

# Impulsivity: why reasonable people do unreasonable things

An interdisciplinary approach to akrasia and addiction

19 April 2014



## Authors

Vera Van Gool

3834115

Judith Smit

3866173

Liberal Arts & Sciences

Utrecht University

## Advisors

Philosophy advisor: Katrien Schaubroeck

Biology advisor: Berry Spruijt

Interdisciplinary advisor: Ria van der Lecq

Image: <http://soberinspiration.blogspot.nl/2012/09/how-to-quit-smoking-in-sobriety.html>

## **Abstract**

People know some actions are wrong to perform, but still perform the actions. It is common knowledge that smoking has many negative consequences, but a lot of people smoke anyway. Why do these reasonable people do unreasonable things? Are these people too weak to resist temptations? Is there a biological basis for this irrationality?

We have discussed these questions using an interdisciplinary approach. Insights from philosophy and biology proved useful since both disciplines deliver insights on irrationality. Philosophers discuss the concept of akrasia as an example of irrationality, which means acting against better judgment. Addiction, using drugs compulsively despite negative consequences, is a focus of biologists when it comes to irrational behaviour. We have integrated the findings on akrasia and addiction to deduce a more comprehensive understanding of why reasonable people do unreasonable things.

We propose a model that shows how unreasonable behaviour arises. Behaviour is motivated by a desire to perform a certain action and representations of possible actions. A desire is influenced by past experiences and is determined by the brain. One can desire to perform an action that has the experience of pleasure as a result. The amount of pleasure an action generates can be overrated, however. Besides desire, long-term goals and deliberations play a role as input for decision making. In akratic and addicted people desire is more dominant than this deliberation, resulting in unreasonable behaviour. We use the description 'impulsivity' to explain unreasonable behaviour.

Our findings can contribute to the academic debate on unreasonable behaviour. Our results could also be used in the search for more effective treatments for addiction.

## **Table of content**

<b>Introduction</b>	<b>4</b>
<b>Chapter 1. Akrasia: knowing and ignoring what is best</b>	<b>6</b>
<b>0. Introduction</b>	<b>6</b>
<b>1. Akrasia</b>	<b>7</b>
1.1 Aristotle's akrasia	
1.2 Contemporary approach to akrasia	
<b>2. Pleasure and avoiding discomfort</b>	<b>12</b>
<b>3. Passions, imagination and malfunction</b>	<b>15</b>
3.1 Passions	
3.2 Imagination	
3.3 Akrasia as a result of malfunctioning imagination	
<b>4. Conclusion</b>	<b>19</b>
<b>Chapter 2. Addiction: the brain fooled by drugs</b>	<b>21</b>
<b>0. Introduction</b>	<b>21</b>
<b>1. Defining addiction</b>	<b>21</b>
1.1 Internal versus external stimuli	
1.2 The role of opioids and dopamine in addiction	
1.3 The economy of behaviour	
<b>2. Liking and wanting</b>	<b>25</b>
2.1 Liking and reinforcement theories	
2.1.1 Pavlovian and instrumental conditioning	
2.1.2 Positive and negative reinforcement theories	
2.2 Wanting and the incentive-sensitization theory of addiction	
<b>3. A neurocognitive approach to addiction</b>	<b>29</b>
3.1 A three-system neural model of addiction	
3.2 The insula	
<b>4. Conclusion</b>	<b>33</b>
<b>Chapter 3. Impulsivity: deliberation overruled by desire</b>	<b>35</b>
<b>1. Overlap, conflicts and common ground</b>	<b>35</b>
<b>2. More comprehensive understanding</b>	<b>37</b>
<b>Conclusion and discussion</b>	<b>40</b>
<b>Bibliography</b>	<b>42</b>

## Introduction

Humans are beings unique for their ability to reason and their use of language. With language humans can give reasons for their motivations. Other animals, lacking the use of a language, cannot do this. The reasons people give for their actions can be plausible, though there are exceptions. Nowadays about one billion people smoke at least one cigarette a day (WHO, 2013). Smoking has many negative consequences. Besides the effects on health there are also political, economical and social consequences. Most smokers know about these negative consequences, but still smoke. Some reasonable people therefore would rather not smoke, but still do it. Why do they give in, when common sense tells them not to?

Unreasonable behaviour, like addiction, is a fascinating phenomenon since humans know they should not perform it. Are these reasonable people too weak to resist temptations? Is there a biological explanation for their irrationality? To answer such questions we will take a philosophical and biological perspective and integrate the insights to answer our research question: *Why do reasonable people do unreasonable things?*

Both philosophy and biology study human behaviour. These disciplines deliver insights on how unreasonable behaviour can be performed by otherwise reasonable people. This justifies the interdisciplinary approach of this project (Repko, 2012, p.84). In philosophy the concept of 'akrasia' forms the centre of debates on unreasonable behaviour. We will describe the philosophical perspective on the concept of akrasia, meaning acting against one's better judgments. Addiction is discussed in biology and can be seen as unreasonable behaviour. In biology, addiction is defined as compulsive drug taking behaviour despite the negative consequences. We will discuss the phenomena akrasia and addiction and compare and integrate the insights we have gained from philosophy and biology in order to gain a more comprehensive understanding of why reasonable people do unreasonable things.

Philosophers see humans as beings capable of reasoning. They assume that reason can be a driving force of behaviour. Insights from philosophy are based on rational deliberations and reasoning. Biologists, on the other hand, use empirical research to search for physical explanations (e.g. in the brain) for behaviour. They do not focus their studies on human deliberation. By empirical research biologists study behaviour. Some non-empirical, also called *a priori*, insights from philosophy cannot be studied by empirical research in biology and vice

versa. Therefore, these disciplines deliver different insights on why reasonable people do unreasonable things. This makes the integration all the more interesting. Will these insights of human behaviour match or will they conflict?

Our project will be based on a literature study on the basis of the two disciplines named above, philosophy and biology. First the perspective of philosophy on akrasia will be discussed in chapter 1. In chapter 2 we will elaborate on the biological perspective on addiction. The sub-questions which will be addressed in the disciplinary chapter of philosophy are:

- ❖ How can akrasia be defined?
- ❖ What do people hope to obtain when acting akratically?
- ❖ What instigates akratic behaviour?

For biology the sub-questions are:

- ❖ How can addiction be defined?
- ❖ What is the role of associative learning in addiction?
- ❖ What role do representations have in addiction?

We will conclude with a chapter containing the integration of the two perspectives. To help our integrative process an interdisciplinary framework on akratic behaviour will be presented in the philosophical chapter. This framework, created in 2008 by Kalis et al., contains the integrated insights from philosophy and neuropsychopathology. The latter discipline uses empirical research to gain insights, just like biology, which makes the framework a means to bridge the gap between philosophy and biology. Therefore this interdisciplinary framework helps us to integrate the philosophical and biological findings on unreasonable behaviour in this project.

We will use integration techniques as presented by Repko in *Interdisciplinary research, process and theory* (2012). We aim to offer a more comprehensive understanding of why reasonable people do unreasonable things. We will do this by integrating the answers to these sub-questions and creating or finding common ground between the philosophical and biological findings. The results of our project can contribute to the academic debate about unreasonable behaviour. Furthermore, one can use this interdisciplinary understanding in further research of addiction and its treatments.

## Chapter 1. Akrasia: knowing and ignoring what is best

### 0. Introduction

When philosophers discuss irrationality, a concept that dominates the debate is “weakness of will”. The discussion on this concept will prove relevant for explaining why reasonable people act against reason and against their better judgment. This is often the case with people who are addicted, who know they do not act in their best interest when they give in to the temptation of drugs.

The will can be understood as a power that backs up one’s judgments and can oppose other motivational influences. Weakness of will then suggests that the will is not strong enough to counter opposing influences and does not have the strength to make an individual act on his or her values. However a more common conception of the will does not leave room for this explanation. This conception defines the will as a general power that leads people to action, also called *volition*. This makes it harder to justify that an act against one’s values can be attributed to weakness of will, because it is still an act (Kalis et al., 2008, p.402).

These various conceptions make weakness of will a controversial concept. Therefore, in philosophical discussions, the concept is often replaced by the ancient Greek word *akrasia*, meaning that one acts against one’s better judgment. Akrasia, as opposed to weakness of will, does not lean on a specific understanding of the will and therefore is a more pliable and useful concept (Kalis et al., 2008, p.402). The term akrasia is therefore the term we use in this project.

In contemporary philosophy, there are two questions at the centre of the discussions about akrasia, the first being “is it possible to act against your better judgement?”. The second concerns the explanation of akratic behaviour (Kalis, 2008, p.403). In this project we assume the possibility of akrasia in human action and aim to illustrate its presence in behaviour. Therefore the main focus of this chapter will be on answering the second question, by explaining akratic behaviour.

In the first paragraph we will give a definition of akrasia as presented by Aristotle. This is illustrated by findings of contemporary philosophers and neuropsychopathologists, who have developed an interdisciplinary framework that explains the various facets of akrasia. We will

discuss the main goal for acting akratically, namely obtaining pleasure or avoiding discomfort, in the second paragraph. We will describe the motivational aspects that instigate akratic action as described by philosopher David Hume in his “Treatise of Human Nature” in the last paragraph. Together these three paragraphs will present the philosophical explanation for why rational people can act irrationally. This disciplinary answer will be contained in the conclusion to this philosophical chapter.

## **1. Akrasia**

The sub-question that needs to be answered first is “How can akrasia be defined?”. The aim of this paragraph is to answer this question, by going back to the origin of the concept. After that this definition will be integrated with contemporary findings in philosophy and neuropsychopathology on the subject of akrasia. This will present a more complete overview of the concept akrasia as it takes both the origin of the concept and contemporary findings into consideration.

### **1.1 Aristotle’s akrasia**

Aristotle is seen as the first philosopher who discusses akrasia in his *Ethica Nicomachea* as Socrates before him denied the possibility of weakness of will. Socrates states that a person can only act in a wrong way because he lacks sufficient knowledge on the subject, hence a person who does possess correct knowledge would never act wrong deliberately (Aristotle, 1145 b 20-30). Aristotle disagrees with Socrates and says that a person can possess correct knowledge, but can act against it nonetheless (Aristotle, 1145 b 20-30). Aristotle writes that if we assume that people possessing knowledge are not easily persuaded to act against their knowledge, this can be contested by practical examples. According to Aristotle, some people with a mere opinion about something can be just as convinced of their perception as they would possessing objective knowledge on the same subject (Aristotle, 1146 b 25-30). An akratic person may have opinions based on his or her moral or reasonable beliefs, according to Aristotle, but fails to act on them as a passion to act otherwise can strongly influence an individual (Aristotle, 1145 b 10-15).

