, they are willing to take bigger gambles regarding said money (Tversky & Kahneman, 1974, 1981). The prospect theory could be seen as a retaliation against the expected utility theory and the norm of humans being rational decision makers.

The somatic marker hypothesis explains very well how emotions have to be accounted for when conducting research in decision-making. However, as a theory of moral decision-making it isn't sufficient. The somatic marker hypothesis gives good insight of how emotional decision-making works and it shows that emotions have a bigger impact on decision-making than rationalist theories originally thought, but it has nothing to do with moral decision-making. The somatic marker hypothesis marked the beginning for newer and
more current theories regarding moral decision-making, even though the hypothesis itself has nothing to do with moral decision-making, and the results of the IGT became the basis for newer theories in the future. However there are other theories which criticize the traditional rationalistic theories of moral decision-making that were important in the creation of these newer theories of moral decision-making.

3.3 The social intuitionist model

Haidt (2001) carried on the arguments and theories of the somatic marker hypothesis by introducing the social intuition model (SIM). The SIM suggest a post hoc problem in moral reasoning: we need moral reasoning to explain why we chose to act like we did, but only after we already have committed the act. Our moral reasoning only justifies the actions we have already made, not actually deciding the action itself. Our brain uses moral reasoning to create justifications for intuitive judgments, which thereby causes the illusion of objective reasoning when decisions are in actuality made emotionally and intuitively (Haidt, 2001).

Haidt (2001) suggested that even though moral reasoning faces the post hoc problem, moral reasoning still has some deciding factors in moral decision-making as well as moral judgments. Haidt (2001) argued that by using moral reasoning for persuasive powers one can persuade someone to see a specific problem in a certain perspective and trigger new intuitive feelings of right and wrong. Haidt (2001) suggested that moral reasoning can be used as a tool for argumentation and for influencing other peoples intuitions about how they feel and how they should act in moral scenarios. Martin Luther King Jr.'s "I have a dream" speech is a great example of using moral reasoning for argumentation. It affected many white Americans to think and feel differently about black Americans by gaining a new perspective on racial segregation (Haidt, 2001).
Haidt (2001) illustrated the functionality of the SIM further by suggesting that our intuitions can be seen as a dog, spontaneously and carelessly doing things, while the dog's tail represents our moral reasoning as it tries to keep up with the movements of the dog's body. Even if the dog's tail is interesting to study, since it is frequently used for communication (as moral reasoning is according to the SIM), one must focus on intuition and emotional processes, which are represented as the dog itself, to understand moral actions.

Haidt (2001) also stressed the importance of emotions in moral decision-making by referring to studies made on psychopathic behavior by Cleckley (1955). Cleckley (1955) suggested that psychopaths differ from normal individuals in their decision-making. By using case studies of psychopaths Cleckley (1955) argued that psychopaths had good intelligence, thought rationally and knew the rules of social behavior but simply didn't care to act according to said rules and, as a result, the psychopaths did whatever pleased them rather than thinking about the consequences which those actions could have on others. Haidt (2001) uses this data to argue that individuals with antisocial personality disorders function similarly to psychopaths because in both cases the VMPFC operates to a lesser effect than in normal controls and this explains the lack of emotional responsiveness in both psychopaths and individuals with antisocial personality disorder. Haidt (2001) argues however that individuals who have lived a life with a functional VMPFC feel sympathy for others and if their VMPFC ceases to function they still understand that they should feel sympathy for others even though they no longer do. Haidt (2001), and Damasio (1994), refers to this phenomenon as acquired sociopathy, where the moral reasoning is intact, but the VMPFC is damaged, and therefore results in the loss of affect. Haidt (2001) concludes that emotions and the VMPFC have a big impact on moral decision-making, similarly to how Bechara and Damasio (2005) concluded that the VMPFC have a big impact on decision-making overall, and that by including the importance of emotions and intuition in moral decision-making theories, it is possible to give
an alternative explanation to the traditional rationalist models regarding the functionality of moral decision-making.

4. A new and unusual perspective on moral decision-making

A scientist called Joshua Greene decided to follow up on the studies made by Damasio (1994). What was so special about Joshua Greene, compared to most other scientists who were studying the subject of moral decision-making at the time, was that Joshua Greene had a degree in philosophy while most other scientists were neuroscientists who had a lot of knowledge of neurology, but not of philosophical dilemmas (Greene, 2014). This made him have a very unique perspective on moral decision-making, and so he implemented moral theories and dilemmas into neuroscientific experiments, resulting in one of the newer theories of moral decision-making. But how was that integration possible and what moral theories were implemented?

4.1 Two moral dilemmas

In 2001 Greene, Sommerville, Nystrom, Darley, and Cohen tested the legitimacy of the traditional rationalistic views of moral decision-making by trying to find an interaction between the neural correlates of emotion and moral reasoning. Greene et al. (2001) applied two different moral dilemmas to their experiment that, even though similar in their nature, are typically solved in two different, and incompatible, ways by most individuals.

The first dilemma that was presented by Greene et al. (2001) is called the trolley dilemma (TD). The TD is a dilemma where a runaway trolley is uncontrollably headed towards five people that have no way to avoid the approaching trolley. If nothing stops the trolley from its current course these five individuals will die. You, as an observer of this event, stand to the side of the path of the trolley and have no way of warning the five
individuals of their impending demise. There's a lever next to you that, if pulled, will divert the trolley to a different path that will save the five individuals from certain death. However, there's a catch. There's one individual who is currently standing in the way of the alternative trolley path, and if the trolley is diverted to the alternative path this lone individual will suffer the same fate as the other five individuals would have done if the trolley course isn't diverted: death.

The decision that needs to be made in the dilemma is whether you will actively kill one person to save five others. Will you divert the trolley by pulling the lever and actively kill one individual to save five other individuals, or will you stand aside and passively watch as the trolley rolls towards its original path and kills five persons (but not actively killing anyone)? Most people seem to feel like the most natural answer is to pull the lever and actively kill one person to save five other persons (Greene et al., 2001).

