University of Cincinnati

Date: 9/15/2017

I, Valentina Petrolini M.A., hereby submit this original work as part of the requirements for the degree of Doctor of Philosophy in Philosophy.

It is entitled:

From Normality to Pathology: In Defense of Continuity

Student's name: **Valentina Petrolini M.A.**

This work and its defense approved by:

Committee chair: Heidi Maibom, Ph.D.

Committee member: Johannes Brandl, AoS

Committee member: Peter Langland-Hassan, Ph.D.

Committee member: Thomas Polger, Ph.D.

27998

UNIVERSITY OF CINCINNATI

McMicken College of Arts & Sciences

Department of Philosophy

From Normality to Pathology: In Defense of Continuity

by Valentina Petrolini, M.A.

A dissertation submitted to the Graduate School of the University of Cincinnati
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Committee Chair:
Prof. Heidi Maibom

Committee Members:
Prof. Peter Langland-Hassan,
Prof. Thomas Polger
Prof. Johannes Brandl

University of Cincinnati, September 2017

Dissertation Abstract

In this project I elaborate and defend a dimensional model of the relationship between mental normality and pathology, which I dub Continuity Thesis (CT). Throughout the dissertation I set out to accomplish two goals. First, I make space for a dimensional approach in philosophy of psychiatry by presenting a model that is compelling and worth taking seriously as a viable alternative to the more popular categorical counterpart. Second, I flesh out a more precise and non-arbitrary notion of what it means to be vulnerable to a mental disorder. This allows me to counter categorical approaches by showing that the cut-off point between normality and pathology is extremely hard to pin down.

The project is divided into three large sections. In the first section ("Background"), I lay the groundwork for a dimensional model of mental disorders. I start by exploring an important historical precedent of CT, namely the psychodynamic account developed by Freud. In the second section ("Strong Continuity"), I start building my dimensional model by proposing to see mental disorders as disruptions of four dimensions of functioning (i.e. salience, confidence, familiarity, and agency). Each of these dimensions represents a different way in which the relationship between individual and environment may be modulated. Mental disorders are thus seen as disruptions of these self-world relations, or as ways in which one's experience of the world can be altered. In the third section ("Meaningful Difference"), I turn more explicitly to the notion of vulnerability and I focus on intermediate cases to uncover their crucial role in the transition from normality to pathology. I discuss a number of case studies where people are imbalanced on one of the dimensions but still fail to qualify as disordered, and I explain what distinguishes them from their pathological counterpart. By introducing the notions of risk and protective factors I also outline a model of how the transition between vulnerable states and full-blown pathology may occur.

Ora questa storia non è inventata, e la realtà è sempre più complessa dell'invenzione: meno pettinata, più ruvida, meno rotonda. È raro che giaccia in un piano.

Now this story is not fictional, and reality is always more complex than fiction: less groomed, rougher, less rounded out. It rarely lies on a plane.

Primo Levi, *Il Sistema Periodico*

To N.

Table of Contents

List of Figures and Tables

Valentina Petrolini

Introduction

Questions and goals

What are the boundaries between mental health and pathology? Which theoretical approach is better suited to explain the relationship between the two?

The relation between normality and pathology has intrigued philosophers for centuries. In the *Meditations* (1641), Descartes famously compares himself to a madman because his dreams closely resemble the experiences reported by delusional subjects. Much later, Wittgenstein characterizes madness as a special form of skepticism that calls into question things that others take for granted – e.g. "Do I live here?"; "Is this a tree?" (1953, §70 & §467). Freud can then be seen as the first to develop a sophisticated account of the interplay between normal and pathological functioning. He suggests that we see mental life as continuous and that we regard psychiatric symptoms as extreme variations of normal psychological phenomena (1938; 1933).

Besides being philosophically interesting, the issue of determining what counts as pathological is obviously critical for psychiatric research and practice. Traditionally, two competing approaches have been put forward to characterize the complex relationship between normality and pathology. On the one hand, *categorical* accounts are committed to the idea that the gap between health and pathology should be conceived of as a difference in kind. In this sense, normal and pathological states would be seen as substances having different chemical compositions or atomic numbers. On the other hand, *dimensional* views maintain that such a gap should be seen as a difference in degree, similar to a spectrum of colors fading into one another. Murphy aptly summarizes the difference between the two approaches as follows:

"A categorical approach treats disorders as discrete phenomena, qualitatively different from normal states in virtue of pathological causal histories. A dimensional system of classification represents disorders as falling between points on an axis, or as a location in multidimensional space, and not as discontinuous categories" (Murphy 2006, p. 345).

In the past thirty years, dimensional models have been getting traction in psychiatry and clinical psychology for a number of reasons. First, studies on non-clinical populations show that psychiatric symptoms are much more widespread than we previously assumed, even in people who never received a formal diagnosis (see van Os *et al.* 2008; Johns & van Os 2001; Rachman & de Silva 1978). These studies also stress the fact that practitioners are often unable to reliably differentiate between clinical and non-clinical subjects exhibiting similar symptoms, thereby supporting the notion of a continuum between normality and pathology. Second, there has been growing interest in the so-called At-Risk Mental States (ARMS) and in their development. Indeed, sub-threshold conditions are becoming increasingly important for prognostic and prevention purposes thanks to cognitive and neurocognitive data suggesting that high-risk subjects may constitute intermediate forms (see Fusar-Poli *et al.* 2013). Finally, diagnostic manuals – e.g. the DSM-5 – have recently adopted dimensional classifications for some disorders (i.e. autism, personality disorders), although their overall framework still follows a categorical model. Other large-scale research initiatives – such as the *National Institute of Mental Health's Research Domain Criteria* (RDoC) – have embraced dimensional models more decisively and across the board.

Despite the growing interest in dimensional approaches in psychiatry, philosophers still overwhelmingly privilege categorical approaches to mental disorders. Interestingly, this preference is shared by researchers working in different paradigms – e.g. analytic philosophy and

phenomenology – although the motivations behind this choice are different. Philosophers working in the analytic tradition are worried that embracing a dimensional approach would not allow clinicians to draw principled distinctions between people that are in need of treatment and people who are not. In other words, seeing normal and pathological states as continuous would risk making the boundary between them too arbitrary or vague (see Murphy 2006). A similar way to express this view is to say that mental disorders should be seen as discrete categories with sharp boundaries separating them from the normal – i.e. as *natural kinds* of some sort (see Samuels 2009). Phenomenologists similarly endorse categorical approaches because of their tendency toward typification. Drawing mostly on Husserl's work, philosophers of psychiatry working in this tradition emphasize the fact that we experience things as belonging to a certain type, or kind. For this reason, most of the available phenomenological accounts of mental illness – such as the one recently proposed by Parnas & Gallagher (2015) – advocate an *ideal type* or *prototype* approach, where exemplars of various disorders stand as the best representative of a certain class (see Fernandez forthcoming, pp. 6-7).

In this project I elaborate and defend a dimensional model of the relationship between normality and pathology, which I dub *Continuity Thesis* (CT). Throughout the dissertation I set out to accomplish two goals. First, I make space for a *dimensional approach* in philosophy of psychiatry by presenting a model that is compelling and worth taking seriously as a viable alternative to the more popular categorical counterpart. Second, I flesh out a more precise and non-arbitrary notion of what it means to be *vulnerable* to a mental disorder. This allows me to counter categorical approaches by showing that the cut-off point between normality and pathology is extremely hard – when not impossible – to pin down.

I use two interrelated strategies to accomplish these goals.

- *Show that intermediate states are frequent.* I introduce a number of intermediate cases between normality and pathology and I show that they are more pervasive than one would expect. This is what I call the argument for quantity (see Chapter Four, §2). By showing that there are multiple gradations between normal and pathological states differing from one another in quantitative terms, I make it more difficult to argue that the difference between the two is categorical.
- *Show that intermediate states are explanatorily relevant.* I show that intermediate cases are not only frequent but also explanatorily relevant, because they shed light on the transition from normality to pathology. By looking at these cases and discussing the role played by risk and protective factors we can better explain the various ways in which a person may cross the threshold to pathology. Broadly speaking, a vulnerability state becomes pathological due to an increase of risk factors or a weakening of protective ones.

This two-pronged argumentative strategy has been repeatedly employed to defend continuity claims throughout the history of philosophy and science. For example, the idea that intermediate forms are too frequent and relevant to be ignored is at the core of Darwin's argument for the continuity of species (1859). More recently, similar strategies supported continuity claims with respect to psychological kinds (Buckner 2016), gestures and language (Bar-On 2013), and delusional beliefs (Bortolotti 2010).

Structure

The project is divided into three sections.

In the first section ("Background"), I lay the groundwork for a dimensional model of mental disorders. I start by exploring an important historical precedent of CT, namely the psychodynamic account developed by Freud.

In Chapter One ("Freud as an Early Defender of the Continuity Thesis"), I show that the core ideas underlying CT were already present in a sophisticated form in Freud's work, as it emerges from his theoretical essays (1910; 1933; 1938) and from the discussion of clinical cases (1895a; 1909; 1911b). In particular, Freud defends two theses that allow us to see him as an early defender of CT. On the one hand, he is committed to the idea that healthy and disordered subjects exhibit deep similarities in terms of mental functioning. I dub this thesis "Strong Continuity". On the other hand, Freud defends the possibility of describing with sufficient precision and generality what goes awry in pathological cases. I dub this thesis "Meaningful Difference". In this chapter I also explore three aspects of Freud's proposal that will work as *desiderata* to develop my own framework:

1) The balance between "*Strong Continuity*" and "*Meaningful Difference*", around which the remainder of the dissertation is organized (see Sections Two and Three). The idea is that a convincing defense of CT should achieve both goals. First, it has to propose a model of mental disorders where the distinction between normal and pathological boils down to a difference of degree. Second, the model has to explain what makes disordered and non-disordered cases importantly different without trivializing the distinction between health and pathology.
2) The etiological intuition about the existence of two determinants of psychopathology: *personal-level* factors (i.e. dispositional) and *environmental-level* factors (i.e. accidental). I develop this intuition in detail in Chapters Four and Five.
3) The endorsement of an *economic* approach, namely the idea that quantitative factors play a crucial role in the transition between normality and pathology. I explore this idea more closely when I introduce the notions of risk and protective factors (see Chapter Four).

In the second section ("Strong Continuity"), I start building my dimensional model by proposing to see mental disorders as disruptions of four dimensions of functioning (i.e. salience, confidence, familiarity, and agency). Each of these dimensions represents a different way in which the relationship between individual and environment may be modulated. Mental disorders are thus seen as disruptions of these self-world relations, or as ways in which one's experience of the world can be altered.

Methodologically speaking, I use of a number of case studies to make my classification more convincing. At this preliminary stage of model construction, some of these cases are taken directly from patients' reports (Reina 2009; Longden 2013), while others are excerpts of autobiographical accounts of mental illness (Saks 2007; Styron 1991; Plath 1963). These examples contribute to offer a more concrete grasp of the patients' phenomenological experience and also make the overall view more intelligible.

This section includes two chapters. In Chapter Two ("What Makes Mental Disorders Continuous with Normal Functioning?") I characterize mental disorders as disruptions of four dimensions – i.e. salience, confidence, familiarity, and agency. Each section starts with a brief characterization of a mental disorder category (e.g. disorders of salience), followed by a detailed discussion of case studies that illustrate the disorder from a phenomenological viewpoint. This discussion makes clear that all these disorders can be seen as opposite disruptions on the same dimensions, with the two extremes of the spectrum representing instances of deficit (hypo) and overload (hyper).

In Chapter Three ("The Nature of the Dimensions and their Relation to Affect"), I take a deeper look at the dimensions introduced in Chapter Two and I characterize them as ways in which we affectively modulate our relationship with the environment. Specifically, I draw on existing dimensional models of affect (Russell 1980; Russell & Barrett 1999; Thayer 1996) to

show that the dimensions aggregate and give rise to *moods*. In this chapter I also discuss the complex relationship between moods, *emotions* and *feelings* and I explain why the dimensions are connected to the former as opposed to the latter. Specifically, I argue that – unlike feelings and emotions – moods have an important unconscious component. Finally, I apply the model to psychiatry and I suggest a few ways in which moods can be seen as pathological.

In the third section ("Meaningful Difference"), I turn more explicitly to the notion of vulnerability and I focus on intermediate cases to uncover their crucial role in the transition from normality to pathology. I discuss a number of case studies where people are imbalanced on one of the dimensions but still fail to qualify as disordered, and I explain what distinguishes them from their pathological counterpart. By introducing the notions of risk and protective factors I also outline a model of how the transition between vulnerable states and full-blown pathology may occur.

This section also includes two chapters. In Chapter Four ("What Makes Mental Disorders Different from Normal Functioning?"), I explore a number of intermediate cases lying on various points of the dimensions introduced in Chapter Two. These cases represent situations of personal-level vulnerability: I argue that there are two ways of being vulnerable to mental disorders as an individual. On the one hand, some intermediate cases may be seen as attenuated versions of their pathological counterpart. Here the distinction between health and pathology hinges on the increase of risk factors such as duration, frequency, urgency, etc. For example, the difference between someone who is overconfident and someone who suffers from grandiosity delusion hinges on how pervasive, frequent, broad and long-lasting this person's convictions are. On the other hand, in some cases instances of vulnerability are counterbalanced by the presence of one or more protective factors (e.g. control, humor, physical and mental strength, etc). In these situations the distinction between health and pathology should be described through the

successful or unsuccessful action exercised by such protective forces. For example, someone suffering from obsessions may have significant resources at her disposal such as a sufficient amount of energy to keep them at bay.

In Chapter Five ("Vulnerable Populations"), I explore the relation between environment and psychopathology and I flesh out what it means to be vulnerable to mental disorders as a population. I start by discussing some recent empirical research that investigates various ways in which the environment can be pathogenic. Then I explore the reasons behind the high incidence of psychopathology in some populations and I suggest that specific forms of disadvantage can be connected to specific disturbances. In the remainder of the chapter I focus on two situations of disadvantage (i.e. migrants and women) to assess whether and how social adversities may contribute to the development of specific disorders (i.e. schizophrenia and depression). The research discussed in this chapter indicates that the transition to pathology is importantly influenced by factors that are *external* to the patients, such as social pressures related to gender or discrimination due to minority status (Berg *et al.* 2014; Gutierrez-Lobos *et al.* 2000). Taken together, Chapters Four and Five thus offer a more complete picture of what it means to be vulnerable to a mental disorder.

In the synthetic table below I summarize the project's structure and the core research questions addressed in each chapter.

Section I – Background	
Chapter One - Where does CT come from? - What are its precedents historically speaking?	I argue that the core tenets of CT have been articulated in a sophisticated form by Freud, and that his model can be taken as an exemplar for the development of a refined version of CT.
Section II – Strong Continuity	
Chapter Two - How can we make a dimensional model of mental disorders more plausible? - What makes mental disorders continuous with normal functioning?	I propose to conceive of mental disorders as disruptions of four dimensions of functioning, corresponding to different ways of modulating one's relationship with the environment: *salience*, *confidence*, *familiarity* and *agency*. I identify four kinds of disorders that encompass many distinguishing features of psychiatric conditions. I show that each of them can be seen as a disruption of the relevant dimension in two opposite ways: deficit (hypo) or overload (hyper).
Chapter Three - What is the nature of the dimensions? - How can we better characterize them and their disruption?	I characterize the four dimensions as affective in nature, drawing on existing models of affect. I suggest that their aggregation gives rise to moods as opposed to feelings or emotions.
Section III – Meaningful Difference	
Chapter Four - What makes mental disorders different from normal functioning? - What does it mean to be vulnerable or at-risk as an individual?	I argue that there are two ways of being at-risk or vulnerable as an individual. In some cases, the distinction between health and pathology hinges on the increase of *risk factors* (duration, frequency, urgency, etc). In other cases, vulnerability reflects the loss or weakening of *protective factors* (control, humor, physical and mental strength, etc).

Chapter Five - In which ways can the environment be pathogenic? - What does it mean to be vulnerable or at-risk as a population?	I explore the role played by environmental factors in the onset and development of psychopathology. I flesh out the notion of vulnerable population using migration & schizophrenia and gender & depression as case studies.

Section I:

Background

Valentina Petrolini

Freud as an Early Defender of the Continuity Thesis

"That no sharp line can be drawn between 'neurotic' and 'normal' people – whether children or adults – that our conception of 'disease' is a purely practical one and a question of summation, that predisposition and the eventualities of life must combine before the threshold of this summation is overstepped, and that consequently a number of individuals are constantly passing from the class of healthy people into that of neurotic patients, while a far smaller number also make the journey in the opposite direction – all of these are things which have been said so often and have met with so much agreement that I am certainly not alone in maintaining their truth"

(Freud 1909, p. 2122).

Abstract

In this chapter I set out to show that the core ideas underlying the Continuity Thesis (CT henceforth) were already present in a sophisticated form in Freud's work, as it emerges from his theoretical essays (see 1895a; 1910; 1933; 1938) and from the discussion of clinical cases (see 1895a; 1909; 1911b). In particular, Freud defends two interrelated theses that allow us to see him as an early defender of continuity. On the one hand, he is committed to the idea that healthy and mentally disordered subjects exhibit deep similarities in terms of mental functioning: the same structures or forces (e.g. id, ego and super-ego) as well as the same mechanisms (e.g. repression) are equally at work in both populations. On the other hand, he defends the possibility of describing with sufficient precision and generality what goes awry in pathological cases. Repression is a case in point: in our mental life, we all "turn something away, and keep it at distance, from the conscious" (1915b, p. 147) to the point that "the maintenance of certain internal resistances is a *sine qua non* of normality" (1938, p. 33). Freud makes clear that repression is beneficial in most cases: indeed, it allows us to avoid distressing thoughts and strengthens the ego in his mediating function between id and external world. However, this common and usually innocuous process becomes problematic for some individuals due to personal traits as well as environmental factors, making it more likely for mental disorders to emerge and flourish. In these

situations, the repressed content usually comes back in a disguised or distorted form and expresses itself *qua* symptom (e.g. delusional explanation or hysterical paralysis).

The chapter is divided into three sections. In **§1** I appeal to Freud's work to show that his account of psychopathology can be seen as an early version of CT; in **§2** I focus on three theoretical aspects of Freud's proposal and I argue that they should be taken as *desiderata* for a refined account of mental disorders; in **§3** I illustrate how the approach works in practice by analyzing one of the clinical cases discussed by Freud (1895a – Elizabeth von R).

§1. An Early Defense of Continuity

In this section I draw on several primary sources – both theoretical essays and clinical cases – to argue that Freud's account of mental disorders can be seen as an early defense of CT. More specifically, I focus on two theses defended by Freud:

a) *Strong continuity*. The same psychological mechanisms, forces and structures are at work both in healthy and mentally disordered subjects. In the sub-section **§1.1**. I introduce Freud's geography of the mind and I show how this grounds the parallel between mental health and pathology. I also briefly discuss the case of dreaming, which nicely illustrates Freud's commitment to a).

b) *Meaningful difference*. Despite a), it is possible to describe with sufficient precision and generality what goes awry in pathological cases. In the sub-section **§1.2**. I discuss a series of essays where Freud directly compares clinical and non-clinical populations in order to detect the differences between the two groups. I also briefly discuss the case of repression, which nicely illustrates Freud's commitment to b).

§1.1. Strong Continuity: Evidence for a)

The evidence for a) is disseminated throughout Freud's writings at various stages of his career, from the early *Studies on Hysteria* (1895a) to the late *Outline of Psychoanalysis* (1938). This strong commitment to continuity arises primarily from the observation that the same psychical mechanisms, forces and structures witnessed in patients can be detected - in a more implicit and attenuated form - in people that are normally regarded as "healthy". As Freud puts it: "Pathology, with its magnification and exaggeration, can make us aware of normal phenomena which we would otherwise have missed" (1933, p. 80). These similarities in mental functioning run deep and uncover a psychical apparatus that can be divided "geographically" into three mental provinces: id, ego and super-ego. To put it roughly, Freud describes the *id* as the core of our being, as that part of our mind that lacks direct communication with the outside world and focuses on a realm of internal perception (e.g. bodily changes, feelings of pleasure and unpleasure). This mental province is also the seat of the instincts originating from our somatic organization and represents the demands of everything that is inherited and present at birth (see 1938, p. 84). The main purpose of the id is thus to give mental expression to the instinctual needs coming from the body. Notably, Freud describes it as a structure that lacks unified will or coherent organization: neither the laws of logic (e.g. law of non-contradiction) nor the notion of time have any significant influence on the id. In fact, contrasting impulses may coexist there side by side without neutralizing each other, and they can be "virtually immortal and preserved for whole decades as though they had only recently occurred" (1933, p. 99). The *ego* develops out of the id and represents the mental agency responsible for our contact with the external world with all its resources, dangers and demands. It is characterized as the seat of thought because it mediates between the instincts coming from the id and the actions affecting the external world. Freud describes this system as interpolating thought between desires and action, as synthesizing and

unifying the pressures coming from different sources into a more or less coherent whole (Ibid., pp. 102-103). The ego's main concern lies therefore in self-preservation and protection against external and internal dangers: for this reason, it is also the seat of anxiety (reaction to danger), repression (flight from an internal threat) and disavowal (flight from an external threat). Finally, Freud characterizes the *super-ego* as a portion of the ego that develops throughout childhood and in particular via identification with the parents: more specifically, it works as a "special agency where parental influence is prolonged" (1938, p. 15). In this sense, the super-ego represents the most recent and superficial part of our mental apparatus, because – as opposed to the id – it is something acquired through nurture and not present from the very beginning of life (see 1933, p. 84). In terms of goals, the super-ego is thus the mental agency that observes, judges and criticizes the ego, thereby representing our moral conscience as well as our self-ideals or expectations.

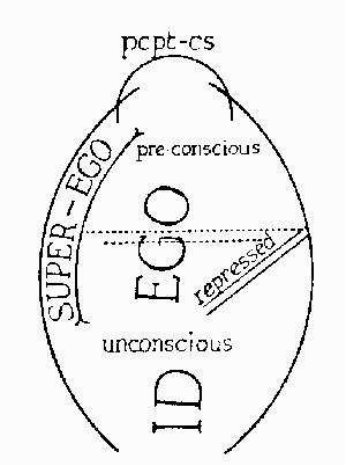

Fig. 1: The Anatomy of the Mental Personality (1933)

This apparent simplification of the psychical realm should not obscure a crucial fact uncovered by Freud. The dynamics among these mental structures do not reflect a situation of harmony, but rather a state of permanent tension and conflict. Indeed, some forces – such as the ones represented by the id - are upward-driving and struggling to emerge, whereas others – such

as the ones at work in the super-ego - are downward-driving and attempting to repress. In this battlefield, the ego plays the delicate role of a mediator: on the one hand, it takes over the instinctual demands of the id and attempts to lead them to satisfaction without neglecting reality; on the other hand, it strives towards the ideals placed on it by the super-ego while trying to avoid punishment. Making use of a series of metaphors, Freud first compares the ego to a constitutional monarch, "without whose sanction no law can be passed and who hesitates long before imposing a veto on any measure put forward by the Parliament" (1923, p. 81). However, elsewhere Freud describes the ego as a slave who "has to serve three harsh masters, and has to do the best to reconcile the claims and demands of all three [i.e. external world, super-ego, and id]" (1933, p. 103). This process of mediation and defense inevitably turns out to be inadequate, as the ego is "fighting on two fronts": on the one side it protects itself against the dangers of the external world, while on the other it puts up with the demands coming from the internal world (1938, p. 87). For our purposes, it is important to stress that this psychical apparatus is *shared* by people that are regarded as healthy as well as by people affected by mental disorders. Freud states the point explicitly: "It is not scientifically feasible to draw a line of demarcation between what is psychically normal and abnormal; so that the distinction, in spite of all its practical importance, possesses only a conventional value" (Ibid., p. 81).

Two more points in support of a) are repeatedly brought up by Freud in his writings. First, the idea that psychical continuity grounds the comparison between normal and abnormal cognition, which plays an important role in philosophy of mind, psychology and psychiatry. Second, this continuity between mental health and pathology is ultimately based on quantitative factors such as the degree of mental energy accumulated or discharged, or the force of the instinctual demands repressed by the ego.

With respect to the first point, Freud claims that the commitment to psychical continuity serves as the main justification for inferring normal functioning from pathological cases. This idea is shared by many contemporary philosophers of psychiatry, who believe that focusing on pathology represents a good way to investigate the mind's normal functioning (see for example Graham 2013; Murphy 2006; Bentall 2004). Yet, the idea that we can learn something about the mind by studying how the mind breaks makes more sense under the assumption that normal and abnormal cognition are importantly continuous. Indeed, without such an assumption nothing prevents us from thinking that pathological and healthy functioning would follow completely different rules and operate under different mechanisms. If this were the case, we would probably be able to discover something interesting about pathology without being able to infer any conclusion or gain any insight about how the mind normally works. Freud makes the point explicit through an analogy: if we throw a crystal to the ground, we notice that it does not break "haphazard" but rather follows specific lines and limits that were implicit in its structure. Similarly, mentally disordered patients show mental "breaches" or "clefts" that can be interpreted as missing links, and make observable "through magnification and exaggeration" some patterns or phenomena that would be otherwise inaccessible (1933, p. 80). Freud fleshes out this claim in several other places, from the *Project for a Scientific Psychology* (posthumously published in 1950 but written during the early years of psychoanalysis) to the late *Outline of Psychoanalysis* (1938). In the *Project*, he directly compares the compulsions experienced by hysteric patients with the "excessively intense ideas" that we all encounter in our everyday lives (1950, p. 405). These two sets of ideas are remarkably similar: for example, they are both hard to dismiss and they often arouse distressing affect in the people who experience them. However, while we tend to regard excessively intense ideas as "the product of powerful and reasonable motives", compulsions "strike us by their oddity" and seem to have special importance only for the person who is

affected by them (Ibid.). This passage shows that the comparison between healthy and mentally disordered populations is fruitful because it also helps to uncover important aspects of normal functioning. Freud defends the point even more strongly in the *Outline*, where he expects the study of neurosis to provide us with valuable contributions to our knowledge of normality. As he puts it: "It may be that we shall thus discover the 'weak points' in a normal organization" (1938, p. 64). In Chapters Fours and Five I elaborate further on the idea that psychopathology provides a deeper understanding of healthy mental functioning by highlighting different kinds of "weak spots" (see the notions of vulnerable individual and vulnerable population).

With respect to the second point, Freud insists that the only meaningful distinction between mental health and pathology should be based on *quantitative factors,* such as the quota of affective tension that a subject can tolerate. This idea is – again – grounded in an underlying continuity between normal and abnormal cognition and arises from the observation that the patients' behavior does not radically differ from the one exhibited by healthy people. Already in the *Studies on Hysteria* (1895a), Freud realizes that we all bear in our consciousness a great number of ideas that have been not "affectively dealt with." Thus, even people who fall ill are able to tolerate this accumulation until "the amount is increased by summation to a point beyond the subject's tolerance" (p. 174). Notably, this breaking-point is highly individualized and depends on a host of factors that are internal as well as external in nature (e.g. personality, environmental influences, traumatic events). As Freud puts it in a later essay: "Each individual has in all probability a limit beyond which his mental apparatus fails in its function of mastering the quantities of excitation which require to be disposed of" (1926, p. 128). This observation allows us to significantly refine the continuity claim and to show the internal complexity of Freud's account. On the one hand, there is no qualitative distinction between healthy and pathological mental functioning because the same pathogenic determinants are present in everyone. Freud

makes the point explicit in the *Introductory Lectures*: "'Being ill' is a practical concept […] But if you take up a theoretical point of view and disregard this matter of quantity, you may quite well say that we are all ill - that is, neurotic - since the preconditions for the formation of symptoms can also be observed in normal people" (1916-17 – XXIII, p. 358). On the other hand, we still need to explain why some people are more liable to developing a mental disorder with respect to others who share a similar psychological makeup and go through similar life-experiences. Freud regards this question as a crucial one: "Why are some people *not* falling ill?" (1924b, p. 254); "What we need and cannot lay our finger on is some factor which will explain why some people are able to subject the affect of anxiety, in spite of its unique quality, to the ordinary workings of the mind, or why others are doomed to break down over this task" (1926, p. 130).

In order to address this issue Freud introduces a different approach to the study of the mind: crucially, he realizes that it is not sufficient to give a *topographical* account of the mental provinces or to investigate the way in which the psychical forces come *dynamically* into conflict. What is missing is an *economical* account, one able to deal with mental events in terms of the intensity of forces running through them and thus having the resources to explain mental phenomena quantitatively. Once adopted this approach, we can regard two subjects as different in terms of the "relative strength of their [mental] forces." For example, in one person the repressed forces may be too strong and hard to keep at bay, while in another the repressing forces may be too weak to resist effectively against the attack (see Freud 1924b & 1938). In these cases we may even witness a similar symptomatic manifestation (e.g. hysterical paralysis) arising for different economical motives: in the first case, the strength of the psychological forces would be primarily responsible; in the second case, the degree of control exercised by the ego would play a more important role. Later in the chapter (see **§2.1**) I take a deeper look at the crucial role played by the economic approach in Freud's account. For the moment, it is sufficient to stress that introducing

the notion of quantitative factors allows him to hold onto continuity – thesis a) – without giving up on the idea that meaningful differences can be detected between mental health and pathology – thesis b). Freud makes explicit his commitment to both theses in passages like the following: "Every normal person is only approximately normal: his ego resembles that of the psychotic in one point or another, in a greater or lesser degree, and by its distance from one end of the scale and its proximity to the other we may provisionally estimate the extent of what we have so indefinitely called the 'modification of the ego'" (1937, p. 389).