Akrasia can be defined as a state of being in which an individual acts against reason

under influence of passions, in other words: it is a state of being in which one acts out of lack of control over oneself (Aristotle, 1145 b 5-15). In English translations of the *Ethica* akrasia is often translated with the term *incontinence*. This is the opposite of the human disposition *continence*, the ability to control oneself and act on reason (Aristotle, 1145 b 5-15). It is however not by choice that an agent acts on these passions, but by mere lack of resistance to them. Otherwise it could be characterised as *self-indulgence*, where an individual chooses to deliberately act on passions (Aristotle, 1146 b 20-25). Self-indulgence in Aristotle's *Ethica* is seen as a vice that should be avoided (Aristotle, 1118 b 1-5). Unlike incontinence, self-indulgence cannot be cured, according to Aristotle. An incontinent person does regret the decisions he or she makes, whereas a self-indulgent person does not. This makes it easier for the incontinent person to correct his or her ways, be it on his or her own account or by persuasion of relatives (Aristotle, 1150 b 30-1151 a 5).

Aristotle further discusses the presence of akrasia in human deliberation by illustrating how it infiltrates logical reasoning about action. In logic there is always a universal premise involved as a starting point. In logical reasoning about action this universal premise concerns an opinion or belief. The second premise concerns the particular, dependent on sensory perception. When the conclusion results in a single opinion, the agent must act on this reasoning immediately (Aristotle, 1147 a 25-30). Aristotle illustrates logical reasoning about action by the following example:

P1 (universal opinion): everything sweet ought to be tasted

P2 (sensory perception): this is sweet

-----

C (conclusion): this ought to be tasted

(Aristotle, 1147 a 25-35)

When a person reasons like this, it should, according to Aristotle, lead him or her to tasting the sweet food. However a person can also hold the opinion that one should not eat sweet things, but he or she does find it pleasant to taste them and has the accompanying appetite for sweet things.



In this case, the universal opinion “one should not eat sweet things” conflicts with the pleasantness and appetite which drives him or her towards the action of eating sweet things against better judgment (Aristotle, 1147 a 30-1147 b 5).

## **1.2 Contemporary approach to akrasia**

Aristotle’s account of akrasia, defined as a state of being acts against reason under the influence of passions, is the leading definition in this project as it is the first philosopher to discuss akrasia in detail. To make this ancient definition applicable in modern times however we must take contemporary findings of akrasia into consideration and relate them to Aristotle’s akrasia.

According to interdisciplinary research done by philosopher Annemarie Kalis, philosopher Andreas Mojzisch, psychologist Tanja Sophie Schweizer and psychiatrist Stefan Kaiser akrasia can be broken down further into categories corresponding to neurological and behavioural symptoms in decision making (2008). These specific categories presented by Kalis et al. can be placed in the overall framework made by Kalis et al. combining insights of philosophy and neuropsychopathology (see table 1). The central element is that both the philosophers and psychiatrists see akrasia as a form of dysfunctional decision making. They have broken down decision making further into three stages, namely option generation, option selection and action initiation. Dysfunctions can be observed in all three of these stages as demonstrated below (Kalis et al., 2008, p.402).

Although these researchers acknowledge that akratic actions are often contrasted with psychopathological findings in philosophical discussions, they try to integrate these conceptions into a common framework. By doing this they do not assume, as commonly done in philosophy, that akratic actions are free whilst psychopathological phenomena are not. Freedom of will is often central to philosophical discussions, but is not, according to Kalis et al., a focus of psychopathological research. The question whether or not akrasia and psychopathological phenomena are free or unfree is therefore not further discussed by Kalis et al. (2008, pp.403 and

409).<sup>1</sup>

Aristotle distinguishes between two types of akrasia, namely *impetuosity* and *weakness*. The first type, impetuosity, is characterised by a lack of deliberation and the immediate following of one's passions (Aristotle, 1150 b 15-25). This can be described as a malfunction in an agent in the process of deliberating possible options to choose from and is called *dysfunctional option generation* (DOG) in the model of Kalis et al. (2008, pp.407-408). The cause for impetuosity, where no deliberation takes place, can sometimes be related to a type of slowness where an individual has difficulty to make plans and for that reason does not contemplate his or her options. This phenomenon is called *cognitive inertia* (Kalis et al., 2008, p.408). Cognitive inertia is a form of *hypogeneration* in the model of Kalis et al., described either as a lack of understanding of the situation, as is the case with cognitive inertia, or a lack of motivation to generate options. Hypogeneration can therefore also result in carelessness to think of multiple options. This is called *apathy* or *accedie* by Kalis et al. (2008, pp.407-408). Thus, impetuosity can be a result of apathy as well. The opposite of hypogeneration is *hypergeneration*, which means that an individual does generate an abundance of options for action to choose from. This often causes problems in the later stages of decision making, especially in option selection, resulting in a *loss of goal directedness*. This means that an increased activity can be observed in an individual, but that the actions lack focus (Kalis et al., 2008, p.409).

Weakness on the other hand, the other type of akrasia as presented by Aristotle, means that a person has deliberated, but acts against the result of the deliberation nonetheless (1150 b 15-25). This means that the agent does know what he or she values, but makes either no decision or the wrong one and it is an example of *dysfunctional option selection* (DOS), also called *decisional akrasia* (Kalis et al., 2008, p.409). Weakness illustrates the subcategory *making wrong decisions* especially well. Under this subcategory an agent can either act according to

---

<sup>1</sup> The contrast between the disciplines involved in this project, philosophy and biology, is similar to the philosophy-psychopathology contrast as stated by Kalis et al. (2008). In biology the question whether behaviour is free or not is not the main focus, as is the case in philosophical discussions. This contrast touches on the immense and important discussion of free will and free action, which unfortunately will not be discussed in this project. With providing interdisciplinary insights this project tries to find and create common ground between insights from biology and philosophy and not focus solely on the views that divide them.

*compulsivity* or *impulsivity*. The first concept means that the agent overestimates future consequences and chooses an overvalued action to prevent these consequences. The agent chooses this action only by considering short-term consequences. This either results in a delayed selection process or a selection of an option that does not reach the agents' ultimate goal. Impulsivity on the other hand is defined as the act in which an agent underestimates risks or has a strong preference for immediate reward. Sometimes both the underestimation of risks and a preference for an immediate reward apply. Impulsivity is considered to be closely related to addiction by psychiatrists whereas compulsivity falls into another category (e.g. obsessive compulsive disorder). Both impulsive and compulsive people have the capacity to generate multiple options, the problem is that these people do not evaluate and select the options correctly (Kalis et al., 2008, pp.409-410). Another form of decisional akrasia or DOS is *making no decision*. This is a form of *ambivalence*, meaning that the agent cannot decide what to do: it is the inability to select an option (Kalis et al., 2008, p.411).

When an individual acts immediately on his or her deliberations by performing a physical reaction, this can be categorised in the framework as *motor impulsivity*. It is hypothesised that an agent who lacks inhibitory control, acts almost immediately on action impulses (Kalis et al., 2008, p.412). Motor impulsivity falls, in the model of Kalis et al., under *facilitated action initiation* which is a subcategory to the main category *dysfunctional action initiation* (DAI). When agents suffer from DAI, this means that they have the intention to perform a certain action, but end up not performing anything at all or something else (Kalis et al., 2008, p.411). Another form of facilitated action initiation would be *hyperactivity*, which manifests itself when an agent acts against his decision at the very last moment, also called *last-ditch akrasia* (Kalis et al., 2008, pp.411-412). Another subcategory that falls under DAI is *impaired action initiation*, resulting in *motor inertia*, meaning the incapability of initiating action. This can be seen as a form of apathy towards possible action (Kalis et al., 2008, pp.411-412).

<b>Decision making stage</b>	<b>Dysfunction</b>	<b>Akratic symptom</b>
<b>Option generation</b> (given by situation at hand, by expertise or simply generated for the first time)	<b>Hypogeneration</b> (lack of motivation or cognitive understanding)	<b>Apathy</b> (not caring/accidie) <b>Cognitive inertia</b> (impairment resulting from lack of elaboration on plans)
	<b>Hypergeneration</b> (generation of an abundance of options)	<b>Loss of goal directedness</b> (increased generation, but not focussed on valued goals)
<b>Option selection</b> (weighing expected value, risk and uncertainty)	<b>Making wrong decisions</b> (agent knows what he/she values, but makes the wrong decision)	<b>Impulsivity</b> (underestimation of risk and/or strong preference for immediate reward) <b>Compulsivity</b> (overestimation of the risk of future events/consequences that an agent aims to prevent by choosing an action in the light of short-term consequences)
	<b>Making no decision</b> (agent knows what he/she values, but makes no decision)	<b>Ambivalence</b> (experiencing divergent emotions towards options, reducing the quantitative capacity to select options)
<b>Action initiation</b> (Cognitive and motor functions combined)	<b>Impaired action initiation</b> (a form of apathy towards taking action)	<b>Motor inertia</b> (incapability of initiating chosen action)
	<b>Facilitated action initiation</b> (acting immediatly/on all action impulses)	<b>Hyperactivity</b> (increased activity, but not focussed) <b>Motor impulsivity</b> (lack of inhibitory control)

**Table 1.** An overview of decision making stages and the dysfunctions and akratic symptoms that can occur in these stages. Based on framework as presented by Kalis et al. (2008).

Concluding, akrasia as defined by Aristotle is a state of being wherein an individual acts against his or her better judgment. Akratic behaviour can further be broken down into various forms that correspond to dysfunctions in the three decision making stages, option generation, option selection and action initiation as presented by Kalis et al. (2008).