The second dilemma that Greene et al. (2001) presented is called the footbridge dilemma (FD). The FD is quite similar to the TD. Five individuals are facing certain death as a trolley uncontrollably heads towards them. However, contrary to the TD, there's no lever nor any alternative path to divert the trolley. Instead, there's one individual that is standing next to you on a footbridge above the trolley rails. The only way to stop the trolley from killing the five individuals that are standing in the path of the trolley is to push the bystander next to you so that he or she falls on to the tracks. The body weight of the individual will be enough stop the trolley from running over the other five individuals that are otherwise facing certain death, but will inevitably kill the person that is pushed.

The decision which needs to be made in the FD is the same as in the TD. Will you actively kill one person to save five others? However, to save the five individuals from dying in the FD you have to actively push another person into the path of the trolley and with that person's body weight stop the trolley in its tracks. In the FD most people will decide not
push the bystander to his or her demise in order to save the five individuals, even though they were willing to pull the lever in the TD (Greene et al., 2001).

What's interesting is that even though both dilemmas ask the same moral question, would you actively kill one individual to save five others, the answer is different depending on the dilemma. The question why it feels morally acceptable to kill someone in the TD but not in the FD is something that has troubled moral philosophers for some time and there are many proposed answers to why this is the case, however one thing that all can agree on is that if a solution to this problem exists, it isn't obvious (Greene et al., 2001).

4.2 A hypothesis forms

Greene et al. (2001) suggested that the crucial difference between the TD and the FD lies in the tendencies to engage emotionally in the moral dilemmas. The thought of pushing someone to their death (as in the FD) is more emotionally disturbing and difficult to imagine doing than pulling a lever (as in the TD), which is not as emotionally disturbing. Greene et al. (2001) argue that some moral dilemmas engage our emotional processing to a greater extent than others, in this case FD being more emotionally engaging than TD, and this emotional engagement affects our judgment when deciding what to do in the presented dilemma.

Greene et al. (2001) show evidence for this claim through a study that is conducted with the somatic marker hypothesis in mind. By referring to Damasio (1994) and Haidt (2001), Greene et al. (2001) hypothesized that brain areas associated with emotion, such as the VMPFC, would be more active when contemplating different moral decisions and that intuitive processes influence moral decision-making. Greene et al. (2001) also argued that the individual would automatically feel disgusted and repulsed when considering moral decisions that go against their emotional input (i.e. gut feeling). Greene et al. (2001) also
hypothesized that when choosing an option in moral dilemmas that goes against the emotional input it should take a longer time to actually choose and go through with that action than the time it would take to decide something that is in line with the emotional input and in turn feels like the right thing to do.

4.3 The scientific experiments of moral decision-making

Greene et al. (2001) conducted two experiments on the basis of these predictions. The first experiment was conducted on participants in a functional magnetic resonance imaging (fMRI) scanner to see which brain areas became active when contemplating three different moral dilemmas. The second experiment measured reaction times in moral decision-making when participants were deciding what to do in the three different kinds of moral dilemmas that were presented in the first experiment. The three different types of moral dilemmas were called moral-personal, moral-impersonal and non-moral dilemmas. Moral-personal covered dilemmas, such as the FD, that are emotionally engaging and presumably activate areas in the brain that are associated with emotional processing. Moral-impersonal covered dilemmas, such as the TD, that require moral reasoning but, if hypothesized correctly by Greene et al. (2001), should show less emotional involvement than the FD. The last type of dilemmas that were presented was called non-moral dilemmas which covered questions that were not considered morally constraining in any way (such as whether to travel by bus or train given certain time constraints; Greene et al., 2001).

By asking participants moral-personal, moral-impersonal and non-moral questions while they were in a fMRI scanner, Greene et al. (2001) could map which brain areas were activated while contemplating on different moral dilemmas and whether there was a neurological difference when considering the three different sort of dilemmas. The results of
the first experiment showed that when participants were considering moral-personal dilemmas, neural areas such as the medial frontal gyrus, posterior cingulate gyrus, and angular gyrus were more active than when considering moral-impersonal and non-moral dilemmas (Greene et al., 2001). These neural areas have been associated with emotional processing in other studies (Maddock, 1999; Reiman, Lane, Ahern, & Schwartz., 1997; Reiman, 1997). The fMRI scans also showed that areas associated with working memory were much less active in contemplation of moral-personal dilemmas than in moral-impersonal and non-moral dilemmas. The neural areas that were less active are called the middle frontal gyrus and the parietal lobe, and both have been associated with working memory before Greene et al. (2001) conducted their experiment (Cohen et al., 1997; Smith & Jonides, 1997).

There seemed to be no significant difference in neural activity when considering moral-impersonal and non-moral dilemmas (Greene et al., 2001).

In addition to the first hypothesis about emotional processing being involved in moral decision-making, Greene et al. (2001) also hypothesized that decisions which go against the emotional gut feeling would take a longer time to decide than making decisions which fall in line with the emotional gut feeling. Greene et al. (2001) formed the second experiment to test this hypothesis. The reaction times were measured between participants who chose to defy their gut feeling, which was categorized as an *emotionally incongruent* response and participants who chose to follow their gut feeling, which was categorized as an *emotionally congruent* response. Greene et al. (2001) used the FD to test this hypothesis since, if one is to believe the results of the first experiment, the FD does evoke an emotional gut feeling of what to do. Participants who chose to throw the bystander in front of the approaching trolley were categorized as making an emotionally incongruent decision while participants who chose not to sacrifice the bystander by physically pushing him or her off the footbridge were categorized as making an emotionally congruent decision.
The results of the experiments showed that making an emotionally incongruent decision in a moral-personal dilemma, such as the FD, did indeed take significantly longer than making an emotionally congruent decision.

Greene et al. (2001) thereby confirmed their two predictions. The evidence support the idea of an emotional influence in moral decision-making. There is, however, one question that arises if one is to embrace the evidence shown in the experiments of Greene et al. (2001): what emotions come in to play in moral decision-making, and how many of our moral decisions are purely emotional?