Finally, Freud's treatment of *dreaming* can be used as an example to illustrate his commitment to a). Indeed, dreams are regarded as theoretically important primarily because they represent an authentic pathological trait experienced by everyone. In the *Five Lectures on Psychoanalysis* (1910), Freud describes dreams as valuable because they exhibit "the greatest outer similarity to the creations of the insane" while at the same time being "compatible with full health during waking life" (III, p. 13). In this sense, dreams are living proof in support of continuity because they show that "repression and surrogate creations are present even under conditions of health" (Ibid., p. 16). Indeed, it is very commonplace for dreams to express contents having the nature of wish-fulfilments - e.g. a dead loved one represented as alive - or to depict repressed contents in a disguised form - e.g. a person towards which we are indifferent receives our affection and thereby replaces the one that we really desire. Elsewhere, Freud talks about the psychical activity during sleep as the most favorable object of study for psychoanalysis because dreams markedly differ from the mental events experienced during waking life (see 1938, p. 38). Here the characterization of dreams as episodes of mental illness becomes more explicit: "A dream, then, is a psychosis, with all the absurdities, delusions and illusions of a psychosis. A psychosis of short duration, no doubt, harmless, even entrusted with a useful function, introduced with the subject's consent and terminated by an act of his will. Nonetheless it is a psychosis, and

we learn from it that even so deep-going an alteration of mental life as this can be undone and give place to normal function" (Ibid., p. 49).

This passage is crucial for two reasons: first, it allows us to see that – according to Freud – we all experience mental disorder to a certain extent, although we do it in ways that appear commonplace or innocent. In particular, we are neurotic because we experience various forms of repression and self-deception (see 1916-17, p. 358); but we are also psychotic because in dreaming we experience deep-going alterations of our mental life that strongly imply the creation of a new reality (see also 1924a & 1924b). Second, the passage suggests that recovery should be possible even for people who experience deep alterations of normal functioning. Freud's reasoning goes as follows: healthy people often experience significant alterations of normal functioning (i.e. dreams) but are able to resume it quite seamlessly and effortlessly when they wake up. By understanding how this transition happens, we should also be able to recreate this spontaneous process through therapeutic effort, and therefore restore the connection between ego and external reality that is lost in psychosis. Indeed, psychotic disorders arise from the weakening of the ego that becomes more or less unable to play its mediating role between internal and external demands.

§1.2. Meaningful Difference: Evidence for b)

As I mention above, Freud's account combines a strong commitment to mental continuity with the attempt to describe as precisely as possible what goes awry in pathological cases. In other words, healthy and mentally disordered subjects exhibit important similarities – i.e. thesis a) – as well as meaningful differences – i.e. thesis b). The evidence for b) comes mostly from a series of essays where Freud directly compares clinical and non-clinical populations, discussing resemblances and differences between the two groups.

One of the earliest sources in this sense is the article “Obsessive Actions and Religious Practices” (1907), focusing on the resemblance between the obsessive actions of neurotics and the religious observances of believers. Freud immediately warns us against regarding the similarity as a superficial one: indeed, the beliefs and actions of neurotics appear deeply continuous with the ones exhibited by religious people. For example, both groups tend to make small adjustments to everyday actions (e.g. eating habits and prohibitions) and they both subject themselves to a range of restrictions and arrangements that often take the form of a *ceremonial*. Freud describes a ceremonial as a sum of conditions placed upon something forbidden so as to make it permissible (Ibid., p. 124). In this sense, celebrating a wedding according to certain rules before having sex and washing one’s hands a certain number of times before going to bed count as ceremonials. Moreover, there are distinguishing psychological traits that accompany the performance of a ceremonial: mostly a sense of conscientiousness and anxiety about “doing things right”, but also an important relationship between actions (or inactions) and guilt. Due to these striking similarities, Freud suggests to regard neurosis as a sort of “individual religiosity” and religion as a kind of “universal obsessional neurosis” (Ibid., p. 128). However, he also points out some important differences between obsessional neurotics and religious people: for instance, the former tend to regard their ceremonials as an eminently private and solitary enterprise, while the latter take great pains to display them as public and collective rituals. Again, neurotics exhibit a high degree of individual variability and idiosyncrasies in performing the rituals, whereas religious people tend to repeat an impersonal series of stereotyped or codified gestures. Notably, the key difference between the two groups seems to lie in the degree of intersubjective agreement upon certain beliefs and actions. Whereas the neurotic’s ceremonials appear “senseless” to others, religious rituals are embedded in a network of socio-cultural practices and symbolic meanings (Ibid., p. 120). Similarly, in the *Project* Freud distinguishes compulsions from excessively intense

ideas by appealing to different degrees of intersubjective recognition: “Excessively intense ideas also occur normally […] We are not surprised at them, if we know their genetic development (education, experiences) and their motives. We are in the habit of regarding these excessively intense ideas as the product of powerful and reasonable motives. In hysterics, on the contrary, excessively intense ideas strike us by their oddity. They are ideas which produce no effects in other people and whose importance we cannot appreciate. They appear to us as intruders and usurpers and accordingly as ridiculous” (1950, p. 405).

Another discussion along similar lines can be found in the article “Mourning and Melancholia” (1917), where Freud explores the emotion of normal grief in order to shed light on pathological melancholia (today known as depression). Once again, the idea is to offer a comparative account of grief and depression showing what the two conditions have in common as well as what makes them importantly different. On Freud’s view, both conditions arise from similar external influences and they are both reactions to a significant loss that can be interpreted literally - e.g. death of a loved one - or ideally - e.g. loss of love. At first glance, grief and depression appear remarkably similar: they are both characterized by “painful dejection”, “abrogation of interest in the outside world” and general inhibition to undertake any activity, even those that were once deemed pleasurable by the individual (Ibid., p. 244). However, depression exhibits a distinguishing feature that most cases of normal grief lack, namely an attitude of self-reproach and low self-esteem that Freud describes as “delusional belittling” (Ibid., p. 246). This important difference suggests that depression and grief may have a similar origin but a different outcome: in other words, they may be the very same process undergoing a normal development in the one case and a pathological one in the other. Freud characterized the common underlying mechanism as follows: the function of reality-testing carried out by the ego reveals the absence of the loved object, causing the sudden withdrawal of all the mental energy previously attached to it.

This free-floating energy then generates a more or less prolonged phase of struggle against reality, where various forms of wish-psychosis may ensue (e.g. dreams about the loved one, phenomena of hyper-familiarity, hallucinations). At this point, a bifurcation occurs: some individuals undergo a normal development and end up deferring to reality after some time; others experience a pathological development where the ego finally identifies with the lost love-object and therefore feels itself "at loss, wounded, hurt, neglected, out of favor, disappointed" (Ibid., p. 251). Depression – like psychosis – is thus characterized as a progressive weakening of the ego, whose energy becomes absorbed in dealing with the loss of a love-object and finally becomes "drained" and "utterly depleted" (p. 253).

In a later essay entitled "The Loss of Reality in Neurosis and Psychosis" (1924b), Freud even characterizes normal psychological functioning as a form of reality distortion and in particular as a compromise between neurosis and psychosis. In the neurotic case, a conflict between the ego and the id resolves in a flight from reality where the ego gains the upper hand and succeeds in keeping the id at bay. However, since the repression is only partially successful, the repressed content ends up coming back as a symptom: for example, in hysteria a body part unconsciously connected with the repressed content becomes paralyzed or painful. In the psychotic case, the conflict between the ego and the id resolves in favor of the id and the ego becomes partially suppressed. Thus, due to the ego's constitutive connection with the external world, psychosis inherently presents itself as a more or less severe loss of contact with reality. According to Freud, normal functioning could be seen as a mere compromise between these two extreme solutions. As he puts it: "A reaction which combines features of both these is the one we call normal or 'healthy'; it denies reality as little as neurosis but then, like a psychosis, is concerned with effecting a change in it" (Ibid., p. 279). Once again, Freud combines a strong idea

of psychical continuity with the effort to distinguishing non-problematic cases from others that are harmful for the subjects who experience them.

The idea that pathology could arise from normal developmental processes gone awry is nicely exemplified by the case of *repression*, described as a mechanism at work both in healthy and disordered subjects despite significant differences. Like Freud's treatment of dreaming illustrates his commitment to a), his view of repression works well in defense of b). In one of his early essays about the neuro-psychoses of defense – i.e. hysteria, obsessions and phobias – Freud characterizes repression as "an act or effort of will", or as the subject's attempt to forget "an experience, idea or feeling which aroused distressing affect" (1894, p. 2). It is clear from this description that Freud does not mean to regard the mechanism of repression as *per se* pathological. Indeed, intentional forgetting often succeeds and the subject manages to resolve the conflict between the ego and a distressing idea by simply eliminating the latter. Yet, at times the ego's attempt to forget an idea does not fully succeed in this task: in these cases, the distressing idea turns into a weaker one that expresses the same content but is deprived or "robbed" of its quota of affect (Ibid., p. 3). In a process similar to the one mentioned above for depression, this free-floating affect is then put to a different use and a variety of symptoms arise as a consequence. Hysterics tend to transform the affect into something somatic (e.g. paralysis), while patients suffering from obsessions and phobias connect it with other psychical contents (e.g. fear of contagion). On this view, the mechanism of repression encounters a pathological development *only* when some kind of conversion takes place. In other words, the repressed content "comes back" in a disguised form and replaces the distressing idea in the subject's consciousness (see also Freud 1926). A more synthetic account of repression is offered in the *Studies on Hysteria* (1895a): here the process is described as originating from the feeling of unpleasure (*Unlust*) caused by the conflict or incompatibility between ideas. For example, the desire of seducing one's mother comes

into conflict with social norms as well as with a direct censorship coming from the parents themselves (see Freud 1909 for a detailed discussion of such a case). This incompatibility then triggers the mechanism of repression proper (intentional forgetting) which acts by excluding the distressing idea from all associative modifications, that is from all the mental processes responsible for "thinking through" and correcting an idea by reference to other ones. By so doing, the idea's content is put out of sight while its whole quota of affect is retained and transformed into something else – e.g. hysterical symptoms, obsessions, phobias. This final phase is where the idea becomes pathogenic, or – as Freud puts it – where "the repressed idea takes its revenge" (1895b, p. 116).

Later in his career, Freud characterizes repression in a more refined form, as a mechanism that does not merely abrogate or annihilate ideas but rather prevents them from becoming conscious (see 1915a, p. 98). He also clarifies that the repressing process does not apply only to ideas but also to affects, movements or actions: the former are prevented from developing, whereas the latter are physically blocked. This dynamic of an affect that is first retained, then free-floating and finally converted into a symptom comes to be known as *anti-cathexis*, a process that withdraws mental energy from the conscious part of an idea while leaving the unconscious part untouched. In the clinical case known as "Little Hans" (1909), an unconscious sexual impulse towards the mother and a related hostile impulse towards the father demand access into consciousness but encounter repression. As a consequence, the affect attached to these ideas becomes free-floating and turns into anxiety, which then attaches itself to another idea connected by association with the repressed one (e.g. "Horses are big and threatening like my father"). The new idea replaces the old one in provoking anxiety and this translates into a state of heightened sensitivity that ends up inhibiting a number of actions (e.g. going out; crossing the street; riding on carriages). In an earlier essay (1895b), Freud discusses the case of a young woman that became

socially isolated because of an incontrollable fear of urinating in public. Once again, this obsession originates from another repressed – and correlated – idea: the patient developed this symptom after a night at the theater, where she saw an attractive man and "felt an erotic desire" in the form of a bodily sensation closely resembling urination. In this case the process of substitution is even more straightforward, as "the desire to urinate had replaced the erotic one" (Ibid., p. 3).

To sum up: Freud characterizes repression as a normal psychological process that becomes pathological only under particular circumstances. Normal repression simply consists in avoiding to think about something that causes us to feel unpleasure by directing our thoughts somewhere else. For example, I feel anxious about the talk I am going to give in a few days and I actively try not to think about it. Pathological repression takes place when we do not completely succeed in "forgetting" the distressing idea and we are reminded of it by "fresh perceptions" (1950, p. 409) – for example, I see a PowerPoint icon on my desktop and I suddenly think about the talk. However, this is not sufficient to transform the initial idea from merely distressing to pathogenic: some process of displacement has to occur, where we unconsciously associate the initial idea to a new one that ends up replacing it while attracting its whole quota of affect and anxiety. In our example, this would happen if I were to unconsciously develop a phobia or obsession towards something only loosely related to my talk – e.g. the blazer that I am supposed to wear that day or the train that is supposed to get me to the venue.

§2. A Refined Account of Mental Disorder: Psychoanalysis as an Exemplar

"Psychoanalysis has warned us that we must give up the unfruitful contrast between external and internal factors, between experience and constitution, and has taught us that we shall invariably find the cause of neurotic illness in a particular psychical situation which can be brought about in a variety of ways" (Freud 1912, p. 238).

In this section I discuss some theoretical aspects of Freud's account in more detail in order to lay down the building blocks of a refined model of mental disorders, one that I elaborate and defend later in the dissertation. More specifically, I focus on three ideas that prove helpful in the construction of such a model but do not require espousing psychoanalysis as a doctrine. This way, those who have reasons to be skeptical about Freud's approach would be able to focus selectively on the aspects of his theory that I take to be exemplar and generally applicable. At the same time, those who are more sympathetic toward psychoanalysis would be able to see the overall project as inspired by Freud's insights on psychopathology. Here I set out to discuss three tenets of Freud's account:

- In **§2.1** I explore the idea that mental functioning can be successfully described by appealing to *quantitative factors* – already introduced in **§1.1**. I also show that an economic approach to mental life allows us to meaningfully compare the psychical forces with the ones operating in the physical realm (e.g. energy, discharge).
- In **§2.2** I focus on etiology and I argue that most mental disorders are caused by two different kinds of pathogenic determinants: *dispositions* (i.e. constitutional factors) and *experiences* (i.e. accidental or contingent factors). This allows us to move towards a more complex view of psychopathology where disorders arise from the interaction between multiple factors and symptoms are almost invariably overdetermined.
- In **§2.3** I take a deeper look at symptoms and I argue that they can be seen as motivated solutions to the patient's problems as opposed to being considered problematic in themselves. More specifically, I show that symptoms can play a crucial role within therapy as long as they are meaningfully connected with the patient's life history and personality.

§2.1 An Economic Approach to Mental Functioning

As I have mentioned in §1.1, throughout his career Freud develops the idea that mental functioning can be described by appealing to *quantitative factors.* This notion can be traced back to one of his earliest works (1894), where he still presents it as a working hypothesis: "In mental functions something is to be distinguished – a quota of affect or sum of excitation – which possesses all the characteristics of a quantity (though we have no means of measuring it), which is capable of increase, diminution, displacement and discharge" (p. 9). Later in the essay, Freud characterizes this quantity by analogy with an electric charge that spreads over ideas or memory-traces and works similarly to the flows and fluids described by physicists (p. 10). A year later, Freud reiterates the same point: "It is impossible any longer to avoid introducing the idea of quantities (even though non-measurable ones)" (1895a, p. 86). The analogy with electricity also undergoes a significant development and starts to be applied to pathological phenomena: "If the tension is excessively high, there is danger of a break occurring at *weak points* in the insulation" (Ibid., p. 203. Italics mine). This passage is particularly interesting for our purposes because we can see an early reference to the idea that every mental apparatus exhibits a series of vulnerabilities (i.e. "weak points") in its organization.[1]

The assumption about the existence of quantitative factors within the mind becomes increasingly important in Freud's theory, to the point that he feels the need to add an *economic* approach to the tenets of metapsychology. As already mentioned above, such an approach sees psychical events not only in terms of their location (*topographical*) or forces (*dynamic*) but also in terms of their intensity, strength and magnitude (see also Kitcher 1995). In Freud's work, the economic point of view becomes particularly important when it comes to comparing healthy and

[1] Such an idea becomes particularly important in Freud's studies on etiology (see for example 1938, p. 64) and will also be extensively discussed later in the dissertation (see Chapters Four and Five on vulnerable individuals and populations).

pathological cases: indeed, the two are often seen as topographically and dynamically similar but economically different. For example, dreams are psychotic forms of wish-fulfilment that can be distinguished from hallucinatory experiences only in virtue of their short duration and of the higher degree of control that the subject exercises upon them (see 1938, p. 49). Notably, Freud also saw the introduction of an economic viewpoint as a way to bring psychology closer to physical sciences, where quantities are successful means to capture and measure a great variety of natural phenomena. If introducing notions such as energy, instincts or cathexis comes at the price of a certain indeterminacy, Freud is ready to point out that a similar degree of approximation affects many other disciplines. As he puts it in the *Outline*: "The basic concepts and principles of the new science (instinct, nervous energy, etc.) remain for a considerable time no less indeterminate than those of the older sciences (force, mass, attraction, etc.)" (1938, pp. 30-31). Therefore, the introduction of the economic approach should be seen as an attempt to carry out a thoroughly materialistic enterprise, where the physical and the psychical realm are subsumed under similar laws (see Makari 2008 for a similar point).

From a theoretical viewpoint, the adoption of an economic approach proves helpful in the investigation of normal functioning and pathology. When describing normal functioning, Freud repeatedly appeals to notions such as "quotas of affect" (1895a, p. 205) or "sums of excitation" (1915a, p. 152) that can increase, decrease, spread over ideas or memories, be displaced or discharged. Such a quantitative approach does not only defend the idea that "in mental life, some kind of energy is at work" (1938, p. 37) but also grounds a particular view of affect, instincts, and control. Freud describes affect in terms of *libido*, a quota of displaceable energy that attaches itself to psychical contents – e.g. ideas, memories – and can join forces both with erotic and destructive impulses (despite being *per se* neutral in quality). The general principle governing libido is to avoid accumulation and facilitate discharge: in this sense, we can imagine some sort of "mental

metabolism" where energy can be stored and released in response to a range of different stimuli (see 1923, pp. 61-62). Every mental structure becomes involved in this process in a way that reflects its peculiar aims and functions. For example, the id exhibits a certain "looseness in displacement" because it regards discharge itself as more important than particular actions or goals. Conversely, the ego is "more particular about object-choice and path of discharge" because of its connection to the ideals promoted by the super-ego as well as to the stimuli coming from the external world (Ibid., p. 64). On Freud's view affect plays a crucial role within thought: indeed, "our ego always entertains purposive cathexes, and often many at the same time" (1950, p. 434). *Instincts* are also characterized quantitatively as "excitations" or "wishful impulses" arising from the relationship between internal and external world: they are "psychical representatives of organic forces" (Freud 1915c, p. 122). Notably, instincts are described as qualitatively similar to one another but as varying with respect to the amount of excitation that they carry. That said, some of the instincts find their satisfaction via an alteration of the internal source (e.g. thirst) while others require a modification of the external world (e.g. getting rid of an enemy by attacking). Finally, *control* arises progressively as a way of mastering stimuli coming from different sources. On the one hand, external stimuli and threats require immediate responses in terms of actions – e.g. flight from a predator. On the other hand, internal stimuli trigger a set of more or less complex activities that aim at changing the world in order to afford satisfaction – e.g. courting a potential partner. The development of control also requires a distinction between primary and secondary processes: the former are animated by drives and primarily seek discharge (pleasure principle); the latter correspond to a regulatory mechanism that subject satisfaction to the resources and constraints available in the environment (reality principle). In what follows, I show that the relative strength of affective processes and instincts as well as the degree of control

that a subject exercises on them play a crucial role in distinguishing between healthy and pathological manifestations.

As I repeatedly stressed above, Freud insists on rejecting any kind of qualitative distinction between mental health and pathology: "Neuroses have no psychic content of their own which is not also to be found in healthy states" (1910, p. 23). On his view, mental disorders arise from acute affects or increases in excitation that fail to find proper discharge in physical and psychical activities. In other words, psychopathology originates from a subject's inability to master a certain amount of excitation crossing over a threshold, or from the failure to sublimate and convert a quota of affect to a different use (1916-17, p. 374-375). Notably, this threshold point is highly individualized although it is possible to individuate general mechanisms at work in specific disorders. For example, when talking about *anxiety* Freud makes clear that we all experience this affective state but that neurotics are affected by it "much more" and "so much more strongly" than others (Ibid., p. 393). He then proceeds to distinguish between realistic anxiety as the reaction to a perceived danger that puts the subject into a state of preparedness, and neurotic anxiety as a general and free-floating apprehensiveness ready to attach itself to any idea.[2] Yet, this distinction is also presented as a quantitative one: "The neurotic will differ from the normal person in that his reactions to the dangers in question will be unduly strong" (1926, p. 127). Similar examples are provided in Freud's account of *compulsions* and *phobias*, characterized economically as "excessively intense ideas" arising with a "special frequency" (1950, p. 405) or as ideas that appear strange because of their "intensity" as opposed to their content (1916-17, p. 399). Elsewhere, Freud discusses a more subtle economic distinction between healthy and hysterical subjects, one that concerns a difference in degrees of *attention*. More specifically, patients would

[2] Notably, Freud points out that such a distinction cannot be applied to young children, as realistic anxiety only arises through education and life experience whereas neurotic anxiety is innately present (e.g. self-preservation instinct). An interesting consequence of such a view is that it makes all children neurotic by definition.

fall prey to an extremely focused – albeit involuntary – attention towards certain aspects of their body or of the environment. Whereas normal subjects who focus attentively on one perception lose the capacity to experience other perceptions only temporarily, for hysterical patients "every idea takes possession of the whole of their limited mental activity" (1895a, p. 230). This remark suggests that some mental disorders could be regarded as continuous – or merely economically different – with phenomena such as focusing or attending to something. This would also help to explain pathological manifestations such as fixation on certain objects (e.g. phobias and compulsions), people (e.g. erotomania, Capgras delusion) or events (e.g. delusion of thought insertion, delusion of reference).

To sum up: the adoption of an economic approach to mental functioning allows Freud to hold on to the continuity thesis, while at the same time identifying two different ways in which normal processes can go awry. In some cases there are mechanisms that undergo a special development and thus become pathogenic (e.g. repression). In other cases, purely economic factors are sufficient for crossing the threshold between healthy and pathological manifestations (e.g. ideas arising with a special intensity or frequency).

§2.2 Dispositions and Experiences: Two Kinds of Pathogenic Determinants

From an etiological viewpoint psychoanalysis offers a complex and multi-factorial account of how mental disorders arise, and by doing so it distinguishes itself from competing views of mental illness as "organic inferiority" (Adler 1907) or "degeneracy" (Janet 1894). Freud repeatedly rejects overly reductionist explanations and insists that different etiological factors have to be active at the same time for a mental disorder to develop. As he sarcastically puts it: "The ideal solution, which medical men no doubt still yearn for, would be to discover some *bacillus* which could be isolated and bred in a pure culture and which, when injected to anyone,

would invariably produce the same illness. Or, to put it rather less extravagantly, to demonstrate the existence of certain *chemical substances* the administration of which would bring about or cure particular neurosis" (1926, p. 136. Italics mine). Yet, Freud's skepticism against a purely organic explanation of mental illness should not be interpreted as a wholesale rejection of the role played by biological factors. Indeed, in his account two different kinds of pathogenic determinants are present: *dispositions* or constitutional factors on the one hand, and *experiences* or accidental factors on the other. The former are described as those elements that "a person brings along with him into his life", whereas the latter are the ones that "life brings to him" (Freud 1913, p. 2623). Among the dispositional factors Freud enlists purely somatic elements (e.g. a high degree of nervous excitability) as well as early childhood experiences or facts related to sexual development. Accidental factors instead include psychological experiences that occurred later in life and significant conflicts between mental agencies (e.g. ego and id). The very idea that accidental factors would play an important role in the development of psychopathology has been introduced by Freud & Breuer in their *Preliminary Communication*: "[Our results] are valuable theoretically because they have taught us that external events determine the pathology of hysteria to an extent far greater that is known and recognized" (1893, pp. 3-4). A few years later, Freud insists that mental disorders should not be treated on a par with cases of "mental degeneracy" but that they could often be seen as motivated responses to traumatic life events (1895a, p. 1). However, Freud also recognizes that external events normally tap into some sort of susceptibility or predisposition exhibited by the patients, one that can be uncovered only by carefully analyzing their habits, personality traits and life-history. The discussion of clinical cases such as Emmy von N. (1895a, pp. 48-105) and Elizabeth von R. (1895a, pp. 135-181) offers a nice illustration of this enterprise (see **§3**).

Therefore, the novelty of Freud's account consists in the idea that mental disorders never have a uniquely identifiable cause but should rather be seen as *constellations* of pathogenic elements. Notably, the particular combination of factors cannot be established in advance and is subject to a high degree of interpersonal variation. As he puts it: "You must know that the same factors always come into operation in the causation and mechanism of every possible form of neurosis; but the chief importance in the construction of the symptoms falls now upon one and now upon the others of those factors" (1916-17, p. 381). Freud thus offers an etiological picture where two general kinds of pathogenic determinants – i.e. dispositions and experiences – are present but also in which every individual exhibits a particular combinations of them. But how does this model work exactly?

Freud fleshes it out more thoroughly in a series of essays focused on etiology, dispositions and onset of neurosis. As early as 1894, he had already acknowledged the importance of childhood experiences for etiology: indeed, clinical observations often showed that patients could trace back the origin of symptoms to one or more events occurred during their early years. Yet, in the article "Further Remarks on the Neuro-psychoses of Defense" (1896), he significantly refines his position by introducing two crucial observations. First, he notices that certain classes of people are more prone to developing mental illness with respect to others and that environmental factors are often responsible for this difference. For example, he points out that women are more frequently subject to sexual abuse during childhood (p. 4) and are often experiencing life-situations characterized by extreme monotony or frustration such as sick-nursing or house chores (see the clinical cases discussed in Freud 1895a).[3] Second, Freud realizes that the way in which

[3] As I mention above, I elaborate on this idea in Chapter Five by developing the notion of vulnerable population. For now, it is interesting to see that Freud repeatedly brings up the connection between the disadvantaged situation experienced by women and their greater vulnerability to mental illness. At the same time, his observations are not accompanied by a social critique but appear as mere descriptions of a *status quo*. For a feminist reading of Freud that makes a similar point, see Mitchell 1974.

the patients relive or remember their experiences is often far more important than the content of the experiences themselves. This remark allows us to characterize the notion of *trauma* in a way that goes beyond the reconstruction of what happened in someone's past and focuses instead on the way in which a person has appraised specific experiences or events. As Freud rightfully points out, someone else could have gone through the same experiences while "remaining unaffected" (1896, p. 10). This point is crucial because it introduces the idea that a subject's personality as well as her attitude towards life events could act as powerful pathogenic determinants. I return to this idea in Chapter Four, where I flesh out the notion of vulnerable individual by focusing on interpersonal differences in terms of risk and protective factors.

In later works Freud continues to develop the idea that pathogenic determinants interact in a complex way, with dispositional elements often going hand in hand with accidental ones. The article "Types of Onset of Neurosis" (1912) represents one of his few attempts at rigorous classification and discusses a list of possible triggers of mental illness ("precipitating causes"). Starting from observation and clinical experience, Freud notices that some patients fall ill because of *frustration*: in many of these cases, a love object is withdrawn and no substitute takes its place, causing libidinal energy to "dam up" due to the lack of discharge (Ibid., p. 232). Most patients treated by Freud at the beginning of his career could be subsumed under this category. Indeed, hysterical women often exhibit a combination of "mental liveliness" and "monotonous life" that makes them particularly prone to accumulate a surplus of mental energy unable to find discharge (see for example Anna O. in Freud 1895a, pp 22-47). Other patients suffer excessively because of the *demands of reality* that exacerbate the conflict between the impulses coming from the id, the repressing forces expressed by the super-ego and the ego's mediating role. The negative consequences of such a conflict are exemplified by another early case discussed by Freud, where the patient is described as "overly hard on herself" and as exhibiting a particular kind of moral

oversensitivity with strong self-deprecation tendencies. In particular, she could not accept the existence of "a whole multitude of indifferent, small things lying between what is good and what is evil – things about which no one need reproach himself." (see Emmy von N. – Freud 1895a, pp. 48-105).[4] In the sense described above, the demands of reality had too strong of an effect on this patient, who took every small event as an occasion for feeling guilty or morally reprehensible. Finally, Freud introduces a third category of patients, namely those whose onset is triggered *only* by a change in economic factors – e.g. a sudden libido increase. As I discuss above, this "crossing of a threshold" plays an important role in drawing the distinction between health and pathology. As Freud puts it: "There is no qualitative difference between determinants of health and those of neurosis […] on the contrary, healthy people have to contend with the same task of mastering their libido - they have simply succeeded better in them" (1912, p. 237).

Despite the attempt at categorization, the boundaries between these groups of patients do not appear clear-cut and the three types of onset often blend into one another. For example, a patient could experience a relative degree of frustration because of her inability to cope with the demands of reality, and this increase of tension could then economically determine her crossing of a threshold into illness. The interplay among different types of factors helps us to clarify the idea of a *pathogenic constellation*, where there is no simple relation between organic dispositions, life events and symptom formation. Indeed, most pathological manifestations emerge through summation: various kinds of dispositions seem necessary but not sufficient for mental disorders to develop, as there must be other "reasons" to bring them about (e.g. traumatic events, mental solitude or isolation). As a consequence, Freud outlines a model in which symptoms are almost invariably *overdetermined*: we cannot speak of a unique "cause of illness" (1912, p. 237) or of a specific "pathogenic excitant" as we would do for an infectious disease (1938, p. 63). The idea of

[4] I discuss this case more thoroughly in Chapter Two, §4.1.

over-determination is nicely illustrated in the following passage, where Freud attempts to trace back the different pathogenic determinants to their origin: “The logical chain corresponds not only to a zigzag, twisted line, but rather to a ramifying system of lines and more particularly to a converging one. It contains nodal points at which two or more threads meet and therefore proceed as one; and as a rule several threads which run independently, or which are connected at various points by side-paths, debouch into the nucleus” (1895a, p. 290).