## 2. Pleasure and avoiding discomfort

The second sub-question that is in focus here reads “What do people hope to obtain when acting akratically?”. We have explained what akrasia is, but now we have to make clear for what reason people act akratically. What is the goal of akratic action? We will consider here what Aristotle sees as the goal of akratic action, namely pleasure and avoiding discomfort. This is a result of his teleological reasoning that characterises his theory. This means that everything in his theory is explained by referring to its end or goal, called *telos* in ancient Greek (Press, 1999, p.72). We

will end this paragraph on a critical note, stating in what way Aristotle's theory on pleasure does and does not match with contemporary science.

Aristotle states that all living beings strive for pleasure or so called *hêdonê* in ancient Greek. Actions of non-human animals and humans alike show that pleasure is desirable to them, proving that it is in essence natural or good as Aristotle calls it (1153 b 30-35). It is essential for humans to rejoice in and be averse to the correct things to be able to have a good life (Aristotle, 1172 a 20-25). Although, he says humans have developed a special fondness for bodily pleasures and have devaluated other pleasures in the process (Aristotle, 1153 b 30-35). In striving for these bodily pleasures a person can become akratic, when he or she strives for these pleasures excessively and against better judgment.

Aristotle distinguishes two types of phenomena that produce pleasure, the one being necessary, the other desirable or worthy of choice. The first type concerns the bodily pleasures, meaning the pleasures involved in eating food or having sexual intercourse. It is in satisfying a craving for these pleasures that an individual can become self-indulgent, when he or she deliberately chooses to follow up on these cravings excessively. The second type is concerned with other pleasure generating phenomena, such as gaining respect, earning a good income and being victorious (Aristotle, 1147 b 20-30). In the pursuit of both these pleasures a person can become incontinent. Aristotle holds that a virtuous and temperate person strives for pleasure and the avoidance of discomfort, but does not do this excessively, whereas an incontinent or even self-indulgent person does (1153 a 25-30).

Ultimately, in Aristotle's opinion, the alluring aspects of pleasure can generate two types of people when they are not acting on reason, but following their passions: ones that have as their main objective to experience pleasure and ones that have as their main goal to avoid pain (Aristotle, 1150 a 25-30). Aristotle explains the appeal of bodily pleasure by stating that it can expel pain and discomfort. This makes it even more appealing than other sources of pleasure, according to him (Aristotle, 1154 a 25-1154 b 5). He further defines pleasures as activities and goals, not as processes. They are the accompanying feelings when humans are exercising their abilities (Aristotle, 1153 a 10).

Aristotle discusses the opinion that pleasure is a process of recovery from an unnatural or

painful state of being, to a natural state. He agrees that this can be true in case of hunger and the want for food. A person experiences the want for nutrition as an unpleasant feeling first, but when a person decides to eat something to recover from this hunger, this is accompanied by a pleasant feeling. However, this process does not apply to all forms of pleasure, Aristotle concludes. This is because other pleasures do not assume the presence of pain, such as the pleasure experienced when studying or the ones involved with other sensory perceptions like smell and sight (Aristotle, 1173 b 5-20).

According to Aristotle, pleasure perfects an activity. He states that every sensory ability corresponds with pleasure, making it the ideal and most pleasant object of one's actions. When a person functions correctly, the ideal object of action would be the most outstanding and inherent object that the ability deployed in the action can have as a goal. The most pleasant activity is then, as Aristotle holds, the most perfect (1174 b 19-25). Pleasure in this, is the phenomenon that accompanies and perfects the activity (Aristotle, 1175 a 4). Aristotle even states that it perfects the lives of people and that it is naturally a reasonable thing to strive for pleasure, because of its perfecting faculty that is desirable in itself (Aristotle, 1175 a 15). However Aristotle's description of pleasure as a goal being able to perfect an action seems somewhat doubtful. In contemporary science it is not common to speak of phenomena as having a 'goal' related to their nature, or inherent to them. Teleological reasoning in biology has largely been put aside for example. This raises the question: is his account of pleasure still acceptable in this day and age?

Although Aristotle's teleology is in some aspects outdated, seeing pleasure as an important influence in motivating beings is in some scientific models still an acceptable theory. The relevance of pleasure for human action becomes apparent in modern research, as stated in the article of Kalis et al. (2008). The involvement of reward or the delay of punishment, which can be described as pleasure and avoiding discomfort, influences the decision making process especially in the stage of option selection (Kalis et al., 2008, p.406).

Another question one might ask is "Does an addict experience pleasure when using drugs?". We aim to answer this question in the fourth chapter where the disciplinary insights are integrated. This results in a more comprehensive understanding of why reasonable people do unreasonable things.

In short Aristotle sees pleasure or the avoidance of discomfort as the main goals for action in humans and non-human animals. Bodily pleasure has the ability to expel pain, which can be the goal of an action as well. This makes bodily pleasures more appealing than other pleasures, because of this extra feature. People are therefore more susceptible to becoming excessive in striving for bodily pleasures against their will, making people akratic. The involvement of pleasure as a reward for behaviour or the delay of punishment influences the individual's decision for action.

### **3. Passions, imagination and malfunction**

Lastly the question “What instigates akratic behaviour?” needs to be addressed. We have considered what akrasia is and what people hope to obtain by performing an akratic act, but now we need to clarify what phenomenon instills akrasia in human deliberation. We will consider Aristotle's and Hume's account of passions as driving forces of human behaviour. These passions, as Kalis et al. state, are the value individuals attach to their options, making the option more or less desirable and this can be determining for the choice of action, without the involvement of (much) deliberation. Thus the following of one's desires above deliberation can be the cause for akratic behaviour. Where Aristotle describes human action in the light of his teleological reasoning as everything having a goal, Hume takes a different approach. His reasoning can be placed in the British empiricist tradition and he explains human action in a coherent framework of causality that has resulted from his account of human understanding (Morris, 2009). This causal and more empirical approach lends itself to be integrated with other empirical studies from, for example biology. In chapter 3 we will come to this integration.

#### **3.1 Passions**

As stated by Aristotle, obtaining pleasure or avoiding discomfort can be seen as a goal for beings to act on (1153 b 30-35). The pleasantness makes it appealing for an individual to strive for pleasure and the passion or urge to experience this pleasure motivates the individual in his or her choice for action (Aristotle, 1147 a 30-1147 b 5). When an individual lacks self-control and gives

in to the strive for these pleasures against better judgment, this is a manifestation of akrasia (Aristotle, 1145 b 10-15). Passion and appetite can therefore be seen as the instigators of akratic behaviour and the experience of pleasure as the goal.

According to the research carried out by Kalis et al. (2008), there are two preconditions for action initiation. One is that the physical cues, also called motor-functions, for actions must be present in the mental system. The second being that the options for action need to have an appeal or affective value, otherwise the agent will not select the option (Kalis et al., 2008, p.405). For example an agent can be hungry and have the option to make a sandwich or eat an apple. The first precondition for either of the optional actions is that the agent must know how to physically perform the actions. The second precondition for the actions is that they must be attractive, where one option is possibly more attractive than the other one. It is stated by Kalis et al. that agents evaluate their desires and assess the affective value they attribute to an option, but that the ultimate choice can be strongly influenced by motivational aspects that go beyond deliberation (Kalis et al., 2008, p.409). The motivational aspects Kalis et al. refer to are all influences that do not involve deliberation, such as instincts and reflexes etcetera.

This is in line with the views Hume holds on what causes human action. In his Treatise he describes what ultimately motivates human action, thus it can be seen as a theory of human motivation for action. He states that everything that comes into existence, including action, must have a cause. He further says that reason cannot be a motive or cause for action on its own (Hume, 2003, pp.57 and 294). Passion or desire always plays a part in the instigation of human action (Hume, 2003, p.294). He distinguishes between 'calm desires' and 'violent passions'. The first can hardly be noticeable or felt by an individual. It presents itself either as an instinct or a general want for good and aversion to evil (Hume, 2003, p.297). A violent passion arises quickly and overshadows consideration and deliberation of the individual. Violent passions produce sensible emotions (Hume, 2003, p.297).

Both calm desires and violent passions have the ability to bring about action. Hume defines strength of will as the prevalence of calm desires above violent passions. Weakness of will can then be described as the triumph of violent passions, that results in a person acting against one's interests and is characterised by the lack of deliberation (Hume, 2003, p.297). The



will in Hume's theory is described as "the internal impression we feel and are conscious of, when we knowingly give rise to any new motion of our body, or new perception of our mind" (Hume, 2003, p.284). The will therefore is a factor in one's conscious mind that can cause bodily motion or bring about different dispositions. Failure or weakness of the will as Hume describes it, namely the fact that violent passions have the upper hand and deliberation cannot take place, corresponds with Aristotle's definition of *akrasia*: acting against one's better judgment due to a lack of self-control. Acting against one's better judgment herein can be characterised as an act against your interests, the lack of self-control is the triumph of violent passions or desires resulting in a lack of deliberation.

### **3.2 Imagination**

The question that we are left with is "Where can passions be placed in the decision making process?". It turns out that vivid representations of possible options can persuade individuals to act in a certain manner. It has become more and more apparent that 'action-representation' and 'goal-representation' are mostly coherent with each other in the decision making process (Kalis et al., 2008, p.405). This coherence of goal and action is consistent with the definition Kalis et al. give to the concept 'option': a possible action and the expected affective value associated with that action (Kalis et al., 2008, p.405). This correlation of goal and action is described by Hume as well, as will become apparent in the following part.

Hume states that humans create representations for optional actions in their mind, in which the goal and action are interwoven (Hume, 2003, pp.61 and 66). In his theory action-representation and goal-representation can be described as forms of 'imagination', a concept used to describe human perception that both contains the 'idea of causation' and explains the cause for human action. The idea of causation is relevant to the definition of imagination and this idea is derived from experience according to Hume (2003, p.65). When circumstances are similar, humans often assume that future events will develop in the same way as past situations they encountered. When an individual has repeatedly experienced pleasure as a result of taking a drug, he or she will expect to experience this again in the future.

Predictions of future events are represented in human imagination, also described as the union of impressions (Hume, 2003, p.66). Hume describes how human perceptions of causes and

effects are made up out of impressions derived from sensory memory (Hume, 2003, p.61). These impressions can result in a belief or opinion. A belief can be defined as a more vivid and immediate impression that accompanies the sensory memory (Hume, 2003, pp.62 and 69). This indicates that beliefs simulate the feelings that past experiences have brought along, in order to predict future outcomes of actions. Beliefs can form the basis for further reasoning, but always find their origin in the impressions, namely sensory memory (Hume, 2003, pp.76 and 86). Beliefs on causation, or cause-effect beliefs as we will call them, endorse the choice for certain actions, because an individual can form a representation of the development of these actions and can decide whether the action is worthwhile or not (Hume, 2003, p.86). Imagination thus includes representations of actions based on cause-effect beliefs and can give rise to action.