4.4 The moral emotions

One year after Greene et al. (2001) conducted their study, Greene and Haidt (2002) sought to clarify the involvement of emotional processing in moral decision-making. Greene and Haidt (2002) suggested that moral psychology is a part of social psychology, which explained the involvement of social cognition and emotions in moral decision-making. However, they also argued that some social-psychological processes were not a part of moral decision-making, and are used specifically to process social information. This doesn't necessarily answer the question asked above however: which emotions and social processes are then involved in moral decisions and which aren't?

Greene and Haidt (2002) argued that the involvement of emotions and social processes in moral decision-making should be seen as a spectrum. Some emotions are more central to our moral lives than others (with compassion, guilt, and anger being examples of emotions that are more central to moral decision-making) but all emotions can be considered moral depending on the circumstances. Greene and Haidt (2002) and Haidt (2003) therefore called the emotions that are more central to morality the moral emotions. Haidt (2003) argued that the moral emotions can be defined as "Those emotions that are linked to the interests or
welfare either of society as a whole or at least of persons other than the judge or agent” (Haidt, 2003, p. 853). He also argued that all emotions are responses to perceived changes, threats or opportunities in the world and that in most cases they are related to the individual that is directly affected by these events (the self). Although it is a spectrum, and all emotions can be classified as moral emotions depending on circumstance, Haidt (2003) suggests that emotions such as contempt, anger, disgust, shame, guilt, gratitude, awe, and sympathy are central to morality.

5. The neural correlates of moral emotions

With the evidence provided by Greene et al. (2001), which supports a view where emotions are indeed a part of the moral decision-making, other researchers within the field started to ask themselves why emotions are a part of moral decision-making. Moll, de Oliveira-Souza, Eslinger, et al. (2002) were not satisfied with the answers that were given by Greene et al. (2001) and decided to study where these emotional connections are located in the brain and how they are connected with rationality. Even if there was a convincing amount of evidence to support the idea that emotion was a part of moral decision-making at the time and the brain correlates of basic emotions had been explored, the neural organization of moral emotions still remained poorly understood (Moll, de Oliveira-Souza, Eslinger, et al., 2002).

Moll, de Oliveira-Souza, Eslinger, et al. (2002) hypothesized that since damage to the VMPFC can result in a lack of empathy, give rise to antisocial behavior and difficulties in dealing with basic emotional inputs (Haidt, 2001; Bechara & Damasio, 2005) the VMPFC would be more active when a normal subject visually perceived stimuli that evoked moral emotions compared to when the subject visually perceived stimuli that evoked basic emotions without moral content, or when visually perceiving non-emotional stimuli. Moll, de Oliveira-Souza, Eslinger, et al. (2002) also hypothesized that basic, as well as moral emotions, would
show overlapping activations in brain regions like the amygdala and insula, which are both involved in emotional processing (Damasio, 1994; Bechara & Damasio, 2005).

5.1 Passive visual task

Moll, de Oliveira-Souza, Eslinger, et al. (2002) conducted a study to see where the neural correlates of moral emotions might be located by using a passive visual task (PVT). As the participants were scanned by a fMRI, their visual task was to look at emotionally charged pictures. Some pictures contained moral content while others didn't. The participants simply needed to look at the pictures to make the researchers able to see the spontaneous brain response triggered by the stimuli without having the participants actively respond to the pictures, hence a passive visual task (Moll, de Oliveira-Souza, Eslinger, et al., 2002). The pictures which were shown to the participants contained six different categories of scenes: emotionally charged pictures, unpleasant pictures, pleasant pictures, visually arousing pictures, neutral pictures and scrambled images.

Since only the spontaneous brain activations of perceiving a visual stimulus were needed for the study, there was no need for the participants to motivate or argue about what they felt about the pictures, but only to look at them as their brains were scanned (Moll, de Oliveira-Souza, Eslinger, et al., 2002).

Immediately after the imaging session, the participants were presented with the same pictures as they had seen while connected to the fMRI. This time the participants were asked to rate the different pictures according to predefined criteria (Moll, de Oliveira-Souza, Eslinger, et al., 2002). The participants had to rate the different pictures emotional response on a scale from "extremely unpleasant" to "extremely pleasant", the level of arousal on a scale from "calm" to "exciting", and moral content on a scale from "absent" to "extreme" (Moll, de Oliveira-Souza, Eslinger, et al., 2002).
The results showed that viewing moral and non-moral unpleasant visual stimuli activate a whole network of brain areas instead of specifically focusing on the VMPFC, as Greene et al. (2001) suggested in their study. Moll, de Oliveira-Souza, Eslinger, et al. (2002) therefore concluded that the effects which happen as a result of moral stimuli cannot be explained on the basis of a single neural area, such as the VMPFC. One has to look for a union of several different brain areas to find all the components that together explain these effects. Moll, de Oliveira-Souza, Eslinger, et al. (2002) suggested that effects such as emotional valence and visual arousal are vital components in the network of brain areas, amongst other components.

5.2 Visual sentence verification task

Moll, de Oliveira-Souza, Bramati, and Grafman (2002) also used a visual sentence verification task (VSVT) in a second experiment to confirm their predictions. In the VSVT the participants were asked to read short statements and to judge whether the statements were, in their opinion, right or wrong. The VSVT is a model which assumes that there are internal representations of different sentences and pictures in our brains. When we hear sentences we have to affirm or negate whether that sentence fits our internal representation of that sentence/scenario (i.e. is right or wrong; Carpenter & Just, 1975).

Moll, de Oliveira-Souza, Bramati et al. (2002) assumed that the VSVT works the same way for moral and non-moral decisions as for simpler sentences such as "the dots are red". Since morals are psychological constructs a distinction between non-moral social violations and moral social violations have been formed and created a psychological basis of what moral as well as conventional transgressions look like (Blair, 1995; James & Blair, 1996). This psychological basis has been worked out for both normal individuals, psychopaths as well as patients with autism (Blair, 1995; James & Blair, 1996). Moll, de
Oliveira-Souza, Bramati, et al. (2002) uses this psychological basis to test how their normal participants react to these short statements and whether their internal representations fits the psychological basis provided by Blair (1995) and James and Blair (1996).

