# Ending the Stigma: How a Causal Deterministic View of Free Will Can Inform Both Healthy and Pathological Cognitive Function and Increase Compassion

by

Cristina Balaita

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Arts
in
Philosophy

Waterloo, Ontario, Canada, 2014

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

## ABSTRACT

Depression is the leading cause of disability around the world, and in Canada, 8% of adults will experience depression in their lifetimes. Nearly half of those with depression will not seek treatment, one of the major barriers being the social stigma associated with depression and other mental illnesses. Some of this stigma results from a mistaken understanding of free will and agency and the degree to which these are compromised in mental disorders. This thesis aims to show that free will in both psychologically healthy and pathological cases can be understood in a scientific causal deterministic way based on recent findings in neuroscience and psychology. The 'will' can be understood in terms of the normal range functioning of mechanisms for control, choice, and valuation. There is no 'free' will that is uncaused, but only relative freedom when these mechanisms are not internally damaged and there is no external coercion. Evidence that depression and mental illness can also be understood in a causal, deterministic way is also presented, and it is argued that this understanding can work back to reinforce the scientific understanding of the will in non-pathological cases. The understanding of free will based on healthy function and that based on pathological function are mutually reinforcing. The thesis concludes by showing that, based on a causal deterministic picture of the will, the stigma surrounding mental illness is unfounded, and that this view can lead to more compassion, understanding, and acceptance of both those with mental illness and the mentally healthy.

ACKNOWLEDGEMENTS

I would like to express the deepest appreciation to my thesis supervisor, Professor Paul Thagard, whose enthusiasm for my project, invaluable comments, and vast knowledge about cognitive science have enriched both this project and my educational experience.

I would like to thank my readers, Professor Mathieu Doucet and Professor Patricia Marino for their helpful and timely suggestions. I would also like to thank the graduate students, staff, and faculty in the Philosophy Department at the University of Waterloo. My experiences as a graduate student have been stimulating, thought provoking, and decidedly positive, not least because of the exceptional teaching, meaningful discussions, and environment of inquiry, diversity, and collaboration in the department. In particular, I'd like to thank Benjamin Nelson for the many coffee meetings in which he helped me clarify my position.

I'd also like to thank Dr. Sujay Patel and his team for the opportunity to do a two-week observership in psychiatry, such that I was able to interact with psychiatric patients and gain a clinical perspective.

Finally, I extend my sincere thanks to my family and friends who helped me stay focused throughout this year with constant encouragement, patience, and support.

Cristina Balaita, University of Waterloo

## Table of Contents

## CHAPTER 1: Introduction

According to the World Health Organization, depression affects an estimated 350 million people around the world and is the leading cause of disability (WHO, 2012). The Canadian Mental Health Association estimates that 8% of adults will experience major depression throughout their lifetimes. Of these, nearly half will not seek treatment (CMHA, 2014). A major barrier to the diagnosis, treatment, and social acceptance of people with depression and other mental illnesses is social stigma (Barney, Griffiths, Jorm, & Christensen, 2006). In the worst of cases, depression is seen as a personal weakness, something one ought to just 'snap out of' or have the 'willpower' to overcome. Conversely, from a medical perspective, mental illness is often associated with a loss of control analogous to the loss of control experienced in diabetes, heart disease, or cancer. A central question in the understanding and treatment of mental illness is the degree of conscious control that people have over their actions. Related to this are questions of free will or agency, where it can be asked if these are compromised in mental disorders. Even in those with healthy cognitive function, recent evidence from neuroscience and psychology has called into question ideas about free will, agency, and conscious control over actions. The implications of these findings could change the way we look at psychological functioning in both healthy and pathological cases, and can further lead to more informed and more compassionate public treatment of people with mental disorders.

There are three main aims of this thesis. First, I will argue that free will can be understood in a scientific, causal deterministic way based on recent findings in neuroscience and psychology. The 'will' can be understood as the normal range functioning of mechanisms for control, choice, and valuation. There is no 'free' will that is uncaused, but rather, only relative freedom when these mechanisms are not internally damaged and there is no external coercion. This view represents a middle ground between the hard determinist

thesis and the compatibilist thesis, which will both be outlined in section 1.1. Second, I will argue that depression and mental illness can also be understood in a causal deterministic way and that this understanding can work back to reinforce the scientific understanding of the will in non-pathological cases. The understanding of free will based on healthy function and that based on pathological function are mutually reinforcing. Third, I will argue that, based on a causal deterministic picture of the will, the stigma surrounding mental illness is unfounded, and that this view can lead to more compassion, understanding, and acceptance of both those with mental illness and those with healthy psychological/neural function.

In this first chapter I will present the conceptions of the will that will be relevant later in the thesis, including philosophical, neuroscientific, and psychological accounts. I will also give a brief introduction to depression and its presentation. While introducing these, I will show that on one end of the spectrum, the end that I will ultimately agree with, there is mounting evidence that there is no free will as traditionally conceived, and that concepts of free will need revision. On the other end of the spectrum, it is argued that the scientific evidence does not change anything with regard to conceptions of free will or behaviours in the world; this is something I will argue against.

In the second chapter, I will argue for a conception of the will that includes mostly nonconscious executive control by the prefrontal cortex, conscious processing important for integration of new inputs, learning, deliberation, and changing plans and values, and normal range functioning of valuation/goal-setting mechanisms such that choices can be made in line with those goals and values. I will argue that there is no real freedom of the will in the sense of being uncaused, but rather, only a limited freedom that comes about from the absence of disease, impairment, or external coercion. After showing that determinism means that there is no freedom from prior causes, I will show that neuroscientific results threaten conscious control and choice as they suggest that decisions are made nonconsciously and that conscious

awareness comes about after the fact. I will then show that the role of consciousness is more limited than previously thought by comparing conscious and nonconscious processing, showing that there is no definite boundary, and by presenting challenges from social psychology that show the large effect of situational factors and biases on our behaviours. I will argue that consciousness can be understood as an emergent property, and as such has various functional roles, but does not play a direct causal role in action. I will further argue against ideas of agency or 'self' that posit a 'controller', and rather argue that plausible views of the self include autobiographical and multilevel mechanisms approaches, which are part of the causal network of brain function. I will then show that while the feeling of agency and control over actions may be important, most of this functioning is nonconscious and includes prefrontal cortex activity and various neurochemical systems including the dopamine system. In the last part of chapter 2, I will address various objections against the interpretation of the causal deterministic view leading to a changed view of free will. I will show that the implications of this view are *increased* control, a decrease in blame, and ultimately an increase in compassion, understanding, and acceptance of others through recognition of our common humanity.

In the third chapter I will argue that depression and other mental illnesses can be understood in terms of the breakdown of mechanisms for executive control, choice, and valuation, among others. I will begin by examining the complex causality of depression to show that depression can be understood in terms of the complex interplay of biological, psychological, and social causal factors. I will then show how the treatment of depression is also based on a causal deterministic view. Where depression involves a breakdown of mechanisms, there is a corresponding loss of control and ability to make decisions; the components of the 'will' break down. Because there is so little control and because there is a breakdown of cognitive processes, the stigma against those with depression is unjustified.

Indeed the difference between healthy psychological functioning and mental illness occurs as a matter of degrees based on the extent to which mechanisms are functioning within the optimal range. Further, depression can be understood causally, in a similar way to other physical illnesses. These factors together indicate that the stigma surround depression is heavily misguided. Finally, the causal deterministic view has implications for autonomy and agency in medical decision-making, not precluding the possibility of choice, values, and preferences, but rather understanding them as causally informed.

### 1.1 Philosophical Perspectives

Free will is a concept that is ill defined and can have many different meanings. It is often related to agency, autonomy, motivation, drive, and the connection between intention and action. In some instances there is an implicit assumption that there is a conscious 'agent' that initiates actions and has the ability to make 'free' choices, that are uncaused by prior factors. Indeed we often have a *feeling* that we are agents and have conscious control over our thoughts, emotions, and actions. The current philosophical debate about free will can be roughly characterized by the following three positions: hard determinism, libertarianism, and compatibilism. While determinism and libertarianism are perhaps the most opposed, the two positions that will be most interesting to compare for the purpose of this thesis are hard determinism and compatibilism as these positions recognize the vast evidence from neuroscience and psychology, but disagree on its implications for free will.

Hard determinism and libertarianism stand opposed as philosophers in these camps disagree on the deterministic nature of the universe (and consequently, cognition). The thesis of determinism broadly asserts the causal nature of physical laws, such that all future events are caused by prior events. Applied to cognition, this leads to a view that the brain also behaves in this causal, deterministic way. If behaviour is a result of brain processes, then

behaviour too is causal; it is always determined by prior factors such as genetics, environmental influences, and past experiences. The dramatic increase in neuroscientific research over the last century points to a picture of the brain that is deterministic in this way; the study of the brain supports determinism. Both hard determinists and compatibilists agree on the thesis of determinism, but disagree on its implications for free will. Libertarians on the other hand, argue against the thesis of determinism.

The hard deterministic view is espoused by philosophers such as Galen Strawson (1994), who argues that any part of our 'self' that we might think could be the basis for 'free' decisions is determined prior factors such as genetics, environmental influences, and past experiences, all of which are beyond our conscious control. For Strawson (1994) in order for an agent to have free will, there would have to be something apart from these inputs, but it is unclear just what that could be. People can still make choices on this view, but the choices themselves are based on values, preferences, and decision-making criteria (what he calls 'principles') that are not caused by the 'agent' but all come about from prior causal factors (Strawson, 1994). Importantly, hard determinists do not think that free will and determinism are compatible; if the universe is deterministic, there is no sense in which we can have free will.

The libertarian thesis is that determinism is not true and that at some crucial moments, people are able to make 'free' choices that determine their character and form the basis of subsequent choices. This view is espoused by philosophers such as Robert Kane (2009), who claim that because of indeterminacies that can enter into neural networks, not all choices are caused. Kane calls these choices 'Self-Forming Actions' (SFAs) and argues that they occur in times when there is a strong value conflict with an indeterminate answer. Libertarianism can be an attractive view because it is in line with the common intuition that we, as agents, are the authors of our actions and can be held morally responsible for them. In this thesis, I will

not be discussing libertarianism beyond this point simply because it is not supported by the large amount of brain science done over the last century. Moreover, the view is not plausible because even if there were indeterminacies in neural networks, this does not show how this would allow conscious control over choices; choices would be in some sense random.

A more relevant and interesting position is compatibilism, the view that we can meaningfully talk about free will, even given the truth of a deterministic universe; free will and determinism are compatible. In this camp are philosophers and scientists that claim that regardless of the truth of determinism, concepts of free will and their consequent actions will remain the same and be unaffected. Philosophers espousing this view include Alfred Mele and Adina Roskies, whose positions, in response to advances in neuroscience, will be discussed in the next section. Other compatibilists frame the free will debate in terms of moral responsibility, claiming that moral responsibility is still possible given the truth of determinism. For example Harry Frankfurt (1969) argues that moral responsibility comes about as a result of making choices or carrying out actions as a result of *wanting* to have done them. Similarly, Susan Wolf (1987) argues that free will understood in terms of moral responsibility is a result of acting in accordance with our 'deep self,' even given determinism.

The position that I find most plausible is one that stands between hard determinism and compatibilism. While I do believe that free will and determinism are largely incompatible, in line with the hard determinist thesis, I argue for a revised understanding of free will such that we may still have some relative freedom, even if there is no ultimate freedom from causal factors. On this view, the will can be understood in terms of the healthy functioning of neural mechanisms and processes, as contrasted by conditions in which these processes break down or in which there is external coercion. This relative free will and determinism are compatible, but importantly, contrary to what some compatibilists argue, this conception *should* change behaviours and attitudes. While the topic of moral responsibility is

not central to this thesis, on the relative view of free will that I espouse, there can be responsibility, but importantly it is responsibility in light of the practical consequences of holding others responsible. This view decreases the degree to which there can be blame and responsibility because it is causal in nature, but does maintain some conception of relative responsibility in line with relative free will.

### 1.2 The Challenge From Neuroscience and Philosophical Responses

The neuroscientific challenge to free will, understood as uncaused choice or conscious control, came into focus largely after the foundational studies by Benjamin Libet. Libet (1993) and his team performed an experiment to determine how voluntary actions arise in the brain. They defined voluntary acts as those that have an endogenous (internal) origin, are not externally constrained or initiated by outside factors, and are felt by the subjects to arise of their own want and at the time that they choose. The experimental design was such that a person was to voluntarily move her finger and report the position of a second timer on a clock the moment she became consciously aware of the intention to do so. Electrical activity in the brain was measured to assess whether the onset of electrical activity would coincide with the conscious intention to initiate action. Readiness potentials (RPs) precede spontaneous voluntary acts, and it was found that the RPs began in the brain between 350-400ms before the subjects became aware of the intention to act (RPs were initiated nonconsciously). Libet (1993) and his team also found that between 100-200ms before the action, the act could be 'vetoed' or triggered by the subjects, and they proposed that the role of conscious will would then be to control the outcome or select between various initiated responses.

After Libet, John-Dylan Haynes (2011) and his team performed a modified version of Libet's experiment. The team's objective was to discover whether the result of a decision

could be found in the brain prior to the subject being consciously aware of it (Soon et al., 2008). The subject saw a letter stream with cards spaced 500ms apart, some providing instructions for which button to press (left or right) and others allowing the subject to choose. FMRI data was used to assess brain activity in real time, and patterns of neural activity were found 10 seconds prior to the conscious choice. The researchers could determine which button the subjects would choose with 60% accuracy before the subject became aware of the choice. They hypothesized that the delay was a result of "the operation of a network of high-level control areas that begin to prepare an upcoming decision long before it enters awareness" (Soon et al, 2008, p. 543). Haynes (2011) claims that there is a "nonconscious causal chain of events [that] can occur outside subjective awareness even before a subject makes up their mind" (p. 161). The implications of these experiments for free will, according to Haynes (2011), is that the outcome of decisions is fully determined by brain activity, much of which happens before a subject is consciously aware of the decision.

Taking a similar approach, but going deeper into the brain, neurosurgeon Itzhak Fried and his team carried out studies on patients that had, as part of a procedure to treat epilepsy, electrodes directly implanted in their brains (Fried et al, 2011). They sought to understand 'self-initiated behaviour' or 'volitional control' involved in the movement of a finger. They were able to record activity in small neuronal populations, and found that 1500ms before conscious awareness of a decision, specific neuronal populations in the medial frontal lobe were active, and 700ms before conscious awareness, they could predict, with a more than 80% success rate, the time point when the decision would be carried out. Fried et al. (2011) similarly concluded that considerable neuronal activity happens before the results of a decision become conscious.

These results show that at least in the case of simple motor decisions, conscious awareness comes some time after the decision to act has been made nonconsciously. They are

a threat to free will because free will is often thought to involve conscious control over actions; if researchers can predict decisions before conscious awareness, only looking at neural firing patterns, then conscious will cannot play the central role it was previously believed to play.

Philosophical objections to these results come from Alfred Mele and Adina Roskies. Both Mele and Roskies argue that the Libet, Haynes, and Fried et al., experiments do not show what they purport to show, namely, that decisions are made prior to conscious experience, beyond conscious control, and that this is a threat to free will. Mele (2013) argues that these experiments involve simple decisions where the people are indifferent or where the outcomes are inconsequential; the findings cannot be generalized to complex decision-making. He argues that there could still be free will in complex decisions, such as moral decisions, where people weigh different actions and at least partially decide on the basis of conscious reasons. Indeed for Mele, the conscious struggle that is involved in choosing between competing choices is important for free will. Mele (2013) also argues that even if the neuroscientific evidence is accurate and conscious awareness comes about after a decision has been made, the decision is still free because conscious deliberation played *some* role. He further argues that it is likely that conscious deliberation can lead to conscious deciding: "perhaps people who consciously struggle with temptation or competing motivation during a process that leads to a decision are more likely to decide consciously than people who unreflectively select among options with respect to which they are indifferent" (Mele, 2013, p. 781). He argues that even if it turns out that the neuroscientific experiments do show that consciousness comes after a decision has been made, "conscious deliberation can [still] play whatever role believers of free will believe it plays in much free decision making" because "conscious deliberation that preceded the decision played an important role in generating the decision" (Mele, 2013, p. 783).

In line with the objections by Mele, Roskies (2006) argues that the findings from neuroscience do not change anything with regard to free will. She points out that challenges to free will exist on a philosophical level, independent of neuroscientific findings. She argues that because these problems already exist, neuroscientific evidence of brain mechanisms underlying action and decision-making cannot help us make progress on the philosophical issues. She further argues that the findings can only tell us that decision-making is influenced by some physical factors, but cannot tell us anything beyond that. While neuroscience does point to the brain as being deterministic, it cannot tell us the status of the universe as being deterministic, and thus cannot tell us whether at a higher level there are indeterminate factors. Either way, what is at issue, for Roskies, is not the thesis of determinism, but rather the consequences of what might happen to our social and legal systems if we abandon a belief in control over actions and have to consequently abandon moral responsibility (Roskies, 2006). Ultimately, she argues that "[a] view of ourselves as biological mechanisms should not undermine our notion of ourselves as free and responsible agents" (p. 421).