### **3.3 Akrasia as a result of malfunctioning imagination**

It has been shown that in Hume's theory the triumph of violent passions can be explained in coherence with Aristotle's views on akrasia. Some of the dysfunctions in decision making, as described by Kalis et al., share their foundations with Hume's findings as well. This becomes apparent in the description of the difficulty some agents have with translating options into action. A result of this problem is that some people do not choose the proper means for their desired effects. In other words they do not select the correct option to act on. For example a person may want to calm his or her nerves, but instead of taking a power nap, which is not bad for his or her health, he or she decides to smoke a cigarette. This can be called *instrumental irrationality* as the person chooses the 'wrong' or less healthy means to the desired end (Kalis et al., 2008, p.407). This takes place mostly in the area of option selection and can either result in impulsivity, compulsivity or ambivalence, which all fall under decisional akrasia (Kalis et al., 2008, pp. 409-411) (see table 1).

Impulsivity can then be explained as a malfunction of imagination in which the overestimation of the predicted rewards, namely pleasure, occurs. The individual can experience violent passions towards this impulsive action, which overshadow any deliberation. Impulsivity is often related to addiction by psychiatrists (Kalis et al., 2008, pp.409-410). Compulsivity can be seen as the overestimation of the predicted risks and as a malfunction of the individual's imagination as well. The agent experiences violent passions towards choosing the action that

aims to prevent the predicted risks and fails to deliberate. Ambivalence only then occurs when the representations of various options by the imagination appear to be of equal appeal, resulting in an indifference towards all options.

Pleasure can be seen as the goal of an action. Passions urge the individual to experience this pleasure and can drive an individual to action. Beliefs on causation, or cause-effect beliefs, are vivid representations of actions in the mind, which simulate the feelings that will accompany this action. Cause-effect beliefs therefore enable the individual to weigh his or her options and the outcomes in the decision making process and make it possible for him or her to decide whether an action is worthwhile or not. The combination of the passions, also called desires, and the decision making process can be summed up in the concept of 'human motivation for action'. The representations that are present in an individual's imagination can be misguided. In the case of malfunction, the individual's representations present a simulation of an action in which certain predicted rewards or risks get over- or underestimates. This results in certain akratic behaviour where the individual selects the wrong options for action to obtain a set goal.

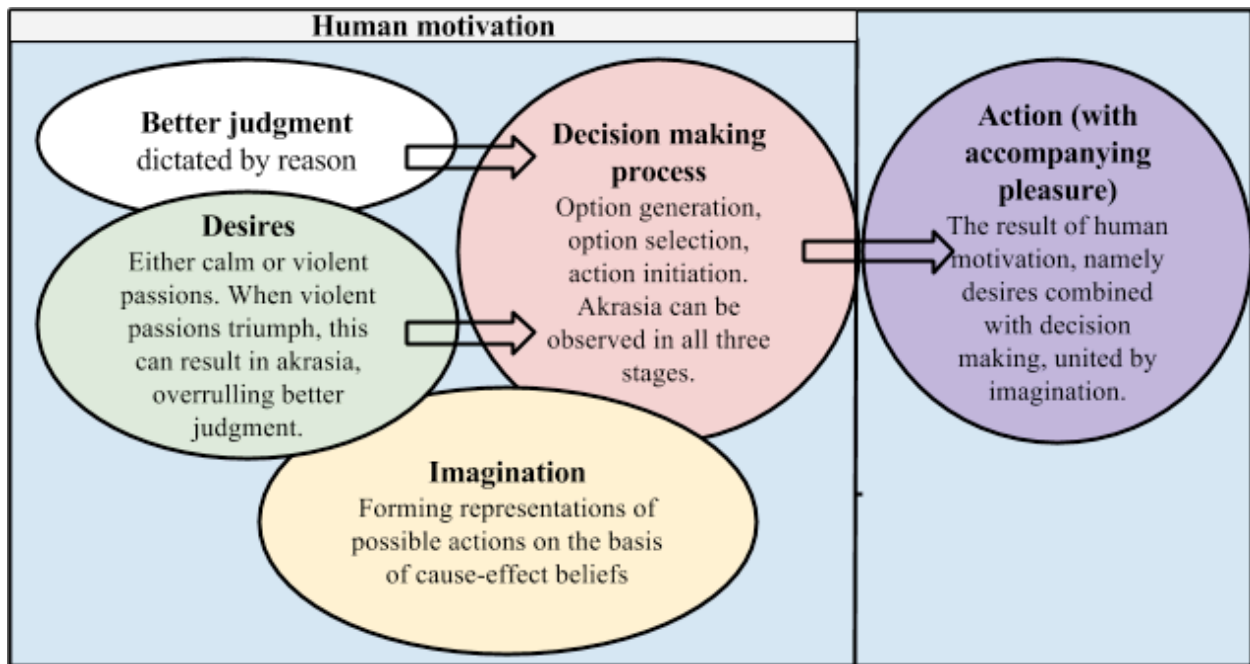
#### **4. Conclusion**

Akrasia can be defined as the lack of control over oneself and acting against better judgment. Kalis et al. (2008) have developed a more fine grained and interdisciplinary framework in which different forms of akrasia are defined. It turns out that akrasia can emerge in three stages of decision making, namely 'option generation', 'option selection' and 'action initiation'. In all three stages akrasia can cause a dysfunction in the decision making process.

The driving forces behind akratic behaviour are passions and representations of possible actions based on cause-effect beliefs. In this process better judgments or deliberation gets overruled. When an individual holds the experience of pleasure as a goal, this is often instilled in the individual by a desire to pursue pleasure, because of its desirability. Positive former experiences with certain actions will influence the individual in the decision making process even more. In akratic behaviour this desire can become more and more apparent and violent to the extent that the agent acts on his desires impulsively, almost without deliberation.

A deciding factor in the choice for action is the experience of pleasure or the avoidance of discomfort as a result of this action. Although Aristotle defines pleasure as essential to a good and human life, akratic people strive for pleasure excessively.

Therefore passions and desires can be seen as the instigators of akratic behaviour. Akrasia can be observed in the process of decision making and akratic people ultimately have the experience of pleasure as a goal. The process of human motivation resulting in (akratic) behaviour is presented in figure 1 below.



**Figure 1** An overview of the process of human motivation leading to action in which akrasia can be observed.

## **Chapter 2. Addiction: the brain fooled by drugs**

### **0. Introduction**

Biologists search, by means of empirical research, for physical explanations for certain phenomena. In biology the hypothetico-deductive approach is used to derive explanations from laws or theories (Repko, 2012). This implicates a hypothesis is formulated which can be falsified by means of experiments. Especially in neurobiology, researchers have great interest in the topic of addiction, since many aspects of addiction are not yet fully understood. Most addicted individuals know they shouldn't use drugs, but they take them anyway. So, addiction is an example of reasonable people doing unreasonable things.

One can be addicted to a range of things, though most research in biology focuses on substance dependency, for example substances like cocaine, opiates and nicotine. Therefore we describe the phenomenon of addiction and relevant theories in the light of drugs.

In this biological chapter we will discuss theories and studies on addiction, that focus on the brain, conditioning and neurocognition. Studies that focus on the brain and conditioning often use animal models to explain human addiction. A neurocognitive approach distinguishes human animals from non-human animals and also studies the humans as a rational being.

In the first paragraph of this chapter we will present a definition of addiction and describes its characteristics. After that, the essential concepts of 'liking' and 'wanting' will be discussed in relation to addiction. We will end this chapter with a paragraph on a neurocognitive approach to addiction and a conclusion of the literature research.

### **1. Defining addiction**

In this paragraph we will answer the sub-question "How can addiction be defined?" by giving a definition of addiction and describing its characteristics and underlying neurological mechanisms. This way a theoretical framework is sketched in which the insights from later paragraphs can fit.

In the book *Neuroscience* Purves et al. (2012) addiction is described as follows: "Addiction (also called substance dependence) is a persistent disorder of brain function in which compulsive drug use occurs despite serious negative consequences for the afflicted

individual” (p.128). This definition includes the negative effects that the compulsive use of drugs have and emphasises the role of brain functions, on which neurobiologists focus.

Researchers are in search of explanations for common characteristics of addictions, for example the ‘withdrawal syndrome’, tolerance, craving and relapses as described by Purves et al. (2012, p.128). The withdrawal syndrome is a common characteristic and emerges when an addict does not take drugs. Symptoms of this syndrome include physical, psychological and motivational aspects, in general the opposite of the positive experience induced by the drug itself. For some drugs a certain tolerance can arise. Tolerance means a reduction in response to the drug when repeatedly used. This can result in a person taking more drugs to obtain the same effect. Another defining characteristic is the intense wanting of drugs, also called craving. These characteristics explain some reasons why fighting addiction is so hard. Therefore relapses can occur long after a person has stopped using drugs.

### **1.1 Internal versus external stimuli**

Humans are influenced by both cues from inside the body and cues from the environment, or internal and external stimuli. An empty stomach can instill a sense of hunger, just as the sight of a tasty hamburger can. The relation between those internal and external cues has been described by McNamara & Houston (1986). Their idea is that whether a person performs an action or not, depends on both how attractive a stimulus is and how much a person ‘needs’ it. In the example of the hamburger: when someone is very hungry he is more easily persuaded to eat a less appetising hamburger than when he is already satisfied.

The same mechanism goes for addiction; both internal and external cues can trigger addictive behaviour. Internal stimuli can be identical to symptoms of the withdrawal syndrome. The sight of someone lighting a cigarette is a good example of an external cue that triggers drug taking behaviour. But how exactly do these internal and external stimuli trigger behaviour that leads to liking, wanting and eventually taking drugs?