The different types of statements were divided into four different categories which were designed to evoke different neural responses (Just as in the PVT): non-moral neutral, non-moral unpleasant, moral, scrambled. (Moll, de Oliveira-Souza, & Bramati, et al., 2002).

These short statements were presented to the participants while they were being scanned by a fMRI and were later presented to the participants once more, but in a randomized fashion, after the fMRI scan where the participants were asked to rate the statements according to similar parameters as in the PVT: moral content and emotional valance (Moll, de Oliveira-Souza, & Bramati, et al., 2002). In the VSVT, the participants were also encouraged to make a short verbal commentary on their score choices regarding each statement.

The results from this study further supported the results of the study conducted by Moll, de Oliveira-Souza, Eslinger, et al. (2002) which was made with the PVT. Not only could Moll, de Oliveira-Souza, Bramati, et al. (2002) conclude that there is no single neural area that alone represents the effects of moral stimuli, but that these moral stimuli activate different neural networks. Moll, de Oliveira-Souza, Bramati, et al. (2002) found that moral decisions associated with unpleasant emotions activate a certain set of neural areas while non-moral social decisions that were associated with unpleasant emotions activated another set of neural areas compared to the first condition. This new data led Moll, de Oliveira-Souza, Bramati, et al. (2002) to believe that the studies made by Greene et al. (2001) did not tell the whole story and that a new theory was needed to explain moral decision-making.
5.3 The event-feature-emotion complexes


The first of the three mechanisms, which is called context-dependent knowledge, is located in an area of the brain called the prefrontal cortex and consists of context-dependent knowledge of events and stimuli happening in one's life (Moll et al., 2005). The term context-dependent knowledge is an umbrella term which covers stimuli from social events and contexts. This includes emotional knowledge and social stereotypes of different social contexts (where neural areas such as the VMPFC are involved). An example would be the different emotions you would feel when you hear the name of a specific person that you might dislike for some specific reason. Maybe those emotions arose from a specific social context, like a party, where that person did something that made you dislike him/her. The umbrella term also covers non-social events and contexts that involve routine tasks, storing long-term goals and thoughts about the future. (Moll et al., 2005).

The second neural mechanism, which is called social semantic knowledge, is located in the superior temporal sulcus (STS) and could be explained as context-independent knowledge (the opposite of first neural mechanism). When interacting with the environment there is an insurmountable amount of information for the brain to process. As a result, the
brain filters the information that is deemed irrelevant to one's personal objective and prioritizes perceptual signs which could be of a social significance for the individuals current objective. These perceptual signs contain knowledge of features and semantics (Moll et al., 2005). An example would be walking down the street to the grocery store to buy food. Even though it is a relatively simple task, it contains an incredible amount of information and stimuli regarding things such as other persons, different objects in the store, cars on the street etc and it would not be uncommon for someone to miss a cyclist or a car when in the pursuit of reaching their objective because that specific cyclist or car is not the personal objective and therefore it is not being attended to. While being able to attend to features and semantics in our environments is a vital part of the social semantic knowledge, it also entails information on how to be able to interact in the social world and make implicit or explicit moral decisions (Moll et al., 2005). In order to interact and reach your personal objectives in the world, you must be able to distinguish between different social as well as functional features such as facial expressions, gaze, body posture, and gestures. The STS stores these social representations and helps out with social decoding (Moll et al., 2005). One could explain the social semantic knowledge as knowledge of everything in the surrounding environment that has no connection with contexts or social situations (like parties or events) and contains information on how to interact with others socially as well as information regarding everything that surrounds us. The social semantic knowledge is referred to by Moll et al. (2005) as social, perceptual and functional features.

The third and last of the neural mechanisms that integrate moral decision-making, which is called motivational states, is located in the limbic and paralimbic structures of the brain and represents the motivational as well as emotional aspects of moral decision-making (Moll et al., 2005). These motivational states include feelings such as sexual arousal, social attachment, hunger and aggression and prone individuals to feel different motives
depending on what emotion the individual is feeling at the time (Moll et al., 2005). For example, if someone feels hungry because it's the last hour before lunch it will affect their motives and decisions differently compared to if it was the first hour after lunch. Moll et al. (2005) are however very clear in distinguishing the central motive states from basic emotions (such as fear and disgust) as basic emotions are a part of context-dependent event knowledge (perceiving a feared object or a disgusting object makes you feel fear or disgust). A good example of this are the somatic inducers which Bechara and Damasio (2005) present in their somatic marker hypothesis. Moll et al. (2005) argue that basic emotions can emerge from moods and motivational states or even induce a motivational state, but can't be a mood, or motivational state themselves. Moll et al. (2005) suggest that these motivational states are a way of motivating the individual to commit, or not to commit, a moral action and can be considered as a drive to actually do what is felt to be beneficial or as the right thing to do. The motivational states are referred to by Moll et al. (2005) as central motive and emotional states.

Combining these three neural mechanisms give emergence to the event-feature-emotion complexes (EFECs) and with the collaboration between these different brain areas the EFECs enables us to make moral decisions (Moll et al., 2005). Moll et al. (2005) argue that with the integration of contexts dependent knowledge, social features, perceptual features, functional features and motivational states we can make moral decisions and it gives us a complete understanding of the processes which are involved in moral decision-making. By combining stimuli and contexts of different events that we are a part of with knowledge of what we perceive in our environments while experiencing these events and the different emotions we feel in those moments we enable ourselves to make moral decisions and take moral positions in regards to the world and ourselves (Moll et al., 2005).
6. The dual-processing brain

The EFECs, which Moll et al. (2005) argue for, does not in itself question or support the results from Greene et al. (2001). However, as more studies have been conducted in order to understand emotional involvement in moral decision-making this `mutual understanding´ has changed and as a result, one can today see the theories of Greene et al. (2001) and Moll et al. (2005) as relatively opposed explanations of moral decision-making.