### 1.3 Plausible Neuroscientific Accounts of Free Will

A plausible neuroscientific understanding of the will involves multiple neural mechanisms for control, decision-making, and valuation. These include both anatomical and physiological mechanisms including the prefrontal cortex (PFC) and dopamine circuits. These mechanisms will be discussed in more detail in section 2.5.1.

A plausible comprehensive neuroscientific model of will comes from Patrick Haggard (2008) who views voluntary actions as a form of decision-making governed by brain circuits. With regard to decision-making or choice, brain areas involved could be "the pre-supplementary motor area (pre-SMA), the anterior prefrontal cortex and the parietal cortex"

(Haggard, 2008, p. 934). These areas both inform decision-making about voluntary action and generate the necessary information a person needs in order to act. Voluntary action can be contrasted with involuntary reflex actions such as the knee jerk reflex, changes in pupil diameter in response to light, or shivering in response to cold which are governed primarily by the spinal cord and autonomic nervous system. Haggard (2008) contrasts voluntary action with simple reflexes, where the latter is determined by external stimuli. Voluntary action involves the cerebral cortex, and has two main subjective experiences: intention and agency. Intention is the plan to do something, while agency is the feeling that an external event was caused by the actions performed. Simple reflexes, on the other hand, do not involve any subjective feeling, and are merely a result of the spinal cord and autonomic nervous system.

A second plausible model comes from Read Montague (2006), who argues that choice and decision-making can be understood from the perspective of computational neuroscience, the brain being an efficient computational device that is flexible and adaptive to new inputs and environmental changes. This fits with an evolutionary perspective, where complex choice making apparatus evolved such as to allow the organism flexibility in light of changing conditions (Montague, 2006). On this model, brains generate choices nonconsciously, where there are some choices that involve no conscious deliberation, and others that involve significant deliberation, accompanied by the feeling of conscious will. The key parts of the model include valuation of states and ideas, goal-setting and maintenance, desires acting as error signals, an understanding or projection of expected consequences, and a view toward the efficient allocation of resources through choice. The "value function tells a machine the value of the current state of affairs…[and] is essential to help the machine decide how to allocate its resources" where resources are limited (p. 55). Indeed Montague argues that "our minds are quite literally valuation machines… [w]e can value separately the past, the present, and the future. We can even value things that did not happen … or have not happened yet"

(p. x). When making decisions, then, the brain takes into account all possible valuations and integrates those with the current state of the organism. In this way, the brain is able to construct models of the future and of the value of future actions, choose the best course of action in the moment (Montague, 2006). This is important because it leads to the creation of goals or desired states that change in accordance with current and future needs. Goals can include basic biological drives such as food, water, or sex, but can also include more complex social goals such as career achievement or theoretical understanding. They are "represented by recurrent patterns of neural activity that are stable for some time and can be distinguished from ongoing background activity" by the prefrontal cortex (Montague, 2006, p. 125). Desires arise to propel the organism toward either basic physiological homeostasis or toward goal attainment, providing an error signal that acts as a guide to action. Indeed goals and the guidance signals that accompany them can be largely nonconscious (and remain so), arising from the nervous system to shift action such that the homeostatic needs can be met (Montague, 2006). When our minds make choices on the basis of goals and values, we are always constrained by biological needs, but interestingly we can choose to act against those biological needs. This ability to value ideas over biological drives is, according to Montague, the basis for agency and will.

### 1.4 Psychological Perspectives

Psychology looks at human behaviour and the internal states that cause people to act as they do. When it comes to free will, particularly defined as the feeling of conscious agency, there is evidence to support the notion that conscious control is largely illusory. For instance, the illusion of control is a concept first named by Ellen Langer (1975). It refers to the tendency to assume influence over events that the person had no control in determining. Psychologist Daniel Wegner (2005), has done work in this area and argues that much of the

time, the feeling that people have that their thoughts are the causes of their actions is mistaken. Indeed the mind attributes authorship in many cases that are not causally related, as a way of making sense of things. He explains his theory of 'apparent mental causation' as having three parts: priority, consistency, and exclusivity. The assumption of agency and causation occurs when there is a conscious thought *prior* to the action, that thought is *consistent* with the action, and it appears to be the *exclusive* cause of the action (Wegner 2005). For example, if I get up to go to the fridge because I am hungry, I assume authorship if I have a thought to get up prior to doing so, if my action of getting up is consistent with that thought, and if I didn't see any other reasons that could have caused me to get up (my thought was the apparent exclusive cause of the action). In this way, the mind attributes authorship in many cases that are not necessarily causally related. Wegner and others performed several experiments to show that people have a subjective feeling of control or agency in cases in which it is impossible for them to control the outcomes. For example Wegner and Wheatley (1999) had study participants sit at a computer with a mouse where they would hear a word 1-5 seconds before the computer cursor moved to select a corresponding picture (e.g. 'swan'). Unbeknownst to the participants it was not them who was making the movement, however participants reported feeling that they intentionally acted to select the swan. In another experiment Wegner, Sparrow, and Winerman (2004) had participants stand in front of a mirror where another person's hands behind them and out of view acted out certain motions (so it appeared to be the participant's hands acting them out). When they heard audio instructions and the hands acted out the instructions, participants had a feeling that they were controlling the hands. Finally, Dijksterhuis, Preston, Wegner, and Aarts (2004) showed that priming individuals with 'I' or 'me' in an experiment where they were to compete against a computer in a word identification task increased feelings of

agency. All of these experiments point to the fact that we may attribute agency in many cases in which we had no control over the actions performed.

### 1.5 Depression

I will argue in chapter 3 that depression and mental illness can be understood in terms of causal determinism and that this understanding then works back to inform healthy psychological/neural function. Based on both of these views, the stigma that comes with mental illness is unjustified; the deterministic view can lead to more acceptance and understanding in both the case of mental illness and in healthy cognitive functioning.

According to the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) released by the American Psychiatric Association (APA), major depressive disorder (MDD) must include either depressed mood or a loss of pleasure or interest, along with a minimum of five other criteria, continually for a minimum of at least 2 weeks. Some of the other possible symptoms include significant weight loss or gain, insomnia or hypersomnia, psychomotor agitation or retardation, chronic fatigue, feelings of guilt or worthlessness, indecisiveness or decreased ability to think/concentrate, and recurrent suicidal ideation with or without a plan (APA, 2013, p. 161).

MDD may present with anxiety, some psychotic symptoms and various combinations of the above symptoms. For example, there can be psychotic depression, postpartum depression, seasonal affective disorder, and bipolar disorder (APA, 2010). Psychotic depression presents with psychosis, which includes delusions and hallucinations (a break from reality, false beliefs, feelings of persecution, seeing hearing things others can't see or hear). Postpartum depression affects 10-15% of women after giving birth to a child. Seasonal Affective Disorder (SAD) usually affects people in the winter months when there is a decrease in scientific sunlight. Bipolar disorder, also called manic-depressive illness, includes

a shift from manic to depressive states. Because of this variability, depression is considered a heterogeneous disorder with many aetiologies and pathological forms, and consequently, variable response to treatments (APA, 2010).

Depression can also vary in severity from mild to moderate to severe and I will show that this coincides with a loss of control and ability to make choices. Mild depression is where there is persistent low mood or lack of interest, but daily function is not fully impaired. Function may be more slow or more of a struggle. Indeed many people with mild depression can mask their symptoms. For mild depression, Cognitive Behavioural Therapy or counselling can provide an improvement. Moderate depression includes a functional impairment in activities of daily living, instrumental activities of daily living, and employment. Activities of daily living (ADL) include bathing, personal mobility, toileting, dressing, personal grooming and feeding, while instrumental activities of daily living (iADL) include managing finances, taking medications, shopping, driving and housework (Foti & Kanazawa, 2008).

Generally a combination of antidepressants and CBT can be effective for moderate depression. Severe depression includes a halt to daily functioning including ADLs and iADL and is generally treated with antidepressants and sometimes Electroconvulsive Therapy (ECT) and other forms of brain stimulation, such as deep brain stimulation (DBS) and Transcranial Magnetic Stimulation (TMS). Severe depression can be recurrent and treatment resistant, causing complete dysfunction in the life of the affected person (APA, 2010).

## CHAPTER 2: Taking a Scientific Causal Deterministic Stance on Free Will

In this chapter, I argue that free will can be understood in a scientific causal deterministic way. I will show that there is no 'free will' in the sense of an agent or 'self' making choices that are uncaused. Rather, the 'will' can be understood in terms of the neural processes responsible for executive control functions, both conscious and nonconscious, and the making of choices in line with the goals and values of the organism. In both of these conscious deliberation plays some role, albeit a limited one, by facilitating a certain type of processing and deliberation that works to align behaviour with goals and intentions. Taking a causally deterministic picture shows that there is no freedom from prior causes, yet freedom can be understood, on this view, in terms of the absence of neural disease or impairment, and as functioning that is not externally coerced. Indeed a relative freedom of the will is had when there is optimal anatomical and physiological functioning of neural mechanisms responsible for control and choice.

### 2.1 Determinism Means That There is No 'Freedom' Apart From Prior Causes

Adopting a causal deterministic view means abandoning any sense of uncaused freedom because all mental states are the result of physical states in the brain, states which are themselves causally determined. Important to this view is an understanding of mind-brain identity, where mental states such as intentions, beliefs, plans, and desires, are physiological states in the brain; they result from the complex interaction of neural firing patterns. There is good reason to think that there is a causal relationship that goes both ways. In one direction, the interaction of neural firing patterns cause mental states, and in the other direction, mental states can be altered by external stimuli and change the interactions between neurons, changing brain structure. Other inputs into the system, such as chemical or genetic factors can also create changes to the system.

As evidence for this, we need not look further than the intake of alcohol, narcotic painkillers, and other recreational drugs that all cause chemical changes in the brain and in doing so cause changes in emotional conscious states (Panksepp, 1998). Specifically, "many recreational drugs such as ecstasy and cocaine increase dopamine levels, leading temporarily to increased positive valence attached to various representations" (Thagard & Aubie, 2008, p. 818). Similarly, in the other direction, there is evidence that psychological changes as a result of cognitive behavioural therapy are reflected as changes in the brain (Goldapple et al., 2004; Linden, 2006). In fact, learning any new skill, such as juggling, causes visible changes in the anatomical structure of the brain (Draganski et al., 2004). Furthermore, evidence from stoke patients, dementia patients, patients with various abnormalities, brain tumours or lesions, shows serious impairment in function and significantly altered mental states (Damasio, 2012).

This shows that there is a causal relationship between brain states and mental states. Determinism means that neural networks are causally closed, and because brain states and mental states are causally related, mental states too are causally determined. On this causal deterministic understanding of the brain and behaviour, the implication is that the 'will' cannot be 'free' in the sense of being uncaused; brain events are always caused by the complex interaction of prior mental states and representations. Furthermore, there is no uncaused 'freedom' because all of the thoughts and actions that we become consciously aware of are products of nonconscious brain processes that are governed by the laws of physics.

### 2.2 Neuroscientific Results Challenge Conscious Control and Choice

Recent advances in neuroscience, as exemplified by the experiments of Libet, Haynes and Fried et al., put pressure on ideas of free will, specifically where it is thought that a

conscious agent of sorts is the initiator of actions. They do so by suggesting that the role of conscious control is limited, as subjects become aware of a decision to act only some time after the decision has been made nonconsciously.

This is clearly a threat to conscious control over actions, because it indicates that we may only have a *feeling* that we are consciously making choices or initiate actions, but that the choices have already been made nonconsciously, beyond our conscious awareness. Accepting that small motor decisions may be governed by factors beyond our control, critics of these experiments, such as Mele and Roskies, can argue that the experiments cannot tell us much about complex, real-world decisions, and it is in these latter decisions that we care about conscious control over choices. Indeed, with regard to the experiments themselves, it is possible that there could be a learning effect or that the choice is so arbitrary that conscious deliberation is not necessary and thus has no causal role. As Mele (2013) argues, it could be possible that even if the experiments show us that neural firing patterns precede conscious activity, this need not threaten our conscious control because our conscious control could still be implicated in complex decision making.

While it is true that the experiments only deal with simple motor decisions, consciousness may not be as central or necessary as Mele believes; this will be discussed in section 2.7.1. Even so, if determinism is true, complex decisions are themselves causally determined and have no way of being free in the sense of being uncaused. Before exploring how conscious control could be implicated in complex decision-making, it will be important to determine what the role of consciousness could be, and to do this, we need to have a plausible theory of what consciousness itself *is*.

## 2.3 The Role of Consciousness

The neuroscientific experiments seem to threaten *conscious* control over choices and actions. What is the role of consciousness in the causation of actions? Before answering this question, it is important to look at what a plausible theory of consciousness could be and then from this, it can become more apparent what role consciousness can play in control over actions or decisions.

### 2.3.1 Conscious vs. Nonconscious States: Dual Process Theories

One way of thinking about what consciousness might be is to differentiate it from nonconscious processing. Dual process theories, such as the one described by Daniel Kahneman (2011), do this by explaining brain function in terms of two systems. The systems do not map directly onto the brain, rather, the systems are used to describe different functions and processing capabilities of the brain. System 1 or 'fast thinking' is nonconscious, automatic, effortless, and seemingly involuntary. According to Kahneman, System 1 contains innate skills including perception of the world, the orienting of attention, and the recognition of objects. Various other skills that we gain through development and experience also become part of System 1 as they become habitual. So far so good: nonconscious processing is automatic and results from habit. System 2, or 'slow thinking', on the other hand, is deliberative and effortful and "allocates attention to the effortful mental activities that demand it, including complex computations…[its] operations… are often associated with the subjective experience of agency, choice, and concentration" (Kahneman, 2011).

The difference between them is that "[t]he automatic operations of System 1 [can] generate surprisingly complex patterns of ideas, but only the slower System 2 can construct thoughts in an orderly series of steps" (Kahneman, 2011). Although system 2 comprises the conscious self that feels in control, the inputs into system 2 come from system 1. This is

consistent with the result of the experiments above, where there is considerable conscious processing prior to conscious awareness of a decision or choice to act. The nonconscious processing is part of system 1 and this processing then becomes a part of conscious processing, of system 2. The inputs from system 1, however, are subject to various biases and cognitive errors. One of the roles of system 2, the system associated with conscious thought, is to deliberately override these errors and alter automatic functions. It does this at the cost of mental effort and the use of sustained attention resources such as increased muscle tensing, heart rate and blood pressure, and pupil diameter (Kahneman, 2011). That there is a *deliberate* overriding of errors makes it sound like there is an agent and is consistent with the folk conception that there is a controller. This will be discussed later, however, it is important to note here, that deliberate only means that the processing is sequential, slow, and requiring mental effort.

Another similar model is the Unconscious Thought Theory by Dijksterhuis and Nordgren (2006). This is a comparison of cognitive processing in terms of decision-making. On this account, "conscious thought can follow strict rules and is therefore ideally suited for if-then decisions like arithmetic, whereas nonconscious thought is better for integrating large amounts of information" (Nordgren et al., 2010, p. 510). Criteria for UTT state that consciousness is limited in capacity, works top-down, leads to more inconsistent weighing of criteria, can follow rules and do arithmetic, and is convergent. Conversely, nonconscious processing has a much larger space and can integrate and organize many variables, leads to more consistent weighing, can only give estimates and cannot carry out complex computations, and is more divergent and broad, leading to more unconventional and creative thinking (Dijksterhuis and Nordgren, 2006).

These theories show that consciousness can be differentiated from nonconscious processing in that it is limited in capacity, requires working memory and attention, is

sequential, overrides cognitive errors, and requires higher mental effort. It can further be the process by which learning occurs; new information or skills are integrated into habitual, automatic functions after a period of conscious attention and practice (e.g. Damasio, 2012).

While there may be some distinction in processing between conscious and nonconscious states, then, conscious states are still the result of the causal integration of nonconscious states. Ultimately, this means that the causal nature of brain/mental states is maintained.