### **1.2 The role of opioids and dopamine in addiction**

To find an explanation for how stimuli can trigger addictive behaviour, we will zoom in on the neurological basis of addiction. In neurobiology a lot of research is done on the neurological

aspects of addiction, but there is a lot still to be discovered. When talking about the cellular aspects that are the basis of the neurobiology of addiction, ‘opioids’ and ‘dopamine’ play a key role. Opioids and dopamine are substances that can function as a neurotransmitter, a substance used for the communication between neurons (cells in the brain). One neuron produces a certain neurotransmitter and another neuron possesses receptors where this specific substance can bind. This is how a signal can be transmitted through the brain (Purves et al., 2012).

Opioids are important substances in the experience of pleasure that accompanies certain behaviour. They act to evaluate behaviour (Spruijt, Bos & Pijlman, 2001). The body produces these opioids and when these substances bind to their receptors this, among other effects, relieves pain. Besides the opioids the body itself produces, there are some drugs that mimic these opioids. These drugs, for example heroin, are called opiates: they can also bind to the opioids receptors (Purves, 2012, p.128).

The dopamine system is thought to play a role in the reinforcement and motivation of drug taking behaviour (Purves et al., 2012). This system is activated after the opioid system. Both synthetic and natural drugs affect the dopaminergic circuit (i.e. the circuit in the brain related to dopamine). Drugs can work directly or indirectly on this circuit, but all addictive drugs increase the level of dopamine in a certain area in the brain, namely the *nucleus accumbens* (Hyman, Malenka & Nestler, 2006). It is thought that dopamine binds the hedonic properties (pleasure) of a goal to the motivation of behaviour (Hyman, Malenka & Nestler, 2006). This means the reward of using drugs becomes associated with a certain behaviour.

For example cocaine works as follows in the brain: it inhibits a substance called DAT (dopamine active transporter), which normally removes dopamine between two neurons. When using cocaine, DAT is inhibited and less dopamine is removed. This results in a higher concentration of dopamine and with that a ‘higher motivational value’ to the use of cocaine (Purves, 2012). A higher motivational value is experienced as an increase in wanting to take cocaine.

Concluding, behaviour is evaluated and reinforced by the opioid system and the dopamine system. Drugs work directly or indirectly on these systems, resulting in excessive pleasure, drug taking behaviour and in some cases addiction.

### 1.3 The economy of behaviour

Individuals act to maximise the benefits and minimise the costs of their behaviour. Spruijt, Bos and Pijlman (2001) describe this as ‘the economy of behaviour’ and state that the ‘common currency’, to express the value of behaviour, is pleasure. With help of this common currency the values of actions are compared and the most valuable action is performed. They state that, under normal conditions, behaviour that results in pleasure is good in the long run (Spruijt, Bos & Pijlman, 2001, p.147). The most rewarding action, resulting in the most pleasure, is normally the most useful action in terms of survival.

Robinson and Berridge (1993) use the terms ‘liking’ and ‘wanting’ in relation to addiction. Liking and wanting are generated by the brain and only the conscious part of it is subjective (Berridge, 2004). Hence the terms liking and wanting apply to the reactions which can be measured in the brain, whether it be conscious or not.

Liking refers to the hedonic aspect, the ‘pleasure’, that is the immediate result of an action. For example: using drugs can give a pleasant feeling. As described earlier, the opioid system is involved in the generation of pleasure. Pleasure is the measurement of the value of an action, the common currency. Wanting drives the behaviour of an animal to obtain or to get access to a goal. An excessive form of wanting is often called ‘craving’. Drug addicts crave drugs. This craving drives them to buying and using drugs, for example. The wanting mechanism follows the liking mechanism. Spruijt, Bos and Pijlman state wanting can be seen as “the imprinting of what was once liked and disliked.” (2001, p.160). Therefore the action that results in the most pleasure is normally the most wanted. But we will point out this is not always the case.

Liking and wanting are separate brain mechanisms, respectively generated in the *lateral hypothalamus* and the *mesolimbic dopamine system* (Berridge, 2004). These separate mechanisms work together to elicit behaviour of animals. It is not clear why separate brain mechanisms evolved for liking and wanting, nor in which order they evolved (Berridge, 2004). Berridge states “liking and wanting normally go together, but they can be split apart under certain circumstances, especially by certain brain manipulations.” (2004, p.16). These ‘certain



brain manipulations' could be a result of drug use. In paragraph 2 we illustrate the effects of drugs on the mechanisms of liking and wanting.

## **2. Liking and wanting**

Certain behaviour, for example using drugs, is followed by a reward. Rewards are experienced as “making things better” (Berridge, in Hyman, Malenka and Nestler, 2006). Organisms pursue rewards, leading to behaviour to obtain a rewarding goal. Mostly a reward follows actions that enhance survival or reproduction. In this way rewards increase the frequency of performing behaviours that are followed by rewards, with that increasing survival. Therefore a reward guides behaviour (Spruijt, Bos & Pijlman, 2001). A reward can be parsed into three psychological components and their underlying neural mechanisms (Berridge, Robinson & Aldridge, 2009): liking, wanting and learning.

Behaviour to obtain a rewarding goal arises via three learning processes, described by Hyman, Malenka and Nestler (2006, p.567) 1) liking the rewarding goal, 2) learning cues that predict the availability of the goal and how to obtain it, and 3) assigning motivational value to the reward (how important is this goal). The three learning processes overlap with liking, wanting and learning, mentioned above. In this paragraph we will discuss liking in more detail. To answer the sub-question: “What is the role of associative learning in addiction?” we will focus on the ‘learning’ component in combination with liking and wanting.

### **2.1 Liking and reinforcement theories**

As explained above, liking and wanting can be distinguished and have different underlying neural mechanisms. Liking is generated in the lateral hypothalamus and refers to the pleasure certain behaviour elicits (Berridge, 2004). For example the sweet taste of food or, a more relevant example for our project, the rush experienced by the use of drugs. Liking can also occur without a person being conscious. It can be measured in behaviour or physiology (Berridge et al., 2009).

Two categories of theories on addiction based on liking can be distinguished, namely positive and negative reinforcement theories. These reinforcement theories are based on

instrumental learning, a certain behaviour is associated with a so called 'reinforcer'. This reinforcer can either be a reward or a punishment. Positive reinforcement means the frequency of a behaviour is increased when it is followed by a reward. In negative reinforcement a removal of an unpleasant stimulus elicits an increased frequency of certain behaviour. In the case of drugs this reinforcer is the use of drugs. This paragraph briefly describes Pavlovian and instrumental conditioning and the positive and negative reinforcement view.

### **2.1.1 Pavlovian and instrumental conditioning.**

Pavlovian and instrumental learning are both forms of conditioning (learning) by association. We will briefly explain the conditioning mechanisms and relevant terms (based on Kirkpatrick & Hall, 2005, pp.152-156)

*Pavlovian (classical) conditioning* relies on forming an association between two stimuli. An *unconditioned stimulus* (UCS) is a stimulus that evokes a certain *response* without learned association. For example the sight of food causes drooling: this process is innate and does not require any learning. On the other hand a *conditioned stimulus* (CS) is a stimulus that is associated with another stimulus by learning. The well known example of Pavlovian learning about the dog and the bell is described below. When a dog is presented food (UCS) it will start to drool (this is called the response). If the moment when food is presented is accompanied by the sound of a bell (CS) an association will be formed between the CS and UCS. After some trials the dogs will respond to the sound of a bell (CS) by drooling, even without the presence of food (UCS).

In short: CS -> response. CS+UCS -> response. UCS alone -> response.

*Instrumental conditioning* is distinguished from classical conditioning by the reward or the punishment that follows a certain behaviour (response). This reward or punishment is called a *positive* or *negative reinforcer*, respectively. In this type of associative learning a certain response is associated with a reinforcer. If a behaviour is associated with a reward, an animal will perform this behaviour more frequently: the response is *reinforced*. This reward could also be the removal of an unpleasant stimulus. On the other hand, if a certain behaviour is punished, it will be less often performed. An often used example of instrumental conditioning with a positive reinforcer

is the skinner box, in which an animal receives food when it presses a lever. One will observe an increase in frequency of lever pressing when the association between a behaviour and a reinforcer (reward) becomes stronger.

### **2.1.2 Positive and negative reinforcement theories**

Positive reinforcement theories view the effect of drug taking as positive reinforcer. Drugs are used because they cause a positive affective state, because one 'likes' it. People react to drugs as if they are rewards. As a result the frequency of drug taking increases.

Negative reinforcement theories however state that drugs facilitate an escape from distress. Drugs counteract the effects of the withdrawal syndrome. Recall from paragraph 1 that the symptoms of drugs withdrawal are in general the opposite of the effects of the drug itself. The reinforcer in the negative reinforcement theories is the reward that counteracts the symptoms of withdrawal. Using drugs result in the removal of an unpleasant stimulus. Another negative reinforcement view states that people take drugs to escape from pain, anxiety and depression that occurred before they took drugs.

We have seen some theories on addiction based on liking. Although the positive and negative reinforcement theories might seem quite plausible at first sight, Berridge and Robinson (1993) argue that both these theories cannot fully explain all aspects of addiction, but stress that reinforcement plays some role in addiction. They give many arguments for their statement. We will shortly summarise some elements of their reasoning (see Berridge & Robinson, 1993, for the full article).

Berridge and Robinson (1993, pp.251-254) state that positive reinforcement theories aren't sufficient explanations of addiction, since drugs cannot provide so much pleasure that it outweighs the negative consequences, like loss of health, job and friends. Also negative reinforcement theories cannot explain addiction. The moment that someone starts to seek for drugs does not correlate with the moment of maximum intensity of the symptoms of the withdrawal syndrome (Wise & Bozarth in Berridge & Robinson, 1993). This does not support the statement of the negative reinforcement views that addicts take drugs to counteract symptoms of the withdrawal syndrome. Also studies show that animals can be addicted to drugs without suffering a withdrawal syndrome at all (Wise & Hoffman in Berridge & Robinson, 1993). On top

of that, both categories of reinforcement theories do not explain the effects of external stimuli on relapses after years of not using drugs.

Instead of the reinforcement theories of addiction that are based on instrumental conditioning, Berridge and Robinson (1993) propose a focus on classical conditioning and wanting. In paragraph 2.2 we will explain a theory of addiction focusing on wanting.

## **2.2 Wanting and the incentive-sensitization theory of addiction**

As described in 2.1, positive and negative reinforcement theories cannot fully explain all aspects of addiction. Therefore, Berridge and Robinson proposed an incentive-sensitization theory of addiction in 1993. In this paragraph we will explain this theory, based on the article of Berridge and Robinson.