In 2004 Greene, Nystrom, Engell, Darley, and Cohen (2004) presented their dual processing theory as a result of the experiments which were conducted by Greene et al. (2001). The dual processing theory suggests that there are two different neural areas that are at work when making a moral decision. The first area is the VMPFC and is associated with emotional inputs in moral decision-making. The second neural area is the dorsolateral prefrontal cortex (DLPFC) and is associated with abstract reasoning and cognitive control in moral decision-making. In the experiments by Greene et al. (2001), the VMPFC was more active than the DLPFC in personal moral dilemmas, and the DLPFC was more active than the VMPFC in non-personal moral dilemmas. With this in mind, Greene et al. (2004) concluded that the VMPFC is responsible for emotional inputs that affect moral decision-making and the DLPFC is responsible for rational inputs that affect moral decision-making.

Greene (2014) explains this dual-processing between the VMPFC and DLPFC as a camera with an automatic setting (representing the VMPFC) and a manual setting (representing the DLPFC) configured into it. The camera's automatic setting is set to automatically focus in on things that appear in the lens while the manual setting makes you able to manually decide what the camera lens should focus on. Greene (2014) suggest that our brain works the same way as the camera where the automatic setting is more efficient, and very fast at being able to focus in on what appears to be of importance. The manual setting is
more flexible and as a result makes us able to choose what the camera should focus on, but it takes much longer to actually take pictures and is not as efficient as the automatic mode (Greene, 2014). The automatic setting (VMPFC) makes us able to make intuitive decisions on the go without thinking very hard about them, which serves us well in our everyday lives since it wouldn't be plausible to consider and waste time about every single decision we make (Greene, 2014). However when we do have to stop and think about a decision and encounter complex and unfamiliar problems (like moral dilemmas) we have to switch into the manual setting (DLPFC) which makes us able to really think about how to go about the encountered problem. Greene (2014) argue that this manual setting is what makes us humans able to adapt into different environments and different situations and is what makes us able to sacrifice one in order to save five in moral dilemmas. This also suggests that the emotional inputs (VMPFC) and rational inputs (DLPFC) are in conflict with each other. Either we are in an automatic setting (VMPFC) or a manual setting (DLPFC).

Greene (2014) suggest that we're able to cooperate, and live in small groups as a results of the automatic setting, while the manual setting allows us to think rationally and consider moral dilemmas. Greene (2014) explains it as

The moral brain's automatic setting are the moral emotions...the gut-level instincts that enable cooperation within personal relationships and small groups. Manual mode, in contrast, is a general capacity for practical reasoning that can be used to solve moral problems, as well as other practical problems. (Greene, 2014, p.15)

Greene (2014) explains the relationship between the different camera settings through a theory called Modular myopia. the Modular myopia theory suggest that our automatic system is created to internally monitor violence. The reason why we feel like it's
wrong to kill someone in the FD is because we're the ones doing it (Greene, 2014). Studies have shown that watching someone else committing an act of violence does not get a negative emotional reaction, and therefore Greene (2014) conclude that our automatic setting only "cares" about if oneself commits an act of violence. The modular myopia theory also explains why it seems ok to sacrifice one to save five in moral dilemmas such as the TD. This internal monitor for violence is blind to harm as a side effect, as in pulling the lever being the main effect in the TD, and therefore does not react negatively to it. At the same time harming someone as a mean to save others, the main effect being pushing someone in the FD, is a direct act of violence and therefore triggers the alarm (this alarm system is discussed in sub-chapter 6.2; Greene, 2014). Greene (2014) also argue that our manual system is rational by nature, and therefore will always be ok with sacrificing one for five. This also further explains the duality between emotions and reasoning. Emotions are gut reactions, emotional responses which respond to violence while reasoning is of a calculated and rational nature and therefore they both are in direct competition with each other of whether to react or not to.

Greene (2014) also talks about previously conducted studies regarding the relationship between the VMPFC and the DLPFC, along with the dual-mode camera, starting with Koenigs et al. (2007).

Koenigs et al. (2007) conducted an experiment regarding the involvement of the VMPFC and emotions in moral decision-making, which supports the evidence provided by Greene et al. (2004). By using fMRI, Koenigs et al. (2007) scanned the brains of individuals with VMPFC damage as they were making decisions regarding difficult moral dilemmas. The authors argued that not only do individuals with damage to the VMPFC have hindered capabilities of making decisions, but they also have difficulties showing social emotions, and as a result become abnormally cold, distant and rational.
Koenigs et al. (2007) referred this abnormally distant and rational behavior as *utilitarian* behavior. Utilitarianism is an ethical viewpoint of maximizing pleasure and minimizing pain (Mill, 1901). It is an impartial moral position that requires individuals to always look from a third person perspective and not to give any special treatment or favor to anyone that we, as individuals, value higher personally than the average man (Mill, 1901). This can be considered to be very constraining and highly demanding for an individual since it often requires very heavy personal sacrifices in order to achieve a universal greater good (Mill, 1901).

Koenigs et al. (2007) suggested that as individuals with damage to the VMPFC becomes abnormally utilitarian, they also start to show emotionally aversive behaviors and do things that normally would feel wrong. Koenigs et al. (2007) even argue that these emotionally aversive behaviors can be associated with psychopathy.

Greene et al. (2004) describe a similar thought process regarding the VMPFC and utilitarianism as Koenigs et al. (2007). Greene et al. (2004) suggested that utilitarian decisions are a product of moral reasoning by arguing that cognitive processes favor utilitarian decision-making.

6.1 Cognitive functions of the dual-processing brain

Suter and Hertwig (2011) argue that moral dilemmas which engage emotional processes result in fast, intuitive gut reactions while moral dilemmas which activate cognitive processes results in utilitarian responses. Suter and Hertwig (2011) suggest that we require cognitive control to be able to override the initial emotional response and activate these cognitive processes. Suter and Hertwig (2011) argue that this cognitive control takes time to establish and so, by experimenting with the time constraints given to participants in studies regarding moral dilemmas, Suter and Hertwig (2011) suggest, and shows evidence for, a
correlation between time and likelihood of having a rational response to the presented moral dilemma. This means that participants with high time constraints responds in correlation to an emotional gut reaction while participants that are given low time constraints tend to respond in correlation to a cognitive response which is of a more rational and utilitarian nature. Suter and Hertwig (2011) conclude that the key lies in how long it takes for the cognitive control to kick in, and that the more time which is given in moral dilemmas the likelihood of having the cognitive control mechanism kick in increases.