### 2.3.2 Conscious vs. Nonconscious States: An Unclear Boundary & the Challenge to Conscious Control From Social Psychology

Clearly, conscious processing can be different than nonconscious processing. As previously mentioned, however, while dual process theories can be indicative of broad differences in processing, their distinctiveness cannot be mapped directly onto the brain. Indeed there is evidence that the boundary between them is not clear; it is not fixed, but rather comes as a matter of degrees. For example, in experiments, the subjective report of conscious perception of stimuli is often used to denote what a subject is consciously aware of. The problem, however, is that subjective reports may depend heavily on subject sensitivity to stimuli and response criteria. For example in Simons and Chabris' (1999) famous 'Invisible gorilla' experiment, they showed short videos to participants of two teams, dressed in black and white, respectively, passing a basketball amongst them. Participants were instructed to count the number of passes made only by the team wearing black t-shirts. In the middle of the video, a woman in a gorilla suit walked through the middle of the screen, paused, thumped her chested and walked away. Surprisingly, about half of participants report not seeing anything unusual, and they are certain when told that the gorilla was not there. This experiment illustrates the 'inattentional bias', where we are often consciously

aware of only that which our attention is focused on and not other aspects of our visual fields (Simons & Chabris, 1999). This experiment shows that consciousness depends on where attention is directed, which itself can change based on the criteria that is used by participants. Consciousness can be broadly directed or can be very narrow and specific, illustrating that it can vary in degrees as attention shifts.

Not only does consciousness depend on sensitivity and response criteria, but there is also evidence that things we are not consciously aware of can influence or alter our behaviours. This evidence threatens free will as it threatens conscious control over actions or choices. The evidence suggests that perception of stimuli can occur without subjective awareness (Merikle et al., 2000). Further, "information perceived without awareness both biases *what* stimuli are perceived with awareness and influences *how* stimuli perceived with awareness are consciously experienced" (Merikle et al., 2000, p. 115). In other words, nonconscious perception of stimuli can influence the way that information is processed and can alter subsequent choices and behaviours. These are called priming effects, and when they come about due to stimuli in the environment, are situational factors that play a potentially significant role in determining how people will act when placed in certain environments. For example, in a famous experiment by Isen and Levin (1972), participants who had found a dime in a telephone booth were 22 times more likely to help a stranger who had dropped her papers than participants who had not found the dime. Darley and Batson (1973) found that participants who were instructed to give a talk about being a good Samaritan, when in a hurry to get to another building to give the talk, were less likely to help a stranger needing medical assistance. Indeed the degree of hurry varied inversely with the number of people that stopped to help; of those in a very hurried state only 10% of people helped, of those in a medium hurry, 45% helped, and of those that were not hurried, 63% helped. In a few field experiments testing the degree to which participants are likely to litter,

Keizer et al. (2008) found that in an environment that was dirty with abandoned shopping carts and graffiti on the walls, 69% of people in a field experiment littered. Conversely, when in a clean environment without graffiti or litter, only 33% of people littered. Latane and Rodin (1968) performed a study where participants, while in the waiting room, heard cries from a nearby room. Though subjects claimed not to be influenced in their actions by there being another person in the waiting room, 70% of participants helped when alone, 70% helped when there with a friend, but only 40% helped when there with one stranger, and only 7% helped when there was one stranger who shrugged off the cries. These experiments and the many others done since show that choices can be significantly altered by factors beyond conscious awareness or control. In doing so, they put substantial pressure on conscious control over choices and decisions.

Beyond the influence of external situational factors, there are internal motivational states, emotional states, or cognitive biases that can also influence the perception of stimuli and subsequently influence behaviour. These biases can shape the interpretation and consideration of evidence. For example, "people are less likely to recall certain kinds of relevant information, less likely to believe unwanted evidence, and less likely to use critical resources to attack conclusions that are motivationally neutral" (Kunda, 1990, p. 671). There is also evidence that unrelated negative emotions, acquired by watching a sad movie, can increase perceived responsibility of another person in a moral scenario, when compared with the judgements of participants not primed with the emotional stimuli (Lerner et al, 2008). Furthermore, Kahneman (2011) cites many cognitive biases in our thinking. These cognitive biases and errors are a result of automatic, nonconscious system 1 processing that comes about before conscious experience. One of the more important is the optimism bias, which generates the illusion of control discussed in section 1.4 (Kahneman, 2011).

The fact that we have implicit biases and that situational factors beyond our conscious control can influence our behaviour, illustrates that many of our actions are a result of processes that are nonconscious. The instability in our behaviour as a result of these biases and as a result of priming and situational factors is humbling in that it shows that decreases the role of consciousness and conscious control; in many circumstances, we may have less of it than we think. Furthermore, as previously argued, nonconscious neural firing patterns are the inputs for conscious processing and mental states, once again showing that there is no freedom apart from the prior causal factors that shape these neural firing patterns and consequent thoughts, emotions, and actions; there is only the relative freedom of having the neural processes function without impairment.

As seen above, the boundary between conscious and nonconscious states is unclear. Given this, it is difficult to know which actions result after conscious deliberation, and which are a result of physiological changes, habit, experience, of cues from the environment. While all of these factors threaten conscious control, they do not negate the role of consciousness altogether. It is still possible that consciousness may play *some* role, albeit a more limited one. Before seeing what the role of consciousness might be, the question of what consciousness *is* will have to be addressed.

### 2.3.3 A Plausible Understanding of Consciousness

In the previous sections we saw that conscious processing can be broadly functionally different from nonconscious processing and that the boundary between conscious and nonconscious states not clearly defined. In order to better understand what role consciousness may itself play, it is important to explore what consciousness is. Although this question is controversial, taking a scientific perspective can help in getting closer to what it might be.

Consciousness is often described as subjective experience that includes attention, various emotional states, and various mental states such as intentions, beliefs, desires, and the like. Because consciousness involves an integration of many mechanisms and processes in the brain, it can plausibly be thought of as an emergent property resulting from these interactions. Emergence is often thought of in terms of complex systems, where the function of the whole system is greater than the aggregate sum of its parts. Findlay and Thagard (2012) argue that "wholes can have emergent properties, which are properties that belong to the wholes, do not belong to any of the parts, and are not aggregates of properties of the parts" (p. 455). Indeed it is the dynamic interaction between parts is what creates these new properties. Findlay and Thagard argue that each of the parts has an identity *tag,* which is used by organizing processes to bring the parts together in various ways. There are also *attachers*, which are processes or forces that hold together the parts, and c*ommunicators* that allow those parts to interact (Findlay & Thagard, 2012). In the brain, there are billions of neurons held together by glial cells, which communicate and interact in various ways. Their complex interactions and the integration of neural firing patterns is what creates emergent properties like conscious mental states.

In nature, an example of emergence in is the behaviour of ant colonies; individual ants do not have full information, they act merely with limited local information, and yet the action of all of the ants together create emergent properties of creating nests, tunnels, defending territories, and the like (Casti, n.d.). Notably, "there is no master ant overseeing the entire colony and broadcasting instructions to individual workers" (Casti, n.d.). In a similar way, individual neurons do not have the full information, and yet the interaction of neurons and neural populations can create various emergent mental states. Just as there is no controller in the ant colony, there is no controller in the brain. The question of a controller will be discussed in a subsequent section. What is important here is that the interactions of

neurons, like the interactions of ants in the ant colony, produce conscious experience, without any overarching direction.

To support the idea that consciousness arises as a result of the interaction of neuronal populations, there is a good deal of evidence to suggest that consciousness is correlated with widespread activation and integration of neural firing patterns across brain areas. For example, neuroscientist Antonio Damasio (2012) provides evidence from sleep studies, anaesthesia, and brain disorders to show that consciousness, as differentiated from these nonconscious states, requires the coordination of many brain sites, with the upper brainstem, thalamus and cerebral cortex being especially important. He presents evidence for convergence-divergence zones (CDZs) that represent the collective firing patterns of many neurons. Those zones coalesce to form higher-order convergence-divergence regions (CDRs). On Damasio's account, consciousness as subjective experience arises from the complex bringing together of chemical stimuli, bodily stimuli, action plans, external stimuli, and 'self' processes. Especially important are the upper brainstem, thalamus and cerebral cortex (Damasio, 2012). As an integration of different brain areas and neural firing patterns, consciousness may work to allow the organism a more comprehensive ability to regulate homeostatic function.

### 2.3.4 Consciousness Plays Only a Limited Role in the Causation/Control of Action

As we have seen above, consciousness arises as an emergent property as a result of the integration of brain states and neural firing patterns. The question that remains is whether consciousness only arises after choices have been made, and so does not have any causal role in action, or whether consciousness can in some way affect the decision-making process. Because of the closed causal nature of the brain and mental states, there is no sense in which neural firing patterns (and consequently mental states) can be uncaused. Indeed, they are all

determined by prior mental/brain states that come about as a result of genetic, environmental and experiential factors. Neural firing patterns can lead to action, but consciousness itself cannot, as consciousness is not its own distinct entity apart from those firing patterns; it emerges from their interaction.

If consciousness is understood as an emergent property arising from the interaction of neural firing patterns and neurochemical interactions, it may still have a few important roles, specifically in learning and decision-making. With regard to learning, conscious states allow the integration of new information with existing representations (Montague, 2006). Learning how to effectively engage in any new activity is a slow and conscious process. Skills such as driving, reading, playing sports, and interacting with others are gained through a slow process of step-by-step conscious learning and linking together of the parts, until the whole process is transferred to nonconscious neural networks. This transference of conscious processing to nonconscious processing starts from birth and endures the lifetime of the organism (Suhler & Churchland, 2009; Damasio, 2012). Once these skills become automatic, certain brain areas decrease in activity and develop a more extensive network such that actions become more rapid and automatic. Consciousness is then directed to monitoring and adapting to obstacles or changes in the environment. This is because attentional deployment and system 2 processing use significant mental resources and are limited (Damasio, 2012). In this way, because of the transferring of skills and processing to nonconscious neural firing, the organism is able to maintain efficiency.

Beyond its role in learning new skills, and monitoring performance change in the environment, conscious states integrate new information with stored experiences and future plans such as to revise or update existing models, goals, behaviours, and rewards (Montague, 2006). The integrative function can also serve as the basis for creativity in problem solving as

mental representations, models, and ideas are combined to find novel solutions (Montague, 2006).

Consciousness can also play some role in decision-making as it facilitates sequential, rule based thinking, and can work to align decisions with the organism's goals (Dijksterhuis & Nordgren, 2006). Particularly when it comes to decisions that are made over long periods of time, conscious deliberative inputs and rules can highly influence the decision that is reached by the organism as different weights and values are considered. The decisions themselves are reached nonconsciously, however, and conscious awareness can come some short time after this. For this reason, conscious deliberation does not have as large a role to play in momentary decisions, but rather has a larger role to play in analysis and future planning (Damasio, 2012). In the moment there is merely monitoring and integration of external inputs. Responses to the environment arise nonconsciously, yet they can be changed as a person engages in conscious deliberation, at the peril of quick response and with the consequence of interrupting smooth functioning.

In these ways, consciousness can have some limited, indirect role to play in the causation of actions through the integration new information, monitoring of the environment, the facilitation of a certain kind deliberative (slow, effortful, sequential) processing, planning for the future, and by playing a role in the revision of plans, models, values, and rewards. It can thus create changes to the system, though these changes do not happen because consciousness is an independent entity, but rather because of the role and function of consciousness that arises from the complex interaction of neural firing patterns (mental representations) and physiological states. The ongoing interaction of these processes produces choice nonconsciously, where consciousness has a role to play in the deliberative processing that goes into making choices, but where awareness of the choice made still only comes about as a result of the nonconscious mechanisms that have produced the choice.

## 2.4. The Role of the Self/Agency in Control

When looking at conscious control over actions, there is often an implicit assumption that there is an agent that is controlling the actions. Indeed we have a *conscious feeling* that we are agents or have a 'self' from which intentions and actions originate. Even in modern psychology, there is an implicit assumption that there is an inner person-like, executive agent, or homunculus, that causes actions to come about by making decisions that are uncaused (see Wegner, 2005 for examples). This agent is then responsible for behaviours that an organism carries out. As Wegner shows, this assumption is problematic; it is unclear just how an inner agent could make uncaused choices. To be uncaused the agent would have to be independent of the laws of physics. Further, if there were to be a little homunculus inside our heads, then presumably there would also be a homunculus inside of the head of that little person, and this would lead to an infinite regress (Wegner, 2005).

Clearly the views of the 'self' that assume (either explicitly or implicitly) that there is an uncaused conscious agent that is the originator of actions, are problematic. Given this, we need a scientific understanding of the 'self' that can be compatible with determinism and causality. The best way to understand the 'self' is by combining an evolutionary perspective with a theory of the self that views it as the complex interplay of different mechanisms.

For the evolutionary perspective, Antonio Damasio (2012) presents a plausible view of the self. At a very basic level, from early in evolutionary history, there is a 'proto-self' that has basic emotional feelings, maps the body of an organism, integrates various sensory inputs, and functions in the regulation of homeostasis and the maintenance of conditions for the organism to flourish. Later, a 'core-self' evolved, which uses the integrations of the 'core-self' to drive in-the-moment awareness and action. Finally, the 'autobiographical self' integrates the bursts of awareness from the 'core-self' into a coherent past, present and future that allows the organism higher-order thinking and the integration of experiences as

subjective. Damasio (2012) uses the metaphor of a symphony that begins without a conductor, and for which the conductor (the self-process) emerges from the music and then functions as a "narrative brain device" that provides biological value by allowing effective life management (Damasio, 2012, p.17). Agency, on this view, is the management of life processes. Much of this happens nonconsciously, as physiological regulation. However, the more complex, higher-order thinking of the autobiographical self could be what maintains some consistency in action and integrates experiences as belonging to the 'agent'. That there is some perceived control over action could be a function of this process of integrating experiences.

For the multilevel mechanisms approach, Paul Thagard (2012) argues that the self is a multilevel system that is a result of the complex interaction of social, psychological, neural, and molecular processes. He argues that causality runs in both directions. For example, in one direction, social factors may create psychological changes that create neural changes that create molecular changes. In the other direction molecular changes may create neural, and so on. All of these processes interact and can influence the others in complex ways (Thagard, 2012). Importantly, social and psychological factors can play an important role in shaping the subjective experience of a self and in changing the self in various ways. This way of looking at the self shows that it is not static, but subject to change on the basis of the interactions between different levels.

A combination of these views illustrates that agency is a combination of mechanisms on various levels that work to manage the life process of the organism. While there may seem to be a controller or self, there is none as an independent entity external to these interactions. The feeling of a controller is a result of the integration of complex processes into the autobiographical self, leading to the illusion that there is something apart from the interactions of mechanisms. However, it is the integration processes into a narrative or

autobiography of past, present, and future that results in agency, and that is responsible for maintaining consistency in action and integrating experiences as subjective (belonging to or being a result of, the actions of the organism).

## 2.5 The Illusion of Conscious Will by an Agent and the Importance of Nonconscious Control

If we accept that the self and agency are a combination of mechanisms and processes, then ideas about conscious will begin to break down. Specifically, a picture of agency and free will that is reliant on full conscious and deliberative decision-making, which controls all subsequent action, becomes less plausible. Nonetheless, as Wegner (2005) points out, "the creation of our sense of agency is critically important for a variety of personal and social processes, even if this perceived agent is not a cause of action. The experience of conscious will is fundamentally important because it provides a marker of our authorship… an authorship emotion" (p. 29). Indeed, the subjective feeling or emotional state that comes along with this authorship is helpful in that it can aid in distinguishing the actions caused by our own organism, as opposed to those that have been caused by others. It can also lead us to take responsibility over our actions and to attribute responsibility to others, forming the basis of a moral system and a functional society.

Importantly, it is not the feeling of conscious will that is the cause of actions, rather the feeling is the result of the integration of neural mechanisms providing feedback to an organism and being integrated as part of the 'autobiographical self'. Because the autobiographical self plays an important role in interpreting the current state of the organism in reference to its past states and future goals, the 'authorship emotion' that Wegner talks about is important for life management. Thus we often cannot help but make inferences of

conscious control and causality even if we may rationally understand that it is not our conscious states that are doing the controlling.

While experiments by Wegner (2005) show that often the feeling of *conscious* agency can be mistaken in many instances, they do not preclude the possibility that in some cases consciousness plays a significant (though perhaps indirect) role, as mentioned in section 1.4, and they especially do not preclude the possibility for *nonconscious* control and agency as a result of the integration of various neural firing patterns in the operation of executive control processes to manage the life and actions of the organism. Indeed as Suhler and Churchland (2009) argue, "nonconscious processes can support a robust form of control … [where] consciousness is not a necessary condition for control" (p. 341). Nonconscious control can be just as important, or more important than what is thought of as conscious control. Indeed researchers are only beginning to understand the extent and capacity of nonconscious processing, where the majority of thought is nonconscious (Suhler & Churchland, 2009).

For example, executive control functions (discussed in more detail in the next section) can work nonconsciously. There is some evidence that when distracted from an important goal or task, "nonconscious processes continue to keep the goal high in priority until resumption of the goal-related action, no matter the interruption by task-irrelevant contingencies" (Suhler & Churchland, 2009, p. 343). This is called the 'Zeigarnik effect' (Zeigarnik, 1927 and Förster, Liberman, & Higgins, 2005 in Suhler & Churchland, 2009) and suggests an important role for nonconscious processes in goal-maintenance and nonconscious control.