§2.3. The Significance of Symptoms

The account discussed in **§2.2** shows that Freud sees the development of symptoms as a complex process where multiple factors are at work and play a causal role. More importantly, symptoms are connected with the patient’s dispositions and life-experiences in a way that often escapes generalization. As he puts it: “The extraordinary diversity of the psychical constellations concerned, the plasticity of all mental processes and the wealth of determining factors oppose any mechanization of the technique” (1913b, p. 123). Psychoanalysis thus presents itself as a practice or method that works effectively on average but whose outcome depends upon a host of factors that cannot be established in advance – e.g. the patient’s motivation to be cured. Within this framework, Freud regards a mere checklist approach to symptoms as deeply wrongheaded: in fact, symptoms turn out to be revealing only when they are meaningfully connected with the patient’s life history and personality. This brings him to forcefully reject the view that curing a mental disorder would amount to removing symptoms from the patient’s experience. The point is defended in several passages: “For the laymen the symptoms constitute the essence of a disease and its cure consists in the removal of the symptoms” (1916-17 - XXIII, p. 358). Similarly, while discussing a clinical case: “It seems to me that we concentrate too much on symptoms and concern ourselves too little with their causes” (1909, p. 2120). Again, in a lecture about delusions

Freud distinguishes psychoanalysis from medical psychiatry by saying that the former regards the content of delusions as significant, whereas the latter focuses only on their external form (1916-17 - XVI, pp. 248-254).

Yet, despite the relentless criticism of medical psychiatry, it is important to point out that Freud conceives of the relationship between psychoanalysis and somatic medicine as harmonious. Indeed, he readily acknowledges that psychoanalysis is currently a super-structure employing psychological concepts that one day will be grounded upon an organic foundation (Ibid. XXIV, p. 389). At the very end of his career, Freud brings up the idea that pharmacological therapy could work as a possible way to exercise "a direct influence, by means of particular chemical substances, on the amounts of energy and their distribution in the mental apparatus" (1938, p. 62). This passage shows – again – Freud's commitment to a materialist view of the mind in which psychological explanations of mental disorders are regarded as fully compatible with neurobiological accounts. In this sense, the common objection that psychoanalysis is a "purely psychological theory that could never explain illness" should be rejected as untenable (Ibid., p. 30). Another commitment to the idea of neural correlates can be found in Freud (1915c), although this is followed by the consideration that most attempts at localizing mental processes in the brain have been unsuccessful: "Every attempt to deduce from these facts a localization of mental processes, every endeavor to think of ideas as stored-up in nerve cells and of excitations as passing along nerve-fibers, has completely miscarried" (p. 107).

Despite the underlying compatibility between psychological and neurobiological explanations, psychoanalysis regards the hope for a one-to-one correspondence between a symptom and a chemical component as utterly misguided. Instead, symptoms should be seen as carrying a meaning that is importantly connected with the patient's experience. More specifically, there seems to be a strong correlation between the degree of idiosyncrasy exhibited by a symptom

and the connection with a specific life-event (Freud 1916-17 - XVII, p. 270). Freud explains this fact by individuating two kinds of experiences that may play a pathogenic role: a) past situations in which the pathological action or thought served a real purpose (idiosyncratic symptom); b) experiences typical of all human beings such as traumas arising from separation or loss (general and frequent symptoms). A good example of a) would be the case of the "tablecloth lady", a young woman who would obsessively stain her tablecloth in a particular way and then call her maid to come see it (see 1907 & 1916-17). Freud connects this highly idiosyncratic symptom with the patient's traumatic experience of her wedding night, when she found out that her husband was impotent and thus was left without a public proof of defloration – i.e. traditionally, a stained sheet. Examples of b) are more commonplace symptoms that differ only quantitatively from normal fears or obsessions – e.g. phobia of dangerous or disgusting animals; overly cumbersome sleep ceremonials. Granted this deep connection between symptoms and experiences, the goal of psychoanalysis cannot be to get rid of symptoms *per se* but rather to uncover the repressed ideas, the motives for repression and the specific pathogenic mechanisms at work in the patient (see 1926, p. 38).

In this sense, Freud can be considered one of the first defenders of the *motivational approaches* in psychiatry, where symptoms are seen as motivated reactions to the patient's problem and not as problems themselves (see Lopez-Silva 2015 for a contemporary defense). This idea has been present since the early years of psychoanalysis and arises from Freud & Breuer's observation that some symptoms vanish after the patient has recollected the trauma and – most importantly – has "put the affect into words" (1893, p. 6). This process comes to be known as *abreaction* and could be described as the discharge of affective energy through language. But why should putting affect into words be therapeutic? Freud's initial reply is admittedly tentative: "Through the cathartic method, strangulated affect finds a way through speech and subjects it to

associative corrections by introducing it into normal consciousness" (1895a, p. 255). The abreactive method has been later abandoned and replaced by others: in particular, Freud slowly realizes that therapy should not focus on the original trauma but on the psychical forces and resistances operating in the present (see 1914, p. 147). However, the idea that speech accompanied by affect would perform some sort of cathartic function remains central within the psychoanalytic method. For example, while discussing the case of Emmy von N., Freud notices that general didactic remarks or pieces of advice – e.g. "There is no need to be afraid of what is new" or "Promise me you won't be obsessed by this animal anymore" – are utterly useless for therapeutic purposes. Indeed, symptoms cannot be tackled in a wholesale fashion and words always need to be accompanied by the proper emotional reaction on the patient's side in order to be effective (1895, p. 60 & 99). Later, Freud insists on the fact that symptoms arise from unconscious processes and cannot be simply communicated to the patient "as pieces of information" (1916-17 – XVIII, p. 279).

In a nutshell: how do symptoms arise? As I discuss above, Freud sees symptoms as compromise formations emerging either from the dynamical struggle between different kinds of mental forces or by summation due to economical changes within the apparatus. This process may take a variety of forms: at times, the unconscious impulses coming from the id succeed in overcoming the resistance put in place by the ego (e.g. hysterical phenomena). Other times, the repression exercised by ego may be too strong and give rise to phenomena of psychical conversion (e.g. obsessions or phobias). In this sense, the formation of symptoms can be seen as a motivated response to the patient's problem because it allows her to partially – although often non-adaptively – discharge some mental energy through displacement. Going back to the example above, staining the tablecloth partially compensates for the fact that the bed was not stained; at the same time, the patient makes the stain public by showing it obsessively to the maid, thereby

relieving the sense of shame connected with her husband's impotence (see Freud 1916-17 - XVII, pp. 261-264).

§3. A Freudian Case Study: Elizabeth von R.

In this final section I set out to illustrate how Freud's account of mental disorder works in practice by examining one of the clinical cases discussed in the *Studies on Hysteria* (1895a): Elizabeth von R. (pp. 135-181). Before delving deeper into the analysis of this case, it is worth stressing that Freud regarded the accurate reconstruction of a patient's history as a crucial element for therapy as well as for the theoretical development of psychiatry.
Here he puts the point neatly:

"I have not always been a psychotherapist. Like other neuropathologists, I was trained to employ local diagnoses and electro-prognosis, and it still strikes myself as strange that the case histories I write should read like short stories and that, as one might say, they lack the serious stamp of science. I must console myself with the reflection that *the nature of the subject is evidently responsible for this*, rather than any preference of my own. The fact is that local diagnosis and electrical reactions lead nowhere in the study of hysteria, whereas a *detailed description of mental processes* such as we are accustomed to find in the works of imaginative writers enables me, with the use of a few psychological formulas, to obtain at least some kind of insight into the course of that affection" (Ibid., pp. 160-161. Italics mine).

In this passage Freud expresses – again – the need for a new level of description for mental disorders, one able to capture the complexity of the patient's experience without being at odds with what happens at the neurobiological level. Despite all their idiosyncrasies and imprecision,

Freud believes that case histories could offer some help in this direction and thus carefully collects them throughout his career (see 1909; 1911b for a few examples).

The case of Elizabeth von R. works as a paradigmatic example to illustrate the theoretical aspects outlined in **§1** and **§2**. Indeed, Freud offers here a fine-grained *etiological* explanation taking both dispositional and accidental factors into account (pp. 160-161); he also makes his commitment to *continuity* explicit while drawing significant differences between normal and pathological cases (pp. 164-165 & 174); and finally explains the emergence of mental illness by appealing to *economic* factors (p. 157 & p. 174). The analysis of Elizabeth's case starts with a short description of the patient's symptomatology along with her character traits and family environment. When Freud first meets her, Elizabeth is a 24-year-old woman suffering from a persistent pain in her legs and exhibiting an interesting form of "painful fatigue": these symptoms prevent her from walking normally and cause her a great deal of distress (p. 135). Freud describes Elizabeth as highly intelligent and ambitious: during her childhood she had a close affective and intellectual relationship with her father, who used to treat her "like a son and a friend with whom he could exchange thoughts" (p. 140). Growing up, her discontent with being a woman grew stronger as she realized that she was not willing to sacrifice her freedom of judgment and her inclinations for marriage. As a consequence, both her sisters ended up getting married while Elizabeth devoted most of her adult life taking care of her sick parents as well as attending various kinds of family business. Following the death of her father – whom she has sick-nursed until the end – Elizabeth started displaying the symptoms described above in a mild form. A few years later, the sudden death of her sister due to a complication during pregnancy marked the definitive onset of Elizabeth's disease. Up to this point, Freud admits that it is very difficult to see a direct connection between the patient's symptoms and her life-history, since the latter seems characterized by "commonplace emotional upheavals" that elicit "human sympathy" but blatantly

fail to explain her disorder (p. 144). However, Elizabeth also seems to fit the general description of the "Typus Hystericus" provided by Freud: an overflowing productivity of the mind, coming from her intelligence and ambition, inevitably clashes against the monotony of her family life and duties – e.g. sick-nursing or house chores (p. 240). Notably, this first part of the analysis already uncovers pathogenic determinants of different kinds. On the one hand there are *dispositional* factors such as Elizabeth's personality and her family's history of nervous illness (p. 140); on the other there are *accidental* factors such as the fact that she had to bear the responsibility of sick-nursing within the family. At this point, Freud has also noticed that some accidental factors are more recurring than others: for example, "sick-nursing plays such a significant part in the prehistory of cases of hysteria" (p. 161). Indeed, this condition – often falling back on women – invariably correlates with a situation of personal neglect, constant worry, lack of sleep, exercise or proper diet (p. 175). Crucially, sick-nursing also brings about the habit of suppressing one's own emotions and diverting one's thoughts away from anything that does not immediately relate to the person that is being assisted (pp. 161-162).

Starting from these observations, Freud formulates a hypothesis about Elizabeth's cause of illness: the symptoms may originate from a conflict between an incompatible idea fended off from consciousness and her sense of duty towards her family. By exploring this intuition Freud succeeds in tracing back the beginning of the patient's leg-pain to two important events: once at a party she spent the whole night with a man to whom she was attracted, just to return home and find her father's conditions worsened. Years later, she got to spend some time with her sister's husband and became fond of him to the point that she started to desire "a husband like him" (p. 154). These two episodes are importantly similar because they display the conflict between Elizabeth's desire of being in a relationship with a man and her sense of duty and guilt towards her family. In both cases, the latter feelings prevail and cause the erotic impulses to be repressed:

Elizabeth cannot bring herself to date the man she met at the party because she decides to take care of her father. Later, she cannot admit to herself that she has fallen in love with her brother-in-law because she does not want to hurt her sister. Finally, the conflict becomes unbearable when Elizabeth finds herself at her death sister's bedside and cannot help thinking: "Now he is free again and I can be his wife" (p. 156). Such an incompatible thought is immediately fended off from the patient's consciousness and thus ceases to enter in association with other ideas. As Freud puts it: "Her love for her brother-in-law was present in her consciousness like a foreign body, without having entered into relationship with the rest of her ideational life" (p. 166).

The psychoanalytic approach to Elizabeth's case allows us to draw a number of conclusions about the etiology and development of her symptoms. Indeed, Freud identifies a both *motive* and a *mechanism* for the patient's hysterical disorder: the former can be described as Elizabeth's need to defend herself against an incompatible idea (i.e. "I am in love with my brother-in-law"), whereas the latter consisted in converting a quota of psychical affect into a physical manifestation (i.e. leg pain and paralysis). Moreover, the symptom itself seems to bear a strong resemblance to the repressed idea, as Elizabeth repeatedly describes herself as "helpless" and unable to "take a single step forward" (pp. 152 & 176). Finally, this case shows that mental disorders often arise by *summation* of partial traumas and that specific symptoms come about after a certain threshold has been crossed. Indeed, as Freud points out, Elizabeth had been able to bear incompatible ideas in her consciousness for some time without serious consequences. Yet, there seems to be a limit to the degree of affective tension that a person can tolerate as well as to the pressure brought about by a new conflict between ideas. Once again, "what we are concerned with is clearly a quantitative [economical] factor" (p. 174). However, since this observation applies to hysterics as well as to healthy subjects, the discussion of Elizabeth's case also offers more

evidence in support of *continuity*: “The view which I have just been putting forward does no more than bring the behavior of hysterical people nearer to that of healthy ones” (Ibid.).

Conclusion

In this chapter I defended an early version of the Continuity Thesis in philosophy of psychiatry by showing that Freud’s account of mental disorder could be seen as an exemplar of such an approach. Generally speaking, Freud rejects the idea that the difference between healthy and pathological subjects could be described as categorical or clear-cut (see 1933; 1938) but at the same time he is committed to highlight some important differences between the two groups (see 1907; 1917; 1924b). More specifically, he outlines a view where mental disorders are characterized as *constellations* of elements that are not pathological in themselves, but become so in virtue of their intensity, accumulation or interaction (see 1912; 1916-17). To put it in Freudian terms, two mental processes can be identical from a topographical and dynamic viewpoint (i.e. involving the same provinces and exhibiting the same conflicts or interactions) while being importantly different from an economical perspective - i.e. displaying a higher or lower intensity, frequency or strength.

Throughout the chapter I focused on three theoretical aspects of Freud’s approach that I deem valuable for the elaboration of a refined version of CT. First, I discuss the idea that mental functioning – as well as malfunctioning – can be described by appealing to quantitative or economical factors. For example, an incompatible idea can be tolerated by a subject until it reaches a certain degree of intensity or frequency that causes her to cross a threshold and develop pathological manifestations. Second, I explore Freud’s suggestions about etiology and I argue that most mental disorders are caused by two different kinds of pathogenic determinants: *dispositions* (e.g. heredity or personality traits) and *experiences* (e.g. traumatic events). This brings about a

complex view that goes beyond the opposition between internal and external factors and regards psychopathology as emerging from the interaction between different determinants. Third, I maintain that symptoms can often be seen as motivated solutions to the patient's problems as opposed to being considered problematic in themselves. Following Freud, I thus reject a checklist approach to symptomatology and I argue that symptoms can play a crucial role within therapy as long as they are meaningfully connected with the patient's life-history and personality. Finally, I offer an example of how the approach proposed by Freud works in practice by discussing the paradigmatic case of Elizabeth von R. (1895a, pp. 135-181). This patient's history proves particularly useful in highlighting the advantages of the view outlined above. On the one hand, the appearance of Elizabeth's symptoms is explained by appealing to different kinds of pathogenic determinants (e.g. character traits and life events) as well as to economical considerations (e.g. accumulation of unbearable ideas). On the other hand, both the motive and the mechanism of her disease (i.e. defense and conversion) are not a unique prerogative of hysterics but appear rather as common mental processes distributed on a continuum of intensity.

Section II:

Strong Continuity

Valentina Petrolini

Are Mental Disorders Continuous with Healthy Functioning? A New Proposal for a Dimensional Model

"A logician would have started by defining what he meant by disease as a whole and then produced individual diseases by sub-dividing the territory whose boundaries he had thus defined. Medicine, being essentially practical and opportunist, proceeded the other way and started with individual diseases. As a result, many of these overlap with one another, and the outer perimeter between disease and health is based on different criteria in different places."

(Kendell 1975, p. 307).

Abstract

In this chapter I put forward a novel approach to psychopathology that aims to describe and classify mental disorders in a dimensional way. This step proves crucial in providing a defense of the Continuity Thesis (CT), because it clarifies what it means for disorders to be extreme variations of normal psychological phenomena. In this sense, the discussion contributes to make space for a dimensional view as a viable alternative to its categorical opponent. I then address the crucial question of how to distinguish between normal and pathological cases in Chapter Four.

I start by identifying four kinds of disorders that encompass many distinguishing features of psychiatric conditions, with a focus on schizophrenia and depression. Then I propose to conceive of these disorders as disruptions of four dimensions of functioning, corresponding to different ways of modulating the relationship with one's environment (i.e. appraisals). These are *salience*, *confidence*, *familiarity* and *agency*. Every dimension may be disrupted in opposite ways, with instances of deficit at the one extreme (*hypo*) and instances of overload at the other extreme (*hyper*). Along the spectrum lies also a multiplicity of intermediate cases: some of them approximate one of the extremes and could be characterized as local imbalances, whereas others can be regarded as healthy or conducive to wellbeing. As we go along it becomes clear that these distinctions amount to over-simplifications, and that the assessment of health and pathology may

be highly individualized as well as varying over time for the same person. Yet, sticking to the simplified version of the model for now helps to better grasp the framework that I set out to defend.

The chapter is divided into five sections. Each section from **§1** to **§4** begins with a brief characterization of a mental disorder category (e.g. disorders of salience). This is followed by a detailed discussion of selected case studies aiming to illustrate the disorder from a phenomenological viewpoint. At the end of each section I briefly introduce a number of intermediate cases, namely situations in which an experience is close to one of the extremes (i.e. local imbalance) or in which there is sufficient balance within a given dimension (i.e. health or well-being). A detailed discussion of these intermediate cases constitutes the bulk of Chapter Four. In **§5** I discuss some important similarities and differences among the disorders discussed in §1 to §4. I conclude by providing some preliminary reasons for characterizing the four dimensions as affective: I then delve deeper into this issue in Chapter Three and explain how the dimensions relate to affective notions such as emotions, feelings and moods.

Introduction

I believe a convincing defense of CT should be two-fold. First, it has to provide a model of mental disorders where the distinction between normal and pathological boils down to a difference of degree. In other words, it has to paint a plausible picture of what it means to be an extreme variation of a phenomenon that is otherwise non-pathological or healthy. Second, CT has to provide some way to distinguish between health and pathology without trivializing this distinction and without neglecting the phenomenological peculiarity that accompanies mental disorder manifestations. In this chapter I set out to achieve the first goal, whereas in Chapter Four I focus on the second.

The list of dimensions I provide here should not be taken as complete: indeed, I am not committed to the claim that every mental disorder could be subsumed under these categories. Rather, the goal of the chapter is methodological. What I propose is a way of thinking about mental disorders in which health and pathology are treated as continuous with one another. One important goal of the chapter is thus to convince the defender of a categorical view that such a classification is in fact plausible. Specifically, the account proposed here aims to explain away the intuitively powerful idea that the transition from health to pathology would have the nature of a "jump" or "switch" (Jaspers 1913/1963; Schneider 1938/1959; Murphy 2006; Samuels 2009). My strategy in defense of CT will be the following: I start by discussing a number of pathological cases and I show that they can be seen as lying on opposite ends of a given dimension – i.e. hyper and hypo. This helps to substantiate the idea of continuity between pathology and normal functioning, where disorders are seen as extreme variations of non-pathological phenomena.

In Chapter Four I then discuss a number of intermediate cases lying on the same dimensions, showing that they either represent attenuated versions of the pathological cases (*type a) cases*) or that they can be seen as at-risk situations where a number of protective factors are present (*type b) cases*). This helps to explain away the powerful intuition that the gap between health and pathology would be categorical in nature. Indeed, both kinds of intermediate cases can be hardly seen as qualitatively different from their pathological counterpart. Provided that a thorough discussion of intermediate cases is the main object of Chapter Four, here I just introduce them briefly as general prototypes instantiating a range of behaviors, dispositions and beliefs.

Generally speaking, intermediate cases can be described as follows:

- *The perfect balance case* (Stoic or Buddhist character). This ideal case represents a situation in which the person enjoys a relative balance in one or more of the relevant dimensions. This person's experience may be characterized as appropriately tuned to the environment as well as to her own feelings and abilities. This "appropriate tuning" may take different forms: for example, the Stoic would be able to assess what is meaningful or important to reach her own goals (salience), or to correctly self-assess her performance in a given task (confidence).
- *The local imbalance case*. This includes all the situations in which a person experiences a certain level of imbalance along a given dimension. Most psychological phenomena are encompassed in this description: indeed, we almost always experience a degree of mismatch between our thoughts, dispositions or expectations on the one hand and the demands or opportunities offered by the environment on the other. For example, most people have difficulty successfully forecasting their own reactions to a variety of real-world situations (see Maibom 2016 & Wilson 2002), and many show a tendency towards confabulation and false memories (see French, Garry & Loftus 2009; Wheatley 2009).
- *The non-pathological extreme case*. This case represents situations that are not commonplace but play a particularly important role in the model proposed here. These are experiences located at the extreme of a given dimension and thus phenomenologically similar to the disordered cases. However, for a variety of reasons they fail to count as pathological. For instance, these may be cases in which the context requires the adoption of extreme psychological measures (e.g. war) or in which such measures are consciously selected in order to serve an adaptive purpose (e.g. sports). For example, a runner may want to promote an unrealistically overconfident attitude towards her performance in order to better train and push her abilities to the limit.

These prototypes are nothing more than crude simplifications and should not be taken as natural kind categories. However, they prove particularly helpful from a heuristic viewpoint. Indeed, they show that the model outlined here succeeds in capturing a great variety of psychological experiences, and they illustrate the underlying continuity between manifestations that appear very different from one another. In the remainder of this chapter I introduce the four dimensions and some of their disruptions in detail. In Chapter Four I explore more closely a number of intermediate cases for each of the four dimensions. The two chapters taken together provide a model of mental disorders that makes the continuity view more plausible and the categorical view harder to defend.

§1. Disorders of Salience

Disorders of salience can be characterized as more or less radical alterations in the sense of what is relevant or important for achieving one's goals or pursuing one's preferences. Disorders of salience affect both perception and cognition, on the one hand by modifying the way in which the world appear to us (e.g. as threatening, as unsafe) and on the other by directing our attention to external cues that gain particular significance (e.g. people's tone or facial expression). The appraisal of salience may be then disrupted in two opposite directions. On the one end of the spectrum, *hyposalience* manifests itself in phenomena such as lack of exploration and anhedonia (typical of depression) but also in instances of extreme fixity of beliefs denoting the inability to revise one's viewpoint in light of contrary evidence (delusions). On the other end of the spectrum, *hypersalience* may take the form of a distressing "interrogative attitude" (Minkowski 1923) where too many aspects of reality appear significant and cry out for an explanation. This phenomenon is particularly striking in patients suffering from delusions of reference, paranoid delusions and positive symptoms of schizophrenia more generally. From

these disruptions we may postulate the existence of an underlying ability that allows us to allocate significance in our environment according to our goals, preferences and dispositions. I dub this phenomenon *appraisal of salience*.

§1.1. Case Studies: Interrogative Attitude and Anhedonia

Schizophrenia and depression strike most of us as very different disorders. Yet, they can both be seen as disturbances where the appraisal of salience is altered in opposite directions. On the one hand, schizophrenic patients (especially in the early psychotic phase) experience a particular "keenness" where environmental stimuli are regarded as highly significant:

"I developed a greater awareness. My senses were sharpened. I became fascinated by the little insignificant things around me"; "Things seemed clear cut, I noticed things that I had never noticed before" (Kapur 2004, p. 15).[5]

On the other hand, patients affected by major depression typically experience a pervasive lack of meaning that permeates everyday activities:

"I felt very still and empty, the way the eye of a tornado must feel, moving dully along in the middle of the surrounding hullabaloo"; "I couldn't see the point of getting up. I had nothing to look forward to" (Plath 1963, pp. 3 & 62).

At the same time, the two groups often employ similar metaphors to describe their experience. For example, Saks (2007) characterizes schizophrenia as a "slow fog, becoming

[5] See Humpston & Broome 2016 for a more comprehensive account of "perplexity" in early schizophrenia.

imperceptibly thicker as time goes on" (p. 35) and associates her recovery with the sensation of the fog being "lifted" (p. 200). Similarly, one patient describes depression as follows: "It is as though a black fog has descended and you are trapped within a black sea of treacle being dragged to a bottomless pit" (Ratcliffe 2015, p. 263).

Hypersalience. Experiences of hypersalience are poignantly described in many first-person accounts of schizophrenia as well as in clinical journals where patients' everyday actions are reported. For example, Minkowski (1923) talks about Paul – a young patient diagnosed with schizophrenia – who exhibits a peculiar "interrogative attitude" with respect to the most commonplace and mundane events:

"When I have to go to the bathroom, I first check the time in order to not stay too long. *It takes a certain amount of my time to look at my watch; I check exactly how the hands are placed* […] Before the bathroom, near the door, there is a feather duster that is used to clean the ceiling. I now find that this duster seems longer than usual; previously, it seemed shorter to me. *I know that an object that has been seen repeatedly ends up looking less remarkable than at first, but I nevertheless wonder whether the duster has become longer or shorter*. Every time, I look at it and ask myself this question. After entering the bathroom, I am not sure of having closed the door. I pull the door several times to make sure that it is shut. Then, since the door does not close well, I examine the crack from top to bottom. *I look to see whether the daylight that filters through this crack is of the same intensity as usual; sometimes, it appears darker or lighter to me*" (p. 273. Italics mine).

Paul's experience of hypersalience clearly expresses the sense of perplexity that haunts many schizophrenic patients, who appear puzzled or disturbed by objects that are commonplace and uninteresting to others, such as clocks, door cracks and feather dusters. In this sense, every

little detail appears significant and is at times interpreted as carrying a deeper (and ominous) meaning – e.g. “I was being messaged by various things I found on the sidewalk: a red elastic band, intact, meant that I would not have to open my wrist and bleed out again” (Hawkes 2012, p. 1109). At the same time, this sense of enhanced salience can be interpreted as a hopeless search for meanings that have been lost: talking about treatment, Hawkes describes the process of re-discovering significance in everyday events such as meals, birthday parties and family celebrations (p. 1110). The experiences of perplexity and interrogative attitude may thus be characterized as exhibiting a peculiar balance between loss of meaning and some items showing up as especially salient.

Hyposalience. The experience of hyposalience that affects many depressed subjects has been beautifully described by Plath in her semi-autobiographical novel *The Bell Jar* (1963). As a 20-year-old English major who won a scholarship to spend a month in New York working for a fashion magazine, Plath cannot help but feeling *empty*:

“*I was supposed to be having the time of my life.* I was supposed to be the envy of thousands of other college girls just like me all over America […] Look what can happen in this country, they’d say. A girl lives in some out-of-the-way town for nineteen years, so poor she can’t afford a magazine, and then she gets a scholarship to college and wins a prize here and a prize there and ends up steering New York like her own private car. *Only I wasn't steering anything, not even myself. I just bumped from my hotel to work and to parties and from parties to my hotel and back to work like a numb trolleybus.* I guess I should have been excited the way most of the other girls were, *but I couldn't get myself to react*”; “The city hung in my window, flat as a poster, glittering and blinking, but it might just as well not have been there at all, for all the good it did me” (pp. 2-3 & 11-12. Italics mine).

This first-person account allows us to get a better grasp of the emotions and moods experienced by depressed subjects affected by anhedonia and flat affect. In particular, Plath describes the stark contrast between things that are "supposed to be" exciting and fitting her interests – e.g. winning a prestigious scholarship, living in New York – and the pervasive sense of emptiness that permeates all her activities. Similarly, the subjects interviewed by Ratcliffe talk about "feeling like a ghost" or "a spectator", where both expressions convey a sense of passivity and incapacity to see things as relevant or worth pursuing (2015, p. 32).

From a phenomenological perspective, the similarity between these two groups of patients may turn out to be closer than it initially appears. Indeed, Ratcliffe (2015) describes the experience of being depressed as "being in a different world" and as something that deeply transforms the overall structure of the relationship with ourselves and our surroundings (p. 15). Similarly, Stanghellini (2000) talks about schizophrenic patients as lacking the sense of "at-homeness" that accompanies most of our everyday experiences (p. 777). Schizophrenia and depression (at least in their interrogative and anhedonic components) may therefore qualify as different ways in which the appraisal of salience can be disrupted. The former implies an overload of significance attributed to objects and events that should be regarded as neutral or commonplace – e.g. the amount of light filtering from the door crack (Minkowski 1923). The latter presents the world as completely devoid of meaning, even in those aspects that are supposed to be significant given one's preferences and goals – e.g. winning a journalism scholarship (Plath 1963).

Intermediate cases lying on the salience dimension can be briefly described as follows. The perfect balance case would be someone whose relationship with the environment is characterized by *engagement* and *flexibility*. Engagement implies the presence of goals, preferences and concerns that drive the person's thoughts and actions. For example, regarding the

relationship with one's friends as important has an impact on how a person assesses her priorities or spends her free time. Flexibility implies the ability to negotiate and modify one's preferences and goals according to the context. For example, a flexible person would understand that he cannot find his favorite dish cooked exactly like at home when travelling to a foreign country. Local imbalances comprise situations where the appraisal of salience is tilted in one direction (i.e. hyper or hypo). On the low end we find phenomena such as conservatism about one's beliefs or habits: examples are forms of fixation or rigidity, where only a few objects, events or people are regarded as worth exploring. On the high end we find instances of neuroticism, where minor events are regarded as urgent and important (e.g. hypochondria) or where idiosyncrasies are magnified rather than downplayed (e.g. Woody Allen cases, see Killmister 2015). One extreme non-pathological case for salience – on the high end – would be a person actively pursuing pleasure, thrill or suspense and particularly susceptible to boredom. Frijda dubs this character type the *sensation-seeker*: "Strong sensations are liked by subjects who can handle them, because they can manage them. They are able to process them and cope with them and are bored with stimuli that are too easy to handle" (1986, pp. 347-349).