Paxton, Ungar, and Greene (2012) found similar results to Suter and Hertwig (2011). By having participants complete a cognitive reflection test before being faced with a moral dilemma, Paxton et al. (2012) prepared the participants cognitive processes for moral dilemmas to see if participants were more utilitarian in the moral dilemmas if they were to use their cognitive processes before actually being exposed to the moral dilemma. Paxton et al. (2012) concluded that participants did indeed seem to be more reflective, rational and utilitarian if they were faced with a moral dilemma after being exposed to the cognitive reflection test. These results go hand in hand with the results of Suter and Hertwig (2011) and one could argue that participants who were exposed to a cognitive reflection test before a moral dilemma already had their cognitive control mechanism primed and ready and therefore able to be much more rational and utilitarian in their moral decisions.

6.2 Emotional processes of the dual-processing brain

In 2012, Amit and Greene conducted a study regarding visual imagery in moral decisions. Amit and Greene (2012) tested whether participants were using a visual cognitive style or a verbal cognitive style in working memory tasks by having participants complete two matching working memory tasks, one visual memory task and one verbal memory task. After categorizing participants to having a visual cognitive style, or a verbal cognitive style the
participants were faced with a personal dilemma and an impersonal dilemma (Amit & Greene, 2012). The results showed that the participants who were categorized as using a visual cognitive style were favoring the right of the individual more (unwilling to sacrifice one individual in order to save five) than participants who were categorized as using a verbal cognitive style, which were more utilitarian (willing to sacrifice one individual in order to save five; Amit & Greene, 2012). After conducting a second experiment on a larger sample size (more participants) Amit and Greene (2012) concluded that the reason why individuals who use a visual cognitive style tend to regard the rights of the individual as more important than saving five others for the greater good (or utilitarianism) is because when using visual imagery the individual tends to visualize and focus on the harmful means more than the beneficial ends and as a result tend to favor the right of the individual rather than maximizing happiness of the greater good, as a utilitarian would do. This suggests that emotional responses of watching the harmful means happening in front of the participants eyes prevents the participant from rationalizing in a utilitarian manner, and to an extent shows support that decisions, especially in personal dilemmas such as the FD, are made based on emotional responses (Amit & Greene, 2012).

Shenhav and Greene (2014) shows evidence which builds on the studies made by Amit and Greene (2012). Shenhav and Greene (2014) suggest that the amygdala have a positive correlation with negative automatic emotional responses to harmful action in personal moral dilemmas, such as the FD, and that these responses directly compete with controlled, cognitive responses which favor utilitarian choices in said moral dilemmas. In their experiment, Shenhav and Greene (2014) suggest that while the amygdala can be seen as an initial alarm response to personal moral dilemmas, and can be considered as an emotional assessment of the dilemma, the VMPFC functions as a sort of bridge between a utilitarian assessment and the emotional assessment of the moral dilemma. Shenhav and Greene (2014)
argue that this makes the participants able to come up with a morally acceptable solution to
the dilemma which considers both the emotional as well as the utilitarian assessment.
Shenhav and Greene (2014) calls this compromise an integrative moral judgment.

With these five studies, we can clearly see that the dual-processing theory has a lot of evidence support its claims. There are however others who disagree with the theory, but who most of all disagree with the terminology which is used in these studies.

6.3 Utilitarian, or just careless?

Kahane, Everett, Earp, Farias, and Savulescu (2015) object to the definition of utilitarian that is used regarding decisions in dual-processing research. Both Greene et al. (2004) and Koenigs et al. (2007) suggest in their articles that individuals who make decisions which defies the emotional inputs are purely cognitive and therefore are considered utilitarian. By referring to the TD as well as the FD, Kahane et al. (2015) argue that only because someone chooses to sacrifice one individual to save five others, it doesn't make them a utilitarian. One can't assume that an individual chooses such an answer because they are utilitarian and looking for a universal greater good, as is the goal of a true utilitarian (Mill, 1901). Kahane et al. (2015) suggest that these so-called `utilitarian´ judgments aren't actually associated with a desire for universal greater good, but rather with an immoral outlook on ethical transgressions and even possibly psychopathic traits. Kahane et al. (2015) also argues that instead of finding similarities between `utilitarian´ decisions in moral dilemmas and a desire for universal greater good, Kahane et al. (2015) found that these `utilitarian´ decisions were associated with a greater endorsement of rational egoism and a low concern for the rest of humanity. This is particularly worrisome since identifying with all people around the globe equally is important for a utilitarian and is considered to be a core aspect of utilitarianism (Kahane et al., 2015). Kahane et al. (2015) conducted several experiments to justify these
In their first experiment, Kahane et al. (2015) recreates the FD and asks regular individuals how wrong they consider the utilitarian choice to be (which is to throw the bystander under the trolley to save five others). The experiment had two variables to validate: If one was to choose to sacrifice the bystander, how much of that choice can be explained by psychopathic behavior and how much of the choice can be explained by reduced empathic concern. Kahane et al. (2015) concluded that it indeed seemed like the ‘utilitarian’ judgments were primarily driven by psychopathic tendencies and much less driven from a reduced empathic concern, even though empathic concern was a factor of the decision. The first study showed the same results as those predicted by Greene et al. (2001) and Greene et al. (2004). Kahane et al. (2015) argue that psychopathic traits and impartial utilitarian concerns do share some common traits such as a cost-benefit analysis plan to guide actions as well as dismissing commonsense morality.