Nonconscious control can come about through the process of learning involving conscious states, as mentioned in section 2.3.4. Specifically, motor and cognitive skills are first learned through a conscious process and later integrated into automatic neural networks (Suhler & Churchland, 2009; Damasio, 2012). Nonconscious self-control and discipline are

also developed as children mature through a process of socialization whereby impulses for action are limited by the prefrontal cortex in response to pressures from parents, educators, and peers (Suhler & Churchland, 2009).

## 2.6 Neural Mechanisms and Brain Areas Involved in Control, Decision-Making, and Valuation

Control, decision-making, and valuation are likely complex processes that involve the integration of many brain areas or neural firing patterns. It is plausible that both anatomical and physiological mechanisms can be implicated in these functions, and that much of this processing is nonconscious.

Anatomically, the prefrontal cortex (PFC) is most plausible candidate brain region, being the region that is associated with the regulation and inhibition of attention, thought, emotion, and action (Arnsten, 2009). The PFC temporarily stores information for immediate use and prevents interference from distractions that might come either from the external environment or internally; this is process or storage and prevention of interference is called working memory (WM). Together, different regions of the PFC "extensively interconnect to regulate higher-order decision making and to plan and organize for the future" (Arnsten, 2009, p. 411). The regulation involves "inhibiting inappropriate actions and promoting task-relevant operations" (Arnsten, 2009, p. 410). The PFC also allows flexibility in behaviour as attention, goals, and decision-making criteria shift and monitors errors and progress in order to indicate when a change in direction is necessary. Indeed the PFC could be heavily involved in maintaining representations (neural firing patterns) of the organism's intentions and goals and further signalling other areas responsible for action to align behaviours with those goals (Miller & Cohen, 2001). On the view presented by Montague too, the PFC is a crucial area for the process of goal-maintenance, and is the place where the formation and

selection of goals takes place. Cognitive control is the ability to choose actions that are in line with those overarching goals in a "cycle of gating, goal selection, goal stabilization, and goal feedback" (Montague, 2006, p. 128). Gating is the blocking of goal-irrelevant information and internal communication of the PFC. Indeed many times the PFC blocks inputs from other areas and maintains internal communication, to help stabilize goals (Montague, 2006).

Because the PFC acts in goal selection and stabilization, it is likely that the extent to which one can maintain goal-directed action even in the face of conflicting desires or drives, correlates with the 'will' or control (Montague, 2006). Similarly, because default brain activity is largely automatic, the extent to which one has conscious deliberative processing capacity such that there is the ability to integrate new information and change existing models and control emotions, thoughts, and actions, can also indicate capacity to choose, apart from automatic processing.

Physiologically, molecular mechanisms include neurochemical changes such as those that involve hormones, enzymes, and neurotransmitters. The dopamine system in the midbrain (that works with the PFC striatum and other areas) is one of the most important for control and for decision-making (Montague, 2006). It is an example of a reinforcement learning system that plays an important role in valuation and choice. It is a neuromodulary system, meaning that it "[plays] a volume-control role…[and a role in the] direct ongoing valuation of experiences, decision-making, and even changes in working memory" (Montague, 2006, p. 105). Other neuromodulary systems also exist using different neurotransmitters and acting on different parts of the brain (e.g. histamine, acetylcholine, serotonin, norepinephrine). The dopamine system located in the midbrain broadcasts information to other parts of the brain including the cerebral cortex is a crucial part of the decision-making apparatus. One of the ways in which it is important is that it gives *ideas* the same status as rewards, such that ideas can shape and motivate action in the same way that

biological drives can. Indeed this allows the possibility of complex and abstract goals (Montague, 2006). To summarize, "[t]he dopamine system simply combines information from other brain regions with [a] reward signal, produces a reward predication error, broadcasts it widely, and teaches the rest of the brain about the new reward" (Montague, 2006, p. 138).

Based on these views, it is clear that control, decision-making, and valuation can be understood in terms of the neural anatomical and physiological mechanisms that interact to produce them. They can be scientific and there need not be anything mysterious about either. It also becomes clear that there is no sense in which there is freedom apart from the genetic, biological, and experiential inputs and their interaction in the complex computational system, which is our brain. The only freedom to be had is relative freedom where the mechanisms are intact and functioning optimally. Thus in terms of control over actions, decisions, and emotions, I agree with Suhler and Churchland (2009) there are two conditions that have to be met in order to maintain it to some degree. Firstly, the brain must be intact and not damaged by trauma or disrupted circuits. Secondly, physiological parameters must be within normal to optimal ranges, such that homeostasis can be preserved. The degree to which control can be maintained, then, depends on the normal to optimal functioning of these mechanisms.

Importantly, while control and decision-making do require conscious states, conscious awareness as such is not always required, and indeed much of this processing and control happens through nonconscious processes. As Damasio (2012) argues, conscious processes do guide nonconscious ones to some degree, through the integration of new knowledge and constant revision of models in line with goals and values of the organism.

## 2.7 Summary, Potential Objections, and Implications

In this chapter, I have argued 'free will' can be understood in terms of (mainly nonconscious) executive control over attention, thoughts, and actions, and in terms of decision-making and choice in line with the goals and values of an organism. In section 2.1, I have shown that determinism and advances in cognitive science that provide evidence for a deterministic picture of brain activity can lead to a scientific view of human cognition, where cognition can be understood in terms of causal processes and mechanisms. On this view, mental states and brain states are different levels of explanation and have a causal relationship, where conscious mental states are produced by the complex interactions of neural firing patterns. The experiments by Libet and others, outlined in section 2.2, show that neural firing patterns precede conscious awareness of simple choices and that conscious states are preceded by nonconscious activity that is beyond conscious control. In this way they pose a threat to the primacy of conscious control over choices. In section 2.3, I argued that there is good evidence to support the idea that consciousness results from the integration of brain states and consequently, is not a separate entity with causal efficacy apart from the brain mechanisms that create it. Indeed conscious states can be broadly distinguished from nonconscious states; however the distinction between them is not very clear. This makes it difficult to differentiate which actions may result from conscious deliberation or processing, and which result from nonconscious processes. Ultimately, consciousness itself may not be causally efficacious in initiating action, but does play a role by facilitating certain types of sequential, slow, integrative, and controlled thinking. Conscious states are important in monitoring the environment and could be important in decision-making, future planning, and in the revision of plans, models, values, and systems of valuation. Problematically, it is often assumed that there is a 'self' or 'controller' that is responsible for initiating actions; this was the topic of section 2.4. In this section I argued that views of the self in evolutionary and

mechanistic terms show how the integration of stimuli and mental representations can lead to this illusion of control by an external self, and yet this is illusory, as the self and agency simply *are* the result of the causal interaction of various levels of mechanisms (genetic, neuronal, psychological, social) and nothing external to them. In section 2.5, I argued that while the feeling of conscious control by an agent can be mistaken, this does not preclude the possibility for nonconscious control, and indeed nonconscious control and processes have a large role to play. Finally, in section 2.6, I outlined some of the anatomical and physiological mechanisms that could be implicated in the 'will' including the prefrontal cortex and dopamine systems. Broadly speaking, the implications of this view are that there is no freedom that is uncaused, but only a relative freedom as seen in relation to the optimal or normal range functioning of mechanisms that comprise the will. Further, relative free will still allows that we make choices (constrained by values, preferences, genetics, experiences and the like) and despite a decrease the role of conscious control and executive function, can lead to *more* control over thoughts, emotions, and behaviours. This revised understanding of free will in a causal deterministic way should also lead to increased compassion and understanding of others; it should change the our attitudes and behaviours. Each of these implications will be discussed as objections are considered and discussed.

### 2.7.1 Response to Objections; Why and How 'Free Will' Needs to Change

As outlined in section 1.2, Mele and Roskies argue that advances in neuroscience do not change conceptions of free will. Mele (2013) argues that complex decisions are different from the simple motor decisions tested by the neuroscientific experiments, that conscious deliberation and processing can play an important role in complex decisions, and that unconscious firing patterns that produce a decision may have come about (at least in part) after conscious processes. On these points I agree with him. Consciousness may indeed play

an important role through the integration of inputs and revision or setting of overarching goals and values to be followed in the decision-making process. However, while wakeful conscious states may be required, it is unclear that conscious attention and explicit deliberation is always necessary because PFC activity in controlling and maintaining goal-states can be largely nonconscious. Evidence from studies of decision-making by Dijksterhuis (2006) show how nontrivial decisions can be made effectively without conscious deliberation. In a study comparing conscious and nonconscious decision-making, Dijksterhuis (2006) asked participants to make purchasing decisions in one of two conditions: one in which they were able to consciously deliberate, and another in which they were distracted and could not deliberate consciously. The decisions involved small household items (hand towels, toasters) and big items (house, car). Interestingly, in both cases the group that did not have time to consciously deliberate made better decisions (based on pros and cons) than the group that had time for conscious deliberation. Dijksterhuis (2006) suggests this could be because consciousness has a limited capacity, making it more difficult take into account all of the relevant information. Conversely, he suggests that nonconscious space is not limited in this way and so can lead to better decision-making when there are many pros and cons to weigh (Dijksterhuis, 2006). As Damasio (2012) points out, however, the nonconscious space likely acts in accordance to previous experience or emotions regarding such purchases, experience which was likely at some point conscious.

The implications of this experiment are that consciousness may have some role to play along the line, if only to allow wakeful states or even to integrate past experiences with goals, values, and plans for how to act or choose in the future. Even complex decisions, however, need not involve conscious deliberation. Mele may allow for this, as he argues that decisions are free so long as consciousness played *some* role. However the role of conscious deliberation is not as necessary nor as extensive as Mele claims. Conscious processing may

have an indirect role to play, but even complex decisions can be made largely through nonconscious processing and without 'conscious reasons' or a 'conscious struggle'.

In emphasizing the importance of conscious deliberation, Mele argues that conscious deliberation which plays some role in the decision allows for free will *as traditionally conceived*, though he shies away from giving a definition of what this might be. If free will is traditionally defined as freedom of the will from causal factors, then this is not plausible because inputs from decisions and for conscious thoughts come about from nonconscious processing. If we understand free will in a causal way, which Mele is inclined to do, then the conception of freedom must be modified such that it is a relative freedom, relative to more un-free states in which neural processes are disrupted or there is external coercion. The concept must also be modified in recognition of the more limited role of consciousness and the fact that consciousness may not always play as substantial of a role as Mele believes.

Though the topic of moral responsibility is beyond the scope of this thesis, an outcome of the causal deterministic view is that there is no possibility for ultimate moral responsibility, the kind of responsibility that may be found if there is indeed full conscious control independent from causal factors. On the deterministic view, all future events are causally determined by prior events, where behaviours are a result of complex processing in the brain with these inputs. All of the inputs, however, are a result of factors beyond the conscious control (genetics, how one was raised, experiences had, mental states), and this is why there cannot be ultimate responsibility to the person.

Thus the concept of 'free will' cannot remain unchanged because a new understanding only allows for a very limited and relative freedom and a will that involves some conscious processes, but involves nonconscious processes to a larger degree. Indeed the requirement that consciousness is involved at some point is not enough to say that an

action was free, there need be an absence of internal or external impediment to the proper physiological function of these mechanisms.

Further the difficulty in distinguishing between conscious and nonconscious states, and the largely illusory nature of conscious will by an agent lend more support to the idea that nonconscious processes are very much a necessary component of any theory of the will, and that conscious deliberation in the moment is not the only requirement.

In line with potential objections by Mele, Roskies (2006) argues that the findings from neuroscience do not change anything with regard to free will. She argues that "even if [the findings] do affect the public's apprehension of the problem, they predict little effect on actual judgements of moral responsibility" (p. 423). The latter problem of how this will impact people's judgements is an empirical question, yet even if there is no change (which has yet to be shown), this does not preclude the normative question of whether there *should* be a change in attitudes and practical responses. I do not believe that the philosophical question can be so easily separated from the evidence presented. Rather, the above evidence, and the evidence from cognitive science that supports the deterministic picture of the brain *should* change public attitudes and behaviours.

As previously mentioned, an outcome of this view is that there cannot be any ultimate moral responsibility or blame, only that which is dependent on a relative freedom that is causal in nature. Consequently, justice systems that are reliant on retribution as a form of punishment are unjustified. Rather, the focus must be on deterrence and rehabilitation, such that the causes of behaviour are well understood and there is a view to altering future behaviour. While there may not be any possibility for ultimate responsibility, responsibility can still be had in a practical sense. The scientific understanding of will still allows that we can still hold ourselves and others responsible for actions performed, even if those actions were beyond conscious control. We do so as a matter of practical consequences in that doing

so can change future thoughts and behaviours and align them with the desirable social outcomes.

Treatment of mental illness too should be oriented toward improving the ability for people to have (conscious and nonconscious) control in line with goals and values, and to engage in deliberation and decision-making. It also means that there need be a greater focus on nonconscious processes and control, and how those can be trained through education and through the lifetime of learning (again, causal ways in which they can be influenced). Responses to others can be changed in that this view decreases the degree to which we can blame others, and should increase understanding and compassion as the causal reasons for behaviour are understood.

Another practical outcome of this view and the way that attitudes should change is that there is a much larger role for social interconnectedness and responsibility. Understanding that we are each the products of firstly our genetic predispositions, but more importantly their interaction with the environment, the societies that we grow up in, and the social interactions and experiences that we have means that there is no sense in which we can separate our 'selves' from these causal factors. While it is true that we may form ideas or values or goals that contravene on learned patterns of behaviour and can go against that which was previously learned, even the new ideas arise due to the complex computations in our brains, the basis for which is those very factors. This also highlights the powerful and important role of education and learning in broadening choice. Particularly important in the future will be education about brain function, control, and decision-making, and the degree to which nonconscious processes can be implicated in these.

### 2.7.2 Two Additional Concerns; Increased Control and Increased Acceptance

There are several other concerns with the proposal that the 'will' can be understood in this causally deterministic way. These include the idea that this kind of view will lead to fatalism and powerlessness over one's present condition, and that this view ignores subjective mental states and will dehumanize. Rather than leading to fatalism or dehumanization, I argue that this view leads to increased control over one's situation, and a more compassionate, accepting attitude toward others.

Firstly, rather than leading to fatalism, this view leads to increased control. Fatalism is the idea that everything is predetermined and so the outcome of events cannot be changed and nothing you do matters to what happens. If future events are causally determined by past events, then future events are in some sense inevitable. Fatalism can lead to a feeling of powerlessness over one's life, such that there is a loss of motivation and a loss of willingness to make an effort to change one's circumstances. It could be argued that taking a causally deterministic approach to the brain and cognition can lead to fatalism and its consequent disempowerment.

I argue, however, that rather than leading to hopelessness, the causal view can, paradoxically, increase choice and control. It can do this because it can lead to a greater understanding one's own causal history. This can lead to a deeper analysis of thoughts and behaviours in the present context based on one's upbringing, experiences, biological factors, predispositions, and future plans. This increased understanding of 'self' is akin to the adage "know thyself", but departs from self-knowledge that is purely internal, and rather is a call to examine all of the causal factors that have led someone to behave in a certain way or to have certain thoughts, emotions, motivational states, perceptions, situational factors, and so on. This includes a big role for environmental and social factors, and may lead to more accurate indications of personality and behaviour than a purely introspective view.

Understanding our embodied and fallible nature opens the possibility to gain tools and skills to modify behaviour. This includes learning coping strategies or tricks that work to help one maintain consistency between intention and action. As an example, we imagine John, who wants to eat cookies, but knows that he should not be eating them for health and weight reasons. If John knows that historically he could not resist the cookies when they were in the house, he could exercise control not by expecting to exercise willpower when faced with the temptation, but by choosing to not put himself in that situation in the first place, by keeping the cookies outside of the house.

A deeper understanding, then, can increase the number of options for choice and the feeling of control over one's life. It can do this by giving people tools that can help increase the ability to align actions with intentions as well as the ability to revise the intentions themselves. With this increased understanding there is the recognition that even if we do not have much *conscious* control, there may be an increased role for nonconscious control and processing, and it is still our organism that is making choices. Thus we can still take ownership of these choices and integrate them into our autobiographical selves. Even if the choices are made in the brain before we become consciously aware of them, recognizing the way that decisions are made can help us to gain exposure to more decision-criteria and facilitate the conscious deliberative, system 2 type, processing.

Another important concern raised by this type of objection is that we are limited by the causal history that produced us in the sense that we must follow a specific path or are 'destined' to do certain things. While it is broadly true that the future is determined by the complex interaction of causal factors, that interaction is so complex that saying one is 'destined' to do one thing or another is overly simplistic. Practically speaking, there are so many factors that determine choices and the computations in our brains, that this kind of concern should not be worrying. Indeed in our daily life, what is more important is to

acknowledge is the incredible computational power of our brains and the ability to change habitual responses and revise plans, models and valuations, despite of and in light of the causal nature of inputs. Recent evidence about neuroplasticity, the ability of our brains to change and adapt, shows that change is possible even later in life (e.g. Doidge, 2007). Of course there are enduring patterns of neural activity that become increasingly difficult to change, but the possibility for change means that there is a case to be made for optimism and against a narrow view of 'destiny'.