The incentive-sensitization theory of addiction can be seen as a ‘neuroadaptationists model’. This is a model that focuses on changes in the brain. Berridge and Robinson state: “It is proposed here that the defining characteristics of addiction (craving and relapse) are due directly to drug-induced changes in those functions normally subserved by a neural system that undergoes sensitization-related neuroadaptations.” (1993, p.249). In their theory drug-induced changes in the brain arise by ‘sensitization’. Sensitization is a term used in neurobiology, which means in short that an amplified response on a stimulus occurs (Purves et al., 2012). In more detail, this means more neurotransmitter (see paragraph 1) is released and this results in a stronger response. Individuals sensitised for a certain stimulus produce more dopamine than non-sensitised individuals. On the other hand tolerance means decreased sensitivity, that results in decreased response to a certain stimulus. Sensitization and tolerance are mechanisms proposed to alter ‘the value’ of a certain behaviour. Sensitization leads to increased value, whereas tolerance leads to decreased value, in case of an abundance of a reward (Spruijt, Bos & Pijlman, 2001).

The mechanism of sensitization is proposed to facilitate the attribution of incentive salience to certain stimuli and actions. Berridge and Robinson define incentive salience as: “the attractiveness of external stimuli, events, places and their mental representations; their ability to capture attention” (p.280, 1993). When incentive salience is attributed to stimuli, they become ‘incentive stimuli’ that trigger wanting. Normally this happens to stimuli relevant for survival, and in general drugs are not essential to survival. Sensitization as a result of drug taking results

in a 'pathological enhancement' in the incentive salience that the nervous system (mesolimbic dopamine system) attributes to the act of drug taking (Berridge & Robinson, 1993, p.249). The neural system of wanting becomes sensitised as a result of repeated drug taking, so drugs become more and more wanted and eventually a craving for them emerges. In addition, classical conditioning of drugs and drug-related stimuli occur. This associative learning results in incentive salience being attributed to drug-related stimuli.

The 'conditioned place preference model' is a good example of learning cues that predict the availability of a drug. This model states that an animal prefers the place where it received drugs over a 'neutral' place (Hyman, Malenka and Nestler, 2006). A study of rats as model organism supports this model (Domjan in Hyman, Malenka & Nestler, 2006). At one location the rats were injected with drugs and at the other location with a saline injection. Later the rats preferred the similar location where they previously received drugs. This explains why an addict should leave his old environment in order to successfully quit drugs. The conditioned place preference model is an example of Pavlovian conditioning (see 2.1.1 for a brief overview of Pavlovian and instrumental conditioning).

Berridge and Robinson (1993) also emphasise that only the neural system for wanting becomes sensitised and not the system responsible for liking. Some drug taking results even in tolerance after repeated use, liking decreases. "Usually a brain 'likes' the rewards that it 'wants'. But sometimes it may just 'want' them." (Berridge et al., 2009, p.67). Drugs fool the brain. The attribution of disproportional incentive salience to drugs can explain this. In this aspect the incentive-sensitization theory distinguishes itself from other theories based on both liking and wanting.

### **3. A neurocognitive approach to addiction**

Now we have seen some basic characteristics of addiction and the mechanisms of liking and wanting the question arises: how do these and other concepts relate to addiction? Many models of addiction based on neurobiology have been published, though these models are often based on animal models and do not describe humans as rational beings. Some neurocognitive models do

focus on aspects like decision making, consciousness and representations, that are the primary focus of philosophical perspectives.

Noël et al. (2013) propose a model of human addiction that does include these cognitive aspects, it focuses on the role of choice in addiction. This neurocognitive model is described in a literature review of the neurobiology of addiction and features three systems: the impulsive system, the reflective system and the insula. In this paragraph we will describe these systems, in specific the role of the insula, to answer the sub-question: “What role do representations have in addiction?”.

### 3.1 A three-systems neural model of addiction

The three systems proposed by Noël et al. (2013) are independent but interacting and different neural regions can be ascribed to each system. Together the systems regulate behaviour, but are imbalanced in case of addiction. Noël et al. (2013) argue that addiction is associated with “disrupted neural systems for will power, which refers to the capacity for choosing according to long-term, rather than short-term, outcomes.” (p.632). This results in bad decision making, choosing to take drugs despite the outweighing negative consequences. They argue the error might be in one, two or all of the systems.

To sketch a conceptual framework we will briefly describe these systems and how their functioning is changed by addiction, table 2 provides an overview of these systems (based on Noël et al., 2013).

System	Function	Neurological region
<i>Impulsive system</i>	Automatic, habitual and salient behaviours, (wanting) Fast, automatic and unconscious	Amygdala-striatum dependent
<i>Reflective system</i>	Controlling basic impulses: decision-making, forecasting the future consequences of a behaviour and inhibitory control. Slow, deliberately and conscious	Prefrontal cortex dependent

<i>Insula</i>	Integrates interoception states into conscious feelings and into decision-making processes that are involved in uncertain risk and reward	Insula
---------------	---	--------

**Table 2** Three neural systems, separate but interacting, that are imbalanced in an addict. Based on Noël et al., 2013. See text for further explanation.

The impulsive system generates basic impulses (wanting) as a result of reacting on addiction-associated information. How this information is used can be explained by associative learning, see paragraph 2. Abused substances like drugs modify the dopamine levels (playing a role in reward) and therefore alter decision making. The process of the impulsive system is fast, automatic and unconscious. It comprises the brain areas the striatum and the amygdala.

The reflective system is located in the prefrontal cortex, that plays a role in decision making. This system plays a role in controlling impulses (wanting), that are generated in the impulsive system. Long-term goals are integrated in choosing an action. Both information from the working memory and the long-term memory are used to choose for certain behaviour. Studies (Bechara et al. in Noël et al., 2013) of patients with damage in a part (ventromedial region) of the reflective system show these patients can no longer observe and act on social conventions. This is an example of what kind of information the reflective system uses in decision making. The reflective system is slow, deliberative and conscious.

Lastly, the insula plays a role in translating the interoceptive (bodily) stimuli to how a person subjectively experiences their physical conditions, like desire. Patients with damage in their insula are more than 100 times more likely to stop smoking than people with an intact insula (Naqvi et al., 2007). The interoceptive representation of smoking is not (well) translated into a conscious experience by individuals with brain damage. This apparently makes these individuals less sensitive to addiction to smoking. These patients with brain damage are characterised by Naqvi et al. as possessing “the ability to quit smoking easily, immediately, without relapse, and without persistence of the urge to smoke” (2007, p.531). The study of Naqvi et al. (2007) emphasises the role of the insula to feel the urge to light a cigarette. The insula and it’s role in relation to addiction will be described in more detail, in the next section of this paragraph.

### **3.2 The insula**

As mentioned before, as one of the three systems of the neurocognitive model of addiction proposed by Noël et al. (2013), the insula plays a role in the conscious urge to take drugs. Research suggests that a crucial role for the insula is translating interoceptive (bodily) signals to conscious feelings (Naqvi & Bechara, 2008). For example the insula plays a role in the translation of physiological symptoms of the withdrawal syndrome to subjective emotional discomfort (Naqvi & Bechara, 2008). This neural substrate, called the insula, has not been paid much attention to in scientific research so far, but is seen more and more as a neural substrate that plays an important role in addiction.

It seems that the urge to take drugs is depending on the sort of drugs (Naqvi & Bechara, 2008). For example the urge of smoking a cigarette feels different from the urge to use cocaine. This can be explained by the activation of the representation of interoceptive effects of a specific drug when an addicted individual is exposed to drug related cues. An example of these interoceptive effects is the airway stimulation when smoking. These representations of interoceptive effects are based on earlier experiences.

Naqvi and Bechara (2008) hypothesise that the interoceptive effects of drug use increase over time, by structural changes in the brain (insula and related areas) induced by the release of dopamine. They propose two learning effects of these changes. First, interoceptive effects of drugs, for example the airway stimulation in smoking, become more positively experienced, more 'liked'. Second, associations between these interoceptive effects and drug related cues, for example a package of cigarettes, are formed. Because of this, these cues trigger wanting. Furthermore Naqvi and Bechara (2008) state these representation of interoceptive effects might play a role in decision making.

But why do addicts still choose to take drugs, despite the fact that this has numerous negative consequences? Three reasons are described by Naqvi and Bechara (2008): 1) the representations of interoceptive effects of drugs are magnified by the increased incentive salience attributed to drugs, resulting in more wanting; this representation might outweigh the representation of the negative effects 2) the positive effects of using drugs are immediate and certain, while the negative effects are delayed and more uncertain 3) addicts have impairments in



decision making (for example assessing risks). Concluding: “these interoceptive representations have the capacity to ‘hijack’ the cognitive resources necessary for exerting inhibitory control to resist drug use.” (Bechara in Naqvi & Bechara, 2008. p.63). Drugs ‘fool’ the brain and this makes rational decision making harder.