§2. Disorders of Confidence

Disorders of confidence represent difficulties in assessing one's performance, abilities or degree of fluency in a certain area. These disorders may be importantly related to anxiety, which can be either too prominent and paralyzing or (conversely) lacking even when the circumstances would require it. On the one hand, a significant loss of confidence translates into phenomena such as pathological self-doubt and lack of self-trust, which are common among patients affected by

major depression and borderline personality disorder.[6] On the other hand, a hypertrophic sense of confidence may give rise to the sense of grandiosity that is typically observed in some schizophrenic delusions, narcissistic personality disorder and psychopathy. Similarly to disorders of salience, these disruptions affect both self-perception and appraisal of the environment. In the hypotrophic case, an individual may feel unable to complete a certain task, or regard the task itself as daunting and completely beyond his grasp. Conversely, in the hypertrophic case someone may experience a sense of subjective omnipotence or project it outwardly and thus perceive people or situations as dull or not challenging enough. Underlying these disorders I postulate the existence of an ability responsible for assessing one's performance with respect to a variety of tasks: I dub this *appraisal of confidence*.

§2.1. Case Studies: Grandiosity and Self-doubt

Hyperconfidence. One paradigmatic example of hyperconfidence can be found in the so-called grandiose delusions, particularly common among people affected by schizophrenia and bipolar disorder. These delusions usually center on the idea of possessing special powers or abilities – e.g. mindreading – and at times take on a specific religious connotation, as when patients claim to have been chosen by God to undertake a certain mission (see Schreber 1903/1955 for a classic case).[7] In a recent first-person account, Reina (2009) briefly describes the onset and development of his delusions of grandiosity. As an architecture student at the University of Michigan, Reina becomes very fascinated by the study of colors and then starts putting together what he calls a "color theory". This roughly amounts to the idea that "one could determine a person's feelings, thoughts and even personality by the color of their clothing" (p. 3).

[6] Items such as "reduced self-esteem and confidence"; "worthlessness"; "feeling of inferior self-worth" are included in the DSM-IV and DSM-5 description of these disorders.
[7] For a recent perspective on religious delusions, see Graham 2015.

Reina quickly becomes convinced that such an ability to interpret people's minds through colors is a special power eliciting in others the utmost admiration. As he vividly puts it:

"In some of my courses, in particular, my introductory architecture drawing course, I was excelling. In that class, I had even become something of a teacher's pet and my instructor, notably a young woman, seemingly fawned over my work. My peers noticed how my teacher appreciated my work as well and began using some of my ideas, such as putting abstract pictures behind the overall composition […] *I began believing that I had a great future as an architect* […] Teachers in my other classes were likewise paying a great deal of attention to me. My physics teacher seemingly singled me out when he lectured and often looked my way during class. As the semester progressed, my teachers began to take an even greater interest in me, and I thought I had found out why. *I had discovered how to charm people into liking me using my color theory.* If I waved around a red pen, the color of power, I would gain some power over those who saw me wave the pen. By using people's natural sensory biases toward these colors, I could control them, to the extent that I could give people a favorable view of myself. I believed that *these new powers were some type of magic*." (pp. 3-4. Italics mine).

This report is particularly revealing because it shows how Reina's augmented sense of confidence brings about a transformed experience of his environment. On the one hand, his conviction of having special powers makes him interpret otherwise commonplace gestures – e.g. the teacher's colorful slides; the classmates' insisting gaze – as directed towards him and as arising from people's fascination with his abilities. On the other hand, these cues feed into the internal perception that he is indeed special and thus contribute to strengthening the delusion of grandiosity until it gets out of control. Reina describes this moment as one in which his magical powers "began to run away with themselves" (p. 4). The world suddenly becomes threatening because every glance, every word, every movement is perceived as directed towards him: people

are now not only fascinated, but want to hurt him out of envy. This is when delusions of reference and paranoid delusions arise out of grandiosity: "I had also begun to notice that I was being followed. My guess was that the university had assigned bodyguards to protect me *because I was such a valuable student with unique magical powers* […] Everyone I passed on the street would give a personal comment about me or insult me" (pp. 4-5. Italics mine). Notably, after some time also inanimate objects acquire the same intentional and threatening character: "Every bump on the wall of my apartment and every cry in the distance were directed toward me" (p. 5). Reina's delusion of grandiosity further develops after hospitalization, as he interprets the various medical tests as ways to assess his supernatural powers: "Apparently I still had to prove to them that I was the next Jesus Christ" (p. 6).

Hypoconfidence. Patients suffering from major depression often experience a pervasive sensation of self-doubt and worthlessness that profoundly affects their everyday functioning. This particular kind of self-doubt – unlike the interrogative attitude reported by schizophrenic patients – is often accompanied by the sensation of being utterly incapable of meeting the demands of reality. Notably, depressed patients seem unable to think that their condition may change for the better, and thus find it particularly difficult to muster resources to cope with more or less difficult situations. This attitude then creates a vicious circle where fewer enterprises are deemed worth undertaking, fewer challenges are embraced, and fewer options are regarded as available. This way, depressed patients often fall prey to self-fulfilling prophecies (see Merton 1948).

In his memoir *Darkness Visible* (1991), William Styron effectively illustrates this overwhelming lack of confidence through a series of real-life examples. During a short trip to Paris he is about to be awarded a generous prize because of his successful career as a novelist. The award is undoubtedly prestigious, and only one American author in history has received it before him. Yet, Styron cannot help but feeling staggeringly inadequate: first he turns down the

invitation to the luncheon in his honor, and then he somehow manages to lose the $25,000 check connected with the prize. He recalls the latter event as follows:

"Did I 'intend' to lose the money? Recently I had been deeply bothered that *I was not deserving of the prize.* I believe in the reality of the accidents we subconsciously perpetrate on ourselves, and so how easy it was for this loss to be not loss but *a form of repudiation, offshoot of that self-loathing* (depression's premiere badge) by which I was persuaded that *I was not worthy of the prize, that I was in fact not worthy of any of the recognition* that had come my way in the past few years." (p. 19. Italics mine).

This passage shows how an external event (i.e. the loss of the check) gets immediately interpreted by Styron as reflecting an internal inadequacy with respect to what the lost object signifies – i.e. "I must have lost the check *because* I did not deserve to receive it in the first place." Notably, this loss quickly becomes a powerful symbol of personal and all-encompassing failure. As Styron puts it: "The check was gone, and its loss dovetailed well with the other failures of the dinner: my failure to have an appetite for the grand *plateau de fruits de mer* placed before me, failure of even forced laughter and, at last, virtually total failure of speech" (pp. 19-20). Plath (1963) describes a remarkably similar attitude: when confronted with a commonplace negative thought – e.g. "I cannot speak German well" – she starts spiraling out of control and adding up all the things she cannot do: "I began with cooking [...] I didn't know shorthand either. This meant I couldn't get a good job after college [...] My list grew longer. I was a terrible dancer. I couldn't carry a tune. I had no sense of balance, and when we had to walk down a narrow board with our hands out and a book on our heads in gym class I always fell over. I couldn't ride a horse or ski, the two things I wanted to do most, because they cost too much money. I couldn't speak German or read Hebrew or write Chinese" (p. 40).

Another crippling aspect of self-doubt that emerges from these first-person accounts is the

perception of *finality*. In other words, depressed individuals attribute to themselves various incapacities or failures without seeing them as something that can be amended, or even slightly improved. Rather, these failures take on a non-modifiable and permanent character that prevent people from attempting to change their attitude in light of new events or situations. Styron stresses this point repeatedly: "In depression this faith in deliverance, in ultimate restoration, is absent. The pain is unrelenting, and what makes the condition intolerable is the foreknowledge that no remedy will come – not in a day, an hour, a month, or a minute. If there is mild relief, one knows that it is only temporary: more pain will follow. It is hopelessness even more than pain that crushes the soul" (1991, p. 62). In this sense, depressive self-doubt produces a somewhat paradoxical effect. A host of uncertainties about ordinary little things contributes to the development of one absolute certainty: namely, that the individual is hopeless and completely unable to cope. As Styron would put it: pain will continue, no matter the circumstances. This situation may also be seen as one in which hypoconfidence fosters hyposalience. Since nothing can be undertaken successfully, no enterprise appears worth undertaking; since improvement is utterly hopeless, no challenge can be overcome; since nothing can be processed with ease, no stimulus is regarded as interesting or worth exploring.

In Chapter Four I discuss a number of intermediate cases connected with the dimension of confidence. Perfect balance cases would be characterized by an appropriate combination of confidence and uncertainty, giving rise to a certain degree of *self-esteem* and *self-assessment*. Being balanced with respect to confidence means being able to assess one's resources in the face of the challenges posed by the environment: for example, understanding whether one has studied enough for an upcoming test. Yet, confidence is also connected with the ability to see difficult events as challenges to overcome rather than as burdens to endure or threats to avoid: such a capacity has been called resilience or hardiness in the psychological literature (see Kobasa 1979).

Local imbalances – on the low end – would be situations where a high degree of insecurity and self-doubt are present and where self-esteem appears compromised. A good example is the so-called "impostor syndrome" affecting several women in various professions and characterized by pervasive feelings of inadequacy with respect to one's status or performance (see Young 2011). On the high end, by contrast, we encounter cases of people systematically overestimating their abilities even in the face of contrary evidence: this has come to be known in psychology as the "overconfidence effect" (see Gigerenzer 1991; Sutherland 2007). Finally, one extreme non-pathological case would be exemplified by situations where people deliberately entertain optimistic beliefs about their abilities or degree of preparation in order to enhance their performance. For example, an athlete may choose to cope with the pressure leading up to a race by telling herself: "You are the best, you will win this!" even if she knows that realistically she does not stand a chance against her opponent.

§3. Disorders of Familiarity

Disorders of familiarity comprise various cases in which one's sense of comfort or security with respect to oneself and the world appears seriously compromised. Like the other cases discussed above, disruptions of familiarity may occur in two opposite directions. At the one end of the spectrum, objects, events or persons that should be familiar for the individual cease to be seen as such and end up generating a great deal of anxiety or fear (*hypofamiliarity*). This is for example what happens in delusions such as Capgras, where patients become convinced that a family member has been replaced by an identical impostor who differs from the significant other in some minor detail – e.g. the way in which he ties his shoes (Frazer & Roberts 1994, p. 557). Notably, cases of hypofamiliarity are not always other-directed but may involve the very nature of a person's subjective experience: this means that not only external objects or events, but also

one's own thought processes or body parts may feel strange and unfamiliar. For example, schizophrenic patients often report experiences such as depersonalization, where thoughts and feelings lack a sense of "mineness" (see Zahavi 2014). Other self-directed cases of hypofamiliarity are those in which one's body parts are perceived as extraneous and at times threatening – e.g. apotemnofilia (desire of having one's own limb amputated). At the other end of the spectrum, some individuals tend to regard objects, events or people that are unfamiliar as commonplace and well-known (*hyperfamiliarity*). This is the case for patients affected by the Fregoli delusion, who insist that all the strangers they encounter are actually one familiar person. A nice fictional example is Philip Roth's character Alexander Portnoy, who believed for years that all his teachers were his mother in disguise: "As soon as the last bell had sounded, I would rush off for home, wondering as I ran if I could possibly make it to our apartment before she had succeeded in transforming herself. Invariably she was already in the kitchen by the time I arrived, and setting out my milk and cookies. Instead of causing me to give up my delusions, however, the feat merely intensified my respect for her powers." (Roth 1969, p. 3). Interestingly, hyperfamiliarity may also be self-directed, as it happens in cases of confabulation or post-hoc rationalization where people "pretend" to have access to their own motives more than they actually do (see Hirstein 2009). Following the cases above, I postulate a further ability that can be disrupted: I dub it *appraisal of familiarity*.

§3.1. Case Studies: Fregoli and Capgras

Hyperfamiliarity. Fregoli delusion is a rare syndrome that usually occurs in the setting of schizophrenia and is often discussed in conjunction with other Delusional Misidentification Syndromes (DMS): perhaps due to the rarity of these conditions, autobiographies and first-person

accounts are hard to retrieve.[8] In a recent paper on the topic, Langdon & Coltheart (2014) characterize Fregoli delusion as "the mistaken belief that a known person is present in the environment in a different guise to his or her typical appearance" (p. 615). This indicates that most people affected by this syndrome acknowledge that the stranger(s) and the familiar person look *physically* different, while they insist in maintaining that they are *psychologically* identical. As opposed to generic persecutory delusions, patients suffering from Fregoli also report being haunted, followed or persecuted by people that are affectively related to them (e.g. mother) or significant for different reasons (e.g. famous person that they admire).

Despite the aforementioned lack of first-person accounts, psychiatric reports on the topic are centered on a bundle of clinical cases exhibiting some significant similarities. First, the known person being misidentified is often closely related to events or preferences that are affectively significant for the patient. For example, in the first reported case of Fregoli delusion a woman obsessed with theater starts to believe that her two favorite actresses are taking the form of other people she encounters in everyday life – e.g. bystanders, doctors, or employers (see Courban & Fail 1927). More recent studies describe women whose husbands had died more or less recently, and who identify fellow male patients or even female doctors with the deceased spouse (see Moriyama *et al.* 2007; Turkiewicz *et al.* 2009). Another interesting case involves a woman who developed Fregoli delusion after her fiancé had refused to marry her: after this event, she "misidentified strangers with an ex-friend who in the past had asked her to marry her but she denied his proposal" (Papageorgiou *et al.* 2002, p. 806). In all these cases hyperfamiliarity appears to be triggered by the memory of traumatic events – i.e. death of a loved one – as well as by current obsessions or disappointments – i.e. theater and break-up. Second, in many of these

[8] It may also be that most people affected by DMS do not reach the level of insight required to write an autobiography or a first-person account. This point is speculative, but it suggests something important with respect to the degree of severity connected with the disorders discussed in this section – i.e. Fregoli, Capgras and Cotard.

cases neurological evidence points towards over-activation of specific brain areas, namely the ones responsible for person recognition or affective response to faces (Langdon & Coltheart 2014, p. 619). These neurological data are still sparse and open to multiple interpretations, but they sit comfortably with the model defended here. Indeed, they suggest that disorders such as Fregoli can be seen as instances of hyper-activation of affective responses to unfamiliar faces or people.

In other – less explored – cases, hyperfamiliarity is directed towards places rather than people: for example, some patients argue that the hospital in which they are living is in fact their childhood home, or that a city they had never visited before is a copy of their hometown. These cases are interestingly similar to their non-pathological equivalents, such as commonplace experiences of *déjà vu* and *déjà vécu* (see O'Connor & Moulin 2010 for a review). However, like Fregoli, they involve specific places that are affectively relevant to the individual and connected with her life history – e.g. childhood home. In this sense, we can see familiarity as being modulated by salience: a particular person or place acquires exaggerated importance for a number of reasons, and this gives rise to an altered experience of one's environment. More specifically, the relationship with one's surroundings becomes filtered by a series of cues connected with the familiar person: small details such as a particular outfit, a way of talking or a certain furniture arrangement can be taken as (misleading) signals of the person's presence.

Hypofamiliarity. As opposed to the cases discussed above, some pathological experiences involve the lack of emotional reaction in the face of people that the patient would normally regard as familiar and affectively relevant. In some cases, family members or friends are thought to have been replaced by strangers, impostors or even sophisticated robots (Capgras delusion). In other cases, parts of one's body are regarded as dead or utterly alien (Cotard delusion) and subjective experiences are characterized by a pervasive sense of detachment (depersonalization).

All these conditions share a common core: on the one hand the patient recognizes that she *should* experience a particular person or object as familiar, but on the other hand she is *unable* to retrieve the appropriate phenomenology. In order to capture this inconsistency some have described hypofamiliarity as a loss of emotional connection to the significance of things (see Ratcliffe 2010) or as a consequence of extreme affective detachment (see Frijda 1986).

Some detailed examples of Capgras and Cotard delusions can be found in clinical reports. For example, Christodoulou (1977) discusses two cases of Capgras in patients affected by depression and paranoid schizophrenia. The first patient is a 60-year-old woman who became severely depressed after one of her daughters got married and moved to a different city. In connection with this event, she developed the delusion that her daughter had been killed and replaced by an impostor. During a meeting arranged by the family to convince the patient that her daughter was alive and well, she reacted this way: "She looked at her daughter carefully, inspected her facial characteristics, asked her to turn around and take off her shoes and walk. After a few minutes' hesitation she decided that 'this person' was only the double of her real daughter" (p. 557). The second patient is a 43-year-old woman who had been unhappily married for years, and at one point developed the delusion that her husband was dead and that "an identical-looking man had taken his place." The patient's behavior was remarkably consistent with the delusion, as "she put on black dressing in mourning of her 'late' husband, refused to sleep with his 'double' and angrily ordered him out of the house, shouting 'go to your own wife'" (p. 558). More recent cases of Capgras are described by Frazer & Roberts (1994). In these reports, one patient maintains that her husband's personality had completely changed and that he had been replaced by a "demonic double", while another is convinced that the same had happened to her son while she was hospitalized (p. 557).

Notably, these patients usually claim to be able to tell the difference between the real person and the duplicate by appealing to some minor detail to which they have privileged access. For example, one patient thought that the impostor could be distinguished from her real husband because he tied his shoe laces slightly differently, while another patient appealed to physical differences (e.g. different-colored eyes) as well as behavioral ones (e.g. her real son would not kiss her). Something similar happens with cases of Cotard, where people experience their body (or mind) as deeply unfamiliar and then attempt to explain these phenomena by focusing on details of somatic (or psychological) functioning. For example, some patients deny having a stomach on the grounds that they "never feel hungry" while others deny having a brain by saying that they "do not think" (see Billon 2015 for other similar cases). Although I do not explore Cotard cases in detail here, I believe they could be easily included among examples of hypofamiliarity and would thus function similarly to the Capgras cases discussed above.[9] The core difference would be that while in Cotard hypofamiliarity appears self-directed and involves one's own body or mind, in Capgras it is other-directed and involves family members or other significant people. Notably, as opposed to what happens in Fregoli, in these cases familiarity appears to modulate salience. Indeed, the overwhelming lack of familiarity experienced by patients may prompt them to regard particular details of the person as extremely significant. In other words, *since* the familiar person is perceived as strange, patients may rationalize and confabulate possible reasons to account for the perceived difference between the original and the impostor.

In Chapter Four I discuss a number of intermediate cases connected with the dimension of familiarity. The perfect balance case would be characterized by *congruence* (Rogers 1959) and appropriate *self-knowledge* as well as knowledge of others. In this sense the balanced person

[9] For a different account explaining Cotard syndrome in terms of familiarity, see Gerrans 2000.

would be able to foster situations and encounters where he feels safe, comfortable and at-ease. Yet, he would be also able to explore the reasons for possible mismatches in these areas. For example, when feeling that a friend is growing distant one may infer that she might simply be busy or rather upset for something that has happened in the recent past. Local imbalances on the low end would then be represented by feelings of estrangement, detachment or dissonance. In these cases, the subject may exhibit an insufficient degree of self-knowledge – as in self-deception – or completely misjudge another person's character (see Wilson 2002; Fine 2006). On the high end we would find phenomena where familiarity is inappropriately attributed to a place or situation (e.g. déjà-vu experiences), as well cases where an excess of familiarity gives rise to confabulation or false memories (see Hirstein 2009). Extreme non-pathological cases would finally be ones where a high degree of detachment or estrangement is achieved in order to endure an unbearable situation – e.g. war, violence or rape. Another example – on the high end – would be the enhanced feeling of familiarity towards strangers experienced after the death of a loved one, often giving rise to hallucinatory phenomena (see Rees 1971).

§4. Disorders of Agency

Disorders of agency could be described as cases where people encounter difficulties in assessing their own degree of responsibility with respect to a relevant action or event. On the one hand, a person may be unsure of whether she initiated an action that others attribute to her, or she might deny having done so despite evidence of the contrary. I call this kind of disturbance *hypoagency*. These cases encompass phenomena such as auditory verbal hallucinations (AVH), thought insertion or alien hand syndrome, where people experience their thoughts or bodies as something acting beyond their control. On the other hand, a person may feel that events that are completely unrelated to her actions (or even thoughts) fall under her own responsibility and thus

experience unbearable guilt as a result. This happens at times with schizophrenic patients, who tend to blame themselves for natural disasters, terrorist attacks or murders committed by others. For example, Elyn Saks talks about being filled with anxiety when reading the newspaper because she would blame herself for every violent crime reported in the area (2007, p. 260). In these cases subjects attribute to themselves a greater degree of agency and control than they actually possess, thereby exhibiting *hyperagency*. Like the other cases discussed above, some of these phenomena present themselves as self-directed (as in thought insertion) whereas others concern external people or events (as in pathological guilt). In both cases, an underlying *appraisal of agency* appears compromised.

§4.1. Case Studies: Pathological Guilt and AVH

Hyperagency. Pathological guilt is a symptom commonly experienced by people suffering from depression, but may also be present – albeit in a different form – in schizophrenic patients. Generally speaking, people experiencing pathological guilt have the tendency to feel responsible for things that they have not done or to feel deeply disturbed by actions and thoughts that are regarded as innocuous by others. These disorders can be characterized as instances of hyperagency because patients fail to properly assess the scope of their (moral) responsibility. This may happen in two ways: either by attributing to themselves actions for which they are in fact not responsible for (e.g. a murder that someone else committed), or by assigning a particularly negative valence to self-generated thoughts and events (e.g. feeling guilty about finding another person annoying). Some patients also highlight a specific bodily phenomenology associated with guilt: “It comes from below, from the belly, like a terrible oppression mounting up to the chest; then a pressure arises like a crime that I have committed; I feel it like a wound

here on my chest, that is my tortured conscience […] then it sucks forth my memories, and I have to think again of all that I have missed or done wrong" (reported by Fuchs 2003, p. 237).

One interesting example of pathological guilt comes from one of Freud's earliest case histories, Emmy von N. (see Freud & Breuer 1893, pp. 48-105). Frau Emmy is a 40-year-old woman who suffers from recurring hallucinations and from a number of tic-like movements, in particular an idiosyncratic "clacking sound" that would come up whenever she is anxious or frightened (p. 54). While analyzing her case at great length, Freud notices that the patient tends to be overly hard on herself and to feel directly responsible "for the least signs of neglect": "If the towels for the massage are not in their usual place or if the newspaper for me to read when she is asleep is not instantly ready to hand" (p. 65). One day, Freud arrives to the patient's house to continue the therapy and finds her in a state of great distress, repeating: "Am I not a worthless person? Is it not a sign of worthlessness what I did yesterday?" Freud cannot recall what happened the day before to justify such a "damning verdict" (p. 70). Despite Freud's repeated admonitions not to feel guilty over small things, Emmy keeps behaving like "an ascetic medieval monk, who sees the finger of God or the temptation of the Devil in every trivial event of his life and who is incapable of picturing the world even for a brief moment or in its smallest corner as being without reference to himself" (p. 66). Notably, after a two-year long therapeutic process, Emmy is able to recover from the majority of symptoms – i.e. hallucinations, tics – but her inclination to torment herself over "indifferent things" never vanishes completely. For this reason, Freud is inclined to regard such a tendency as a fundamental aspect of her character (p. 84).[10]

This pervasive presence of guilt feelings in some psychiatric disorders has been explored by several authors in the past couple of decades. For instance, Frijda connects guilt with the

[10] For a fuller account of self-loathing as a core component of melancholia (depression) in Freud's thought, see Radden 2009.

notion of control: "[Guilt feelings] may provide an explanation for one's misery, an explanation that provides an aspect of *controllability,* some shred of it, in the morass of helplessness; it permits acts of contrition and efforts at paying penance" (1986, p. 431). In this sense, disorders of hyperagency may arise as an attempt to control and thus to justify or explain – at least to an extent – feelings of worthlessness and helplessness. Ratcliffe (2010) instead characterizes the kind of guilt found in depression in terms of *depth*. As opposed to a circumscribed feeling of guilt about a specific action or event, depressed subjects tend to experience guilt as an "all-encompassing way of being" (p. 9). Two aspects distinguish deep forms of guilt from shallow ones: irrevocability (no sense of contingency is preserved) and inescapability (no reparation is possible). On this view, depressed patients feel guilty *both* because they believe that there is something essentially wrong with them, *and* because they feel irrevocably responsible for everything that happens. Ratcliffe (2015) collects a number of patients' reports in support of this idea: "The reason my life is so awful at these times is *because I am a terrible, wicked, failure of a person*"; "Everything that goes wrong in my life is *directly my fault*; I caused it by not doing things I should have done, or doing things I shouldn't have done" (p. 135. Italics mine). With respect to the dimensions discussed in this chapter, hyperagency appears to interact both with confidence and salience. On the one hand, a diminished sense of confidence may play a role in over-attributing guilt to oneself in the face of negative events (e.g. "It happened to me because I am bad person"). On the other hand, feeling overly responsible may contribute to assigning exaggerated importance to otherwise trivial events (e.g. "He must be mad at me for that remark I made yesterday"). Yet, in some cases hyperagency may also correlate with hyperconfidence: for example, in Reina's case a strong sense of confidence gives rise to augmented agency – e.g. the idea of being able to read other people's thoughts (see above, pp. 7-8).

Hypoagency. Disorders of hypoagency can be characterized as situations in which a person "loses grip" over her own thoughts or actions, thereby experiencing them as alien and beyond her control. One extreme example is the occurrence of auditory verbal hallucinations (AVH), often associated with schizophrenia but also common in a variety of other disorders – e.g. psychotic depression, manic-depressive disorder. People experiencing AVH often describe the voices as exhibiting a markedly "alien" character and as differing sharply from first-person inner speech (although there are exceptions, see Longden 2013). Generally speaking, some voices utter descriptive statements – e.g. "She is opening the door" – while others issue commands or evaluations that are often negative and threatening both in tone and content – e.g. "You should get that knife and cut yourself"; "You are a failure." In recent years, AVH have received great attention in the philosophy of psychiatry and have featured in a number of debates about agency (see Proust 2006), inner speech (see Langland-Hassan 2008) and ownership (see Maiese 2015). Within this debate, most authors distinguish between a sense of *agency* or authorship (i.e. X is caused by me, I am the author of X) and a sense of *ownership* or mineness (i.e. X is mine, X is part of my experience). In cases of AVH authorship and ownership importantly come apart, as patients experience voices as alien – thereby denying authorship – but still as occurring within their bodily or mental boundaries in some significant sense – thereby preserving ownership.[11]

In her vivid first-person report about the experience of "voice-hearing", Eleanor Longden (2013) recalls the first appearance of this phenomenon during her early college years. She describes her younger self as struggling with severe anxiety and worries about the future, but also as exhibiting a strong tendency towards suppressing her feelings. The first voice makes its appearance one evening while Eleanor is going home after a class: she characterizes it as neutral, similar to her own voice but narrating all her actions in third person, like a running commentary –

[11] I cannot explore the debate about agency and ownership in detail here, but see Proust (2013 – Chapter 12) for a recent overview of the main positions.

e.g. “She is leaving the room”; “she is opening the door”. In the following weeks voices grow in number and intensity, becoming more persistent and menacing: in particular, *they* start threatening Eleanor and make her comply with a series of bizarre tasks with the promise of “getting her old life back”. These tasks are experienced by Eleanor as some sort of “Labors of Hercules” over which she has absolutely no control, but that she nonetheless feels forced to carry to completion. She describes them as initially quite small (e.g. pull out a few strands of hair) but then as progressively more extreme (e.g. harm yourself) or violating social norms (e.g. pour a glass of water on the head of the instructor during a lecture). Notably, she experiences overwhelming feelings of powerlessness because she lacks the resources to exercise any form of control over the voices. Her agency appears so compromised that at one point she attempts suicide by trying to drill a hole in her head in order to get the voices out.

The second part of Longden’s account is devoted to her process of recovery, which begins once she gets in touch with the UK-based *Intervoice* movement, founded in 1988 by psychiatrists Marius Romme and Sandra Escher. The tenet of this therapeutic movement consists in claiming that voices should be treated as experiences rather than symptoms, and that the content of the voices often provides important insights on the person’s life story and personality. The primary goal of this approach is not to get rid of the voices *per se*, but to accept them while learning a series of coping strategies focused on “taking power back” from them. The turning point towards recovery therefore consists in realizing that voices may be appropriate responses to traumatic life experiences – e.g. childhood abuse – but also ways to get in touch with one’s repressed emotions. For Eleanor this was clearly the case: during therapy she realizes that many of the voices – especially the more aggressive ones – were mirroring her hidden emotions: “Whenever I repressed anger (and that happened very often) the voice sounded frustrated.” Another patient describes this phenomenon in greater detail:

"When I went out, the voices sometimes said to me: 'She is going out again' and that felt like criticism. It was none of their business because I had no work so I could go out. But then I realized that I had the same thought myself. There were more examples: When the voices said: 'See how awful she looks', it happened on days when I felt myself pretty awful. But they always made such exaggerated statements. By exploring this I started to realize that in a certain way the voices expressed my own thoughts. It is rather strange, but they are your own thoughts about an emotion" (reported by Romme & Morris 2013, p. 263-264).