Lastly, the objection about fatalism negates the fact that all of the actions of our organisms (from speech to body language to behaviours) make a concrete impact in the world and influence other people. There is still great power to be had in this condition of limited freedom. Thus there is also a responsibility to learn all we can about our cognitive function and how that is related to the particular cultural context we are in. This learning process can again give us *more* freedom. Recognizing that we are working within the bounds of our genetic predispositions, experience, knowledge, physiological states and the like can help us to generate intentions to learn more and to gain more exposure to different fields, which can then increase computational power of our brains, corresponding to an increase in control and decision-making ability.

The second objection is that the neural understanding of cognitive function negates the importance of subjective states (e.g. happiness, sadness, etc.) and dehumanizes. For former, it is important to recognize that the neural level of explanation is compatible with, and only a different *level of explanation*, of the mental states/psychological phenomenon. They are both describing the same thing. Indeed as Thagard and Aubie (2008) point out, a neural view can actually give us "explanations of many aspects of what it is like be happy, for example that happiness has positive valence, varying intensity, onset and cessation, and that it is discriminable from other emotions" (Thagard & Aubie, 2008, p. 830).

Not only is there a worry that subjective states will be negated, but also that advances in neuroscience and discussions of the causal mechanisms of the brain, will dehumanize. For example, Peter Strawson (1964), though convinced that a belief in determinism would have no effect, thought that if it did, it would cause us to adopt an objective attitude that would mean 'human isolation' and an inability to participate fully in interpersonal relationships; it would mean not recognizing others as agents capable of making choices.

I argue that instead of dehumanizing, a belief in determinism actually leads to an increase in compassion whereby we gain a recognition of our common humanity: the extent to which we are all governed by mechanisms and are the products of our genetics, environments and experience. It also leads to humility and objectivity in approaching others (and oneself) as it is evident that we are all subject to cognitive biases and errors, possible emotional distortions, possible troubles with emotional regulation, stress, and so on.

Because adopting a causal deterministic perspective does not negate mental states, it means recognizing that each person has a rich mental life. Understanding cognition in terms of the processes that comprise it also means understanding the very human needs for recognition, love, fulfillment, and so on. This is because these things are necessary for healthy psychological functioning and participation in society. Further, the recognition of a limited agency is not precluded. We still make choices through the complex computational systems, which are our brains. While it is true that the choices are constrained by genetics, experiences, and values interacting with inputs to the system from social and situational factors, choices are made nonetheless.

Rather than increasing isolation and dehumanizing, then, a causal deterministic view can increase compassion, respect, and understanding while maintaining a robust sense of agency through the exercise of (both conscious and nonconscious) cognitive control, decision-making, and model/value revision.

## CHAPTER 3: Depression and the breakdown of mechanisms

In this chapter, I will examine the relationship between mental illness and control or choice, focusing on depression. Ultimately, I will argue that the understanding of mental illness in terms of causal determinism can work back to inform healthy psychological functioning by providing more support for a causally deterministic picture of healthy brain function. Depression and other mental illnesses occur as a matter of degrees where healthy functioning can be similarly impaired. Examining depression and its related symptoms gives more evidence to support the model of will in the previous chapter. On this model, the will involves optimal functioning of (conscious and nonconscious) prefrontal cortex executive control functions, conscious processing that is responsible for integration, learning, goal-direction, and revisions, and proper functioning of the valuation and goal-setting mechanisms themselves. In the previous section I argued that there cannot be any 'freedom' in terms of an uncaused agent making choices in a way that is independent from prior causes. Each choice or decision is preceded by neural firing patterns and decision-making criteria or values that are determined by prior factors including genetic predispositions, environmental influences, how one was raised, and social influences. While there cannot be any true freedom, there can be a limited and relative freedom when the neural and physiological mechanisms responsible are intact and functioning within the normal range and there is no disruption due to disease, stress, or external coercion. The difference in freedom can clearly be seen with the increased severity of mental illness as it can cause a decrease in control, decision-making, and valuation systems, compromising the ability to match actions with intentions and impairing the very intentions and valuations on which actions may be based.

Ultimately, the causal deterministic view of both impaired and healthy functioning can show that the stigma of difference or the conception of those with mental illness as

'others' is mistaken and not justified. This can lead to a more compassionate view of both those with mental illness and those with healthy psychological functioning.

### 3.1 Depression Can be Caused by a Variety of Interacting Biopsychosocial Factors

Depression is caused by the complex interaction of biological (e.g. genetic, disease) and psychosocial factors (e.g. interpersonal, environmental factors). Indeed the biopsychosocial model for mental illness is most prevalently used when trying to sort out causality of depression and other mental illnesses (Engel, 1980). This means that mental illness can result from the complex interaction of some or all of these factors. This is compatible with the multi-level model of the self that was presented in the previous section, where each level of functioning can be affected and interacts with and changes the others. Furthermore, that depression can be understood in terms of this complex interplay of causal factors gives more reason to believe that the causal deterministic picture is accurate, where there is no 'freedom' apart from this causality. The consequence of this is that there can only be a limited causal 'freedom', which is present in the absence of disease or impairment of neural networks and firing patterns.

#### 3.1.1 Genetics

Depression sometimes runs in families, and so it is possible that there is some genetic component that is independent from the environment a child is raised in. While a genetic component is likely and may contribute to risk, the interaction with other complex factors (environmental, developmental, psychological, and social) may contribute to substantiating that risk and initiating the course of the illness (Cowan, Kopnisky, & Hyma, 2002). Current research that explores interactions between genes and the environment is based on research in other areas that has shown that diseases like cardiovascular disease, which have some genetic

components, can be propagated by social factors such as diet/lifestyle (Cowan et al., 2002; Lang et al., 2012). The risk for depression also increases when there are changes to genes that encode proteins that are involved in the regulation of stress signalling pathways (Arnsten, 2009).

There can also be epigenetic mechanisms implicated, where epigenetics is the study of the effects of environmental factors on protein expression in the cell. "Epigenetic mechanisms can provide a potential pathway by which early experience can have lasting effects on behaviour" (Insel & Quirion, 2005, p. 2222). At one time, only about 15% of genes are active in the cell at one time, the rest controlled various mechanisms including methylation. For example, an external stimulus could reduce methylation, activating genes and changing genetic expression, leading to behavioural outcomes. Genetic variations have been shown to shape predispositions to vulnerability and resilience when faced with a stressful situation (Insel & Quirion, 2005).

### 3.1.2 Causality and Comorbidity with Biological and Mental Disorders

Evidence suggests that depression can be caused by, cause, or be comorbid (coexisting) with other biological and mental disorders. Some causes implicated in the onset of depression or its related symptoms can include deficiency of vitamin $B_{12}$ and folate (Tiemeier et al., 2002), and low blood sugar or iron deficiency, which can lead to chronic fatigue (Onder et al., 2005). Illnesses that may be causally related with depression include cancer, Parkinson's disease, HIV/AIDS, thyroid conditions, heart disease, diabetes, and stoke. These illnesses can lead to more severe depressive symptoms and depression can lead to more difficulty in adapting to and treating the medical condition (Cassano & Fava, 2002).

Depression can also be comorbid or caused by other psychiatric disorders such as anxiety disorders including Generalized Anxiety Disorder (GAD) and Post-Traumatic Stress Disorder (PTSD), panic disorders and Obsessive Compulsive Disorder (OCD) (APA, 2010).

### 3.1.3 Psychological and Social Stressors

While depression may have no trigger, very commonly, stressful life events or trauma can trigger depressive symptoms and episodes (Arnsten, 2009). Stress and its resultant anxiety can cause a decreased ability to engage executive control mechanisms, undergo conscious deliberation or make choices in line with the goals and values of the organism. This decreases relative freedom, where freedom is defined by the optimal functioning of these main processes. As a result of acute psychosocial stressors, there is impairment of the prefrontal cortex (PFC), which normally has connections with other brain areas in order to regulate attention, emotions, behaviours, and thoughts and maintain them in line with the organism's goals and values (Arnsten, 2009). Further, chronic stress can lead to structural changes in connectivity in the PFC, and changes in the dendrites in the hippocampus and amygdala, together leading to a disruption of working memory, attention, longer term memory, and emotional circuits. It can be very damaging because it "weakens the structures that provide negative feedback on the stress response and strengthens the structures that promote the stress response" leading to chronic trouble in regulating this response, and higher susceptibility to mental illness (Arsten, 2009, p. 418).

Physiologically, the stress response can alter the neurochemical environment in the brain, leading to increased levels of neurotransmitters like noradrenaline (NA) and dopamine (DA), which under optimal conditions help to regulate thoughts, emotions and behaviours. High levels of these neurotransmitters can impair PFC connectivity and function (Arnsten, 2009). Studies have shown that emotional distractions, social stressors (e.g. public speaking

or performing calculations in front of a panel of strangers), imagining stressful situations, or even studying for stressful exams, can impair PFC functions including working memory, cognitive flexibility, and attention regulation (Arsten, 2009). This emphasizes the extent to which environmental factors can cause impairments and lead to depression.

Importantly, the formation and function of the PFC can be affected by chronic stress during the developmental years, and this can in turn affect the susceptibility to mental illness later in life (Arsten, 2009). For example, traumatic events or maltreatment early in life can increase the risk for depression (APA, 2003).

Clearly stress can strongly affect both those with healthy psychological functioning, and those with less healthy or pathological functioning, and indeed can be a major cause of the transition from the former to the latter. Because there is decreased PFC function, stress causes a loss of executive control functions and ability to engage flexible decision-making circuits that involve conscious processing to make good choices in light of the presented data. This is particularly perilous if there is new information that needs be incorporated into the decision or new valuation schemas need to be taken into account. This is because cognitive function switches from slow, system 2 type processing mediated by the PFC, to the automatic and habitual function of system 1 that is associated with 'primitive' brain systems including the amygdala and subcortical structures related to it (Arnsten, 2009).

Ultimately the above-presented evidence shows that psychosocial stressors can cause impairment of PFC activity and executive control functions that are important for control, choice, and valuation, and can thus lead to the onset of depressive episodes. In less severe cases they can impair function in healthy people and even change the structure of the brain if exposure to stress is chronic. Either way they lead to a shift away from the functions of the brain that allow flexibility and changes to automatic processes. This shift can clearly happen to different extents depending on the severity of the stressors and individual responses to

them. The fact that it can happen to different degrees highlights the ways in which healthy psychological functioning and mental illness too occur in gradients, mental illness being a deviation from optimal structural and physiological functioning.

### 3.1.4 Other Psychological and Social Causes

One of the causes of depression is that a person has poor coping strategies or problem-solving skills in the face of stressful life events such as the death or illness of a loved one, relationship breakup/divorce, interpersonal conflicts, work related stress, and financial pressures (APA, 2010). A lack of coping strategies means that people can be easily 'overwhelmed' by the stressor, and not have the tools to effectively manage it or reduce its impact on mood or daily function (APA, 2010).

Depression may also be a result of cognitive biases. For instance, Gotlib and Krasnoperova (1998) provide evidence that depression is associated with biases in attention and memory toward negative information. Specifically, it has been found that while in the general population there is an optimism bias and propensity to overestimate positive future outcomes, in those with depression, positive future outcomes are underestimated (and may be more realistic) (Strunk, Lopez, & DeRubeis, 2006). These biases may constitute an increased vulnerability for depression or increased likelihood of repeated depressive episodes (Gotlib & Krasnoperova, 1998). Further, as previously mentioned, cognitive biases are largely nonconscious and beyond conscious control. They can be overcome in normal functioning by conscious deliberative processing at the cost of sustained attention and cognitive resources and through the function of the PFC in as it monitors and corrects cognitive errors (e.g. Kahneman, 2011; Arnsten, 2009).

Depression can also be associated with rumination, where negative, self-critical, self-doubting, and worrying thoughts are repeated over and over, are intrusive, and are difficult to

escape (Siegle & Thayer, 2004). Rumination may be associated with positive (reinforcing) feedback loops in the brain, and with impairment of the brain structures that normally regulate or shut off those feedback loops. The former include emotional processing centers such as the amygdala and hippocampus, while the latter means impairment of prefrontal cortex activity (Siegle & Thayer, 2004). There is evidence that depressive symptoms and the onset of depressive episodes can be predicted by the degree to which people engage in negative rumination or worry (Nolen-Hoeksema, 2000).

Depression is also correlated with low self-esteem or feelings of worthlessness, possibly as a result of bullying, psychological abuse, childhood or other trauma. There is strong evidence to support a vulnerability model, whereby low self-esteem plays a part in the onset of depressive episodes (Orth & Robins, 2013). It is speculated that vulnerability can be transformed into MDD through the effects of social isolation and rumination that may occur as a result of low self-esteem (Orth & Robins, 2013).

Cultural factors can also lead to an increased risk for depression. For instance, in North America, there is a very high emphasis on 'happiness' and on maintaining a perpetual state of elevated mood. There is some evidence to suggest that highly valuing happiness may sometimes be self-defeating and lead people to disappointment at their levels of happiness, and consequently increase depressive symptoms (Mauss, Tamir, Anderson, & Savino, 2011). Similarly, this emphasis on happiness has caused many to make this state of elation a goal to strive towards. Evidence from research on goal setting indicates in depression and other psychiatric disorders, there is a bias toward abstract goals (rather than concrete goals) (Watkins, 2011). If goals are abstract and difficult to attain, this can lead one to become trapped, unable to let the goal go and unable to achieve it, proliferating depressive symptoms.

Overall it is clear that depression comes as a result of the complex interplay of a variety of causal biological, psychological, and social factors. The freedom associated with

the normal functioning of anatomical and physiological processes erodes as PFC activity is impaired, further compromising executive control functions, conscious and nonconscious, and undermining the degree to which an person can engage in conscious deliberation and processing for learning and to establish or maintain the gaols for decision-making.

The role of psychosocial stressors also comes into clear view on this causal account as it becomes clear that these stressors change the brain. One of the most debilitating symptoms of depression, as it pertains to control or choice, is indecisiveness or decreased ability to think or concentrate. As mentioned above, this can be a result of the impairment of PFC activity resulting from both biological and psychosocial stressors.

### 3.2 Treatments

The treatment of mood disorders also shows that depression is causal in nature and that external interventions are needed to increase control and decision-making ability. In the treatment of mild to moderate depression, there are two broad routes: psychotherapy and pharmacotherapy. In more severe cases, other treatment options are also explored, including Electroconvulsive Therapy (ECT), Deep Brain Stimulation (DBT), and Transcranial Magnetic Stimulation (TMS). The ideal endpoint of these treatments is *remission* of symptoms such that the person is able to function at pre-episode levels and only has minor and infrequent residual symptoms. *Recovery* is when remission is maintained for more than two months (Craighead & Dunlop, 2014).

The prognosis for depression can vary depending on a variety of factors including the complex causality, stressors, effectiveness of treatments, social support, etc. Prognosis is inversely correlated with the chronic nature of the disease; long-lasting, chronic depression is more difficult to treat and prognosis is not as promising for those with chronic and recurrent depression (Blom et al., 2007). While only 30-40% of patients will achieve remission of their

symptoms (assuming ideal conditions present in clinical studies), about two-thirds of patients will see some positive improvement (DeRubeis, Siegle, & Hollon, 2008). It is interesting to note that while patients respond differently to treatments, single modality treatments (psychotherapy alone or medication therapy alone) are "essentially identical" in terms of overall outcomes (Craighead & Dunlop, 2014). Indeed combination therapies can be more effective, as will be discussed below.

Looking at which kinds of treatments are effective for depression lends more support to the view that depression can be viewed in a causal way and as the breakdown of various mechanisms. The aim of treatment is to restore function and cognition and to teach patients skills and strategies to manage emotions and behaviours. This coincides with an increase in control and decision-making ability and an increase in the relative freedom that I've mentioned in previous sections.

### 3.2.1 Psychotherapy

Currently one of the most used psychotherapies is Cognitive Behavioural Therapy (CBT), which works on the basis of changing cognitive appraisal associated with emotionally salient events (Beck et al., 1979). It is generally administered over an extended period, where patients participate in 12-16 weekly or biweekly sessions. Because those with depression tend to have negative appraisals and thought patterns, CBT helps people re-interpret their interactions with others and environmental stimuli in a more positive way. Psychoeducation is usually a part of CBT, and refers to the education provided to those with mental illness regarding the prognosis, course, treatment options, coping strategies, tools, and the like, that can aid in their recovery. As a part of CBT, it is used to help people identify self-sabotaging behaviours or behaviours that worsen depressive symptoms in order to change their behaviour in those situations (APA, 2010).