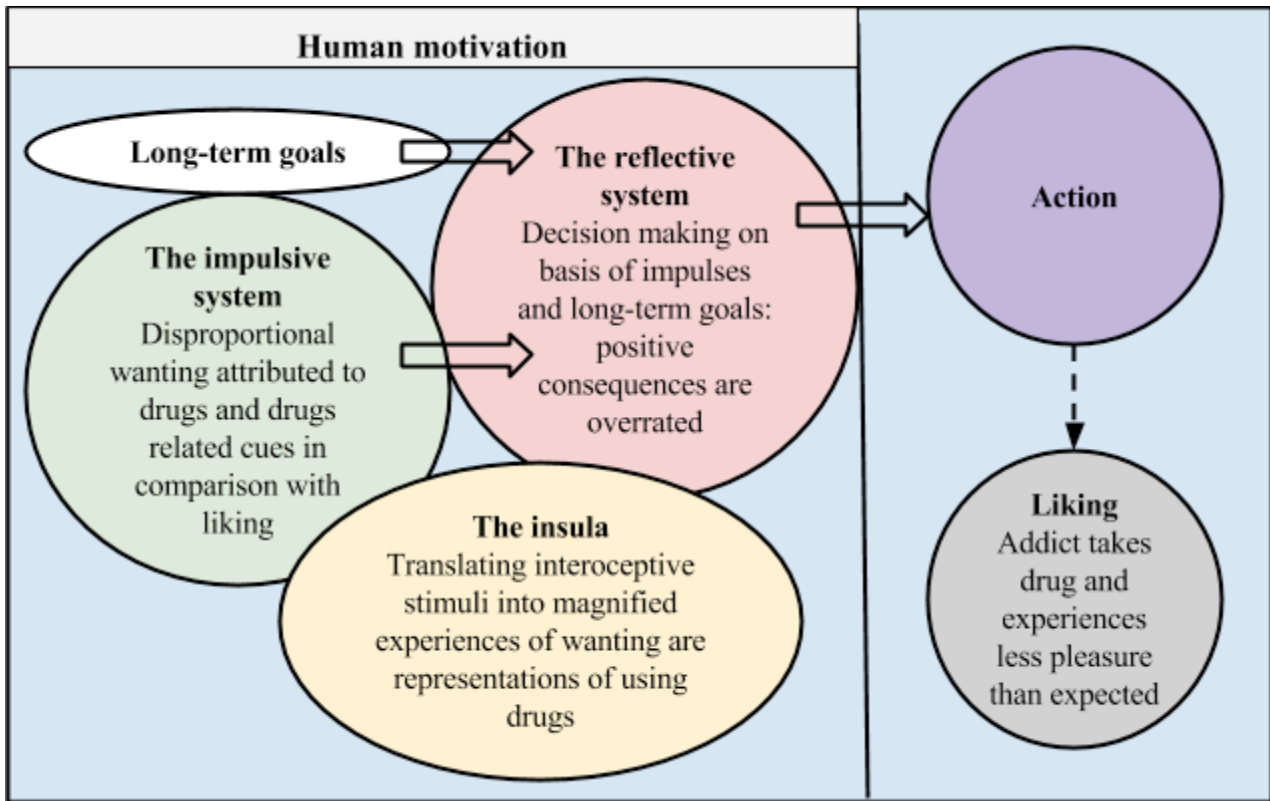
#### **4. Conclusion**

A drug addicted person does things he knows he’d better not do. Drugs are taken, even though the positive effects mostly do not outweigh the negative consequences. Some defining features of addiction are the withdrawal syndrome, tolerance, craving and relapses. A lot of researchers work on the topic of addiction to explain these defining features.

Both liking and wanting play a role in addiction. In paragraph 2 we described why reinforcement theories, that are based on liking, cannot fully explain all aspects of addiction. The incentive salience theory of motivation focuses on intense wanting, or craving, and provides a more complete explanation of addiction.

These insights are used to illustrate the three-systems neural model of addiction. The impulsive system generates basic impulses (wanting). These impulses are disproportionately strong in addicted individuals, since they are sensitised to drugs and drug related cues. With this sensitization the brain predicts a high reward as a result of using drugs, as wanting increases, but liking does not change or even decreases. Next, the reflective system controls these impulses and integrates long-term goals in decision making. But an addicted individual is characterised by an impaired ability to make decisions. The positive immediate effects are certain in contrast to the long-term effects that are uncertain. Lastly, the insula plays a key role in translating interoceptive stimuli to (subjective) experiences. With this, the magnified impulses generated by the impulsive system can be experienced subjectively. In figure 2 the three systems in an addicted individual are shown.

**Figure 2** An overview of the three-systems neural model of addiction as proposed by Noël et al. (2013) that results in an action, and what this means for an addicted individual. The



impulsive system, reflective system and the insula are located respectively in the amygdala/ striatum, the prefrontal cortex and the insula.

Concluding, humans are not adapted to drugs and the necessity of drug taking is overrated by the brain looking at the pleasure an individual will experience. ‘Overrated by the brain’ means that the brain attributes disproportional value to the use of drugs, which is expressed in a craving for drugs. As a result, the representations of bodily effects of drugs are magnified and therefore not accurately assessed. These insights explain why an addicted person *knows* one shouldn’t take drugs but still *does* it.

### **Chapter 3. Impulsivity: deliberation overruled by desire**

There are evident differences between the explanation of akrasia by philosophers and the biological explanations of addiction, but there are overlapping aspects as well. First the overlapping and conflicting aspects per concept in the created disciplinary models will be discussed (see figure 1 and 2). After that, the overlap and conflicts that have emerged in the comparison of the philosophical and biological concepts will be integrated as found or created *common ground* (Repko, 2012). This results in the integration of the disciplinary insights and a more comprehensive understanding of why reasonable people do unreasonable things.

#### **1. Overlap, conflicts and common ground**

Desire in philosophy and wanting in biology play a role in decision making. In philosophy desires are described as the driving forces behind human action and form an important influence on one's decision making process. They urge individuals to act and can either be calm or violent. Calm desires are hardly noticeable to the individual and are characterised by being instinctive, whereas violent desires generate perceptible emotions. The violent desires instigate akratic behaviour, they urge the individual to act against his or her interests. Biologists state that individuals are driven by wanting. Due to drug induced changes, the brain of an addicted individual generates magnified wanting for drugs and drug related cues. Both these definitions of desires and wanting include a conscious and unconscious aspect. They are both described as the input for decision making and therefore are driving forces behind human action.

There are no conflicts between the two concepts of desires and wanting. Therefore common ground can be created by finding a common word to replace the disciplinary terminology and to point out the common meaning (Repko, 2012, pp.335-336). The technique used to find this concept is called redefinition to be able to include similar concepts in one word (Repko, 2012, p.336). We chose to use 'desire' as a concept instead containing both the concept of desires and wanting. The term 'wanting' as it is interpreted in common language, might suggest that an individual has a choice to define what he or she wants. Choice is however not involved in wanting, nor in desire. Desire is intuitively understood as a more involuntary urge

and we want to appeal to this intuition as it makes for a clear common ground concept (Repko, 2012, p.329). **Desire** as defined by us is the conscious or unconscious urge that causes actions.

The decision making process for philosophy and biology are important in describing how a person proceeds to act. The decision making process as described in the interdisciplinary framework of philosophers and neuropsychopathologists consists of three stages. These stages are option generation, option selection and action initiation. In all three stages at least one form of akrasia can be observed, resulting in a dysfunction in a particular stage. Biologists emphasise that decision making takes place on the basis of wanting and long-term goals. Due to the increased wanting and impaired decision making as a result of drug use, the positive consequences of an action get overrated.

The insights from philosophy and biology show a different focus on decision making, but the insights do not conflict. By using redefinition we define **decision making** as the process, divided in three stages, in which the input consists of desire and deliberation. Deliberation involves, among other things the consideration of long-term goals.

Imagination, as defined in philosophy, consists of vivid representations of possible actions that can be called cause-effect beliefs. These are beliefs on the causal relationship between an act and a result. These representations let the individual experience the action and outcome mentally beforehand on the basis of which the individual decides whether to act on this or not. These cause-effect beliefs are based on past experiences which enable the simulation for similar, projected actions and outcomes. In biology the insula is described as a part of the brain that translates bodily stimuli into representations. The insula plays a role both in the experience of craving and in the representation of experiencing the use of a specific drug. The insula has not been paid much attention to in research so far. For this reason and the fact that explaining representations is not our main goal, we discuss representations only briefly. Further research could shed more light on the role of representations in human motivation.

In both philosophy and biology representations are important in human motivation. The concept of representations as used in philosophy and biology overlaps in the sense that biologists and philosophers acknowledge that representations influence the decision making process. Desire and representations influence each other. Desire makes the simulations of possible actions

and outcomes more vivid and representations can enhance the desirability of the simulated action. We conclude that **representations** are simulations that influence decision making and thus human motivation.

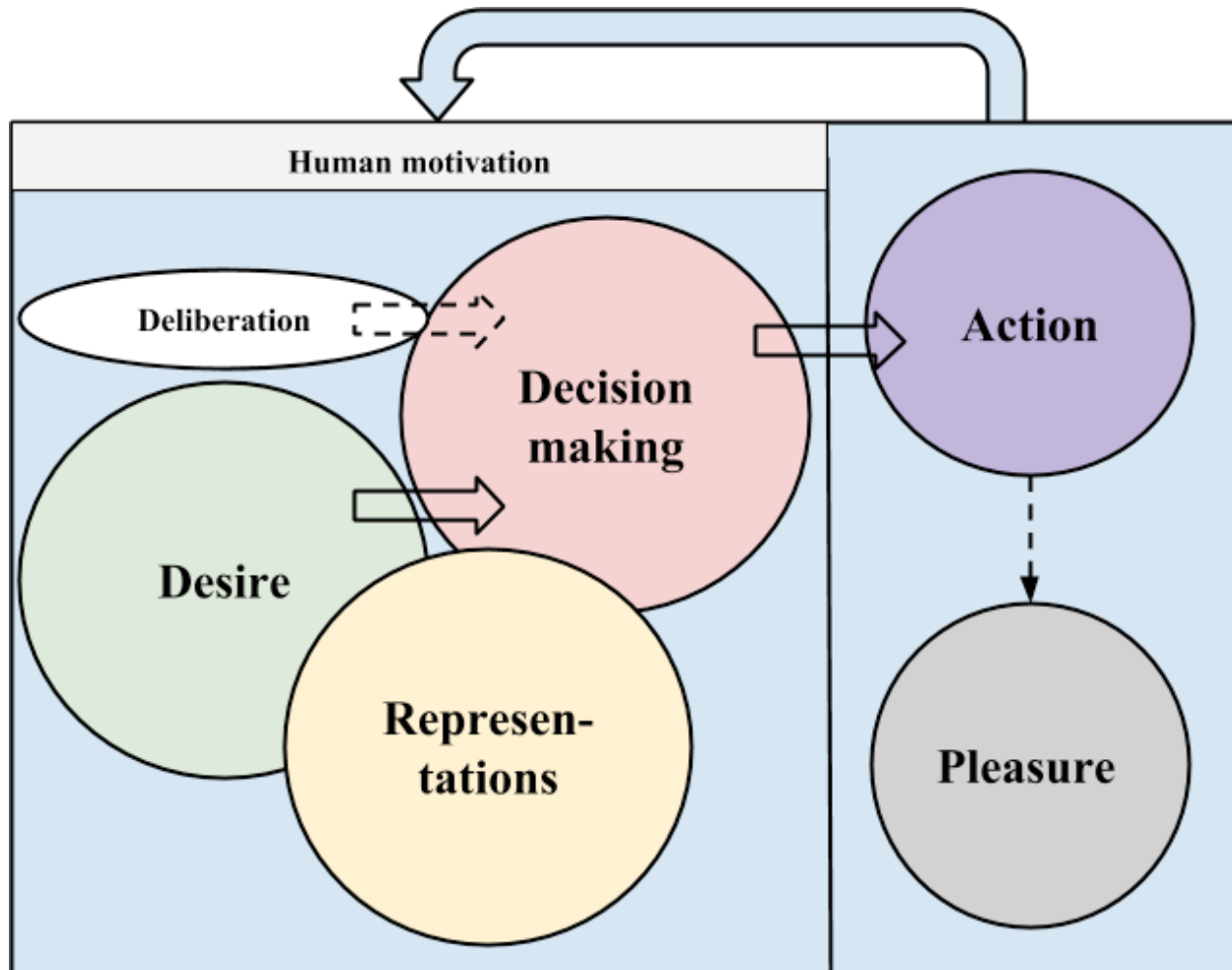
Decision making can result in action. In philosophy the action that results from human motivation is sometimes accompanied by pleasure. The experience of pleasure or the avoidance of discomfort is seen as an essential goal for animals in general, this makes life worth living. Therefore the pursuit of pleasure is often the desired objective for performing an action. Akrotic individuals strive for pleasure excessively and against better judgment. According to the biological view an action can result in a reward, that is experienced as liking. Liking functions as a tool to evaluate behaviour. It is seen as a positive reinforcer, which means it increases the frequency of performing certain behaviour. However this cannot fully explain addiction, since some drugs are addictive but one develops tolerance for them and one does not derive much pleasure from taking the drug. Biologists state desire and pleasure are separate manifestations and one can desire to perform an action, although this action does not result in proportional pleasure. **Pleasure** in philosophy and biology is defined in the same way, namely as the pleasant feeling that can accompany an action.