With the results of their first study, which was created as a baseline, Kahane et al. (2015) conduct a second study which goes deeper into the definitions of utilitarian judgments and anti-social behaviors in moral dilemmas. With being able to see the similarities between psychopathic behavior and impartial utilitarian concern, Kahane et al. (2015) asked themselves what the core differences are. Kahane et al. (2015) argue that a true utilitarian has a concern for humanity as a whole, an impartial concern for distant strangers, compared to the exclusive egoist concern and minimal altruism of a psychopath and conducted a second experiment to confirm these predictions of differentiated behavior. By having normal participants complete an online survey, Kahane et al. (2015) tested the participants according to *The Identification with All Humanity Scale* (IWAH), their stance towards psychological, ethical and rational egoism, willingness to sacrifice others if it's to their own benefit, and asking about hypothetical anonymous donations. The results showed a tendency where
utilitarian judgments were associated with lower impartial concern for the greater good and an increased endorsement of rational egoism (actions are only rational if they serve one's own self-interest) and impartial utilitarian judgments were associated with a higher impartial concern for the greater good, and a decreased endorsement of rational egoism (Kahane et al., 2015).

Kahane et al. (2015) conclude that the term `utilitarian´ judgments are used wrongly by Greene et al. (2004), Koenigs et al. (2007) amongst the other researchers of the dual-processing theory. As one can see in the studies conducted by Kahane et al. (2015) there is a clear difference between a utilitarian and an anti-social individual with damage to the VMPFC which, according to Kahane et al. (2015), needs to be taken into account when conducting experiments regarding moral dilemmas.

7. Conflicting opinions

Moll and de Oliveira-Souza (2007) suggest that it is not as easy as Greene et al. (2004) and Koenigs et al. (2007) make things out to be. Moll and de Oliveira-Souza (2007) argue that even though damage to frontal brain areas, like the VMPFC, has been documented to cause behavioral impairments, such as social inadequacy and moral violations, it is hardly the only brain area that is involved with such behavior. By referring to the fMRI scans that Koenigs et al. (2007) acquired in their experiment, Moll, and de Oliveira-Souza (2007) argues that areas such as the frontopolar cortex (FPC) and the DLPFC are active along with the VMPFC. Moll and de Oliveira-Souza (2007) suggest that since connectional, functional and behavioral aspects of the FPC are different from that of the VMPFC, this could have important implications for the results of the experiment which Koenigs et al. (2007) conducted. Moll and de Oliveira-Souza (2007) also argue that the FPC was not only active in the experiment which Koenigs et al. (2007) conducted, but also in the experiments of Greene
et al. (2001). Moll and de Oliveira-Souza (2007) suggest that this new insight questions the
relevance of the FPC in these moral dilemmas, and that it renders the claim that the VMPFC
alone is responsible for emotional input invalid, which in turn would make the dual
processing theory false.

Moll and de Oliveira-Souza (2007) also argue that the dual processing theory
does not have sufficient scientific evidence. In order to make the dual processing theory
viable, a *double dissociation* would need to be found between the VMPFC and the DLPFC. A
double dissociation would show that selective VMPFC damage increases utilitarian choices
and that selective DLPFC damage leads to emotional choices. In other words, Moll and de
Oliveira-Souza (2007) argue that an experiment is needed where one participant with damage
exclusive to the VMPFC would make more utilitarian choices while a second participant with
damage exclusive to the DLPFC would make more emotional choices.

Instead of using the dual process theory Moll and de Oliveira-Souza (2007)
suggested that the VMPFC and FPC might be necessary to understand which moral
sentiments are socially acceptable. with that being said, Moll and de Oliveira-Souza (2007)
thorized that emotion and reason work in unison instead of conflict, by referring to the
EFECs, which suggest a collaboration between emotions and reasoning (Moll et al., 2005).

7.1 Two different approaches to decision-making

Other more recent studies favor Moll and de Oliveira-Souza (2007) viewpoint
on the neural correlates of moral decision-making. Pessoa (2008) suggest that the field of
moral decision-making support a view of *functional specialization*, which means that brain
areas can be conceptualized as either affective or cognitive. Pessoa (2008) argues that this
view is wrong. The only reason why the field supports such a view is because it only consists
of lesion studies. Most studies conducted by Greene et al. (2001) and Koenigs et al. (2007) (as
well as other studies) focused in on specific brain areas in their experiments, and therefore naturally conclude that the affected brain areas are the key to explain the functionality of how emotions and reasoning work. However, Pessoa (2008) suggest a different view. By referring to previously conducted studies, such as that of Moll et al. (2005), Pessoa (2008) argues that reason and emotion work together as cognitive-emotional behaviors and that they originate from a network of brain areas. None which should be specifically tied to either emotion or reason. Pessoa (2008) argues that the interaction between the different networks of brain areas explain our behaviors and since networks that are associated with reason and networks that are associated with emotion integrate information between each other, it is logical to conclude that both networks work in unison to form decisions by exchanging information between each other.

Siegel (2015) follows the research done by Pessoa (2008). Siegel (2015) argues that emotions can be found throughout the entire brain, and not in a specific brain area. Siegel (2015) suggest that the limbic region is specialized to have wide-ranging effects on meaning and values of different stimuli and the center of social cognition. This is the same region which Moll et al. (2005) found to be representing central motives and emotional states.

Not everyone agrees with the model which is provided by Moll et al. (2005), Pessoa (2008) and Siegel (2015) however. Evans and Stanovich (2013) suggest that the criticism of dual-process theories is unjustified. Evans and Stanovich (2013) argue that the approach of Moll et al. (2005), Pessoa (2008) and Siegel (2015) is too generic and needs to be specified in order to actually give any meaning. Evans and Stanovich (2013) argues for a different kind of dual processing theory that supports type 1 and type 2 processing. Type 1 processing consists of rapid autonomous processes while type 2 processing consists of default responses (how we naturally act) unless consciously stopped and distinctive higher order reasoning processes. Evans and Stanovich (2013) refer to Dennett (1996) who suggest that
these different types of processing work in a hierarchical manner where the reflective, higher level control (type 2 processing) controls the lower level of type 2 processing (default responses) and the type 1 processes. Evans and Stanovich (2013) also specifically specifies that the type 2 processes are linked with working memory and supports hypothetical thinking.

Lindquist, Wager, Kober, Bliss-Moreau, and Barrett (2012) suggest that two different approaches have risen in regards to the interaction between emotion and reason. Some scientists argue in favor of a locationist approach which tends to assign emotion and reason to specific brain regions, just as Damasio, Haidt, Greene and Koenigs, while other scientists favor a constructionist approach which suggests that emotions are emerging from a general brain network rather than a specific brain region, Like Moll, Pessoa, and Siegel.