The treatment proposed by Romme and Escher appears particularly interesting for our purposes because it focuses on various coping strategies to regain control over the voices (see Romme & Escher 1993). Indeed, it could be seen as a way to enhance agency in people that experience a significant diminution in their power of controlling their mental events. In a more recent paper, Romme & Morris (2013) describe recovery as a process of progressively gaining control over the voices by creating a dialogue with them, while at the same time setting boundaries and avoiding being overwhelmed. One of the patients talks about the process this way: "I was already able to talk back to my voices with my thoughts, but I learnt to make a specific time of day, the evening, when I would focus, and simply tell the voices 'later' if they came at another time" (p. 263). With respect to the dimensions discussed here, Romme and Escher's approach appears to counter hypoagency by strengthening familiarity. The more the patient learns to incorporate the voices in her experience and to treat them as legitimate (or at least revealing) aspects of her personality, the more agency over them is restored.

In Chapter Four I discuss a number of intermediate cases connected with the dimension of agency. The perfect balance case would be characterized by an appropriate assessment of personal responsibility, and connected with abilities such as *self-attribution* and *self-ascription*. These capacities allow a person to correctly determine the scope of her thoughts and actions, and

also to reliably distinguish between self-generated and other-generated input. For example, a balanced person would be able to tell the difference between being upset because of something he did or because of something his significant other did. On the low end of the spectrum we then find common phenomena such as distraction, dreaming and daydreaming, where the sense of agency over one's thoughts and feelings appears considerably diminished. Indeed, when we are distracted (or dreaming) we are not sure about the origin of our thoughts and we do not feel able to exercise a sufficient control over them. Local imbalances on the high end would be instances where people tend to over-attribute responsibility to themselves or to feel guilty for something they have not done: false confessions are a particularly striking example (see Gudjonsson 2003). Finally, extreme – albeit non-pathological – cases involving a loss of agency would be meditation practices and mystical experiences (Thompson 2014), as well as hypnosis (Dienes 2012).

§5. Connections among the Disorders

While discussing the case studies, it has become apparent that some disorders are connected with one another and that they interact in a number of ways. In some cases the lack of confidence appears to foster a lack of salience (e.g. depression) whereas in other situations a loss of agency can contribute to a loss of familiarity (e.g. AVH). Moreover, familiarity and salience influence each other: when a person becomes particularly salient, some details of the environment may show up as unduly familiar (Fregoli). Conversely, when a known person appears disturbingly unfamiliar we may look for specific aspects of the situation to account for the mismatch (Capgras). Finally, agency importantly interacts both with confidence and salience: a low sense of confidence may result in one over-attributing responsibility to oneself and in assigning exaggerated importance to trivial events (e.g. pathological guilt).

Setting specific examples aside, I now draw attention to some general connections among the kinds of disorders discussed above. Disorders of familiarity appear related to disorders of competence and confidence: while an enhanced sense of competence makes things more graspable and less unfamiliar, an increased or regained familiarity positively affects one's confidence. Frijda makes a similar point in his study on emotions: "Fear of the unfamiliar decreases or disappears through general increased competence or self-confidence" (1986, p. 273). Disorders of agency and familiarity are also closely related: in particular, it looks like a diminished sense of agency would fuel a lack of familiarity. Indeed, the patient affected by the alien hand syndrome surely exhibits a disturbance in his sense of agency (i.e. he cannot exercise control on what his hand does) but familiarity may be compromised as well (i.e. his hand appears utterly alien). Another connection that can be drawn is the one between disorders of agency and competence: barring cases of hyperagency, it looks like an enhanced sense of control over one's actions and thoughts would contribute to strengthening one's sense of competence in a given area. This seems to be the case for people who achieve mastery over a discipline or acquire a set of practical skills – e.g. the more I can control the car's movements and make decisions about what to do in a wide range of circumstances, the more I feel competent and confident about my driving abilities. This preliminary discussion hints at the fact that disorders of competence, familiarity and agency may be significantly interconnected, and that positive correlations may be found among them. For example: when familiarity increases, competence and agency are enhanced; when agency increases, competence tends to follow the same path.

Disorders of salience seem rather to indicate a more pervasive disruption of the relationship between individual and environment: in this sense, they can be regarded as "running through" all the other cases. To be more specific: assessing one's level of competence, regarding an object as more or less familiar, or seeing an action as beyond one's control already imply a

minimal connection between goals or intentions on the one hand and possibilities for action offered by the environment on the other. For example, in order to assess one's competence a person has to see something as a task worth undertaking or (on the contrary) as a waste of time, or as an enterprise exceeding her abilities. All things considered, a disorder of salience should be regarded as more basic with respect to the others discussed above. Indeed, whereas competence, familiarity and agency concern more or less circumscribed areas of experience, salience characterizes the more generalized sense of relating to the environment in a *meaningful* way. Salience would therefore be what allows a person to regard a particular event as worthy of attention (relevance), as requiring a more or less immediate response (urgency) or as particularly attractive as opposed to aversive (valence) – see also Frijda 1986, pp. 205-208.

Frijda (1986) describes a similar phenomenon when he talks about events and objects impinging or touching on the individual's *concerns* (p. 277). On his view, stimuli coming from the environment are "appraised" or "coded" in a certain way based on a variety of factors: the overall current situation, past experiences, the subject's abilities of coping, and so on (see also Lazarus 1991). This is why – for example – small gestures of forgetfulness such as showing up late at a meeting or putting too much sugar in one's partner coffee leave some people indifferent and others infuriated. The same actions may be appraised as minor infractions or as revealing signs of neglect. As Frijda puts it: "A given event is more serious to one person than to another because it relates to more of his concerns or to more focal concerns; it has more farther reaching meaning" (p. 291). In other words, concerns play an important role in shaping people's *sensitivity* which in turn is reflected in different appraisals of similar situations. As a result: "Some people have a stronger tendency to perceive events as threatening, or as infringements, or as damaging generally than have other people or than they themselves have at other moments" (Frijda 1986, p. 300). This idea of sensitivity (or vulnerability) becomes thus particularly important in explaining

pathology, because we can think about pathological constellations as combinations of "weak points" along several dimensions – e.g. low sense of familiarity, leading to low sense of confidence and consequently to heightened salience of specific details that corroborate the initial sense of unfamiliarity (Capgras).

I explore the affective nature of the dimensions more thoroughly in the next chapter. As I mention above, dimensions are ways of modulating one's relationship with the environment that comprise a variety of affectively-laden aspects: e.g. motivations, concerns, commitments, goals and preferences. In the discussion of disorders – §1 to §4 – affective notions have played a particularly important role. For example, disorders of salience qualify as situations where a person's concerns are dramatically reduced in intensity and in number (e.g. anhedonia) or as cases where too many things appear puzzling and interesting (e.g. interrogative attitude). Disorders of confidence are also importantly related to affect because they indicate a disruption in the person's ability to assess her degree of efficacy or fluency. Finding a task easy to process and being able to carry it out give rise to affective phenomena with a positive valence – e.g. arousal, excitement – whereas encountering a challenge that we perceive as beyond our grasp make us distressed and frustrated (see Russell 1980). In this sense, disorders of confidence include both feelings of omnipotence disconnected from one's real abilities (e.g. grandiosity) as well as systematic underestimation of one's potential despite clear evidence of the contrary (e.g. self-doubt). Disorders of familiarity appear connected with affect in two main ways: first, both in Fregoli and in Capgras the misidentifications involve people who are affectively significant for the individual – e.g. husband, daughter. Second, in some cases particular affective experiences in the patient's everyday life seem to ground the misidentification. For example, the patient whose daughter has moved away thinks that *she* has been replaced by an impostor, while the patient obsessed with theater feels persecuted *by her favorite actresses*. Finally, disorders of agency are

related to affect more indirectly: notably, the feeling of control that a subject has over a situation heavily depends on the physiological and psychological resources available. On the one hand, when one feels overwhelmed or lacking energy, the sense of agency appears importantly diminished (e.g. AVH). On the other hand, patients affected by pathological guilt appear to invest a disproportionate amount of energy in thinking about small things that they might have done wrong – e.g. Emmy von N. In the next chapter I explore the nature of these appraisals in more detail, and I argue that their aggregation gives rise to moods.

One final remark: some of the disruptions discussed in this chapter – such as Capgras or pathological guilt – do not dovetail neatly with the current classifications of mental disorders, such as the ones put forward in the DSM-5 or in the ICD-10. Indeed, they may be experienced by patients diagnosed with very different disorders such as depression, schizophrenia or manic-depressive disorder. Yet, this does not represent a problem for the account. In fact, the local disruptions discussed in the case studies describe specific experiences (or symptoms) whereas full-blown mental disorders are characterized as constellations (i.e. set of points along different dimensions). The same local disruption (e.g. hypersalience) may thus feature in different constellations (e.g. schizophrenia and obsessive-compulsive disorder). Similarly, in the diagnostic manuals the same symptoms are listed as belonging to different disorders – e.g. delusions in schizophrenia and psychotic depression.

Conclusion

In this chapter I outlined a model of mental disorders where healthy and pathological cases are represented as collections of points along four dimensions of functioning – i.e. salience, confidence, familiarity and agency. These dimensions represent different ways in which the relationship between individual and environment may be modulated (i.e. appraisals). On this

view, mental disorders are thus seen as disruptions of appraisal relationships and as ways in which one's experience of the world can be altered.

The introduction of this model has some important implications for the defense of the Continuity Thesis (CT) in philosophy of psychiatry. Indeed, it contributes to making a dimensional account of mental disorders both more precise and plausible. With respect to the checklist approach endorsed by the DSM and other classificatory manuals, the model outlined here has two main advantages. First, it relates disorders to abilities that we all have (e.g. appraisal of salience) and that come to be disrupted under particular circumstances. Second, it connects more closely with the individual's psychological makeup and life experience because appraisals are rarely correct or incorrect *tout court*, but always working in relation to one's goals, concerns, preferences, etc. In this sense, a person who has hallucinations of his dead spouse (a case I discuss in Chapter Four) may not count as disordered because the symptom can be seen as adaptive and conducive to her well-being. The analysis of intermediate cases becomes therefore crucial to better understand the transition from vulnerability (or local imbalance) to pathology. In Chapter Four I focus on the notion of *at-risk individual* and on a number of intermediate cases to offer another argument against the categorical view. Taken together, these two chapters thus strengthen CT by making the dimensional view more plausible and the categorical view harder to defend.

Valentina Petrolini

The Nature of the Dimensions and Their Relation to Affect

"We experience our emotions first and foremost as the world that we inhabit, rather than as internal constituents of our psychology. And the world is not experienced as something set apart from our projects, from our sense of who we are and what we strive for. The world, as experienced, embodies what we take to be significant, what we care about, what we identify with and what drives us." (Sartre 1939/1962, xv)

Abstract

In Chapter Two I identified four kinds of disorders that encompass many distinguishing features of psychiatric conditions as currently classified. I then proposed to conceive of these disorders as disruptions of four underlying dimensions of functioning: *salience*, *confidence*, *familiarity* and *agency*. Here I take a deeper look at the nature of these dimensions and I characterize them as affective ways in which we modulate our relationship with the environment. Despite their affective nature, these dimensions do not neatly map onto affective processes as we usually describe them. They cannot be identified with emotions or feelings; they are closely related to moods, but there is no one-to-one correspondence between them: in other words, a subject cannot be characterized as being in a "salient" or "confident" mood. However, it is the aggregation of these dimensions that gives rise to moods as we normally encounter them. This means that a certain combination of salience, confidence, familiarity, and agency is what makes up affective states such as being elated, depressed, etc. In this chapter I disentangle the complex relation between dimensions and affect by clarifying the connection between salience, confidence, familiarity, and agency on the one hand, and feelings, emotions, and moods on the other hand.

The chapter is divided into four sections. In **§1** I connect the dimensions with affect by showing how their aggregation gives rise to *moods*. In doing so, I draw on existing models of affect that embrace a dimensional approach (Russell 1980; Russell & Barrett 1999; Thayer 1996)

and I present my proposal as a similar but finer-grained attempt to characterize moods. In **§2** I explain why the aggregation of the dimensions gives rise to moods as opposed to feelings or emotions. Specifically, I argue that – unlike feelings and emotions – moods have an important unconscious component that better accounts for the world-directed character of the dimensions. Roughly speaking: the dimensions (aggregating in moods) have the main function of informing us about "how the world appears to us" at a given moment. In **§3** I discuss some important differences between my view and the one recently put forward by Ratcliffe (2008 & 2005). Like Ratcliffe's *existential feelings*, the dimensions I describe here can be seen as "background orientations" or "ways of being in the world" (Ratcliffe 2008, p. 41). However, I resist the idea of characterizing them as feelings and I rather take them to be ways in which we structure self-world relations (see Berninger 2016). In **§4** I briefly apply the model to psychiatry and I suggest a few ways in which moods can be seen as pathological.

§1. Dimensions and Moods

There is a clear sense in which our current affective state influences the way in which the world in general, and things or events in particular, appear to us. Depending on how we feel and how much energy we have, different aspects of our environment show up as more or less interesting, challenging, familiar and controllable. This connection between our mood and the way in which reality appears to us is integral to our everyday experience: when we are irritable we are more likely to be offended by a neutral remark coming from a colleague, when we are in love we see everything through rose-colored glasses, etc. Despite its pervasiveness, the relationship between affect and world-experience has been overlooked by philosophers until recently. By focusing almost exclusively on emotions, philosophers (especially those working within the analytic tradition) have neglected the complex and subtle ways in which moods

influence our lives and thoughts (see Wong 2016 for a similar point). By contrast, in the past thirty years psychologists and neuroscientists have proposed various models to explain the origin of moods. Some of these models are *dimensional* in nature and they see moods as arising from the intersection of two axes: a certain degree of arousal combined with a certain valence (Russell 1980; Russell & Barrett 1999); or a certain degree of energy combined with a certain degree of tension (Thayer 1996). In what follows I briefly review two of these models and I show how the dimensions introduced in Chapter Two provide a similar but finer-grained characterization of moods.

§1.1 Dimensional Models of Affect

Arousal-valence. In a series of interrelated papers, Russell (1980; 2003), and Russell & Barrett (1999) introduce and flesh out the notion of core affect. Core affect can be characterized as the encounter between two axes on a Cartesian plane: *valence* refers to a qualitative feeling of pleasure or displeasure, while *arousal* refers to a quantitative feeling of energy. Examples of positive and negative valence are fairly obvious: one may feel accomplished after meeting a deadline at work, or disappointed because things did not turn out quite as she expected. Examples of low and high arousal may include feeling sluggish on a Monday morning, or feeling psyched before going to a concert with friends. Core affect thus includes states of positive valence and high arousal (excitement), negative valence and high arousal (anxiety), positive valence and low arousal (calm) and negative valence and negative arousal (fatigue). On this view, core affect is always present: in any given moment every person lies at exactly one point of the affective circumplex (see Fig. 2 below). Despite being always present, core affect has multiple causes that are neither easily detectable nor directly controllable: for example, one may feel fatigued after a bad night's sleep (low arousal + negative valence), or chipper thanks to a sunny day (high arousal

+ positive valence). In this sense, core affect lies in the background of consciousness and becomes salient only under particular circumstances, such as when it changes abruptly or lasts for an unduly long time (see Russell 2003, p. 148). There is a special connection between the notion of core affect and the one normally referred to as mood: one might simply say that a mood is nothing but core affect prolonged over time. As Russell & Barrett put it: "Core affect is assessed by asking how one is feeling right now. When extended over moderate lengths of time, core affect becomes a mood and is assessed by asking how one generally felt during that period" (1999, p. 815). Typically, moods lack a specific object and tend to refer to the world as a whole. This has been described as the global effect of mood: "When one is in an anxious mood, one is not afraid of any particular object; instead, one is afraid of whatever one has encountered or may encounter. That is, everything appears to be a possible threat to one's well being. When one is in a depressed mood, one is not sad about any particular tragic event; instead, the whole world looks gray and uninteresting" (Wong 2016, p. 180).

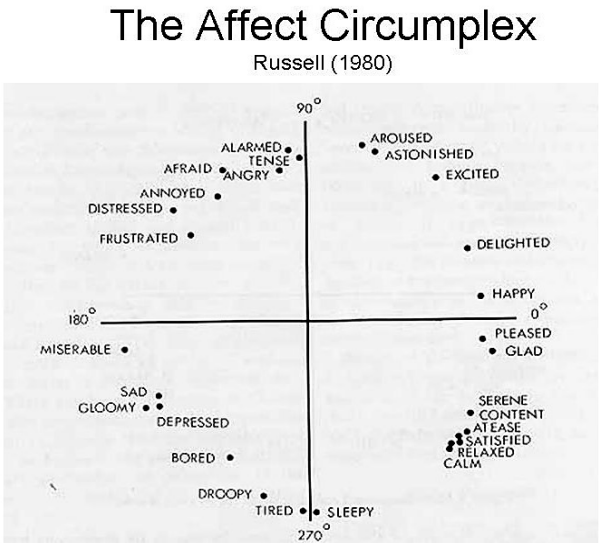

Fig. 2: Russell's affective circumplex (1980)

Energy-Tension. The model of affect proposed by Thayer (1996) bears many similarities to the arousal-valence approach just discussed. On his view, moods stand "at the core of our being" and amplify meaning attribution by enhancing or reducing pleasure in our lives (p. 4). They do so by working as filters through which our activities are evaluated: for example, if we feel tired we are more likely to see going to a party as a nuisance as opposed to something to be excited about. Compared to Russell & Barrett, Thayer focuses more on the idea of moods being indexes of physiological and psychological functioning. He employs the analogy of a thermometer that informs us about our energy levels (p. 20): a core function of mood would thus be to engage us in the right task with the right amount of energy (Wong 2016, p. 184). A great part of Thayer's work consists in exploring how biological processes such as food ingestion, exercise and circadian rhythm have an impact on mood. The idea that moods would have a core physiological component also explains why people are only "dimly aware" of their functioning: most of the time, they operate automatically and outside of conscious awareness (Thayer 1996, pp. 111 & 215). This aspect is important because it suggests that the effects of mood can be regulated through a process of "cognitive override" (pp. 39-40), which consists in consciously acting against what a mood suggests. For example, according to Thayer, we should resist the tendency to make plans before falling asleep at night, when we are more likely to be worn-out, fatigued and slightly depressed – i.e. more tense and tired.

The model of affect proposed by Thayer is also dimensional and identifies moods as points at the intersection of two axes (see Fig. 3 below). The first dimension, "energy-tiredness", resembles Russell & Barrett's arousal and indicates the quantity of physical or mental energy available to the subject. For example, going to the gym seems easier after a good night's sleep and a healthy breakfast than after having pulled an all-nighter to meet a deadline at work. This is because in the former scenario we would feel physically and mentally more energetic and less

tired. The second dimension, "calm-tension", indicates the relationship between our energy level and the one required to complete a given task. Typically, we experience calm when our energy level meets or exceeds the one required by the task at hand: for example, if we are studying for an exam and we have planned accordingly, we have enough time and resources to satisfactorily go through all the material. On the contrary, we experience tension when our energy appears insufficient in relation to the action that we have to perform: if we have procrastinated until the night before the exam, we now have to complete a huge amount of work in a short time and we probably lack the energy to do so. Thayer describes four basic moods arising from the encounter of these two dimensions: calm-energetic, calm-tired, tense-energetic, and tense-tired (pp. 18-20). Calm-energetic would be an ideal mood that combines a sense of focus and confidence with the perception that the task at hand is manageable and under our control. A tense-energetic mood is also perceived as positive in many situations and corresponds to the sense of being "in the zone", therefore vigorous but slightly anxious. Calm-tired rather implies a sense of relaxation united with a certain satisfaction, such as when we go to sleep after a long but productive day. Finally, tense-tired characterizes the prototypical "bad mood", in which our energies are depleted but our levels of tension and anxiety are still high.[12]

[12] It is worth noting that the two dimensions introduced by Thayer are not – in fact – completely independent from one another. Indeed, an increase in energy often causes a decrease in tension (e.g. eating a snack also has a calming effect) while a decrease in tension may cause an increase in energy (e.g. when we feel reinvigorated after a walk outside). See Wong 2016 for a more detailed discussion of this point.

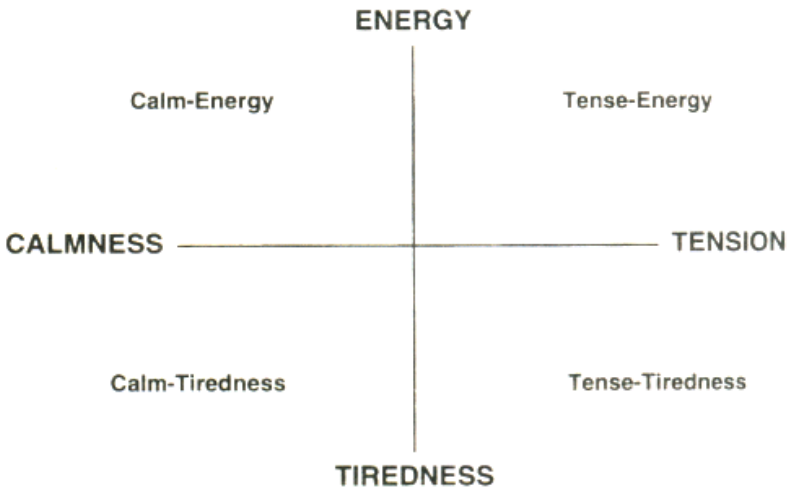

Fig. 3: Thayer's model of moods (1996)

The models of affect proposed by Russell (2003; 1980) and Thayer (1996) are particularly useful for our purposes. First, they characterize moods as dynamic configurations that unfold over time and change according to the individual's resources and the environment's demands. Second, they regard moods as complex processes that arise from the encounter of more basic dimensions (i.e. arousal-valence; energy-tension). In what follows I propose a dimensional model where the dimensions of salience, confidence, familiarity, and agency aggregate to give rise to different moods. More so than the alternatives discussed here, this four-dimensional model helps us to better understand how moods shift and give rise to new self-world relations. Indeed, I suggest that a combination of salience, confidence, familiarity and agency is primarily responsible for our seeing things and events in one way or another. In this sense, while the models discussed above highlight the subjective aspect of moods (i.e. how I feel), the account I develop below uncovers their world-related aspect (i.e. how the world appears to me).

§1.2 Moods: "How I Feel" and "How the World Looks"

Yesterday I woke up expecting my morning to go as usual: I had a work plan in place for the day and a couple of meetings lined up; I was feeling quite energetic after a good night's sleep and a nice breakfast. While I was getting ready to go to the library I received a call from my sister, who sounded very distressed and wanted to discuss a complicated family issue that had many ramifications for me and my loved ones. Needless to say, after hanging up with her my mood had changed significantly. How would I describe this change more precisely? In Russell & Barrett's terms, I moved along the axes of arousal and valence by experiencing an increased activation as well as a markedly negative sensation. Figuratively speaking, I was now lying on the upper left corner of the affective circumplex, an area characterized by high arousal and negative valence (see Fig. 1). The distinction between energy and tension (Thayer 1996) also helps me to pin down the mood change that I experienced. Indeed, I suddenly transitioned from a state of calm-energy to one of tension-tiredness: from being quietly ready to meet the challenges of my day, I became overwhelmed by a problem that I did not quite know how to solve. Despite feeling more anxious, I also felt that my energy level was declining as if it had been completely expended during the phone call. As Thayer would put it, I felt *both* more tense and more tired, because I perceived a danger along with a scarcity of resources to deal with it (p. 82).

The notions of arousal-valence and energy-tension aptly describe the mood change I experienced: specifically, they contribute to explain how an external event influenced the way I felt. This is undoubtedly a key aspect of mood: indeed, we normally connect "being in a mood" with a pervasive way of feeling – e.g. more anxious, more fatigued, less enthusiastic. Yet, there is another – equally essential – aspect of mood that these models fail to capture: being in a certain affective state importantly modifies the way in which the world appears to us. In fact, being in a mood simultaneously changes "how we feel" and "how the world looks": after my sister's call, I

did not simply *feel* different, but I started *relating* to my environment in a different way. My attention and thoughts were now overwhelmingly directed toward the problem that I had discussed on the phone, to the point that other things became uninteresting or at least less important. Probably due to the decreased energy level, the tasks I had to complete for the day looked now more daunting and intrinsically difficult. When I was on the bus on the way to work I also noticed that the city looked less familiar and comforting, as if to further underline my long distance from my loved ones. My feeling of impotence translated into the impression that things were beyond my control and that there was little I could do to manage them successfully.

This world-directed aspect of mood fails to be captured by models that employ subjective notions like arousal-valence or energy-tension. By contrast, the dimensions introduced in Chapter Two – i.e. salience, confidence, familiarity, and agency – may be more successful in explaining what happens when a mood comes about or changes. By describing different ways in which we relate to our environment, these dimensions describe how the world appears to us when we are "in the grip" of different moods. Going back to the example, there is a sense in which *salience* decreases after talking to my sister because I suddenly become uninterested in things that were important before – e.g. the amount of work I have to complete for the day. At the same time, salience increases locally because I become more focused on aspects of the environment that I would have otherwise overlooked: e.g. I now check my phone several times a day to see if my family members need help or comfort. *Confidence* instead decreases both with respect to the problem at hand – which I do not know how to handle – as well as more generally, because I also become slightly more insecure and jittery when I am at work or hanging out with friends. *Familiarity* and *agency* also decrease significantly: a wider range of events appear to be outside of my control and my relationship with the city itself changes to become a painful reminder of the fact that I am far away from my family. The four dimensions thus track how the relationship with

my environment changes along with my affective state. These two aspects of mood can be seen as two sides of the same coin. On the one hand, the way we feel in a given moment affects how the world appears to us – e.g. when we are fatigued, tasks themselves appear more difficult. On the other hand, the way in which the world appears to us also influences our feelings in multiple ways – e.g. on a sunny day we might find ourselves chipper, and then experience a sudden mood change when it starts raining. Talking about moods as arising from the aggregation of the four dimensions thus uncovers a key (and often overlooked) aspect of moods themselves. Being in a mood does not (only) mean *feeling* in a certain way, but also *relating* to the world differently: things and events may acquire or lose importance, as well as appear more or less difficult, known, and controllable. For this reason, models that emphasize subjective feelings (such as arousal or energy level) cannot fully account for the range of experiential changes that moods bring about. In particular, they do not explain why things appear different – e.g. more difficult – when we are in the grip of a mood. By contrast, the dimensions of salience, confidence, familiarity, and agency focus on self-world relations and are therefore better equipped to describe the changes that come about when a mood arises or shifts.

The view proposed here has two further advantages. First, introducing the four dimensions makes it easier to account for the nuanced differences among moods as well as for the dramatic affective changes typical of some mental disorders. In fact, the models proposed by Russell & Barrett and Thayer succeed in distinguishing between different moods only to a certain extent. A popular objection against Russell's model is that affective states that are phenomenologically very different (such as anger and terror) occupy the same space on the circumplex – i.e. high arousal and negative valence.[13] Similarly, Thayer repeatedly stresses the difficulty of distinguishing between high-tension and high-energy states (1996, p. 89). The

[13] See Marraffa & Viola 2017; Ellis & Faw 2012 for recent versions of this objection.

introduction of further dimensions allows us to describe moods more accurately and to distinguish among them with a greater level of precision. Depression can be taken as a case in point. Besides being a state of low arousal with negative valence (Russell & Barrett 1999, p. 161) and a tense-tired state (Thayer 1996, p. 224), a depressed mood presents itself as a certain combination of diminished salience, low confidence and low agency. When in the grip of depression, one sees things as uninteresting (low salience), tasks as more difficult to complete (low confidence) and events as falling out of control (low agency). Similarly, in non-pathological cases, the world looks different when we are in an irritable mood as opposed to a terrified one. Although both states are high in arousal and negative in valence, dimensions such as salience and confidence help us to disambiguate between the two. An angry mood usually implies a narrowing down of things and events that draw our attention, with our mind fixating on the perceived injustice or offense (low salience). By contrast, terror amplifies the range of things or events that are important to us because we tend to see everything as potentially harmful or threatening (high salience).[14] Moreover, anger is often felt as empowering and connected with action or fighting behavior (high confidence), whereas fear tends to correlate with avoidance and submissiveness (low confidence). In terms of agency, the two states can be quite similar because they both imply a loss of control over one's actions or the situation more generally. The model I propose thus has the advantage of distinguishing more precisely among affective states that are conflated in other dimensional approaches.

Finally, this account also allows us to correct a common misconception, namely the idea that moods "do not concern *what* we represent, but *how* we represent it" (Wong 2016, p. 180).

[14] In this example, as in many others, talking about low and high salience *simpliciter* does not seem sufficiently precise. In fact, what happens here is that while salience decreases *globally* (i.e. we become fixated on a few things), it markedly increases *locally* (i.e. the things we fixate on acquire an exaggerated importance). Although I cannot delve deeper into this matter here, I am planning on expanding the account as to include these finer-grained considerations. I would like to thank Johannes Brandl for bringing up this point.

This surely seems to be the case in many situations, where moods "color" our experience without altering it dramatically: again, when we are tired we may notice that some activities appear more difficult and less enticing. However, there are cases in which what we represent and how we represent it may be hard to disentangle. Take a patient affected by the Capgras delusion, who comes to believe that her spouse has been replaced by an impostor. In this case, a pervasive change in *how* something is perceived brings about a change in *what* is perceived. In other words, the fact that the spouse is seen as strange and unfamiliar causes the patient to see him as someone else – i.e. an impostor.[15] This is precisely what moods can do, also in non-pathological circumstances: if I go back to a place where I used to live but that I left a long time ago, I might realize that I see it differently. This change in how I perceive my old house easily translates in a change in what I perceive: for instance, I now see my old home as a building among many others as opposed to a place to which I am attached.[16]

§1.3 Dimensions and Concern

Besides focusing on the world-directed aspect of moods, there is another important way in which the dimensions relate to affect. Indeed, they highlight different ways in which things and events matter to us and motivate us to act. In other words, they uncover patterns of *concern*. The notion of concern has been widely discussed in studies on affect. In brief, it is the idea that affective states do not arise in a vacuum, but are connected with the individual's overall situation, including her physiological state, previous experiences, current goals, preferences and tastes.