Part of the basis of CBT and psychoeducation is a scientific understanding of emotion. Emotions can be understood in neural and mechanistic terms. For example, Thagard and Aubie (2008) advance a neural theory of emotional consciousness. On this model, conscious emotional experience is a result of the interaction of brain areas in working memory. Emotions "arise from a combination of neural representation, somatic perception, cognitive appraisal, and working memory" (Thagard & Aubie, 2008, p. 811). Emotions *are* the "neural processes that represent the cognitive and somatic state of the organism" (Thagard & Aubie, 2008, p. 811). Importantly, emotions involve both somatic perception (physiological/bodily responses) and cognitive appraisal.

Cognitive appraisal is the interpretation and evaluation of the emotions in the context of the current state of the organism, past experiences, and current/future goals. Indeed appraisal may be the differentiating factor between physiologically similar states, particularly when it comes to complex social emotions like shame, pride, or guilt, for which goals or aims can be social in nature. For example, a feeling of anxiety in a stressful social situation (e.g. a first date, public speaking) can be interpreted in a negative way in terms of fear, but can also be interpreted as excitement and looked upon with positive anticipation. In these two cases physiological state is very similar, but the appraisal is very different. As previously mentioned, people with depression have negative biases in attention and memory, this can cause negative appraisal of emotionally salient events (Gotlib & Krasnoperova, 1998). As such, one of the strategies used in therapy is the changing of that appraisal. The process of changing emotional responses is called emotional regulation, and it can be done both with the input of conscious processing and nonconsciously. According to James Gross (1998), there are five possible points throughout an emotional reaction that can be modulated. These are the selection of the situation (people, places), change of the situation, deployment of attention, changing the understanding of the situation (framing, appraisal), or finally through

altering the emotional response through various actions including exercise, music, and medications (Gross, 1998).

The fact that the appraisal can be changed and the knowledge that this appraisal is part of the way in which emotional assessments are made, is evidence to support my argument that an understanding of brain function in a deterministic way can *increase* control. This is because knowing exact mechanisms can help one, through conscious and nonconscious processing, intervene at certain points and effectively create a change in interpretation of events and in the psychological or behavioural response to them.

Apart from Cognitive Behavioural Therapy, there are several other types of psychotherapy including interpersonal therapy (IPT) which involves psychoeducation and counselling about interpersonal relationships, behavioural activation therapy (Jacobsen et al. 2001), behavioural marital therapy, problem-solving focused therapy, and short term psychodynamic psychotherapy (APA, 2010). The basis of all of these types of therapies is counselling and psychoeducation, whereby the patient is able to learn strategies and gain tools to overcome and improve maladaptive thoughts, behaviours, and relations. On average, there is little evidence that one psychotherapy is better than another; when it comes to evidence based psychotherapies, all are about equal in terms of efficacy (APA, 2013). The choice of treatment generally depends on patient choice, availability of competent/trained therapists, relevance of the treatment modality to the symptoms (Craighead & Dunlop, 2014).

Interestingly, all of these therapies operate on the basis of a scientific view of emotion and with the aim to alter automatic or habitual responses through the function of the PFC. Similarly, by providing psychoeducation about brain function and coping strategies, they change the inputs to the system and increase the number of choices available to patients, as well as the capacity of the decision-making apparatus to function optimally. That these therapies are so effective is good evidence to support the view that psychoeducation

(understanding of brain function) can lead to increased control and better decisions, even when it is based on a causally deterministic account with a view to alter thinking patterns and the inputs that the computational system runs on.

When depression becomes severe enough, psychotherapy is no longer effective, and antidepressant medications are used, at least in part to regulate the physiological and neurochemical environments and bring patients up to a level at which psychotherapy can be helpful in increasing choice, control, and valuation.

### 3.2.2 Antidepressant Treatments (ADM)

According to the American Psychiatric Association's (2010) treatment guidelines for depression, antidepressants can be grouped into several classes based on structure and function. There is no evidence of superior efficacy in any of the groups, so treatment indication depends on various other factors such as safety, family history of success or ineffectiveness with a medication, personal history, cost-based considerations, patient choice (placebo, therapeutic optimism), side effects, and possible interactions with other medications. First line antidepressants are Selective Serotonin Reuptake Inhibitors (SSRIs), Serotonin-Norepinephrine Reuptake Inhibitors (SNRIs), mirtazapine, or bupropion (APA, 2010). Second line treatments, used only if first-line treatments are ineffective due to their increased health risks, include monoamine oxidase inhibitors and tricyclic antidepressants. Most antidepressants take at least 4-8 weeks to become effective (to eliminate depressive symptoms). Dosage may be increased every 2-4 weeks up to the maximum recommended dosage (APA, 2010).

Combination therapies, combining both psychotherapy and antidepressant medications may be on the whole more effective. Patients that derive most benefit from combination therapies are those with chronic MDD. Several models exist for combination

therapies: additive, synergistic, and adherence models (Craighead & Dunlop, 2014). Additive therapies work because individual patients may not respond to single modality treatments (only antidepressants or only psychotherapy) and so engaging in both may increase remission rates. The synergistic model proposes engaging in one modality to supplement another. For example, pharmaceutical interventions may be necessary in order to bring people up to a level of functioning to allow successful professional intervention with CBT or other therapies Further, patients with depression may have hippocampal damage, preventing creation of new memories with CBT, pharmaceuticals can help with this in order for the patient to be able to do CBT. Patients may also not want to take antidepressants at first, and the CBT could help them recognize this need. The adherence model is simply starting one of the treatments in order to be able to do the other. This means either starting psychotherapy to be able to adhere to pharmaceuticals or taking pharmaceuticals to be able to adhere to psychotherapy (Craighead & Dunlop, 2014).

While combination treatments may help to decrease the severity of depression or lead to remission, there are several barriers to treatment or factors that predict poor outcomes. The strongest predictor of poor outcomes with MDD is low socioeconomic status (Galea et al., 2007). Others include old age, limited education or lower intelligence. In addition, childhood trauma, abuse, neglect, and adversity lead to poorer outcomes, and there is a significantly worse prognosis when MDD is comorbid with high levels of anxiety (Craighead & Dunlop, 2014). These barriers clearly illustrate the importance of causal environmental impacts on depression and how these factors may constrain choice and control.

### 3.2.3 Treatments for Severe Depression

When depression is severe, neither psychotherapy nor antidepressant medications are effective. In these cases the neural structures and function is so heavily impaired that there is

no ability to choose or to control emotional or behavioural states. When this happens, more aggressive therapies are used, including Electroconvulsive Therapy (ECT) and Deep Brain Stimulation (DBS), and Transcranial Magnetic Stimulation (TMS). ECT is a therapy that is used after all other options have been tried. It involves electrodes applied to the scalp which induce a seizure in the patient for one minute, changing electrical activity in the brain (Lisanby, 2007). DBT involves continuous electrical pulses from electrodes that are implanted directly into the brain from a generator implanted in the chest. There is some evidence that stimulating subgenual circuit white matter can change activity in limbic-cortical circuits associated with depression (Mayberg et al., 2005). TMS is also a promising new therapy for depression that uses magnetic fields produced by a coil that affects the cerebral cortex in the area where it is applied. It stimulates neuronal firing, is associated with changes in blood flow, it changes the release of neurotransmitters including dopamine, and can possibly even alter the expression of certain genes (Gershon et al., 2003)

### 3.2.4 Summary and Discussion of Therapies

On the basis of the above discussion of treatment options for MDD, it is clearly the case that causal factors are at play and that the biopsychosocial factors outlined earlier have an overwhelming role to play in the treatment of MDD. Where MDD causes impairment in cognitive function and executive control functions, there can also be decreased ability to make choices to seek treatment in the first place, or to adhere to a treatment regime. This effect is only compounded by the stigmatization and blame that those with MDD and other mental illness must endure.

Furthermore, the relative freedom and ability to match intention to improve function with adherence to treatment can be impaired and impeded by fatigue and psychomotor retardation, the long wait to treatment efficacy, demoralization when treatments are

unsuccessful, and the adverse side effects of medications, including weight gain, upset stomach, and headaches. These latter side effects and symptoms can also decrease ability to make exercise control and to choose to act in ways that could be beneficial.

### 3.3 Depression Involves a Breakdown of Mechanisms and a Loss of Control with Increased Severity

The main question to be answered about depression with regard to free will, is in what sense can choices made while in the grips of depression or mental illness be said to be free? Psychiatric disorders cause a breakdown in various neural mechanisms and in this way decrease even the relative freedom that is had in healthy psychological functioning. The neural mechanisms that break down during depression include executive control functions mediated by the PFC, both conscious and nonconscious. There can also be a breakdown in the ability to engage in conscious processing, meaning that some of the various functions of conscious processing can be impaired. These include monitoring of the environment and adaption to changing conditions, the integration of inputs, deliberation, reflection, learning, future planning and the ability to revise plans, models, values and rewards. There can also be a breakdown of the degree to which the brain (consciously and nonconsciously) integrates past experience, present actions, and future goals into the autobiographical self, and as such a breakdown of the feeling of agency or control over one's behaviour and choices, and over the ability to influence external events. Furthermore, the ability to shift between conscious and nonconscious processing may be impaired as there is a shift to nonconscious, habitual, and automatic brain circuits. The ability to make choices or decisions evaluated in the context of the values and goals of the organism may also break down both as a result of impairments to decision-making neural circuits and as a result of a breakdown of mechanisms for valuation or goal-setting.

Two of these are particularly important: the decreased conscious processing and shift to nonconscious automatic and habitual circuits, and the breakdown of valuation mechanisms. In the first, there is a decreased ability to form new ideas that can then veto the previously formed ideas and habits (Montague, 2006). This means that the mind goes further and further on 'autopilot', relying habitual and constrained patterns of thought and behaviour. These states can also cause fixation on certain goals and the inability to let them go (the basis for rumination), or make it such that it is very easy to become distracted from goals and aims. These states can also interfere with valuation and the ability to create and maintain goals (Montague, 2006). The second, involving the breakdown of mechanisms for valuation, means that valuation may not be assigned properly or that the option with the highest value may be maladaptive ones such as constant sleep or decreased physiological functioning, as other actions are perceived by the brain to require too much effort or too many resources. If valuation mechanisms break down, then there really is no choice; the values upon which any choice could be made are entirely beyond control of the PFC and associated structures (Montague, 2006).

When addressing the question of how much choice is had while a person is depressed, the answer likely depends on the severity of the depression and the degree to which cognitive function, particularly PFC activity related to decision-making, valuation, and control over actions is impaired. In psychiatric clinical practice, there are several ways of assessing this kind of impairment through the Mental Status Examination (MSE), which is an examination of the mental state of the patient not based on what they say, but based on the patient's presentation (e.g. Zimmerman, 2014). In the MSE both insight and judgement are assessed. Insight involves the degree to which persons are aware of their illness and its implications. Judgement involves the ability to make or appreciate different choices and outcomes and to choose reasonably based on them. Both of these outcomes can be indicative of the extent to

which cognitive mechanisms have broken down or are impaired. Various aspects of cognition can also be tested formally through tests like the Mini-Mental State Exam (Folstein M.F., Folstein S.E., & McHugh, 1975) and the Montreal Cognitive Assessment (Nasreddine et al., 2005), again giving indication as to which mental processes are disrupted and to what degree. Future tests could also be developed to test executive control functions, the ability to make choices independently of, and to veto habitual responses or thoughts and impulses that arise in consciousness (Montague, 2006).

The disruption or impairment of cognitive mechanisms can lead to paralysis in making choices or acting and indeed depression is often described as "walking through water" where there is a global decrease in the ability to function and activities of daily living become enormously difficult to do (Foti & Kanazawa, 2008). This impairment becomes more and more severe as depression severity increases. Exploring several case studies can show how executive control and choice/action/valuation mechanisms progressively break down as the severity of depression increases. They can further show the complex interplay of bio-psycho-social factors that cause depression, and that further act to influence treatment outcomes and prognosis. The case studies are based on real cases encountered in psychiatric practice through my own completion of an observership. All identifying information has been changed.

For a case of mild depression, we can consider Mark, a successful 24-year-old student studying engineering at university. Mark has been having difficulty getting out of bed in the mornings, and wants to sleep all the time. He's lost appetite over the last few months and has lost 20lb. He has trouble sleeping at night and constantly feels fatigued during the days. He reports that "everything is off" and that "nothing is fun anymore." He has stopped going out with his friends and is increasingly socially isolated. He has lost the motivation to work and has trouble finishing his assignments or concentrating in class. He can still get to class, but

everything is a struggle and he feels like he's constantly overwhelmed. He feels that he has lost the feeling of purpose in his life and he feels a sense of hopelessness about the future, that his degree will take him nowhere and that his future is empty and bleak.

From this case it is evident that mild depression involves persistent low mood or lack of interest, where daily function is not fully impaired. Function may be more slow or more of a struggle, but there can still be some ability to choose and control actions, despite low mood fatigue, and an inability to think clearly. As previously mentioned, for mild depression, Cognitive Behavioural Therapy or counselling can provide an improvement, as there is an intervention to help change appraisals or maladaptive thinking patterns or behaviours.

For a case of moderate depression, we can consider Jonathan, a 35-year-old male who describes functional impairment in his ability to take care of himself and his young children. Like Mark, he has trouble getting out of bed and always feels exhausted, claiming that "everything feels impossible to do." He has started missing days from work, and has not attended any social events. His children and wife are concerned as he finds it very difficult to find the energy to engage in activities of daily living such as showering or going to buy groceries. Jonathan's trouble started when he had an emotional 'breakdown' at work. He feels that he has been treated badly and unfairly singled out over the last few months. He feels overpowering anxiety when it comes to leaving the house and feels that he cannot bring himself even to call or contact anyone from his place of employment.

From this case, it is clear that moderate depression includes a functional impairment in activities of daily living, instrumental activities of daily living, and employment. Generally a combination of antidepressants and CBT can be effective for moderate depression, ADs to bring the patient up to a level that he can at least choose to seek therapy or see a councillor, and CBT to help him overcome the stress and anxiety associated with his workplace, such that he will be able to begin ADLs and gradually return to work.

For a case of severe depression, we can consider Dorthy, a 70-year-old mother and grandmother. She has been in hospital for 6 months, suffering from Major Depressive Disorder that is resistant to treatment. She has been taking antidepressants for several years with no change in her symptoms. She reports that she has severe anxiety. On mental status examination she has shaking in her arms and legs. She is difficult to interview because she only gives one or two word answers, saying that she feels hopeless and that she sees no reason to keep going. She just wants it all to end, but doesn't know how to make it stop. She cannot get out of bed on her own, and only does so upon continual persistence from nursing staff. She has difficulty performing a psychological test, depression and anxiety preventing her from thinking clearly. She has nearly no delayed recall (recall of words after 5 minutes) and although she shows some insight into her condition, her judgment is impaired.

Severe depression clearly includes a halt to daily functioning including ADLs and iADLs. People with severe depression are largely paralyzed and unable to exercise executive control or to make choices. At this level of severity there is no freedom in the slightest and external interventions need be more aggressive including antidepressant therapy and sometimes ECT and DBS, due to its recurrent and treatment resistant nature.

These cases illustrate that with the increased severity of depression, there is also increased impairment in carrying out the ADLs. The degree of control and decision-making ability also decrease with increased severity, where in the most severe case it is very difficult to make any choices at all. These cases also show that depression can have a varied causality, onset and presentation, the degree of impairment will depend on specific people what which causal factors are contributing to their depression and constraining their freedom. For Mark, depression could be caused by social isolation and a breakdown of goal-setting mechanisms or inability to accurately view the future. For Jonathan part of the causality is from his being treated unfairly at work. For Dorthy, there may be a biological cause that makes it such that

her depression is resistant to treatment and recurrent. Some of the constraints that can influence control, choices, and valuation in these and other cases include fatigue/exhaustion (caused by decreased sleep or oversleep), stress and anxiety, decreased self-esteem, loss of hope and future orientation, and a breakdown of goals or aims.

For an additional case, we can consider the symptoms of suicidal ideation and hopelessness as they are manifest in MDD. Jane is a 25-year-old female suffering from post-partum depression. She describes her pregnancy as the "happiest time in her life" but after her daughter was born, she was been overcome by persistent low mood and suicidal ideation to kill herself. In hospital, she is in tears and visible emotional pain: "I'm done" she says as she recoils under the blankets, "who would want to live like this? The moment I leave this hospital I'm going to kill myself, I'm just waiting for the chance." Since the baby was born several months ago, she has been getting very little sleep and has a complete loss of appetite. Her relationship with her partner has deteriorated and she feels alone and without any social support. Furthermore, both her and her partner are under substantial financial pressure since they both have lost their jobs and they cannot afford to pay the rent at their current apartment.