8. Discussion

There has been a paradigm shift in the field of moral decision-making. From originally believing that rationality and reason are the cornerstones of moral decision-making, and having a very rigid outlook on the mind, we now have a field full of emotional input and unpredictable outcomes which suggest that really anything can happen depending on the current mood (Moll et al., 2005). While some of the earlier theories of emotional decision-making (the somatic marker hypothesis and the social intuitionist model) argues for emotional influences in moral decision-making, as well as decision-making overall, and begs to question the traditional, rationalistic theories (the expected utility theory and Kohlberg's stages of moral development) the later theories of moral decision-making suggest a overruling emotional input that overwhelsm the cognitive ones and that there is a struggle between emotional and rational inputs (locationist approach and the dual-processing theory), other theories suggest a collaboration between the different inputs that together form moral decision-making (constructionist approach and the EFECs).
However in-between these two approaches of moral decision-making there seem to be a lack of mention to what non-emotional based moral decision-making is. While supporters of the dual-processing theory, such as Greene et al. (2001), Greene et al. (2004), Greene (2007) and Greene (2014) amongst others followers of the dual-processing theory argue for the involvement of emotions, the authors gives very little thought to what moral actions without emotion is defined as and assumes that it is the same as acting rationally. Kahane et al. (2015) show evidence which suggests which shows that Greene's assumption is wrong. While actions without emotional involvement seem, cold, calculated and non-affective it doesn't necessarily mean that those actions are made out of reason and logical conclusions, but rather from rational egoism.

There also seems to be a lack of research regarding how group thinking and social conformity would affect the outcome of the decision (such as the experiments of Asch, 1946). For example if a participant, who were in the same room as several others who were all a part of the experiment, where to choose in the FD (where it seems very prominent to choose not to push the individual off the bridge) but have several others choose to push the individual to their death, would the participant follow the group through social conformity and agree with pushing the individual even though it would go against their emotional inputs? This would give a clarification of how strong the emotional inputs are and if we are able to overrule them, even though we might not feel like we want to, or if the force is so strong that it would feel so wrong as to be able to defy the rest of the group and the social conformity in order to follow our emotional inputs.

The experiments which Greene et al. (2001) conduct partly explain the importance of emotional inputs in the FD. However, there seems to be a lack of mention of why our decisions are, at least partly, emotion-based. One could argue that the reason why a participant won't push the individual off the bridge in the FD is purely egoistic rather than
empathetic (like how Kahane et al., 2015 argue that pushing the individual isn't necessarily an act of utilitarianism). Since it is through own volition, one could argue that the participant is protecting themselves from committing a terrible crime and that the participant then justifies their egoistic action through empathy for the individual (similarly to how Haidt, 2001 explains the dog and its rational tail). To kill someone, with their own hands, is a horrible act to think of and it is doubtful that anyone (who is normal) would consider killing another person to be an easily committed act. The feeling of killing someone is repulsive, and even if the person which is about the be killed can be considered `bad´ or have committed terrible crimes, it doesn't make killing him/her easier. This would argue that the choice of not killing an individual isn't necessarily moral or done for empathetic reasons, but for protecting oneself from traumatic experiences and made for egoistic reasons. In other words, killing someone (or pushing them to their death as in the FD) isn't necessarily avoided because of empathy for the individual, but because of egoistic reasons and because it feels wrong to oneself.

A way to confirm such a theory would be to have an alternative TD where the participant can steer the trolley on a side path to kill someone that the participant has a personal relationship with (such as a parent, child, relative or close friend). This would force similar emotional reactions as in the FD, but the personal involvement in the act of the killing would be lesser than that of the FD (pushing a lever, not pushing an individual). One could still argue that a person wouldn't commit to the act because it's their relative, and therefore making it an act of egoism but arguably it wouldn't be the same pressure from the ego since there's less personal involvement in the act even though killing someone the participant cares about is against their emotional inputs.

9. Conclusion
The discussion about the relationship between emotion and reason has branched to a more general level than just moral decision-making. This is a topic that covers many fields and is now more a question of what relationship between emotions and reason is in general. We can see that there are two different approaches that argue for how emotions work. Followers of both views still argue about the nature of the relationship between emotion and reason to this day.

While there are quite a bit of evidence which supports a dual-processing approach to moral decision-making, the general critique of the dual-processing theory is that there are more processes behind moral decision-making than just the VMPFC and DLPFC. In other words, scientists seem to be in agreement that the VMPFC and DLPFC are both a part of moral decision-making, but followers of a constructionist approach, and EFECs, argue that there are more to moral decision-making than just those two areas. It is also however important to note critique from authors such as Evans and Stanovich (2013) who does mention that the constructionist approach doesn't actually bring up any concrete alternative neural functions, and that the constructionist approach is quite generic in their claims. Therefore a legitimate counter-argument to the constructionist approach would be: what other functions are involved then?

With this said, it is hard to favor one side or the other. Everyone can agree that the VMPFC and DLPFC are a part of moral decision-making, and that there is an extensive amount of research which support that, but it's hard to neglect the critique given by the constructionist approach even though the constructionist approach itself is of a very generic and vague nature.

In conclusion, we can see a field which was once dominated by religious followers and philosophers has in recent years turned into a hot topic for debate among
scientists. A topic which consists of people who argue for a traditional, rationalistic set of ideas regarding moral decision-making, people who argue that emotions are more emergent than, and to some extent control, reasoning and a third faction consisting of individuals who argue for the importance of emotional input and gut feelings in moral decision-making, but suggests a collaboration between emotion and reasoning.

What do we base our moral decisions on? A question, which the answer seems to be quite obvious at first, that proves to be very tough to give a concrete answer to. Are we truly doing what we desire? Or are we doing what seems to be moral and righteous?

Whether you believe in a rationalist explanation where humans are reasoning creatures or believe that humans are emotional animals that follow our desires doesn't really tell much of a story. It seems to have become a question of perspective rather than a fact. A world which was once black and white has turned into nuances of grey.

References