[15] One may rightfully object that in both cases being in a mood leads to forming judgments that one would not otherwise have made: e.g. that working on a new paper is a bad idea (in the tired case); or that this person is not actually our spouse (in the Capgras case). Yet, the judgments formulated here are different in one important respect: in the former case the mood simply leads to a different *evaluation* of the situation (i.e. I should not do this today), while in the latter it brings about a *different* description of reality (i.e. this person is not the one I know). I would like to thank Peter Langland-Hassan for allowing me to clarify this point.

[16] The example is adapted from Ratcliffe 2008, p. 151.

Appraisal theorists of emotion see the notion of concern as central: for an event or object to be noticed or seen as relevant, it has to "touch upon" or "impinge on" an individual's concerns. Frijda expresses the point this way: "Being alone is not the same as being alone after one's partner died; no food is different from no food when food was expected; threat when there is a way to escape is different from when there isn't" (Frijda 1986, p. 268).[17] More recently, constructionist theorists of emotion such as Barrett & Bar (2009) have also stressed the connection between affect and concern. In particular, they underscore the importance of previous experiences that allow us to categorize incoming data in an affective manner: "People see with feeling […] They do not wait to evaluate an object for its personal significance until after they know what the object is. Rather, an affective reaction is one component of the prediction that helps the person see the object in the first place" (pp. 1325 & 1331).

The four dimensions introduced in Chapter Two and discussed in **§1.2** contribute to clarify the notion of concern by uncovering different ways in which things or events matter to us. Generally speaking, the dimensions modulate our relationship with the environment by helping us pursue our goals in the face of various demands and challenges. *Salience* allows us to pursue our goals by parsing the environment into things that are relevant and others that can be (at least momentarily) disregarded. This process promotes a flexible interaction between individual and environment, where actions are prioritized or postponed according to what is taken to be relevant in a given moment. Salience may thus be experienced as the sense of something being important, interesting or urgent: if we are in need of sleep we see our couch as particularly enticing; if we are recovering alcoholics, the bottle of gin on the shelf acquires special relevance. Yet, experiencing low or high levels of salience has less to do with the particular objects one focuses

[17] For a similar view that includes both physical and intellectual actions, see Stocker 2004. "Without certain forms of care and concern we would not 'parse' the world at all, much less as we do; without such care and concern, nothing would be salient, intellectually or otherwise" (p. 135).

on and more to do with the scope of one's attention. When salience is (too) low, subjects tend to be focused on few things that end up occupying a great part of their resources: this is what happens with obsessions, delusional fixation and depression. When salience is (too) high, subjects are distracted by too many stimuli that are considered equally urgent or relevant. Examples of this are manic episodes or the pathological doubt typical of some cases of schizophrenia, where ordinary aspects of reality are seen as deeply puzzling and in need of an explanation (see Minkowski 1923).[18] *Confidence* helps us assess what we can manage in the current situation, by providing us with a rough indication of our chances to succeed in a certain task. This is another way in which individuals modulate their interaction with the environment: they assess what they are good at and make decisions accordingly. This process applies to a wide range of situations, from motor actions – e.g. deciding whether to jump across a creek – to more intellectual scenarios – e.g. taking up a new challenge at work because we perceive ourselves as competent and up to the task. *Familiarity* guides us in evaluating what is comfortable or known with respect to our environment. Like the other dimensions, familiarity impacts behavior to a significant extent: we move around differently in our home and when we are hosted by someone; we do not interact in the same way with our best friend and with someone we just met at the bar. Finally, *agency* allows us to assess what is controllable and what we are responsible for, thereby facilitating processes such as self-ascription. This form of interaction with the environment thus helps us distinguish between actions and events that are caused by us and those that are not. The sense of agency may apply to simple movements and words – e.g. we perceive ourselves as having lifted a box, or as having talked to our spouse – as well as to more complex ones – e.g. we are trying to recall what happened during an accident to figure out who is responsible between us and the other driver.

[18] For an extended discussion of this case, see Chapter Two, §1.1.

Although all these assessments may seem "cold" or cognitive, they are in fact closely related to affective states and to moods in particular. As I mention above, one of the core functions of mood is to engage us in the right task by matching the level of energy we possess with what the environment demands (see Wong 2016). The dimensions contribute to this matching process in a special way, by helping us assess whether our energy level is sufficient to deal with the current situation and by guiding our decisions accordingly. Salience helps us flag meaningful things and events, confidence and agency help us identify those actions that we can successfully perform or control, while familiarity highlights the situations that make us feel comfortable. As I stress above, a certain configuration of the dimensions (similar to a certain combination of arousal-valence and energy-tension) indicates a particular mood.

§2. Feelings, Emotions or Neither?

In the previous section I argued that the four dimensions are affective in nature and I showed that their aggregation gives rise to moods. In what follows I want to resist the idea that salience, confidence, familiarity and agency should be rather characterized as emotions or feelings. In a recent paper, Carruthers (2016) discusses a variety of phenomena connected to *salience* – i.e. surprise, interest and curiosity – and qualifies them as "epistemic emotions" (pp. 9-12). Similarly, Proust (2013 & 2009) argues that processes related to *agency* should be characterized in terms of "noetic or metacognitive feelings". Following the psychological literature on the topic, Proust also refers to processes connected to *confidence* – e.g. the tip-of-the-tongue phenomenon – in terms of "feelings of knowing" (see Dunlosky & Metcalfe 2009; Koriat 2000). In a similar spirit, Dub (2015) has proposed to characterize some psychiatric disorders – delusions in particular – as responses to "cognitive feelings" such as *unfamiliarity* (p. 22). Here I offer some reasons to reject these characterizations.

§.2.1 Unlike Feelings

Philosophers typically talk about feelings as conscious and qualitative phenomenological states exhibiting an important bodily component (see Sizer 2006). The extent to which the body is involved varies considerably: for example, our finger may feel itchy after a mosquito bite or we may feel nauseous all over because of food poisoning. In some cases, feelings are used to describe global sensations where the reference to the body becomes metaphorical. For example, we may say that we feel under pressure because of a looming deadline or that we feel like we are drowning in a relationship. Despite the obvious differences, all these examples share an important feature: they all qualify as cases of feeling in virtue of their being *conscious* and having a *specific phenomenology*. Indeed, it would be nonsensical to claim that we feel itchy or nauseous without realizing it, and it would be odd to maintain that we are unable to tell the difference between the two sensations. To put it in Jamesian terms: if we were to rob a feeling of its conscious component and its experiential qualities, nothing would be left.

Given this fairly standard characterization, I turn to the question of whether salience, confidence, familiarity and agency should be described as feelings. In ordinary language we sometimes talk as if they were: for example, Trump's victory in the recent elections *feels* salient to me as a foreign person working in the United States. I also *feel* confident about submitting my paper on time; my grandmother's house *feels* familiar because I spent a lot of time there as a child; I *feel* I am in control when I drive. However, this use of language might be misleading. In fact, in cases like the ones just mentioned the two defining features of feeling appear to be missing: there is no conscious awareness and no specific phenomenology associated with salience, confidence, familiarity, and agency. Let me tackle these two aspects in turn.

First, as opposed to itches, nausea and global sensations like (metaphorical) drowning, the dimensions appear to operate unconsciously in most situations. Due to their world-directed

component, they track how our relationship with the environment changes along with our affective state. This means that in normal circumstances we directly *experience* the world as being in a certain way (e.g. more threatening) without being consciously aware of doing so. For instance, the day after Trump was elected, I found myself approaching people and situations more warily before I could consciously realize I was doing so. The news implicitly affected my relationship with my surroundings and with others, and it took some time and self-reflection to perceive this as a distinct feeling caused by that event. It was more like a subtle and all-encompassing change that shifted my way of experiencing things: my foreign accent just seemed more pronounced, some people's way of looking at me appeared more hostile, some passing comments more aggressive. For a different example, take familiarity: when I move to a new city I perceive my surroundings as being deeply unfamiliar for some time. I do not know my way around, I keep getting lost and even everyday activities like grocery shopping appear challenging. Although this sense of unfamiliarity certainly influences my behavior – e.g. I might walk around more tentatively – it is not something I am explicitly aware of like a sudden itch or a wave of nausea. It is a way in which I directly experience the world and relate to it. Yet, unusual circumstances may prompt us to become consciously aware of the way in which the dimensions shape our experience. After a few weeks in the new town I might notice "after the fact" that my perception of the surroundings has changed: my neighborhood now looks welcoming and safer, I recognize the people living in my building, I am comfortable walking home at night, and so on. Acknowledging that an important change has taken place, I retrospectively become conscious of how things looked different and less familiar some time ago.[19] These examples suggest that moods directly affect our world experience, to the point that the way things and events appear to

[19] The unconscious character of the dimensions is not – strictly speaking – Freudian. According to Freud most of the unconscious material remains hidden *by definition*, whereas we are able to become aware of the dimensions' functioning (at least retrospectively). In this sense, the dimensions are more similar to the *preconscious* material that is in principle capable of entering consciousness (see Freud 1933, pp. 95-96).

us may significantly change when we enter a different mood. In this sense, I am proposing that something similar to cognitive penetration might apply to these affective states (i.e. "mood-penetrability of experience").[20]

Second, the dimensions differ from proper feelings because they do not possess a specific phenomenal character. Indeed, in virtue of their being inherently conscious, feelings such as itchiness or nausea present themselves as qualitatively distinct from one another. There may be some room for confusion among them, as when we are unable to tell precisely whether we are sick or just very tired. Yet, in most cases feelings can be interpreted quite easily and work as cues guiding our behavior: if I feel itchy in my finger for a mosquito bite, I scratch; if I feel itchy all over after having eaten raw fish, I head to the pharmacy to buy antihistaminic medication. Salience, confidence, familiarity and agency do not seem to work this way: they affect our interaction with the world without being properly felt. This lack of a specific phenomenological character is related to the unconscious, automatic and pre-reflective nature of these processes. As I point out above, the dimensions are ways in which our experience is structured: things, events or people directly appear as more or less interesting, difficult, familiar or controllable. Going back to depression, the feelings of fatigue and sluggishness are signals indicating that we have been already approaching the world with a particular outlook. For some time we have been regarding things and events as uninteresting (low salience), too difficult to pursue (low confidence) or control (low agency).[21] The aggregation of dimensions is thus in place before feelings can arise and it somehow sets the stage for their development. We would not be able to *feel* anything if we were not already *experiencing* the world in a certain way.

[20] See Banerjee *et al.* 2012; Bubl *et al*. 2010; Meier *et al*. 2007 for some empirical studies on how affective states influence perception and judgments. See also Firestone & Scholl 2016 for an extensive criticism.

[21] The fact that dimensions (and the resulting moods) operate unconsciously in most situations might also explain why mental disorders such as depression are so difficult to treat. Indeed, it is common for patients to reject treatment because they do not consciously realize that they are suffering. I would like to thank Marco Viola for this suggestion.

The idea that affective processes can be experienced before they are felt may seem outlandish. Yet, a growing number of empirical considerations support its plausibility. Pioneering work in this sense has been conducted by Damasio and his collaborators in the past twenty years (see Damasio 1996; Tranel & Damasio 1993). In a series of experiments on patients affected by neurological and memory impairments (e.g. prosopagnosia) these researchers show that unconscious affective states impact behavior without the subjects being aware of it (see also Zajonc 1980). A famous case is the one involving David, a patient affected by severe amnesia who lives in some sort of "perpetual present" because his memory span does not last more than 40 seconds. Despite his inability to recall distant and recent facts, David's behavior in the experiments shows that he has in fact preserved some implicit affective memory. In particular, when he is asked to choose among different people with whom he has interacted in the past, he reliably picks the ones who were friendly towards him as opposed to the ones that were hostile (see Tranel & Damasio 1993 for more details). This experiment persuasively suggests that affective abilities – such as the preference towards specific people – may be governed by mechanisms outside of conscious awareness. According to Damasio, this applies equally well to non-pathological cases: "We often realize quite suddenly, in a given situation, that we feel anxious or uncomfortable, pleased or relaxed, and it is apparent that the particular state of feeling we know then has not begun on the moment of knowing but rather sometime before" (1999, p. 36).[22] Although Damasio wants to hold on to the notion of feelings when describing these phenomena, he introduces a distinction similar to the one I draw above. On the one hand, there are background and unconscious affective processes that implicitly govern our behavior and may survive neurological damage. These are the processes that I call "moods" and that Damasio calls "unconscious feelings" (p. 53). On the other hand, there are more conscious and explicit affective

[22] See Tamietto *et al.* 2009 for more recent evidence in this direction.

processes that exhibit a peculiar phenomenological quality: Damasio refers to these as “feelings of feelings” (p. 285), whereas I label them “feelings” *simpliciter*.

§2.2 Unlike Emotions

In the previous section I motivate my resistance against the idea that salience, confidence, familiarity and agency should be characterized as feelings. Yet, some authors – such as Carruthers (2016) – describe similar processes as emotions. I believe this is equally misleading. Indeed, the moods resulting from the dimensions characterize ways of relating to the world that are unlike emotions in two important respects. First, while emotions usually arise in connection to a specific eliciting event, moods unfold over time without presenting a clear onset-peak-termination structure. Second, these longer-term states act as the backdrop against which specific emotions occur.

With respect to the first point, it is helpful to distinguish between affective states that are *phasic* as opposed to ones that are *tonic* (see Frijda 1986, pp. 41-43). Some affective states – such as emotions – are better described as *phasic* because they arise in response to a specific event and exhibit a clear temporal development. Specifically, it is possible to distinguish a phase in which the emotion kicks in (onset), a phase in which it reaches a maximum level of intensity (peak) and finally a phase in which it wanes away or subsides (termination). This seems to aptly describe what happens when I am angry at a colleague who I believe offended me. Right after my colleague’s snappy remark, I feel irritated; rage mounts as I realize that this was not his first offense and that I have not defended myself properly, and finally it vanishes later in the day when more important stuff distracts me. In this case there is little room for confusion about the event that has elicited the relevant emotion, as I am aware all along that my colleague’s remark caused the anger that I am experiencing. Other affective states – such as moods – behave differently:

they are *tonic* because they tend to be less circumscribed and always present in the background, often changing in intensity due to external and internal events. When I experience a certain mood – e.g. fatigue – it is difficult to pin down a specific eliciting event: it might have arisen for purely biological reasons (e.g. coming down with a cold) as well as for psychological or environmental ones (e.g. being under pressure at work). Moods thus appear intertwined with a complex set of factors including biological parameters (e.g. lack of sleep) as well as motives, goals, and interest (e.g. desire to quit one's job). Moreover, the moods that I experience evolve over time but do so in a less clear-cut way with respect to emotions like fear or anger. For example, I might realize that I feel quite energetic after working out in the morning only to experience a sudden fatigue as soon as I set foot in the office. In this sense, moods appear more difficult to interpret and control than emotions, and this becomes especially relevant in pathological cases. Indeed, a defining feature of moods (low ones in particular) seems to be one's inability to pinpoint what causes them and one's difficulty in taking action to snap out of them.

With respect to the second point, I argue that the affective states emerging from the aggregation of the dimensions constitute the backdrop against which specific emotions occur. When someone experiences a depressed mood characterized by low levels of salience, confidence and agency, even simple actions such as getting out of bed might appear too daunting or effortful. These general dispositions to act translate into an increased or decreased likelihood that specific emotions will occur: experiencing happiness or surprise may be difficult or even impossible in a depressed mood. Being in a mood would thus facilitate or hinder the emergence of certain emotions: when we are irritable we feel angry more easily; when we are relaxed we are less likely to experience sadness. These mood variations would then have a significant impact on our actions and decisions: when feeling particularly energetic and happy, we may decide to take on a new

challenge at work; if we experience constant stress and frustration we may end up reconsidering long-term career choices.

Taking stock: in **§2.1** I offer two reasons to resist the characterization of the dimensions as feelings. In a nutshell, I argue that – unlike feelings – the moods arising from the dimensions are *experienced* without being *felt*. In **§2.2** I turn to emotions to show that this notion also fails to capture the nature of the dimensions. Taken together, these reasons contribute to debunk the idea that salience, confidence, familiarity and agency could be successfully described as feelings or emotions. Rather, the dimensions aggregate to give rise to the pervasive and long-term affective states that we know as moods.

To get a clearer idea, think about a pyramid from unconscious to conscious with the dimensions, (and the resulting moods) lying at the bottom, and with feelings and emotions standing at the top – see Table 1 below. On this view, dimensions and moods are ways of structuring our experience that are commonly unfelt but may come to the fore in particular situations. By contrast, feelings and emotions are commonly felt but come into play only "after the fact", as a consequence of *unfelt* changes that have begun some time prior and continue to unfold (Sizer 2006, p. 133).

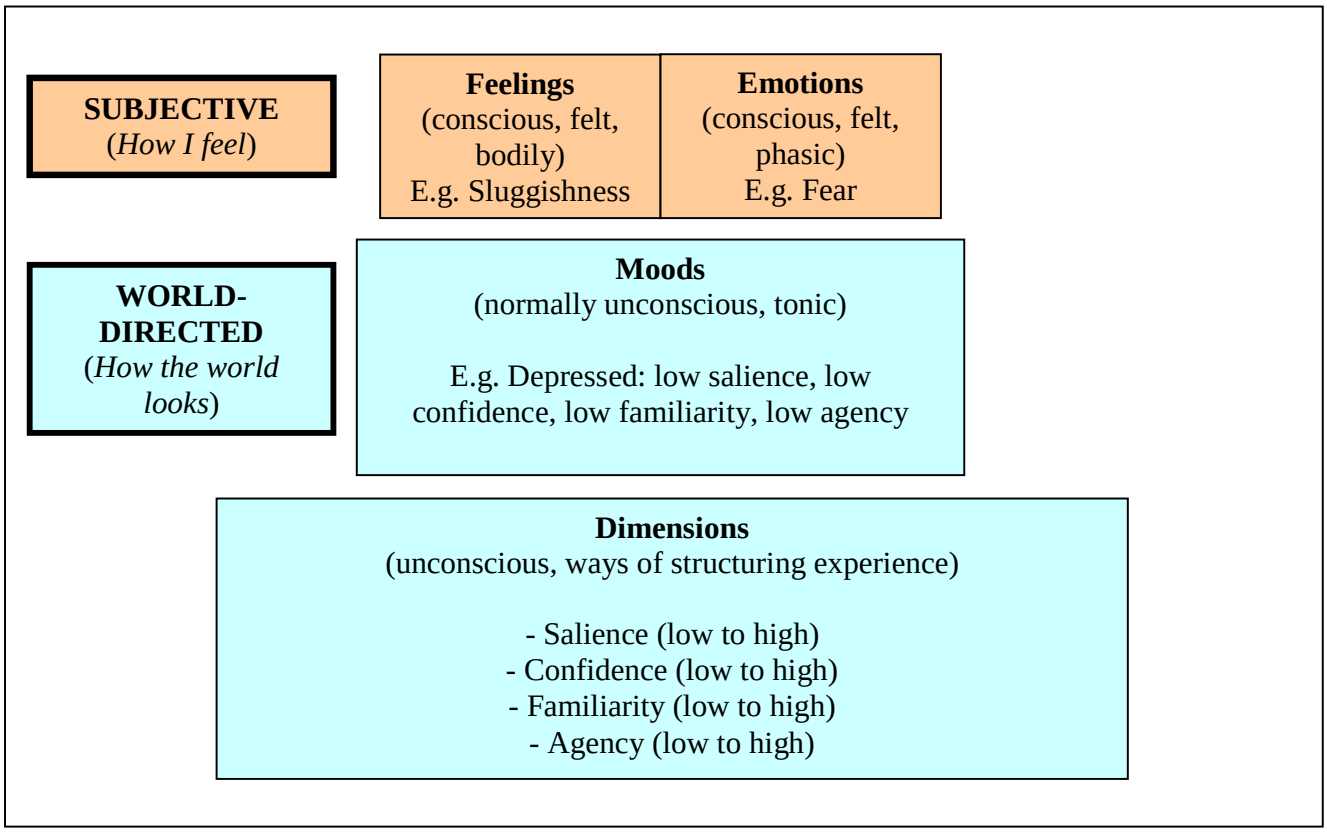

Table 1: Sketch of the proposed model of moods, feelings and emotions

§3. Ratcliffe and Existential Feelings

The existing account that most closely resembles my proposal is the one recently defended by Ratcliffe (2008 & 2005). In his book *Feelings of Being* (2008), Ratcliffe spells out the notion of "existential feelings" (EF henceforth) and focuses on their role in the development of psychiatric disorders. He characterizes EF as "ways of being in the world" or as "background orientations" that help subjects structure their sense of belonging to a shared reality. However, being feelings, EF are also bodily states of which we have some awareness (p. 2). Ratcliffe is not preoccupied with offering an exhaustive list of EF and rather insists on the idea of there being a variety of them, from short-lived (e.g. feeling momentarily disconnected or detached from a situation) to more enduring and pervasive ones (e.g. feeling fatigued due to depression or illness).

Notably, EF are ways of expressing the relation between us and the world, a relation that can be pervasively disrupted in psychopathology. As Ratcliffe puts it: "The world as a whole can sometimes appear unfamiliar, unreal, distant or close. It can be something that one feels apart from or at one with. One can feel in control of one's overall situation or overwhelmed by it. One can feel like a participant in the world or like a detached, estranged observer staring at objects that do not feel quite 'there'" (p. 37).

The view I outline in this chapter has a lot in common with Ratcliffe's. First, and most importantly, both proposals explore the key role that affective states play in structuring and modulating our relationship with the environment. As I explain in **§1.2**, the four dimensions – aggregating in a mood – help us track "how the world appears to us" to us in a given moment. After receiving some painful news our bodily feelings and psychological state change, and *at the same time* our entire experience of the world is transformed. This is similar to saying – as Ratcliffe does – that our affective states cannot be distinguished from the way in which we experience objects or events. Feeling "a nail being driven into the stomach" simply *is* the apprehension of a state of affair external to the body – e.g. that a loved one has died (2008, p. 27). Both EF and the dimensions thus exhibit a peculiar *double identity*: on the one hand, they inform us about how external things and events are experienced (world-directedness); on the other, they tell us something about our subjective state (through feelings or emotions). Both processes can thus be described as affective ways in which our experience of the world is structured.[23] If after reading the news I start thinking that another terrorist attack in Europe is highly likely, my feelings of terror and my mood of dread significantly modify the way in which I interact with the world. Situations that I would have regarded as innocuous or even pleasant (such as a weekend

[23] For another view of feelings mediating self-world relations, see Goldie 2009.

getaway) suddenly appear threatening and ominous. The affective change has transformed the way in which I experience the world – e.g. as something potentially dangerous and frightening.

Another important similarity between my proposal and Ratcliffe's is that we are both convinced that these changes in these existential orientations – EF or moods – play a key role in explaining how mental disorders come about. For instance, experiences of depression cannot be satisfactorily characterized by appealing exclusively to internal feelings – e.g. worthlessness – or ways of seeing the world – e.g. as deprived of anything interesting. Rather, depression presents itself as a simultaneous change in one's affective state and world experience in which the two aspects are inextricable from one another (see **§1.2**). In Capgras delusion, the feeling of unfamiliarity and unsettledness experienced in the other person's presence translates into seeing her as someone else. In the discussion above I refer to Capgras as a case in which a change in *how* something is perceived triggers a change in *what* is perceived. The spouse seems so different that she comes to be seen as someone else (i.e. an impostor) or not even as a person (i.e. a robot or alien). Again, what has changed is not simply our affective state or the person's features, but our global way of relating to her. Ratcliffe puts it this way: "An experience of the spouse that did not offer up the usual possibilities would involve her appearing somehow incomplete, lacking" (2008 p. 158. See also pp. 148-149).

Despite these similarities, my proposal differs from Ratcliffe's in one important respect. Although we both maintain that affective states play a key role in modulating self-world relations, we characterize these states differently. In particular, Ratcliffe insists that *feelings* are the relevant category when it comes to capturing different ways of experiencing such relations (2008, p. 37). In holding this position, Ratcliffe parts ways with Heidegger who famously argued that *moods* are what constitute our sense of belonging to a shared world (1927/1962, p. 173). Notably, Heidegger does not characterize moods as feelings in any relevant sense and refers to

them as "existential modes" (Ratcliffe 2008, p. 55). In his view, moods are conditions of possibility that open up the experience of the world, making it possible to encounter things, events and other people as "mattering". In arguing that moods (as opposed to feelings) are the relevant category when it comes to characterize self-world relations, my proposal sides with Heidegger over Ratcliffe.

Like Heidegger, I maintain that moods better capture the multiple ways in which we relate to the world because of their unconscious component. As I stress above (see **§2.1**), we are in a mood even when we are not aware of it and we do not fully realize that we are relating to the world in a certain way until significant perturbations occur.[24] These are the situations in which moods come to the fore and become "conspicuous" (Heidegger 1927/1962 p. 186). Seeing our neighborhood as threatening all of a sudden makes us realize how safe we used to feel; seeing a family member as a stranger prompts us to look for the affective component that is missing, etc. This view also sits comfortably with the one defended by Russell (1980) and Frijda (1986), where moods are seen as background states that are always present but become noticeable only when they are particularly unusual or intense. As Russell puts it: "Intense core affect can be the focus of consciousness, but milder core affect is typically a part of the background of the person's conscious world" (2003, p. 148). Similarly to Heidegger, I also see moods as the backdrop against which specific emotions or feelings occur. This is connected to the idea that moods highlight pattern of concern and ways in which things matter to us. In other words, moods set up the world in such a way that specific and object-directed emotions become possible by making us *care* about some thing or other. In a famous example, Heidegger claims that it is only possible to experience fear for a being "whose own being matters to it" (1927/1962, pp. 176). This seems very plausible: someone who does not care about his or her own life – e.g. a kamikaze terrorist or

[24] See Elpidorou 2013 for a discussion on Heidegger's account of moods that emphasizes this point.

a suicidal patient – probably experiences fear to a lesser extent than someone for whom survival is valuable.

What is then lost in Ratcliffe's view that feelings – rather than moods – are the affective states mediating self-world relations? In my opinion, he fails to account for the fact that affect may (and often *does*) influence our experience of the world while lying outside of conscious awareness. Our interactions with the environment are always affective in nature because they highlight what matters to us, what we desire or want to pursue. Even our simplest actions can be seen as affective in this sense: I get up and leave my bed in the morning because I am hungry and eating matters to my survival; or I do it because I want to see a friend and this person is important to me. It is difficult to pinpoint a conscious, phenomenally circumscribed feeling at play here: rather, a complex set of goals, preferences and dispositions make these actions or decisions possible. This is exactly how the dimensions work: depending on our location along them, our assessments of what is interesting (salience), doable (confidence), comfortable (familiarity) and controllable (agency) change and give rise to particular moods that – in turn – facilitate or hinder the experience of feelings or emotions.[25]

There are other points of departure between my account and Ratcliffe's that I cannot discuss in detail here. An important one concerns our ability to effectively regulate moods. Heideggerian-inspired accounts – such as Ratcliffe's – often make use of the notion of *thrownness* (*Geworfenheit*) to depict our interaction with the world as one in which situations are something "we find ourselves into" and are "handed over to us" (Elpidorou 2013, p. 4). Along these lines, moods are something that "assail us" and reveal the fact that we exist in a "worldly situation that is not of our making" (Ratcliffe 2008, p. 48; Ratcliffe 2013, p. 4. See also Ratcliffe forthcoming). This view appears compelling because it aptly describes a range of everyday

[25] For other authors who are sympathetic to Ratcliffe's account but use notions such as "moods" or "affect" in this context see Sass (2004) and Stanghellini (2004).

situations: more often than not, we find it difficult to snap out of a mood or to switch from one mood to another. In some pathological cases – such as depression – this passivity becomes particularly remarkable: despite their best efforts and wishes, depressed patients feel stuck in their predicament and the very possibility of being in a different mood seems impossible to them (e.g. "I could not get myself to react" – Plath 1963, p. 3). Yet, this view of moods also appears too restrictive because it overlooks a range of situations in which we are able to successfully exercise various forms of *regulation*.[26] People routinely attempt to change their mood by acting on purely physiological factors, such as one's amount of sleep, caffeine intake and blood alcohol content (see Thayer 1996). They also adopt a variety of psychological strategies to achieve the same goal: one may decide not to attend a party because there is a person one does not wish to see; one may focus on something positive in order to counter anxiety; or one may deliberately change a topic of conversation that has become uncomfortable. All these strategies, along with many others, are known as *reappraisals* and indicate that mood regulation might be more ubiquitous and successful than Ratcliffe suggests.

To sum up: in this section I discussed Ratcliffe's account of EF to clarify what our views have in common and what sets them apart. Both Ratcliffe and I maintain that affective states play an important role in modulating self-world relations and that changes in these orientations contribute to explain how mental disorders arise. However, Ratcliffe prefers to characterize these phenomena as feelings whereas I argue that the notion of moods better captures these unconscious, pre-reflective ways of experiencing the world.

[26] Appraisal theorists (Frijda 1986; Lazarus & Folkman 1984; Gross 2002) have repeatedly emphasized this aspect of mood regulation. For a view of moods that is inspired by Heidegger and at the same time sympathetic to appraisal theories see Elpidorou 2013 (in particular p. 34, footnote 27).