In Jane's case, the depression is plausibly a result of a combination of hormonal changes and psychosocial stressors. The constrains on her freedom include the enormous burden of stressors including the birth of her daughter and lack of sleep, the relationship strain with her partner, the financial pressures from a loss of employment, and poor coping strategies. It is possible that she has negative bias toward the future, unable to see positive events occurring. She may not have access to positive representations or outcomes, and furthermore her cognition is clouded by overwhelming emotion and an inability to appreciate options for action. Her freedom is thus internally constrained by the breakdown of PFC mechanisms for executive control and decision-making and externally constrained by the

possibilities open to her, her external environment playing a large role in propagating her depressive symptoms and her symptoms of hopelessness and suicidal ideation.

This latter case is interesting because depression sometimes presents with recurrent suicidal ideation, characterized by a state of hopelessness and worthlessness where a person feels she cannot go on living. Risk factors for suicide are many and include a family history of suicide or mental illness, demographic factors, other psychiatric and physical illness, and psychosocial factors including low social support and relationships, loss of employment or unemployment, decrease in socioeconomic status, stressful life events, and childhood trauma (APA, 2003). Other risk factors include psychological and cognitive features such as decreased self-esteem, impulsiveness, aggression, shame and humiliation, severe anxiety, decrease in executive function, constricted thinking, and polarized thinking (APA, 2003).

Protective factors that can prevent suicidality include a sense of responsibility to family, friends, or community, positive coping skills, problem-solving skills, social support, life satisfaction, religiosity, and a positive relationship with therapists (APA, 2003). Both risk factors and protective factors indicate that suicidal ideation, as a part of depression, can be understood in a causal deterministic way. Of course it is the complex interaction of these factors in particular situations that can lead to one to attempt or carry out the suicide.

Based on the above considerations, it becomes more clear that depression can be understood in a causal deterministic way based on the complex interplay of biological, psychological, and social factors. It also becomes clear that the will, understood as normal functioning of executive control, decisions, and valuation processes is heavily impaired with increasing severity of depression.

### 3.4 A Deterministic View Should Decrease the Stigma and Increase Compassion

The stigma associated with mental illness includes seeing people with mental illness as indulgent, using it as a dishonest excuse, pathetic, or weak (Byrne, 2000). It also involves seeing those with mental illness as 'other' or different. As a consequence the experience of mental illness includes feelings of shame, self-blame, isolation, social exclusion, and discrimination (Byrne, 2000). It often leads to secrecy, which leads to more isolation, which can then further exacerbate depressive symptoms. It can cause immense suffering, be dehumanizing and alienating as those with mental illness are denied, as Johnstone (2001) aptly puts it, "moral entitlements to the things that other so-called 'normal' persons may take for granted: [daily] function, a sense of place in an intersubjective world, empathic connection with reciprocating others, peace of mind, happiness, participatory citizenship" (p. 200). This stigma needs to be overcome if people with mental illness are to have their rights respected such that they are no longer "marginalized, invalidated and/or ignored" (Johnstone, 2001, p. 202). The prevalence of mental illness in the general population means that this problem must be addressed; according to the Canadian Mental Health Association, nearly 20% of the Canadian population, about seven million people, will have a mental illness in their lifetime. Of these, 8% will experience MDD (Canadian Mental Health Association). These figures make the fight against stigma all the more relevant and necessary.

Similar to what I argued in section 3.3, those with mental illness are not free due to both internal and external pressures (Johnstone, 2001). Internally, disrupted cognitive functioning and unregulated emotional processing create confusion, fear, and feelings of worthlessness or self-doubt. Externally, social isolation and attitudes of suspicion and fear from others intensify feelings of isolation, increasing psychosocial stressors and exacerbating symptoms. This social isolation leads to those with mental illness being treated as 'different'

or ‘other’, as less than human and as not worthy of the same respect and understanding as those with healthy psychological function.

Some of the stigma and othering of those with mental illness is a result of false beliefs about the causal nature of conscious will and agency. There is an implicit assumption that there is a conscious agent directing actions and making choices somehow ‘freely’. In this way depression can be seen as a weakness, something that one should just have the ‘willpower’ to overcome. Those with depression might also be seen as lazy and not trying hard enough to exercise ‘willpower’ to improve their condition. Those that are mentally healthy may also say that “depression is a choice” and advise those with depression to “choose to be happy”.

Some of the stigma associated with depression also comes from ‘positivity’ movements. One example is *The Secret,* a book by Rhonda Byrne (2011) where one’s thoughts are said to fully direct the course of one’s life. If one has positive thoughts and ‘wills’ a positive outcome, than that positive outcome will happen. Problematically, the implicit message is that if a person doesn’t have positive thinking, then that person is to blame for negative outcomes. While thinking positively or framing events in terms of learning opportunities may be helpful (e.g. emotional reappraisal of emotionally salient events), the assumptions of causal efficacy and ‘willpower’ are dangerous and harmful, blaming the victims for not ‘choosing’ to be happy.

The stigma and view of those with mental illness in these above mentioned ways is clearly troubling and problematic because it negates the complexity of the illness and merely leads to futile blaming of those with mental illness. While in healthy psychological functioning regulation of mood and an optimistic outcome may be more easily influenced by intentions to be happy, in those with mental illness intentions are not efficacious in the same way. Furthermore, this stigmatization can lead to an internalization of feelings of guilt or worthlessness on the part of those with mental illness, where they are unable to ‘just choose,’

and indeed have immense trouble engaging in basic activities of daily living. The inability to initiate action where there is an expectation to "just have willpower" can cause shame, frustration, and ultimately more harm to those with depression.

One way to overcome the stigma of difference is through a causal deterministic understanding of cognitive function in both those with mental illness and those with healthy mental functioning. Similar to what I argued in section 2.7.2, the causal deterministic view can lead to more compassion, understanding, and acceptance as it becomes clear that we are all governed by mechanisms that are causally shaped by genetic, biological, psychological, environmental, and experiential factors. The main difference is that in the case of mental illness, there is a breakdown of mechanisms, such that processes are largely automatic and habitual. In the case of healthy cognition, the complex and robust computational ability of the brain functions optimally and allows for control, valuation, and choice apart from habitual or previously learned thoughts or behaviours, and indeed allows for a modification of previous responses. In both cases there is no freedom from prior causes, there is only more of less of a breakdown of mechanisms, which happens on a continuum. This is important because relative freedom that may allow for practical responsibility and agency breaks down in degrees to the point where there cannot be even relative freedom and consequently no responsibility. Both the questions of what counts as optimal function and to what degree the breakdown occurs, need be empirically informed as they form the basis of relatively free agency and responsibility.

This understanding can decrease the stigma associated with mental illness through a recognition that cognitive mechanisms can and often do break down for many people in the population. Even cognitively healthy people are subject to cognitive biases and impaired judgements on account of social stressors, non-optimal physiological function (e.g. due to lack of sleep, poor nutrition), substance use, mere inattentiveness or lack of willingness to

devote cognitive resources various problems. The transition from mental health to mental illness can occur for anyone given certain conditions at each of these levels as various risk factors that are beyond conscious control interact and contribute to causality.

Knowing that anyone can develop a mental illness as a result of the complex interplay of biological, psychological, and social factors, is a call for humility. It shows that in this way, mental illness is not radically different from other physical illnesses. In the same way that biopsychosocial factors interact to increase and ultimately cause depression, risks for cardiovascular disease can increase as a result of biological and genetic factors, or as a result of psychosocial factors including lack of exercise, poor eating habits, and smoking (Lang et al., 2012). Similarly, in the same way that low socioeconomic status, low education, poor work prospects and unemployment, and disability can lead to disease, these causal factors are also implicated in and increase the risk of mental illness (Galea et al., 2007). While mental illness can develop for anyone, there are populations that have a higher risk due to psychosocial factors. The fact that mental illness does discriminate to some degree against those that are less fortunate, we have all the more reason not to stigmatize it.

When it comes to motivation and intention to change, there should be a recognition that change is not a matter of 'willpower' but rather a question of changing inputs to the system. This means, for instance, that change requires creating an environment without temptations, learning new coping skills, creating new habits, and building new relationships, rather than the abstract 'willpower.' As previously mentioned, a causal view of the self also means getting rid of a 'controller' and rather recognizing the self as a process resulting from the complex interplay of neural processes and mechanisms, where changes on one level (including the social and psychological levels) cause changes on all levels of cognitive function.

A causal view means looking forward toward treatment that involves rehabilitation of those with mental illness, as opposed to isolating them from society through discriminatory practices or through permanent hospitalization. A wonderful example of a model for rehabilitation that includes social integration and meaningful employment is the town of Geel in Belgium (Wells, 2014). In this town, hundreds of foster families adopt those with serious mental illness as part of their homes. This has been a tradition for centuries, "[w]hen numbers were at their highest in the late 1930s, there were 3,800 psychiatric patients living with families in Geel, a town at the time of only 15,000" (Wells, 2014). Those with mental illness have gainful employment, including work in houses, in the fields, bicycle shops and also are part of social activities including bands and sports teams. Additionally, the "Geel psychiatric hospital provides sport therapy, work therapy, cooking and even gardening therapy for mental illness" (Wells, 2014). This is a remarkable model: there is acceptance by the community, there is no judgement, those with mental illness have homes, friendships, and meaningful experiences. There are still visits from psychiatrists, but professionals only come by every two weeks for check-ups and with medications, the rest is up to the foster families. Clearly in this town there is compassion, acceptance, and understanding. The view toward rehabilitation can be taken up by "any community … with motivation, education and effort…What Geel has always done and still does today is: (a) acknowledge and accept the human needs of the boarders and (b) respond to those needs rather than acting on unfounded or exaggerated fears" (Goldstein & Godemont, 2003). Based on a rehabilitation approach, even if symptoms do not improve, quality of life can still be greatly improved. This approach is based off both an understanding of causality based on the complex interaction of biological, psychological, and social factors, and an understanding of the common humanity of those with mental illness, leading to the breakdown of the stigma of distancing.

Taking a causal view toward understanding of onset, prognosis, and treatment, means that the complexity, prevalence, and prevention of depression can be better understood. It also means that the role of education in reducing the stigma is very important. This education can involve public campaigns, psychoeducation for patients' families, friends and the community, former and current patients acting as advocates, and indeed psychologically healthy individuals acting as advocates. Recent campaigns to decrease the stigma bring hope that this is possible. Bell Canada's campaign "Let's Talk" raised over 5 million dollars for mental health initiatives and includes Olympian Medalist Clara Hughes and other prominent spokespeople. There were over "109,451,718 text messages, mobile and long distance calls, tweets, and Facebook shares" (Bell Canada, 2014). This shows that a discussion about mental health is beginning; this is the time to provide education about causality and to reduce the stigma, such that we can increase compassion, understanding, and acceptance.

### 3.5 Implications for Autonomy

The topic of free will has strong implications for autonomy. In medical ethics, the principle of autonomy states that people have a right to self-governance based on their own chosen path and goals, without "controlling influence by others and … [without] limitations, such as inadequate understanding, that prevent meaningful choice" (Beauchamp & Childress, 2008, p. 58). In order for someone to be considered autonomous, that person needs to have capacity or competence in decision-making, be rational, well informed, and not externally coerced. This still allows that people can have idiosyncratic beliefs, even if these beliefs are strange or uncommon, such as the case with Jehovah's witnesses and their refusal of blood transfusions. It is easy to see how an internal breakdown of mechanisms or an external coercion could constrain autonomy or the ability for self-determination, or choosing what is best for one's organism. It also becomes clear that in no case can there be full autonomy,

defined as that which is not causal, because all values, goals, and decision-making criteria are causally related to prior factors and the interaction of those prior factors through computational networks. Autonomy then, even in those with normal range cognitive function, is constrained by the interplay of genetics, biological factors, psychology, environment, family, and social influences. On the causal view, then, autonomy is limited in the same way that freedom of the will is limited; there is no sense in which there is ultimate freedom, yet there can be relative freedom when neural functioning is optimal or within a normal range.

Even if full autonomy is limited in this way, there must still be a recognition of preferences, values, desires, and the like. This is because while preferences and values can change, some are so deeply ingrained that changing them would be difficult, even through the work of conscious PFC deliberation and processing. This is particularly true of long-standing patterns of behaviour or values upon which a person has invested and built over the years. One such factor could be religious beliefs, where there can be a good deal invested over a relatively long period of time. For these reasons, autonomy and personal choice must still be respected, though with the recognition that all choices and values are determined by prior factors and the complex computations or the reflection that may (or may not) have gone into forming those beliefs and values.

In the case of mental illness, like depression, the capacity to consent to medical treatments and to exercise this type of autonomy is impaired (Hebert, 2009). Depression in particular can affect one's ability to make choices in the case where thinking is clouded or where there is persistent suicidal ideation. When determining functional capacity, physicians need to test five criteria to assess functional capacity or competence: that the patient can express a choice, that it be a 'reasonable' choice, that it be based on 'rational reasons', that the person shows an ability to "understand the information necessary to make a decision" and

that the person shows "an ability to appreciate one's situation and its consequences" (Hebert, 2009). Further, one of the key criteria is not just understanding the situation and issues at hand, but also the "ability to manipulate information rationally to make a decision" (Hebert, 2009). Clearly, those in the grips of MDD may not fit this criteria not least because they may not have functional capacity or the sufficient operation of cognitive mechanisms such as to be able to make a decision or think clearly. The impairment experienced by those with MDD and other mental illness need not be all-encompassing, but certainly does increase with increased severity of mental illness.

As previously discussed, there are degrees to which faculties for executive control, decision-making, and goal/value alignment are impaired by mental illness and even in healthy psychological functioning. Because of this, autonomy and capacity are not all or none, but rather vary with mental or physiological states, where there can be differences in the ability to recognize and weigh competing alternatives and to understand the consequences of action or inaction. In medicine and particularly in psychiatry, the role of caretakers, then, is not to choose for patients, but rather to help bring patients to a level whereby they can, even if it is mainly through nonconscious control with input from conscious processing, understand decision-criteria and outcomes, deliberate and make choices aligned with overarching goals and values.

The importance of the free will question to autonomy and capacity then, is in recognizing all of the causal factors that have led someone to where they are, in recognizing potential factors that can cause impairment of dysfunction of those choice mechanisms, and further in aiding persons to re-gain those abilities.

### 3.6 Summary and Conclusions

In this chapter I have argued that depression too can be understood in a causally deterministic way. Indeed it is easier to see how depression and other mental illnesses are causal in nature than to see how healthy psychological function is. A view of depression and mental illness in this way, however, can work back to inform our view of healthy psychological functioning. It can do this through the recognition that mental illness is but a breakdown of the processes, mainly governed by prefrontal cortex activity, involved in executive control, choice, and valuation. The main difference between healthy and pathological functioning is that in the former case there is complex computational ability including functional benefits from conscious states and processing. These conscious states allow for learning and integration of new inputs into the system and additionally allow for goal creation, maintenance, and change. The will, then, can be understood as the ability to overcome previously defined maxims for action or habitual responses or valuations; it is the ability to maintain action aligned with intention, despite instinctual drives. While there cannot be ultimate freedom from causal inputs to the system, there can be relative freedom when there is normal range psychological functioning such that brain function and computational ability are not impaired. The study of depression tells us that this freedom comes as a matter of degrees; not only is it more impaired as there is an increase in severity of mental illness, but it can also be heavily diminished by biopsychosocial factors including stress, fatigue, hunger, biological illness, and the like, in the healthy.

In chapter 1, I gave an introduction to the problem, outlining the main philosophical, neuroscientific and psychological perspectives on the question of free will and the implications of neuroscientific research on free will. In chapter 2, I argued that in light of evidence from cognitive science, a causal deterministic picture of cognitive function is plausible, where mental states are brain states; changes in the brain are reflected in mental

states and changes in mental states are reflected as changes in the brain. The challenge from neuroscience to free will suggests that the role of conscious control is limited, as subjects become aware of decisions only some time after the decision has been made nonconsciously. The experiments by Libet and others suggest that neural firing patterns precede conscious awareness of choices made and lend more support to the causal view, and less support notions of free will involving conscious control.

Examining the role of consciousness, it quickly becomes clear that the boundary between conscious and nonconscious states is not a clear one, and yet that conscious and nonconscious processing can be broadly distinguished. Importantly, conscious processing involves slow, sequential, and rule based computations in a limited space and with mental effort, whereas nonconscious processing is more automatic, habitual, and prone to error or bias. Challenges from social psychology further stress that we are subject to biases, priming, and situational factors, decreasing the possibility of full conscious control and illustrating that conscious attention and error correction are limited. A plausible theory of consciousness sees it as an emergent property based on the dynamic interaction of neurons where it can play a role in the causation of action not directly, but through its role as the slow, deliberative processing. This processing allows some of the more important roles of consciousness including its role in learning and transferring conscious skills to nonconscious processing, its role in reflection, analysis, and future planning, its role in monitoring and adapting to the environment and integrating new information with existing models, and its role in the revision of plans, models, rewards, and systems of valuation. While consciousness plays an important role in cognition, being an emergent property from complex interactions, it does not as a separate entity cause action. Rather, neural firing patterns and computations that result from prior inputs cause action, with conscious states being an important part of those computations.