Although pleasure is described alike in both disciplines, we identify a conflict between the disciplinary theories on how pleasure is related to desire. In some philosophical theories desire is seen as an urge to experience pleasure as a result of an action. Biologists propose, on the basis of empirical research, that desire and pleasure can be seen separately. Acting on a desire does not always lead to pleasure. Whereas philosophers ascertain an indirect relationship between desire and pleasure, biologists do not observe this. We have created common ground by extending the biological theory (Repko, 2008, p.368). This theory does not assume the causality between desire and pleasure as some philosophers do, but leaves room for this causality. Therefore individuals do not act on desire *and* pleasure, but only on desire that is, in most cases, followed by pleasure.

## **2. More comprehensive understanding**

The combined common ground concepts lead to a more comprehensive understanding of the problem as described in our research question: “*Why do reasonable people do unreasonable*

things? ". We have developed a model that gives an overview of the process of human motivation and how this leads to behaviour (see figure 3). Below, we explain how things can go awry in this process resulting in unreasonable action.



**Figure 3** This model represents the found and created common ground on why reasonable people do unreasonable things. The left box contains human motivation, and the right box comprises of the outcome of this motivation.

The relations between the common ground concepts are explained in figure 3. Humans are driven by desires, deliberation and decision making. Both desire and deliberation provide input for decision making. Representations are also involved in decision making as it makes vivid simulations of possible actions. When human motivation leads to unreasonable behaviour desire has overruled deliberation in an individual and representations have presented an overly positive

simulation of the action and outcome making it more appealing. Acting unreasonably can be described as **impulsivity**, since the person acts on desire and short-term overestimated outcomes rather than long-term goals considered in deliberation. Performed actions and the (possible) experience of pleasure can form the input for future actions and thus influence human motivation, starting the cycle all over again.

In unreasonable behaviour the pleasure behaviour elicits is not proportional with desire, the pleasure is disappointing. This makes such behaviour irrational. Concluding, unreasonable behaviour can be characterised as impulsivity: deliberation overruled by desire.

## Conclusion and discussion

In our project we created a model to answer our research question: “*Why do reasonable people do unreasonable things?*”. The insights from philosophy on akrasia and the findings from biology on addiction were integrated in this model.

Humans are animals, unique for its capacity to use reason and ability to explain their motivation for certain behaviour. Though humans can act unreasonably as well. Unreasonable behaviour performed by reasonable people can be explained by acting on a desire. This desire influences and is influenced by representations and past experiences with similar actions the individual has performed. We showed this desire has a neurological basis. The brain is hypersensitive to drugs and drug-related cues. The desired goal of the action is mostly pleasure, although desire and pleasure are separate manifestations. Therefore, an action does not have to result in pleasure that is in proportion to a person’s desire. The human ability to deliberate is not as strong as this desire in cases of unreasonable behaviour. This means that an individual does not act deliberately but unreasonably.

We found the answer to our research question by focussing on two phenomena, namely akrasia and addiction. Some forms of akrasia are an example of addiction when a person is aware of his or her addiction and acts on this addiction, but would rather not. An addicted person acts akratically when he or she overestimates the consequences of using drugs and underestimates the risks involved. This corresponds to the description of impulsive behaviour, namely behaviour in which a person acts on desire and short-term overestimated outcomes rather than long-term goals considered in deliberation. Ultimately, impulsivity as the common denominator in both disciplines was deduced from the insights to provide a more general explanation for unreasonable behaviour of reasonable people.

In this project we did not aim to provide aids for curing akrasia nor for addiction, it sketches a theoretical framework for a deeper understanding of these phenomena. However, further research could use our findings in the search for more effective treatments of addiction. Besides that, our



answer to the research question can contribute to the academic debates on akrasia, addiction and unreasonable behaviour in general.

Although philosophers and biologists have a different way to research the world, finding common ground between and integrating the insights from these disciplines showed the gap is not as big as one might expect. While philosophers discuss topics that are not researchable empirically, biologists can provide empirical explanations that cannot be explained by *a priori* statements. When these different approaches of philosophy and biology are combined this proves fruitful. This is illustrated in our research since a more comprehensive understanding of irrationality, that plays a role in both disciplines, was formulated.

The social sciences, for example psychology, could also contribute to our research topic since those sciences discuss phenomena like irrationality as well. When social sciences are involved in future research this will generate an even more comprehensive understanding on how unreasonable behaviour can arise. They can add a social dimension to the explanation unreasonable behaviour.

With the integration of insights of different disciplines, we come closer to explaining phenomena like unreasonable behaviour and finding pieces of the puzzle of human motivation for action.

## Bibliography

### General:

- Repko, A.F. (2012) *Interdisciplinary Research: Theory and Process*. Los Angeles: SAGE.
- WHO (2013). *Tobacco*. [ONLINE] Available at: <http://www.who.int/mediacentre/factsheets/fs339/en/>. [Last Accessed 1 April 2014].

### Philosophy:

- Aristotle (1973). 'Ethica Nicomachea (Nicomachean Ethics)'. *In*: Richard McKeon (ed), *Introduction to Aristotle*. 2nd ed. Chicago: The University of Chicago Press. pp.346-594
- Aristotle (2009). *Ethica Nicomachea*. 7th ed. Groningen: Historische Uitgeverij.
- Hume, D. (2003). *A Treatise of Human Nature*. 3rd ed. Mineola, New York: Dover Publications, Inc.
- Kalis, A., Mojzisch, A., Schweizer, T.S., & Kaiser, S. (2008). Weakness of will, akrasia, and the neuropsychiatry of decision making: An interdisciplinary perspective. *Cognitive, Affective, & Behavioral Neuroscience*. 8 (4), pp.402-417
- William Edward Morris (2009). *David Hume*. [ONLINE] Available at: <http://plato.stanford.edu/entries/hume/#UniIma>. [Last Accessed 17 april 2014].
- Press, G.A., (1999). 'Brief Summary of Aristotle's Writings'. *In*: Popkin, G.H. (ed), *The Columbia History of Western Philosophy*. 1st ed. New York: Columbia University Press. pp. 72-74.

### Biology:

- Berridge, K.C. (2004). Motivation concepts in behavioral neuroscience. *Physiology & Behavior*. 81 (2), pp.179-209
  
- Berridge, K.C., Robinson, T.E., & Aldridge, J.W., (2009). Dissecting components of reward: 'liking', 'wanting', and learning. *Current Opinion in Pharmacology*. 9 (1), pp.65-73
  
- Hyman, S.E., Malenka, R.C., & Nestler, E.J., (2006). Neural Mechanisms of Addiction: The Role of Reward-Related Learning and Memory. *Annual Review of Neuroscience*. 29, pp. 565-598
  
- Kirkpatrick, K., & Hall, G. (2005). 'Learning and memory'. *In: J.J. Bolhuis & L.-A. Giraldeau (ed), The behavior of animals. Mechanisms, function, and evolution*. 1st ed. Oxford: Blackwell. pp.146-169
  
- McNamara, J.M., & Houston, A.I., (1986). The Common Currency for Behavioral Decisions. *American Naturalist*. 127 (3), pp.358-378
  
- Naqvi, N.H., Rudrauf, D., Damasio, H., & Bechara, A., (2007). Damage to the insula disrupts addiction to cigarette smoking. *Science*. 315 (5811), pp.531-534
  
- Naqvi, N.H., & Bechara, A., (2009). The hidden island of addiction: the insula. *Trends in Neuroscience*. 21 (1), pp.56-67
  
- Noël, X., Brevers, D., & Bechara, A., (2013). A neurocognitive approach to understanding the neurobiology of addiction. *Current Opinion in Neurobiology*. 23 (4), pp.623-638
  
- Purves, D., Augustine, A.J., Fitzpatrick, D., Hall, W.C., LaMantia, A., & White, L.E. (2012). *Neuroscience*. 5th ed. Massachusetts U.S.A.: Publishers Sunderland.

- Robinson, T.E., & Berridge, K.C., (1993). The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Research Reviews*. 18 (3), pp.247-291

- Spruijt, B.M., Bos, R. van den., & Pijlman, F.T.A., (2001). A concept of welfare based on reward evaluating mechanisms in the brain: anticipatory behaviour as an indicator for the state of reward systems. *Applied Animal Behaviour Science*. 72 (2), pp.145-171