§4. Dimensions, Moods and Pathology

What does the model proposed here tell us about the boundary between normality and pathology? One important feature of this approach is that it allows us to track how moods shift as people move along the dimensions. As I explain in **§1.2**, affective changes bring about different ways of relating to the world that eventually crystallize into moods. In the example discussed above, my initial mood was characterized by a combination of high salience (e.g. being focused on different aspects of my job), confidence (e.g. being sure about completing a task on time), familiarity (e.g. knowing the problem well) and agency (e.g. being in control with respect to the task at hand). After having received upsetting news from my sister, all these parameters have dropped significantly. I have a hard time concentrating because my thoughts keep wandering in one direction (diminished salience); I become less confident about being able to complete work on time and I also become more estranged from the task (diminished confidence and familiarity). All these factors contribute to my sense of the situation being beyond my control (diminished agency). By assessing where one stands on each dimension in a given moment and by tracking these changes it is thus possible to observe the impact of an event and the resulting affective configuration. This is valuable both for ordinary and pathological cases, because it clarifies what it means for moods to be ways of relating to the world that shift in response to significant events.[27]

However, this still tells us little about what makes a particular configuration pathological as opposed to another. Some further observations are therefore in order. First, an *extreme local imbalance* on one or more of the dimensions may work as an indicator of an underlying disorder.

[27] Here one may rightfully object that the assessments provided by the dimensions are still too coarse-grained. For instance, a subject may experience high and low familiarity simultaneously toward different objects: e.g. my philosophy books as opposed to my mom's calculus books lying on the same shelf. Although these finer-grained assessments might be helpful in some cases (e.g. to account for localized delusions), we need not delve too deep at this stage. I would like to thank Marco Viola for bringing up this point and providing a version of this example.

In fact, as I discuss in Chapter Two, some disruptions can be primarily characterized by a loss (hypo) or overload (hyper) on a specific dimension. These cases include various delusions – such as Capgras (hypofamiliarity) or grandiosity (hyperconfidence) – as well as symptoms connected to other disorders – e.g. self-loathing in depression (hypoconfidence) or auditory hallucinations in schizophrenia (hypoagency). In this sense, a configuration exhibiting an extreme local imbalance may offer a reliable indication of pathology. Second, there are configurations where a subject appears *stuck* at one end of the spectrum with respect to all the dimensions. These are cases in which pathology does not simply consist in lying on one extreme but also in being unable to regulate one's mood and eventually modify it. A paradigmatic example discussed in this chapter is major depression, represented in the model as the aggregation of extremely low points on all four dimensions. Indeed, depression can be seen as a combination of low salience – i.e. nothing appears interesting or worth exploring – low confidence – i.e. self-loathing – low familiarity – i.e. things and events appear unknown and unsafe – and low agency – i.e. tasks and situations seem overbearing. All these local imbalances are already aversive in themselves and would probably be associated with unpleasant emotions or feelings (e.g. fear or anxiety). Yet, there is an additional pathological component in this case: the fact of lying at the extreme low of every dimension creates a sense of being stuck and makes it almost impossible to entertain the possibility of change. This is (again) related to the world-directed component of moods emphasized above: the aggregation of the dimensions is not experienced as a subjective feeling, but as a way in which the world looks to the subject. This makes it in turn difficult to act directly on a mood to modify it: in a low salience mode I do not perceive myself as suddenly uninterested or unmotivated; I see things and events themselves as having lost the enticing pull they used to have. Similarly, I do not experience myself as incapable of completing a task but rather the task itself as objectively more difficult. Third, there are cases in which pathology may be

characterized by a sudden switch between two extremes on one or more dimensions. A paradigmatic example here is bipolar depression, where a subject tends to alternate between hypo and hyper manifestations. With respect to salience, one may go from a depressive phase in which excitement and motivation are impossible to retrieve (hypo), to a manic phase where too many things or events appear interesting and soliciting immediate action (hyper). Similarly, patients affected by bipolar depression may transition from self-loathing to grandiosity (confidence) and from feeling impotent to omnipotent (agency).[28] Again, part of the pathological nature of this disorder comes from the fact that subjects experience extreme local imbalances; however, a specific pathological aspect is added by the problematic oscillation between two extremes. The model defended in this chapter thus contributes to better understand some important features of pathological moods. In some cases, these configurations are characterized by *extreme imbalances* on one or more of the dimensions (e.g. Capgras delusion). Others represent situations in which the subject is problematically *stuck* at one extreme (e.g. unipolar depression), while still others exhibit a *sudden switch* between extremes (e.g. bipolar depression). Further developments of the model may then allow for a more precise characterization of what counts as a pathological mood.

Conclusion

I conclude the chapter by flagging a potential problem with my account. Earlier in the discussion, I show that a four-dimensional structure allows us to characterize moods at a higher level of precision with respect to other similar models (Russell 1980; Thayer 1996). Yet, this might not be enough. In particular, one might worry that the model proposed here would still fail to distinguish between different disorders as well as between ordinary and pathological cases. The objection appears *prima facie* justified, as it is likely that some disorders that are currently

[28] For a first-person account of this disorder, see Jamison 1996.

classified as different would exhibit a similar configuration in the model. In other words: like anger and fear occupy the same space in Russell's circumplex, there may be configurations falling in the same region while denoting different disorders. For example, AVH and delusion of control would exhibit a similar configuration, characterized by enhanced salience towards certain stimuli (e.g. inner speech) combined with diminished confidence, familiarity and agency. One strategy to address this objection would be to further flesh out the difference between the two disorders in quantitative terms, for instance by characterizing the loss of agency as more or less severe in one case or the other. A different strategy would be to bite the bullet and suggest that similar configurations in the model reflect similar lower-level disruptions, even when this turns out to be at odds with current diagnostic classifications. In this sense, two disorders that appear superficially and nosologically different could be grouped together in virtue of their similar configuration that would in turn reflect – by hypothesis – a similar neurological profile.[29] Yet, the same objection can be raised when it comes to distinguish between ordinary and pathological cases that exhibit similar configurations. For example, unipolar depression and grief would probably appear very similar and hard to disentangle in this model. Since the distinction between normality and pathology is the main object of the dissertation, I discuss the point more thoroughly in Chapter Four where I take a closer look at the notion of *vulnerability* and I introduce the notions of *risk* and *protective factors*. For now, it is sufficient to point out that the model introduced here allows us to identify patterns of vulnerability, namely configurations that resemble pathological ones and are at higher risk of becoming pathological themselves.

In this chapter I explored the role that the dimensions (aggregating in moods) play in modulating our experience of the world. Specifically, I argued that salience, confidence, familiarity and agency are background orientations that structure the way in which the world

[29] Murphy 2006 (chapter 3) adopts a similar strategy.

appears to us in a given moment. For instance, being in a state of high or low salience determines the range of things and events that catch our attention. Different degrees of confidence give us a sense of what we can successfully undertake; different degrees of familiarity allow us to assess what is known or safe in our environment, and different degrees of agency indicate the extent to which we are in control with respect to a situation. Similarly to other dimensional models of affect, a particular mood then arises from the aggregation of these dimensions. While Russell and Barrett's core affect emerges from the intersection of arousal and valence, and Thayer's mood arises from the encounter of energy and tension, on my view moods originate from the combination of salience, confidence, familiarity, and agency. Moods can thus be characterized as points in a four-dimensional space, with every individual lying in one portion of the space in any given moment.[30]

[30] The fact that the model employs four variables makes its visual representation less straightforward than the alternatives. For this reason I do not develop a graphical model here, although I plan on doing so in future work. For a three-dimensional model, see Hobson 2001.

Section III:

Meaningful Difference

Valentina Petrolini

What Makes Mental Disorders Different From Normal Functioning? Explaining (Away) Discontinuity

"In psychiatry we want to know what is actually wrong with people, and we want to predict what will happen to them and how we can intervene to prevent or mitigate negative symptoms. To do this we need to know the structure of the human mind and what can go wrong with it"

(Murphy 2006, p. 344)

Abstract

In Chapter Two I claim that a convincing defense of CT should achieve two goals. First, it has to propose a model of mental disorders where the distinction between normal and pathological cognition boils down to a difference of degree. This means that the model has to paint a plausible picture of what it means to be an extreme variation of a non-pathological phenomenon. To do so I introduced four dimensions of mental functioning – i.e. salience, confidence, familiarity and agency – and I showed that they may be disordered in two opposite ways (i.e. hyper or hypo).[31] However, a good defense of CT also needs to achieve a second goal, namely it has to provide a non-arbitrary way to distinguish between normality and pathology. In other words, the model has to explain what makes disordered and non-disordered cases importantly different without trivializing this distinction and without neglecting the phenomenological peculiarity that accompanies disordered experiences.

In this chapter I set out to accomplish this second goal by exploring more closely a number of intermediate cases lying on various points of the dimensions outlined in Chapter Two. As I mention previously,[32] some intermediate cases may be characterized as *local imbalances* because they approximate one extreme of a given dimension. For example, someone who systematically overestimates his abilities would come close to grandiosity on the confidence

[31] See Chapter Two, §1-§4.
[32] See Chapter Two, §1.2.

dimension. These local imbalances thus represent situations of vulnerability to mental disorders that are crucial to explain the transition from normality to pathology. By comparing these intermediate cases with their disordered counterpart, it becomes possible to draw important distinctions between pathological and non-pathological manifestations.

The issue is complex and requires the introduction of further notions. In what follows, I argue that there are two different ways of being *vulnerable* to mental disorders. On the one hand, some intermediate cases may be seen as attenuated versions of their pathological counterpart: here, the distinction between normality and pathology hinges on the increase of *risk factors* (e.g. duration, frequency, urgency, etc). Going back to the example above, the core difference between someone who is overconfident and someone who suffers from grandiosity delusion lies on how pervasive, frequent, and long-lasting this person's convictions are. On the other hand, some intermediate cases represent situations of vulnerability that are counterbalanced by the presence of *protective factors* (e.g. control, humor, physical and mental strength, etc). In these situations the distinction between normality and pathology should be described through the successful or unsuccessful action exercised by such protective forces. For example, someone suffering from obsessions may have significant resources at her disposal such as a sufficient amount of energy to keep them at bay. In both cases, external events (especially traumatic ones) may importantly interact with risk and protective factors and give rise to pathological constellations that would otherwise fail to emerge.

The chapter is divided into two large sections. In **§1** I render the notion of vulnerability to mental disorders more precise by introducing the two kinds of intermediate cases mentioned above. Some of these cases (*type a) cases*) qualify as attenuated versions of their pathological counterpart, while in others the imbalance is counteracted by one or more protective factors (*type b) cases*). At the end of this section I introduce *type c) cases*, namely situations in which there is a

significant interaction between risk and protective factors. In **§2** I discuss a number of examples in order to illustrate in detail how intermediate cases relate to the four dimensions identified in Chapter Two. This section also serves the purpose of exploring the relationship between vulnerable and pathological cases, by focusing on the role played by risk and protective factors.

§1. Two Types of Vulnerability

The notion of vulnerability is commonly employed in psychiatry and encompasses a wide range of conditions. In some cases, an individual is deemed vulnerable to a disorder because she belongs to a group where said disorder appears at a higher rate with respect to the general population. In this sense, women are considered more vulnerable to depression than men (see Kessler 2003).[33] In other cases, the degree of risk is assessed by looking at a number of factors ranging from the incidence of the disease in the individual's family to cognitive measures such as discrepancies in self-perception (see Morrison *et al.* 2006). Psychiatrists currently employ several labels to refer to subjects at high risk of developing mental disorders: "At Risk Mental State" (ARMS), "Ultra High Risk", and "Clinical High Risk" are only some of them. Despite the attention that the notion of vulnerability has garnered, a precise characterization is still lacking and different research groups end up employing different measures (see Fusar-Poli *et al.* 2013 for a review).[34]

My goal in this section is to propose a more precise account of what it means to be vulnerable to a mental disorder. More specifically, I show that it is important to distinguish between three different types of vulnerability that in turn correspond to different notions of risk.

[33] I take a deeper look at this environmental-level vulnerability in Chapter Five, while in this chapter I focus on different types of personal-level vulnerability.

[34] This recent meta-analysis conducted by Fusar-Poli and colleagues (2013) focuses on At-Risk Mental States (ARMS) from a neurocognitive perspective. The findings they discuss suggest that high-risk individuals who convert to psychosis show more severe neurocognitive deficits than those who do not develop a psychotic disorder over time. In this sense, high-risk individuals could be treated as intermediate cases of different severity.

Generally speaking, a person's degree of vulnerability indicates the adequacy of her resources in facing environmental challenges. Lazarus & Folkman (1984) stress the fact that these resources may be physical – e.g. sufficient amount of sleep – as well as psychological – e.g. motivation towards a goal (pp. 50-51). They also distinguish between factors that are stress-inducing and thus increasing vulnerability (i.e. risk factors) and factors that are stress-reducing and thus diminishing vulnerability (i.e. protective factors). The notions of vulnerability that I propose here build on such a characterization. On my view, *risk factors* can be seen as obstacles that prevent a subject from dealing effectively with a stressful situation, whereas *protective factors* are resources that the individual can mobilize in order to cope with it. Depending on the role played by these factors, different kinds of vulnerability to mental illness emerge and thus different groups of intermediate cases should be distinguished.

a) Vulnerability & Risk Factors

These are intermediate cases that differ from their pathological counterpart in a way that can be seen as straightforwardly quantitative. In other words, they can be seen as attenuated versions of a pathological condition. For example, the experience reported by a subject who is very sad and a clinically depressed patient may be quite similar: they both show little interest for activities that others (or their former selves) regard as pleasurable, they both suffer from low mood and fatigue in carrying out everyday tasks, etc. In these cases, the core difference between intermediate and pathological manifestations lies in the presence or absence of a number of risk factors that are quantitative in nature.

One important risk factor is *duration*, which indicates how long a symptom, thought or feeling extends over time. Back to the example above, having the blues for a couple of days differs from experiencing a comparable degree of sadness for months in a row. Another risk

factor is *frequency*, indicating the number of times in which a symptom presents itself to consciousness. Again, having obsessive thoughts a few times per month differs from having the same thoughts hundreds of times per day (see Rachman & de Silva 1978). *Intensity* and *urgency* also play a key role: while the former can be characterized in terms of strength or depth, the latter indicates how soon a problem needs to be dealt with, a thought or feeling acted upon, etc. The notion of intensity is notoriously hard to pin down: indeed, a psychological manifestation may be regarded as intense in virtue of its *strength*, measured by the amount of resources that the subject has to mobilize in order to cope with it. In this sense, an obsession qualifies as intense if dealing with it requires an amount of effort that leaves a person unable to carry out any other task. Yet, a psychological phenomenon may qualify as intense also in virtue of its *depth*, for example by causing someone to radically revise her view on things. In this sense, a delusional idea may be intense because it slowly transforms a person's relationship with the environment and gives rise to new interpretations of events and situations. As Frijda puts it: "Emotional intensity is not a unitary concept. What affects one index of intensity need not affect another one. Events that cause no acute upset may keep nagging the subject for days" (1986, p. 290). Finally, the *scope* of a symptom may act as a risk factor because it concerns the number of things or aspects of a situation that the individual regards as relevant. For example, it might be easier to control intrusive thoughts when they are restricted to a specific domain and harder as they start applying to more objects, people, or events.

These examples suggest that in some cases the transition between health and pathology could be explained by an increase in risk factors. In §2 I discuss a number of such cases in more detail. For now, a metaphor may be helpful to understand this first notion of vulnerability. A vessel struggles to stay above water while fighting against a number of dangers that may cause the water level inside the ship to increase. If the vessel encounters a storm, it is crucial to assess

how long the threat will last (duration), how likely it is for other storms to occur soon (frequency) and how strong will this event be in comparison to others or relatively to the vessel's conditions (intensity). It would also be important to assess how soon the storm will start damaging the ship's vital components (urgency) and how much effort and resources would be needed to repair it after the threat has ended (scope). Crucially, the breaking point will be reached when one or more of these factors cause the internal water level to increase to the point of overflow. Sticking to the metaphor, the level of threat posed by the storm also importantly depends on the overall conditions of the vessel. If the lifeboats have been lost during the last expedition, even a mild storm may qualify as dangerous; conversely, the ship may have been recently repaired and prove more resistant in the face of adversities. In our terminology, the current state of the vessel represents a person's vulnerability while the various characteristics of the storm are the risk factors. The overflow point represents the emergence of a mental disorder through a process of *summation*, where the risk factors keep increasing until a certain threshold is reached. As Lazarus & Folkman put it: "'Oh God, yet another thing' is the final cause of the breakdown" (1984, p. 111). [35]

b) Vulnerability & Protective Factors

These are intermediate cases that differ from their pathological counterpart due to the role played by one or more protective factors. As opposed to *type a)* cases, it is not the increase in risk factors that facilitates the transition to pathology. Rather, it is the weakening of the protection coming from counterbalancing forces that causes the subject to reach an individual breaking point. For example, a person may exhibit a number of idiosyncratic obsessions but still fail to

[35] These cases can also be compared with some at-risk situations in somatic medicine. For example, a person suffering from high blood pressure exhibits an attenuated version of a pathological condition, which may then arise by the mere increase of quantitative factors.

count as pathological because she is able to laugh about it. If – for whatever reason – this person loses her sense of humor, her obsessions may spiral out of control and massively impair her functioning. One of the most important protective factors is thus a person's ability to *control* a challenging situation by employing a number of coping strategies. For example, one may attempt to avoid situations where the threat is present or – when this is not possible – alter the situation by adopting a different behavior. Alternatively, one may direct attention elsewhere or change one's perspective on the issue – e.g. by downplaying its importance.[36] Another important protective factor is the amount of *strength* that the subject has at her disposal to handle the relevant challenges. Notably, the notion of strength here comprises both physical resources – such as the ones afforded by sleeping or eating – and psychological ones, such as the ability to focus, solve problems and regulate emotions. The idea that some mental disorders may arise through processes in which the exhaustion of strength plays a key role has been repeatedly defended by Freud. As he puts it: "Each individual has in all probability a limit beyond which his mental apparatus fails in its function of mastering the quantities of excitation which require to be disposed of" (1926, p. 139). Generally speaking, a sufficient degree of physical and psychical energy protects the subject from feeling overwhelmed by external or internal demands. In this sense, the notions of strength and stress are closely connected, with stress acting as a force that wears out the individual's resources. For this reason, phenomena such as psychotic outbreaks often occur when the patient is going through particularly stressful times – e.g. being appointed for a prestigious but demanding job (see Schreber 1903), or going up for tenure (see Saks 2007). Other protective factors may come into play in preventing intermediate cases from becoming

[36] One might worry about how to quantify an apparently elusive notion such as *control*. Indeed, factors like frequency or duration lend themselves more easily to being operationalized and are already employed by a variety of clinical tools (DSM included). In what follows I refer to control as the cognitive ability that allows a subject to employ one or more coping strategies to deal with stressors. See Gross 2002 for a comprehensive classification of these strategies.

pathological. For instance, *humor* has already been mentioned as a powerful means to keep distressing thoughts and emotions at bay. This result may be accomplished in a variety of ways: by promoting detachment from something that the subject dreads (e.g. pre-surgery or pre-exam jokes), by magnifying a negative personality trait (e.g. through self-irony or self-deprecation), by voicing thoughts that are repressed or unacceptable (e.g. sexist or racist jokes). Finally, various instances of *abreaction* and *discharge* act as protective because they allow the subject to release mental or emotional tension – e.g. by venting, talking to a friend, writing in one's diary – as well as physical one – e.g. through sports, crying (see again Freud 1914; Freud 1905).

In *type b)* cases the transition between normality and pathology can be explained by the weakening of one or more protective factors. In §2 I discuss a number of these cases in more detail. Metaphorically speaking, this second kind of vulnerability may be represented as a medieval fortress whose inhabitants attempt to defend themselves from a siege. As the attack unfolds, there are a number of countermeasures that the inhabitants can take. First, it is important to understand the enemy's strategy and to make decisions about how to better react (control). Assuming that the threat could not be ignored, the inhabitants have to establish what can be done about it and where exactly the relevant resources should be mobilized. Some of these protective measures would involve rendering the fortress stronger and more resistant against the attack, while others would focus on preserving the human and material resources available for the longest possible time. For example, soldiers may arrange a watch system to ensure that everyone gets a minimum amount of sleep, or they may regulate the distribution of food and water to the population (strength). We can also imagine that, in order to release the stress caused by the threatening situation, the inhabitants would engage in activities directed towards abreaction – e.g. composing songs or telling jokes about the siege – or discharge – e.g. getting drunk. If the situation becomes serious enough and the siege continues over time, the defenders would become

weaker and have fewer resources at their disposal to counteract the enemy's advance. Some of the protective measures might cease to be effective (e.g. joking about the siege is no longer fun), and others may become plainly unavailable (e.g. the city runs out of wine). On the brink of disaster, the inhabitants may decide keep fighting despite lacking the resources to sustain the attack any longer, thereby suffering *defeat*. Notably – as opposed to the vessel's metaphor – the breaking point here gets reached when the defenses put in place wear out and the inhabitants are unable to protect themselves. In our terminology, the current state of the fortress represents vulnerability while the various defenses employed by the inhabitants are the protective factors.[37] In these cases, the transition between health and pathology arises from an imbalance between environmental demands and the person's resources to cope with them. The breakdown point thus represents the emergence of a mental disorder through a process of *exhaustion* where the protective factors progressively weaken until a certain threshold, which varies from individual to individual, is reached.

c) Interaction between risk & protective factors

As I explain above, type a) and b) cases represent situations in which the threshold to pathology is crossed *mostly* via an increase in risk factors or via a weakening of protective ones. However, there are also cases in which *both* types of factors play a major role in bringing about the transition to disorder. On the one hand, an increase in risk factors may render some of the protective factors ineffective or unavailable. For example, if an obsessive thought increases in frequency, intensity and urgency, exercising control over it becomes more difficult and detachment strategies – such as humor – may prove impossible or even counterproductive. On the

[37] These cases may also be compared with some at-risk situations in somatic medicine. For example, a person engaging in a risky behavior in absence of protection is more vulnerable to the development of certain diseases – e.g. HIV infection.

other hand, the weakening of protective factors may cause an increase in risk factors. For example, a recovering alcoholic who does not manage to avoid her circle of drinking friends would probably end up thinking about the next drink more frequently and experience the thought as more urgent. *Type c)* cases should thus be seen as more complex because the causal interaction between risk and protective factors is harder to disentangle. However, one may be tempted to regard them as always reducible to a) and b) cases. Despite the more complex causal chain, c) cases can still be divided into ones crossing the threshold via summation (i.e. increase in risk factors) or via exhaustion (i.e. weakening of protective factors). Going back to the examples, an obsessive thought may take over because acting upon it becomes too urgent (summation) *or* because the subject cannot control it successfully (exhaustion). Similarly, the alcoholic may relapse because she thinks about the next drink too frequently (summation) *or* because she cannot handle the pressure of her friends any longer (exhaustion). Yet, in these complex scenarios it is hard to predict how exactly the breaking point will be reached. By looking at the existing vulnerabilities and resources, more than one route to pathology appears to be open.

Despite being in principle possible to reduce c) cases to a) and b) cases, it would not be profitable to do so in most situations. Indeed, the interaction between risk and protective factors may be hard to disentangle and contextual contingencies play a crucial role in determining the nature of the breakdown. A casual walk in front of one's favorite bar may kindle more frequent thoughts about drinking (i.e. rise in risk factors) as well as the urge to text old friends (i.e. weakening of protective factors). Similarly, a message from someone that we have not seen in a long time may trigger a backlash of emotional intensity (i.e. rise in risk factor) but also make us realize that we are not ready to joke about it yet (i.e. weakening of protective factors). In this sense, talking about "the final cause of the breakdown" may not be the right way to capture the complexity contained in the vulnerability profiles. By introducing *type c)* cases, I thereby sidestep

the need to univocally specify a final cause for each intermediate case and I acknowledge that in many situations it would not be epistemically (or practically) feasible to do so.

Before moving on to the next section, I introduce a special class of factors that can act as risk or protective depending on the circumstances, on the individual's personality, etc. Following Lazarus & Folkman (1984), I dub them *dual factors*. One important dual factor is *anticipation*: having more time at one's disposal to deal with a problem may be stress-inducing or stress-reducing depending on a variety of contextual elements. For example, if I am getting ready for a job interview, the thought of having enough time in front of me is usually relaxing because I believe I can prepare better and collect more information about my potential employer. Yet, if the waiting time becomes too long – e.g. the interview is postponed – I may become more anxious because I suddenly have more opportunities to think about how things can go wrong. Lazarus & Folkman put the point nicely: "Given time, people can reflect, suffer or grieve; they can also avoid the problem, think about it, take action or make efforts to gain self-control" (1984, p. 98). Similarly, *imagination* has a powerful dual nature: on the one hand it helps subjects to think through problems and simulate various scenarios – e.g. I think about questions that may come up during the interview. On the other hand, it may magnify potential problems or issues – e.g. I visualize me blacking out completely after the first question. Whether imagination acts as risky or protective depends on external circumstances (e.g. how competitive the interviewing process is), on my current emotional state (e.g. if I am anxious for other reasons that affect my performance) or on general personality traits (e.g. I may be prone to pessimistic fantasizing). A nice illustration of the duality of imagination can be found in Billy Wilder's movie *Seven Year Itch* (1955), where the protagonist Richard Sherman indulges in daydreaming about seducing his new neighbor (Marilyn Monroe) but invariably imagines the disastrous consequences that would derive from that. Throughout the movie, Richard vindicates imagination as one of his most defining character

traits: “It’s just my imagination. Some people have flat feet. Some people have dandruff. I have this appalling imagination”. Another dual factor is *ambiguity*: a situation that is not completely clear may elicit a great degree of anxiety (e.g. when a loved one does not come home at the expected time) but also allows for a variety of interpretations, some of which are positive or neutral (e.g. he might have run into a friend). Again, ambiguity fosters imagination in a way that may enhance stress (e.g. via paranoid thoughts) or reduce it (e.g. via self-reassuring thoughts). The dual aspect of these factors comes to the fore also when we think that situations of anticipation, imagination or ambiguity are often actively pursued and regarded as pleasurable. People engage in fiction, watch suspenseful movies, go on roller coaster rides and keep up romances without making things too explicit. All these activities indicate that a certain degree of anticipation, imagination or ambiguity is not only tolerated but actively sought for in many situations.

To take stock: I propose to render the notion of personal-level *vulnerability* more precise by distinguishing between three ways of being at-risk of developing a mental disorder. *Type a)* cases comprise situations where an individual exhibits an attenuated version of a pathological condition. Whenever one or more of the risk factors increase the transition to pathology becomes more likely or even inevitable. Going back to the example above: a person who experiences sadness ever more frequently, with higher intensity and wider scope might eventually cross the threshold for clinical depression. *Type b)* cases comprise situations where an otherwise pathological condition is countered by the presence of one or more protective factors. In these cases the transition to pathology occurs whenever the protective factors become insufficient for dealing with the current demands or stressors. For example, a person who is usually able to laugh about her obsessions might go through a particularly stressful time during which she lacks the mental and physical energy to do so. *Type c)* cases are those in which risk and protective factors

causally influence each other. In these cases the threshold to pathology may be crossed in different ways but it is not immediately transparent which one is more probable.[38] Although this tripartite distinction helps to shed light on different ways of being vulnerable, I do not follow it too strictly in the remainder of the chapter. Indeed, in the next section I turn to the bulk of the argument in defense of CT, which consists in rejecting categorical views through the discussion of a series of intermediate cases.

§2. Intermediate Cases

In this section I confront categorical views by showing that the core difference between normal and disordered cases can be explained by appealing to quantitative processes, such as the increase in risk factors or the weakening of protective ones. This represents a key move in defense of CT: so far, the supporter of a categorical model has no reason to reject the distinction between different kinds of vulnerability I propose in §1. However, she may still regard normal and pathological states as importantly distinct from a phenomenological, functional and possibly ontological viewpoint. So, where does the disagreement comes from?

Generally speaking, proponents of categorical views are committed to the idea that the gap between health and pathology should be conceived of as a difference in kind, similar to the one between substances having different chemical compositions or atomic numbers. Conversely, proponents of dimensional views maintain that such a gap should be seen as a difference in degree, similar to a spectrum of colors fading into one another. However, for a categorical view to pass muster it is not sufficient to show that pathological experiences differ significantly from

[38] As I mention above, in Chapter Five I introduce yet another type of vulnerability to mental disorders. As opposed to the ones discussed here, this sort of vulnerability can be found at the population level and reflects the fact that some groups are more exposed to a variety of stressors due to structural injustices.

normal functioning. In fact, a dimensional account needs to do so just as well in order to avoid the charges of arbitrariness and vagueness. In this sense, both approaches take it as common ground that normal and pathological phenomena should be regarded as different in a number of ways – i.e. phenomenologically, in terms of well-being. Yet, they importantly disagree on how to characterize such a difference.