Implicit in many accounts of agency is the idea that an 'agent' or 'self' is directing processes and acts as a 'controller'. This view is not plausible, however; a scientifically informed view of agency sees it as resulting from the complex interaction of multilevel mechanisms including social, psychological, neural, and molecular processes. On this view, causation works in both directions such that even social and psychological mechanisms can influence neural and molecular mechanisms. The integration of mechanisms leads to the narrative of a self and gives rise to the subjective experience of agency. This subjective experience may be often mistaken, however, particularly when there is an assumption of conscious agency.

On the view of consciousness presented above, particularly in light of challenges from neuroscience and social psychology, the role of conscious agency is limited, and rather nonconscious processes become increasingly important. Through the transference of conscious to nonconscious processing, nonconscious mechanisms for control and choice are trained from childhood through to adulthood. Neural mechanisms for (mainly) nonconscious control, choice, and valuation include both anatomical and physiological aspects. Structurally, the prefrontal cortex (PFC) is of central importance for control, decision-making, future planning, and the maintenance of behaviours in line with the overarching goals of the organism. Physiologically, the dopamine system is of central importance as it is involved in valuation, rewards, and the motivation of action. In order for there to be a limited freedom of the will, understood as the complex interaction of these mechanisms, it is important that these mechanisms are functioning optimally or within the normal range.

Objections to the impact of the neuroscientific evidence of the revised understanding of free will, as presented by Mele and Roskies, are not plausible. Both argue that neuroscientific results make no difference and that conceptions of free will can be maintained unchanged despite of the evidence. In the former case, Mele is not right that there must be

conscious deliberation or processing in complex decisions in order for there to be free will. Evidence suggests that conscious deciding and deliberation are not always involved and rather that nonconscious processes have a large role to play in decision-making and choice. Further, beyond the role of consciousness, there is a requirement for relative free will that there is healthy, non-pathological function of the various mechanisms and a lack of external coercion. In the latter case, Roskies' view that the scientific understanding of free will changes nothing regarding our understanding of free will and subsequent behaviours, is also mistaken. Specifically, understanding the causal nature of cognition can lead to increased control as there are more possibilities for behavioural change. It should also lead to increased understanding and compassion for others through the recognition that all are governed computations that result from causal inputs on the basis of biology, environment, and past experiences. In this way it should change the way we treat others (and ourselves) to include more kindness and understanding and less blame. Because the view leads to an understanding of responsibility in this causal way, it leads to views of punishment that are based on practical outcomes including deterrence and rehabilitation. Similarly, it leads to an increased need for social responsibility and emphasizes the important and powerful role of education in broadening choice, control, and decision-making ability. Other objections that this view leads to fatalism and is dehumanizing can also be addressed in a similar way. Rather than leading to fatalism and disempowerment, the causal deterministic view leads to increased control and empowerment. Rather than being dehumanizing, this view leads to increased compassion and understanding.

In chapter 3, I explored the question of free will from the perspective mental illness, focusing on depression. Depression can be understood in a causal way based on the complex interaction of biopsychosocial factors. These include genetic predispositions, causality on the basis of other physical or psychiatric disorders, the important role of psychological and social

stressors in inhibiting PFC function, a lack of coping strategies, numerous cognitive biases, neural feedback loops and rumination, low self-esteem, and cultural factors. Based on this it is clear that depression is complex, but more importantly can be understood in a causal deterministic way. The numerous treatments for depression including psychotherapy, antidepressant medications, and more aggressive therapies including ECT and DBT, all work toward showing the causality and importance of interventions (rather than willpower) to improve outcomes. Indeed depression can be understood as a breakdown of mechanisms including executive control by the PFC, conscious processing and all of its related functions, shifts between conscious and nonconscious processing, and mechanisms for valuation and decision-making. As depression becomes more severe, the breakdown of mechanisms also increases, leading to less and less control or ability to make choices, and more automatic or habitual functioning where values and goals are also disrupted.

Based on this causal understanding of depression and mental illness, the stigma often associated with depression, particularly that which sees those with mental illness as needing to merely exercise 'willpower,' is unjustified. Moreover, the causal deterministic view shows that depression and healthy psychological functioning are separated as a matter of degrees, based on the normal range or out of range functioning of mechanisms. Though there is no ultimate freedom apart from causal factors, freedom or agency, understood as the absence of disease or impairment, can be constrained in many ways, even in those with otherwise healthy psychological functioning. Constraints on this agency can include aging and the increase of mental habits, context and situational factors, traumatic and stressful events, and other factors such as fatigue and hunger. Agency can also be constrained by external coercion or pressures. Furthermore, depression and mental illness do not discriminate; in the same way that biopsychosocial factors increase the risk for other biological diseases, they also increase the risk for mental illness. Rather than 'willpower,' then, treatment modalities

should include a view toward changing the inputs to the system and improving processing. This also means rehabilitation and inclusion of those with mental illness into communities; it means recognition of our common humanity.

The free will question is also important to conceptions of autonomy in medical decision-making as autonomy is constrained by mental illness. The causal view means that no one has full and independent autonomy from prior factors, but allows that people have individual goals, values, and priorities. As such the goal of healthcare professionals is to broaden choice and increase the optimal functioning of these mechanisms.

Overall, I have argued that free will can be understood in a scientific, causal deterministic way based on recent findings in neuroscience and psychology. The will can be understood as the normal range functioning of mechanisms for control, choice, and valuation. There is no free will that is uncaused, but the only freedom is relative and is had when these mechanisms are not internally damaged and there is no external coercion. I have also argued that depression and mental illness can be understood in a causal deterministic way and that this understanding can work back to reinforce the scientific understanding of the will in non-pathological cases. The understanding of free will based on healthy function and the understanding based on pathological function are mutually reinforcing. I have also argued that, based on a causal deterministic view of the will, the stigma surrounding mental illness is unfounded, and that this view can lead to more compassion, understanding, and acceptance of both those with mental illness and the mentally healthy.

## References

American Psychiatric Association (APA). (2003) *Practice Guideline for the Assessment and Treatment of Patients With Suicidal Behaviors*. Am J Psychiatry ; 160(Nov suppl):1–60 [G].

American Psychiatric Association (APA). (2010). *Practice Guideline for the Treatment of Patients with Major Depressive Disorder*. Arlington, VA: APA. 3rd ed.

American Psychiatric Association (APA). (2013). *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: APA. 5th ed.

Arnsten, A. F. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nature Reviews Neuroscience*, *10*(6), 410-422.

Barney, L. J., Griffiths, K. M., Jorm, A. F., & Christensen, H. (2006). Stigma about depression and its impact on help-seeking intentions. *Australian and New Zealand Journal of Psychiatry*, *40*(1), 51-54.

Beauchamp, T. L., & Childress, J. F. (2008). *Principles of biomedical ethics*. Oxford university press.

Beck, A. T. (1979). *Cognitive therapy and the emotional disorders*. Penguin.

Blom, M. B., Spinhoven, P., Hoffman, T., Jonker, K., Hoencamp, E., Haffmans, P. M., & van Dyck, R. (2007). Severity and duration of depression, not personality factors, predict short term outcome in the treatment of major depression. *Journal of affective disorders*, *104*(1), 119-126.

Byrne, P. (2000). Stigma of mental illness and ways of diminishing it. *Advances in Psychiatric Treatment*, *6*(1), 65-72.

Byrne, R. (2011). *The secret*. Simon and Schuster.

Canadian Mental Health Association (CMHA). (2014) *Fast Facts about Mental Illness*. Retrieved from Canadian Mental Health Association website:

http://www.cmha.ca/media/fast-facts-about-mental-illness/#.U4Tajl7cndt

Cassano, P., & Fava, M. (2002). Depression and public health: an overview. *Journal of Psychosomatic Research*, *53*(4), 849-857.

Casti, J. L. (n.d). Emergence in an ant colony. *Encyclopedia Britannica, Inc*. Retrieved from http://www.britannica.com/

Cowan, W. M., Kopnisky, K. L., & Hyman, S. E. (2002). The Human Genome Project and Its Impact on Psychiatry*. *Annual Review of Neuroscience*, *25*(1), 1-50.

Craighead, W. E., & Dunlop, B. W. (2014). Combination Psychotherapy and Antidepressant Medication Treatment for Depression: For Whom, When, and How. *Annual review of psychology*, *65*, 267-300.

Damasio, A. (2012). *Self comes to mind: Constructing the conscious brain*. Random House LLC.

DeRubeis, R. J., Siegle, G. J., & Hollon, S. D. (2008). Cognitive therapy versus medication for depression: treatment outcomes and neural mechanisms. *Nature Reviews Neuroscience*, *9*(10), 788-796.

Dijksterhuis, A., Bos, M. W., Nordgren, L. F., & Van Baaren, R. B. (2006). On making the right choice: The deliberation-without-attention effect. *Science*, *311*(5763), 1005-1007.

Dijksterhuis, A., Preston, J., Wegner, D. M., & Aarts, H. (2004). *Effects of subliminal priming of natural and supernatural agents on judgements of authorship*. Manuscript submitted for publication.

Doidge, N. (2007). *The brain that changes itself: Stories of personal triumph from the frontiers of brain science*. Penguin.

Engel, G. L. (1980). The clinical application of the biopsychosocial model. *Am J Psychiatry*, *137*(5), 535-544.

Findlay, S. D., & Thagard, P. (2012). How parts make up wholes. *Frontiers in physiology*, *3*.

Frankfurt, H. G. (1969). Alternate possibilities and moral responsibility. *The Journal of Philosophy*, 829-839.

Fried, I., Mukamel, R., & Kreiman, G. (2011). Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. Neuron, 69(3), 548-562.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*, *12*(3), 189-198.

Förster, J., Liberman, N., & Higgins, E. T. (2005). Accessibility from active and fulfilled goals. *Journal of Experimental Social Psychology*, *41*(3), 220-239.

Foti, D., & Kanazawa, L. (2008). Activities of daily living. *Pedretti's Occupational Therapy: Practice Skills for Physical Dysfunction*, *6*, 146-194.

Galea, S., Ahern, J., Nandi, A., Tracy, M., Beard, J., & Vlahov, D. (2007). Urban neighborhood poverty and the incidence of depression in a population-based cohort study. *Annals of epidemiology*, *17*(3), 171-179.

Gershon, A. A., Dannon, P. N., & Grunhaus, L. (2003). Transcranial magnetic stimulation in the treatment of depression. American Journal of Psychiatry, 160(5), 835–845.

Goldstein, J. L., & Godemont, M. M. (2003). The legend and lessons of Geel, Belgium: A 1500-year-old legend, a 21st-century model. *Community mental health journal*, *39*(5), 441-458.

Gotlib, I. H., & Krasnoperova, E. (1998). Biased information processing as a vulnerability factor for depression. *Behavior Therapy*, *29*(4), 603-617.

Haggard, P. (2008). Human volition: towards a neuroscience of will. Nature Reviews Neuroscience, 9(12), 934-946.

Haynes, J. D. (2011). Beyond Libet: Long-Term prediction of free choices from

neuroimaging signals. In Characterizing Consciousness: From Cognition to the Clinic? (pp. 161-174). Springer Berlin Heidelberg.

Hébert, P. C. (2009). *Doing right: a practical guide to ethics for medical trainees and physicians*. Oxford University Press.

Insel, T. R., & Quirion, R. (2005). Psychiatry as a clinical neuroscience discipline. *Jama*, *294*(17), 2221-2224.

Lang, T., Lepage, B., Schieber, A. C., Lamy, S., & Kelly-Irving, M. (2012). Social determinants of cardiovascular diseases. *Public Health Reviews*, *33*(2), 601-622.

Langer, E. J. (1975). The illusion of control. *Journal of personality and social psychology*, 32(2), 311.

Libet, B. (1993). Unconscious cerebral initiative and the role of conscious will in voluntary action. In *Neurophysiology of consciousness* (pp. 269-306). Birkhäuser Boston.

Linden, D. E. J. (2006). How psychotherapy changes the brain–the contribution of functional neuroimaging. *Molecular psychiatry*, *11*(6), 528-538.

Lisanby, S. H. (2007). Electroconvulsive therapy for depression. *New England Journal of Medicine*, *357*(19), 1939-1945.

Mauss, I. B., Tamir, M., Anderson, C. L., & Savino, N. S. (2011). Can seeking happiness make people unhappy? Paradoxical effects of valuing happiness. *Emotion*, *11*(4), 807.

Mayberg, H. S., Lozano, A. M., Voon, V., McNeely, H. E., Seminowicz, D., Hamani, C., Schwalb, J. M., & Kennedy, S. H. (2005). Deep brain stimulation for treatment-resistant depression. *Neuron*, *45*(5), 651-660.

Mele, A. (2013). Unconscious decisions and free will. *Philosophical Psychology*, *26*(6), 777-789.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual review of neuroscience*, *24*(1), 167-202.

Montague, R. (2006). *Why choose this book?: how we make decisions*. New York: Dutton.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695-699.

Nolen-Hoeksema, S. (2000). The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. *Journal of abnormal psychology*, *109*(3), 504.

Onder, G., Penninx, B. W., Cesari, M., Bandinelli, S., Lauretani, F., Bartali, Gori, A. M., Pahor, M., & Ferrucci, L. (2005). Anemia is associated with depression in older adults: results from the InCHIANTI study. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *60*(9), 1168-1172.

Orth, U., & Robins, R. W. (2013). Understanding the Link Between Low Self-Esteem and Depression. *Current Directions in Psychological Science*, *22*(6), 455-460.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. Oxford University Press.

Roskies, A. (2006). Neuroscientific challenges to free will and responsibility. *Trends in Cognitive Sciences*, *10*(9), 419-423.

Siegle, G. J., & Thayer, J. F. (2004). Physiological aspects of depressive rumination. *Depressive rumination nature theory and treatment*, 79-104.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception-London*, *28*(9), 1059-1074.

Suhler, C. L., & Churchland, P. S. (2009). Control: conscious and otherwise. *Trends in cognitive sciences*, *13*(8), 341-347.

Strawson, P. (1964). Freedom and resentment. *Perspectives on Moral Responsibility*, 67-100.

Strunk, D. R., Lopez, H., & DeRubeis, R. J. (2006). Depressive symptoms are associated with unrealistic negative predictions of future life events. *Behaviour research and therapy*, *44*(6), 861-882.

Thagard, P., & Aubie, B. (2008). Emotional consciousness: A neural model of how cognitive appraisal and somatic perception interact to produce qualitative experience. *Consciousness and Cognition*, *17*(3), 811-834.

Thagard, P. (2012). The self as a system of multilevel interacting mechanisms. *Philosophical Psychology*, (ahead-of-print), 1-19.

Tiemeier, H., van Tuijl, H. R., Hofman, A., Meijer, J., Kiliaan, A. J., & Breteler, M. M. (2002). Vitamin B12, folate, and homocysteine in depression: the Rotterdam Study. *American Journal of Psychiatry*, *159*(12), 2099-2101.

Wang, L., LaBar, K. S., Smoski, M., Rosenthal, M. Z., Dolcos, F., Lynch, T. R., Krishnan R. R., & McCarthy, G. (2008). Prefrontal mechanisms for executive control over emotional distraction are altered in major depression. *Psychiatry Research: Neuroimaging*, *163*(2), 143-155.

Watkins, E. (2011). Dysregulation in level of goal and action identification across psychological disorders. *Clinical psychology review*, *31*(2), 260-278.

Wegner, D. M. (2005). Who is the Controller of Controlled Processes?. *The new unconscious*, 19.

Wegner, D. M., Sparrow, B., & Winerman, L. (2004). Vicarious agency: experiencing control over the movements of others. *Journal of personality and social psychology*, *86*(6), 838.

Wegner, D. M., & Wheatley, T. (1999). Apparent mental causation: Sources of the experience of will. *American Psychologist*, *54*(7), 480.

Wells, K. (2014, Mar. 9). Psychiatric community care: Belgian town sets gold standard. Canadian Broadcasting Corporation. Retrieved from http://www.cbc.ca

Wolf, S. (1987). Sanity and the Metaphysics of Responsibility. *1987*, 46-62.

World Health Organization (WHO). (2012). Depression (WHO Fact sheet N°369). Retrieved from World Health Organization website: http://www.who.int/mediacentre/factsheets/fs369/en/

Zeigarnik, B. (1927) Das behalten erledigter und underledigter handlungen. *Psychologische Forschung*. 9, 1-85.

Zimmerman, M. (2014) *Interview Guide for Evaluating DSM-5 Psychiatric Disorders and the Mental Status Examination*. East Greenwich, RI: Psych Products Press.