There are different versions of the categorical view that one might distinguish; here I do not focus too much on these differences, but I tackle them all at once. Let me thus briefly mention some typical versions of the categorical view, although this list is not supposed to be exhaustive. Some authors insist that some mental disorders should be regarded as *natural kinds*: in this sense, psychiatry would be similar to biology or chemistry in its attempt to uncover a natural structure shared by a class of objects. Just like "copper melts at 1083° and Syrian hamsters have a sixteen day gestation period", mental disorders would exhibit some determining properties that allow us to group them together and distinguish them from normal functioning (Cooper 2005, p. 46). This approach has the clear advantage of facilitating explanation and inductive inferences, as well as to bring psychiatry closer to other sciences (see Samuels 2009 for a similar argument focused on delusions). Other authors – such as Murphy – defend the idea that categories are better suited than dimensions in representing mental disorders as underlying destructive processes that share a *common cause* (2006, p. 318 & 357). As he puts it: "[A] categorical approach treats disorders as discrete phenomena, qualitatively different from normal states in virtue of pathological causal histories. A dimensional system of classification represents disorders as falling between points on an axis, or as a location in multidimensional space, and not as discontinuous categories" (p. 345).[39] Phenomenological accounts of mental disorders also tend

[39] Murphy also grants that some disorders may be dimensional in nature, and suggests that the debate on classification may turn out to be "a matter of convenience" (p. 356). Although I cannot discuss the details of his

to support a qualitative distinction between healthy and pathological experiences (see Ratcliffe 2015 & 2010). On these views such a distinction is often cashed out in *transcendental* terms: for example, Ratcliffe presents normal sadness as the inability to experience specific emotions – e.g. enjoying a meal – whereas depression involves the loss of a whole "possibility space" (2010, p. 12). Thus, while normal sadness is experienced as a "contingent psychological state", depression presents itself as a *structurally different*, "all-encompassing way of being" (p. 9).[40] Finally, in the clinical literature many adopt the *checklist approach* put forward by the diagnostic manuals (i.e. DSM-5 and IC10), where experiencing a certain number of symptoms for a given duration or with a certain frequency puts a person beyond the relevant diagnostic threshold. Kendell & Jablensky (2003) describe this approach as based on the assumption that mental disorders are clearly separated from normality: "Mental disorders are separated from one another and from normality by natural boundaries (zones of rarity)" (pp. 4-5). This point is crucial for our purposes because it stresses the absence of intermediate cases or at least their theoretical irrelevance. Crucially, the claim that most of the categorical approaches share is that intermediate forms do not threaten the validity of discrete categories "because they are uncommon compared with the defined conditions" (Kendell & Jablensky, p. 8).

In this section I counter such a claim by discussing a number of intermediate cases distributed across the four dimensions of functioning introduced in Chapter Two. The goal is to provide an *argument for quantity* against categorical views: by presenting a structured list of examples, I show that intermediate cases are not comparatively infrequent but rather pervasive and systematic. Once enough cases have been discussed, it becomes apparent that some categorical models are untenable. For instance, the boundaries between health and pathology are

proposal here, I believe our models are compatible in a number of ways. For instance, they both emphasize the "explanatory story" from normality to pathology (p. 369).

[40] See Chapter Three, §3 for more details on Ratcliffe's view.

not discrete (*contra* Samuels), phenomenological experiences vary in degree but not in kind (*contra* Ratcliffe) and zones of rarity are hard to come by (see Kendell & Jablensky 2003). However, more ecumenical models – such as Murphy's – are open to the idea that some disorders may be dimensional in nature. Murphy explicitly accepts this possibility: "Suppose that changes in the value of some variable were manifested behaviorally *not stepwise but smoothly*, so that there was *no obvious cut-off* either way. But suppose that we nonetheless felt that some values of the behavioral variable were sufficiently deviant to merit a diagnosis. Then we would have a dimensional disorder, since now the clinically significant values would *shade into normal ones without any break*" (p. 360. Italics mine). In this sense, the difference between the model I propose here and the one put forward by Murphy may boil down to a terminological one.

In what follows, I discuss a series of intermediate cases lying on the four dimensions of functioning previously introduced – i.e. salience, confidence, familiarity and agency. On the one hand, I show that every intermediate case can be meaningfully related to a pathological counterpart. On the other hand, I argue that the core difference between intermediate and disordered cases can be explained by appealing to quantitative factors. In particular, it is the increase in risk factors and/or the weakening of protective ones that causes an individual to cross the threshold between normality and pathology.

§2.1. Salience: Neuroticism and Sadness

Hyper. Notable intermediate cases lying on the salience dimension are different forms of *neuroticism*, a personality trait described as the tendency to experience anxiety, moodiness, worry or frustration (see Goldberg 1990). Instances of neuroticism include mild obsessions (e.g. having a crush on a famous person), paranoid thoughts (e.g. being slightly hypochondriac) as well as various idiosyncrasies (e.g. having one's clothes or books arranged by color or size). In this

sense, neurotic individuals exhibit a form of exaggerated attention to details that are usually overlooked by others (e.g. how objects are arranged on a desk) and act upon them in ways that are considered irrational (e.g. not going out on an 'unlucky' day of the week). These manifestations can thus be seen as imbalances in the salience dimension, and more specifically as cases of hypersalience. Although neuroticism – like all the major personality traits – appears to be normally distributed in the general population, some individuals score particularly high on the scale.

Paradigmatic examples are the characters impersonated by Woody Allen in his movies, who take neuroticism as a core feature of their personality. In a recent paper, Killmister (2015) characterizes "Woody Allen cases" as situations in which attitudes that are utterly unreasonable become personality-defining. As she puts it: "The classic Woody Allen character is one who is driven by desires and anxieties that he knows to have no justification, and yet who takes these desires and anxieties to be reasons to live his life in a certain way" (p. 2). Killmister's analysis uncovers one key aspect of neuroticism: these individuals are not just unreflectively driven by unreasonable thoughts or desires, but often regard their own reasons as irrational or objectionable. For example, someone may realize that her fear of flying is completely unjustified, while still taking this fear as a reason for acting in a certain way – i.e. taking the train even if it is more expensive and time-consuming (p. 11). In some cases, neurotic individuals might fall prey to some kind of *tunneling effect* that makes them locally recognize their failures but keeps up the illusion that things will work out next time.

An extreme case of this kind of neuroticism is exemplified by Allen's character Allan Felix in the movie *Play it again Sam* (1972).[41] A San Francisco-based, recently-divorced movie

[41] Readers might notice that in discussing intermediate cases I often refer to fictional or quasi-fictional characters. Indeed, one might worry that the choice of these examples would undermine the empirical nature of my proposal. Yet, I believe that the cases presented here are particularly illustrative and vivid, to the point that they – in fact –

critic, Allan is eager to meet women and go out on dates but constantly has to fight against his pervasive feelings of insecurity and anxiety. Getting ready to go on a blind date, he starts speculating about what might happen and inevitably envisions catastrophic scenarios: "I wish she had seen me before. I hate to be there on a blind date with a girl that first sets eyes on me. What if she's disappointed? What if she laughs or screams?" Among many other idiosyncrasies, Allan has an obsession with Humphrey Bogart who vividly shows up in his imagination to offer advice on how to treat women. Bogart acts as Allan's more confident and nonchalant counterpart, ironically mocking his ineptness in romantic matters. In particular, Bogart oversimplifies situations that Allan has made unbearably complicated out of worry or anxiety. When Allan asks him: "What's the matter with me? Why can't I be cool? What's the secret?", Bogart replies: "There is no secret, kid. Dames are simple. I never met one that didn't understand a slap in the mouth". Crucially, this Bogart obsession does not have a calming effect on Allan, but rather inevitably makes him more insecure. This happens either because he sets his own bar too high – e.g. "Bogart was [also] short. That did not seem to bother anybody" – or because he always ends up on the losing side of the comparison – e.g. "You know who's not insecure? Bogart"; "I am not Bogart, I never will be". Allan's character is paradigmatically neurotic: he keeps engaging in actions (e.g. aggressively approaching women) while being convinced that the reasons behind them are irrational or irreconcilable with his personality – i.e. "This is what Bogart would do". As Killmister would put it, he does not simply act irrationally but takes such irrationality as personality-defining (2015, p. 2).

How may one distinguish between Woody Allen cases and pathological manifestations of hypersalience, such as the interrogative attitude exhibited by some schizophrenic patients? The

contribute to making the account more intelligible. In a later development of the model, these cases could be replaced by patients' reports or structured interviews where the imbalances on the various dimensions could be assessed more precisely.

two situations appear similar in some respects, since neurotic individuals over-attribute salience to objects and events that are trivial from other people's perspective. This preoccupation over insignificant details is shared by those patients – such as Paul (see Minkowski 1923) – for whom everyday objects such as feather dusters or door cracks take on a personal and ominous meaning.[42] Moreover, both neurotic and schizophrenic subjects regard their obsessions as distressing sources of anxiety. Paul's trip to the bathroom takes a long time because he is distracted by all sorts of doubts about the objects around him, while Allan's preparation before a date turns into a raving monologue about loosely related topics. Despite these similarities, significant distinctions can be drawn between the two cases by appealing to the role played by risk and protective factors.

On the one hand, some factors that act as protective in Woody Allen cases appear severely compromised in schizophrenic patients. One paradigmatic example here is *humor*: while self-irony about one's own obsessions becomes an important part of the neurotic personality (as Allen's success as a comedian clearly shows), some kind of humor impairment has been reliably associated with schizophrenia. Recent studies provide evidence in support of this point: schizophrenic patients have a hard time distinguishing between humorous and non-humorous texts (see Ivanova *et al*. 2014), as well as reacting appropriately to humorous stimuli in general (Falkenberg *et al*. 2007). By contrast, neurotic individuals often exhibit a striking tendency towards sarcasm and self-irony that may act as a buffer against pathology. In our example, the more Allan indulges in his own idiosyncrasies and obsession, the more he appears – paradoxically – able to exercise some control over them. In this sense, humor becomes a way in which the neurotic affirms his own self-determination or defines his personality as non-standard. Humor may also act as protective for some individuals by promoting *detachment* from the

[42] See Chapter Two, §1.1. for more details on this case.

stressful situation. Similarly, the ability to exercise irony on one's own character flaws indicates some degree of *self-reflection*. While Paul describes his actions as if they were completely normal – e.g. "It takes a certain amount of my time to look at my watch; I check exactly how the hands are placed" – Allan is painfully aware of the gap between him and others – e.g. "Why can't I develop that attitude? […] Why is it always so complicated? […] Why didn't I see it coming?" This higher degree of self-reflection probably translates into a better *control* over one's situation. Indeed, the neurotic still has the option of doing something about his obsessions (e.g. joking about them; using them as reasons for action) whereas the schizophrenic patient does not have such possibility. Lacking both humor and self-reflection, Paul cannot help but adopting a passive attitude towards his obsessions and thus being painfully subject to them.

On the other hand, the difference between intermediate and pathological cases may be explained by an increase in risk factors. For instance, the neurotic's tendency towards hypersalience appears narrower in *scope* with respect to the interrogative attitude exhibited by schizophrenic patients. While Allan's obsessions are restricted to a specific domain (i.e. his romantic life), Paul's questioning extends to objects completely unrelated with one another and lacks a specific focus (i.e. the clock, the door, the feather duster). However, although a difference in scope may be helpful to draw a distinction in this case, such a rule cannot be applied generally. Indeed, schizophrenic delusions can be remarkably circumscribed, as it happens in cases of erotomania where patients believe that a famous person is secretly in love with them (see Jordan *et al.* 2006). Conversely, neurotic obsessions may be generalized and cover a wide variety of aspects in a person's life – e.g. food, sex, cleanliness, tidiness, etc. Other risk factors such as *frequency* or *urgency* may also be significant in distinguishing between intermediate and pathological cases of hypersalience. Indeed, part of Paul's impaired functioning may derive from the fact that his obsessions are more frequent as well as more urgent, to the point that they cannot

be ignored and he feels forced to act upon them. Notably, increased frequency and urgency may in turn foster a weakening in protective factors. In particular, when obsessive thoughts become too many and hard to ignore, both control and humor are likely to diminish. The individual's resources may become insufficient to deal with increasingly stressful demands (control decreases) and there may be no strength left to distance oneself from the situation to laugh about it (humor decreases). Finally, a similar effect may be produced by *duration*: some obsessions may not be intense nor frequent, but simply last long enough to constantly absorb one's resources up to a breaking point of exhaustion.

Hypo. At the opposite end of the same dimension, sadness clearly qualifies as an intermediate instance of hyposalience. Similarly to its pathological counterpart (i.e. depressive anhedonia), sadness often includes a significant loss of pleasure or motivation in everyday activities and may be accompanied by physiological symptoms such as change in appetite or sleep patterns. The issue of reliably distinguishing between depression and normal sadness is notoriously thorny, to the point that past editions of the DSM included a number of "exclusion clauses". For instance, up to the DSM-IV the definition of Major Depression acknowledged that people who recently experienced the loss of a loved one were to be exempted from diagnosis even upon meeting all the symptomatic criteria (Bereavement Clause). This is one of the few places in which the DSM has recognized that the observation of symptoms does not suffice to produce a correct diagnosis and that the context needs to be taken into account. In their book, Horwitz & Wakefield (2007) take the Bereavement Clause as a starting point to criticize the overly inclusive criteria proposed by the DSM and flag the related risk of medicalizing normal loss responses. They proceed to present a number of cases in which intense sadness can be seen as the normal and at times adaptive response to life stressors – e.g. loss of a valued job, divorce, loss of status. Notably, a reliable distinction between these cases and clinical depression cannot

be drawn solely at the level of symptoms. As Horwitz and Wakefield put it: "Normal sadness can be intense; can be accompanied by sleeplessness, lack of concentration, change of appetite, and so on; can be impairing and distressful; and can last for two weeks, as the criteria demand" (p. 15). On their view, three criteria are helpful to distinguish between ordinary and pathological forms of sadness. First, normal sadness is triggered by a "discrete life event" involving a significant loss (e.g. death of a loved one); second, the level of distress exhibited by the person is "proportional" to the nature of the loss; third, the course of symptoms closely follows the presence, absence or persistence of cause (pp. 27-29). Yet, drawing a satisfactory distinction between ordinary and disordered cases may be harder than Horwitz & Wakefield are willing to admit. With respect to the first criterion, studies indicate that depressive disorders – like normal sadness – are often triggered by life events involving loss or trauma. For example, grief evolves in depression in 10 to 20% of cases (Bonanno & Kaltman 2001) and childbirth frequently gives rise to clinically relevant depressive manifestations (see Williamson & McCutcheon 2004 for a review). The second criterion (i.e. proportionality to the cause) appears difficult to assess without taking into account specific details of the person's life situation: for example, the death of a pet may count as a devastating loss to someone whose social life is otherwise seriously impaired. The third criterion is also problematic, as the chronic presence of stressors in a person's social and financial life have been proven to impact depression rates significantly. For example, marital struggle and dissolution (Simon 2002) and socio-economic status (Costello *et al.* 2003) have been reliably associated with the development and exacerbation of depression. Contrary to Horwitz & Wakefield's suggestion, in many cases environmental factors can be considered additional triggers of psychopathology rather than conditions grounding the exclusion from clinical status.[43]

[43] In Chapter Five I explore this issue in more detail to substantiate the idea that environmental factors may create or increase vulnerability to pathology.

Does the model proposed here do a better job than Horwitz & Wakefield's at distinguishing between sadness and clinical depression? In terms of risk factors the two conditions appear remarkably similar: indeed, both sadness and depression can persist over long periods of time due to the presence of a chronic stressor (*duration*). Moreover, both conditions may exhibit high intensity and absorb a great part of the person's physical and mental resources (*strength*). However, it makes sense to think that in most cases depression and sadness would importantly differ in terms of *depth* and *scope*. While depression casts a gloomy light on the person's overall existence, sadness tends to be more localized and connected with some specific trauma or negative life event. This difference might also be reflected in the fact that depressed patients often talk about pleasure having completely disappeared from their life (Styron 1991, p. 38), while sad individuals usually recall pleasurable events vividly but are unable to reproduce the same feeling due to the circumstances. The dual factor of *imagination* also seems to play a negative role in depressed cases: indeed, the painful feeling of finality described in first-person accounts may originate from the subjects' inability to depict a future different from the present. Styron characterizes this sensation poignantly: "The pain is unrelenting, and what makes the condition intolerable is the foreknowledge that no remedy will come – not in a day, an hour, a month, or a minute. If there is mild relief, one knows that it is only temporary: more pain will follow" (1991, p. 62). Plath describes her future in a similar way: "I could see day after day after day glaring ahead of me like a white, broad, infinitely desolate avenue"; "I saw the years of my life spaced along a road in the form of telephone poles, threaded together by wires. I counted one, two, three…nineteen telephone poles, and then the wires dangled into space, and try as I would, I could not see a single pole beyond the nineteenth" (1963, pp. 65 & 67). One way to distinguish intermediate from disordered cases would thus be to assess the person's outlook on her imagined future and see whether imagination acts as a risk or protective factor. In terms of purely

protective factors, the degree of physical and mental *strength* exhibited by the individual plays an important role. Intermediate cases may be seen as not having (yet) crossed the point of exhaustion and therefore as able to mobilize some resources in pursuit of goals, whereas pathological cases would be characterized by a state of energy depletion where carrying out small everyday tasks becomes impossible. Again, Plath and Styron describe this experience vividly: "I guess I should have been excited the way most of the other girls were, but I could not get myself to react" (Plath 1963, p. 3); "I fell onto the bed and lay gazing at the ceiling, nearly immobilized and in a trance of supreme discomfort. Rational thought was usually absent from my mind at such times, hence *trance*. I can think of no more apposite word for this state of being, a condition of helpless stupor" (Styron 1991, pp. 17-18). Finally, the degree of *abreaction* or *discharge* might make a significant difference between intermediate and pathological cases. Indeed, some people may be able to release tension effectively through various forms of social activities, talk-therapy or physical movement, whereas others may lack the environmental or psychological resources to do so.

§2.2. Confidence: Impostor Syndrome and Overconfidence Effect

Hypo. An intermediate case lying on the confidence dimension is the so-called *impostor syndrome* (IS henceforth), characterized as a collection of feelings and beliefs about one's inadequacy that tend to persist even in the face of contrary evidence (see Young 2011). This phenomenon is widespread among successful and over-achieving individuals who should have plenty of evidence at their disposal to counter self-doubt and be reassured about their talents. Yet, individuals affected by IS tend to systematically dismiss every piece of evidence pointing toward the fact that they are worthy of their success. Commenting on the promotion they received, they might say that they just got lucky; talking about their impressive achievements, they claim: "If I

made it, anybody can". In her book, Young dubs this phenomenon *intellectual fraudulence*: "You believe you are somehow 'fooling' other people into thinking you're brighter and more capable than you 'know' yourself to be. Deep down you feel like an impostor, fake, and fraud" (2011, p. 71). Kay & Shipman (2014) also stress this point repeatedly in their study on the "confidence gap" between successful women and men across various professions. While their work started as a collection of success stories of women across the United States and the United Kingdom, it soon acquired a different focus:

"To our surprise, as we talked with women, dozens of them, all accomplished and credentialed, we kept bumping up against a dark spot that we could not quite identify, a force clearly holding them back. Why did the successful investment banker mention to us that she did not really deserve the promotion she'd just got? What did it mean when the engineer who'd been a pioneer in her industry for decades told us offhandedly that she wasn't sure she was really the best choice to run her firm's new big project?" (pp. 3-4).

For reasons that still need to be fully explored, women suffer from IS more than men do, and this may reflect on other gender gaps in the workplace. For instance, a study conducted by *Hewlett-Packard*'s Human Resources reports that women tend to apply for a promotion only when they believe to meet 100% of the qualifications, whereas men apply when they think they can meet 60% of the job requirements (reported by Kay & Shipman 2014, p. 10). In this sense, hypoconfidence may connect with reduced risk-taking behavior and consequently with shrinking opportunities in the workplace (p. 19).[44]

Is there a way to pry apart cases of IS and cases of self-doubt characteristic of depression, such as the ones exemplified by Styron (1991) and Plath (1963)?[45] The two situations appear similar in many respects: the women interviewed by Kay & Shipman report recurring feelings of

[44] See Briñol, DeMarree & Petty 2010; Estes & Felker 2011 for more studies on this phenomenon.
[45] See Chapter Two, §1.1.

inadequacy and believe they do not deserve the recognition they are given. Similarly, Styron interprets losing the check as a sign of his worthlessness with respect to the literary prize he had won (p. 19), and Plath starts feeling "dreadfully inadequate" for not knowing more languages (p. 40). Indeed, the same thoughts, feelings and beliefs may be ascribed to individuals affected by IS and depression: "I feel like a complete failure"; "I do not deserve any of the recognition I am getting"; "It is just a matter of time before I am exposed as a fraud". Yet, it is possible to draw some distinctions between the two cases by appealing to the role played by risk and protective factors.

On the one hand, the self-doubt experienced by depressed individuals usually differs in *scope* with respect to the one reported by IS sufferers. While the latter struggle with their feeling of professional ineptness, the former tend to take *any* episode of inadequacy as a sign of their all-encompassing failure. Both Styron and Plath make this point particularly vivid: "I was not worthy *of the prize*. I was in fact not worthy *of any of the recognition* that had come my way in the past few years." (Styron 1991, p. 19. Italics mine). "I could not speak *German* well […] I started adding up *all the things I couldn't do*" (Plath 1963, p. 40. Italics mine). Other candidate risk factors are likely to differ between the two situations: the self-doubt experienced by depressed patients may last longer (*duration*), be triggered more easily (*frequency*), or take up a greater amount of time and effort (*intensity*). On the other hand, the difference between intermediate and pathological cases lies in the presence or absence of some protective factors. For instance, people affected by IS may be able to cope with their feelings of inadequacy by acquiring a more objective view on their accomplishments (*detachment*) or by being self-ironic about their own insecurity (*humor*). Some may also consciously adopt a number of coping strategies including self-reassuring beliefs – e.g. "You can make it, you're good at this" – cognitive changes – e.g. "It is not the end of the world if someone else gets the job" – or attentional deployment – e.g. "Now

that I have turned in the application I can focus on other tasks". All these strategies help to gain *control* on the paralyzing thought of being a failure and also restore some confidence connected with action. If my past record shows that I have good chances of getting a promotion, I might as well try this time. If I can laugh about my constant fear of being rejected, I can finally see that it is not justified after all. If I can psych myself up enough to turn in my dossier, I enhance my chances of actually succeeding. Notably, these protective factors appear unavailable to depressed patients. Many of them feel irremediably trapped with no possibility of detaching themselves from the situation: "It is as if a black fog has descended" (Ratcliffe 2015, p. 263); "I've gone around for most of my life as in the rarefied atmosphere under a bell jar" (Plath 1963, p. 130). Laughing often becomes impossible – recall Styron's "failure of even forced laughter" (p. 19) – and coping strategies are at times attempted but completely ineffective – Plath keeps repeating to herself to no avail: "I was supposed to be having the time of my life" (p. 3).

Hyper. At the opposite end of the confidence dimension there are cases of people that tend to overestimate their abilities with respect to their own and others' performance. This phenomenon has come to be known in psychology as the *overconfidence effect* (see Pulford 1996 for a review). Early studies on the topic focused on specific abilities, such as driving: for example, Svenson (1981) surveyed 161 Swedish and American undergraduates, asking them to compare their driving skills to other people. The results show a striking tendency towards overconfidence, with 93% of the American sample and 69% of the Swedish sample placing themselves in the top 50% (p. 146). Other studies investigate more general abilities, such as self-assessment on academic performance. The results are – again – striking: 68% of the faculty surveyed at the University of Nebraska rated themselves in the top 25% for teaching ability, and more than 90% rated themselves as above average (see Cross 1977). Similarly, 87% of students interviewed at Stanford University rated their academic performance as above the median (see

Zuckerman, Ezra & Jost 2001). In their seminal work on this phenomenon, Kruger & Dunning (1999) then uncovered a positive correlation between tendency towards overconfidence and lack of skill. Roughly put, the poorer the participants scored in a number of domains (i.e. humor, grammar, and logic), the more they rated their performance as being above average. The researchers suggest that this phenomenon may reflect a metacognitive deficit, where people who are incompetent also lack the means to detect their own mistakes. As Dunning & Kruger put it: "Incompetence does not only cause poor performance but also the inability to recognize that one's performance is poor" (p. 1130).[46]

Another extreme but non-pathological case lying on the confidence dimension is the *pep-talk* employed by athletes to focus on their strengths and counter performance anxiety or fear of failure. During training or before an important game, it is common for athletes to psych themselves up by repeating exaggerated remarks (e.g. "You're the best! You can do this!") or by projecting unrealistic expectations on the upcoming performance (e.g. "You know you're going to win! They are afraid of you"). This motivational strategy have been shown to have positive effects in many situations, enhancing confidence and self-efficacy (see Vargas-Tonsing 2009). A paradigmatic example in this sense is the boxing champion Muhammad Alì, who famously claimed: "I am the greatest. I said it even before I knew I was". This case is particularly interesting for our purposes because it represents a situation lying very close to one extreme (i.e. grandiosity delusion), despite being non-pathological and even adaptive in some circumstances.

What distinguishes then moderate and severe cases of overconfidence from delusions of grandiosity such as the ones reported by some schizophrenic patients (see Reina 2009)?[47] Again, drawing such a distinction based on the content of the beliefs or on the feelings involved seems wrongheaded. Indeed, people who tend towards overconfidence and those affected by grandiosity

[46] For more studies on the overconfidence effect see Sutherland 2007; Gigerenzer 1991.
[47] See Chapter Two, §2.1. for more details on this case.

similarly believe to be more skillful than others in a number of domains. Moreover, the feeling of omnipotence derived from perceived or actual success may not differ substantially: schizophrenic patients often claim to have been chosen by God to bring salvation to the world (Schreber 1903), while John Lennon famously said: "The Beatles are bigger than Jesus" (1966). Yet, the distinction between intermediate and pathological cases becomes easier to draw once we turn to the role played by risk and protective factors. Like in hypoconfidence cases, *scope* plays an important role: ordinary people appear overconfident in selected areas (e.g. driving), and individuals that are extremely talented in one domain (e.g. John Lennon or Muhammad Alì) do not necessarily transfer their confidence to other aspects of their life. Conversely, delusional patients experience grandiosity as pervasive and as applying to every domain: for example, Reina interpreted every gesture, glance, word and object as referring to him *because of* his magical powers (2009, pp. 4-5). The degree of *intensity* also appears to differ: in most cases, ordinary overconfidence does not possess depth because it does not impact people's decisions or worldview to a significant extent. It also fails to possess strength because ordinary people's resources are not constantly mobilized to deal with the consequences of overconfidence. On the contrary, patients affected by grandiosity perceive their special powers as something that needs to be understood, explored and developed further: "My teachers began to take an even greater interest in me, and I thought I had found out why. I had discovered how to charm people into liking me using my color theory" (Reina 2009, pp. 3-4). At the same time, they perceive this phenomenon as deeply distressful: "Everyone I passed on the street would give a personal comment about me or insult me" (p. 5). Notably, extreme non-pathological cases may also score very high in terms of intensity. For example, Muhammad Alì's conviction of being "the greatest" probably played a crucial role in some of his life-defining decisions, such as accepting to fight against the much-stronger Foreman despite being at the end of his career. Ultimately, the action

of protective factors may be the only element separating extreme non-pathological cases from disordered ones. One important difference may lie in the ability to *control* one's overconfidence and express it full-force only when it really matters – e.g. right before a match – as opposed to being passively subject to it – e.g. "My magical power began to run away with themselves" (Reina 2009, p. 4). Moreover, the ability to use *humor* appropriately may turn a stressful situation in an advantageous one. During the legendary fight against Foreman, Alì took punches for seven rounds and looked like he was going to be defeated by knockout. But, as Foreman reports: "About the seventh round, I hit him hard to the jaw and he held me and whispered in my ear: 'That all you got, George?' I realized that this ain't what I thought it was" (Gast 1996). The option of employing humor or self-irony as protective seems unavailable in pathological cases, where people are incapable of detaching themselves from their experience and observe it from an external perspective: e.g. "Every bump on the wall of my apartment and every cry in the distance were directed toward me" (Reina 2009, p. 5).

§2.3. Familiarity: Post-Bereavement Hallucinatory Experiences & Dissociation

Hyper. One intermediate case lying at the high end of the familiarity dimension is the group of phenomena that has come to be known as Post-Bereavement Hallucinatory Experiences (PBHE). Although their nature remains somewhat elusive, these phenomena are usually taken to encompass "a heterogeneous group of disturbances of perception and thought process, ranging from hallucinations, pseudo-hallucinations, illusions, and felt-presences" (Castelnovo *et al.* 2015, p. 271). PBHE are often experienced by family members or close friends during – but also after – the mourning period, and may take the form of visual or auditory hallucinations, felt presence of the lost person, and attempts to communicate with her through talking or rituals. In his pioneering study on the psychology of grief, Parkes (1972) distinguishes between the phenomenon of "felt

presence" (or illusion) and full-blown hallucinations. The former experience is characterized by feeling the presence of the dead person without directly perceiving him or her. In these cases, subjects would say things like: "Spiritually, he's near"; "I still feel that he's around"; "He's not anywhere in particular, just around the place" (p. 61). The latter experience instead includes a strong perceptual component, as one subject reports: "He's with me all the time. *I hear him and see him* although I know it's only imagination" (p. 62. Italics mine). The subjects interviewed by Parkes report visual hallucinations as being particularly frequent: "One widow was resting in her chair on a Sunday afternoon when *she saw her husband*, quite clearly, digging in the garden with only his trousers on; another *saw* her husband coming in through the garden gate; a third *saw* her dead father standing by her bed at night" (p. 62). The first quantitative study on PBHE has been conducted by Rees (1971) on a sample of Welsh widows (n = 227) and widowers (n = 66). The goal was to assess the presence and frequency of hallucinatory experiences following the death of a spouse, while excluding metaphorical references (e.g. "I can see him with my mind's eye"), dreams and half-waking perceptions. The numbers resulting from these interviews are surprisingly high: "Almost half of the people interviewed had hallucinations or illusions of the dead spouse – i.e. 46.7%" (p. 38). Among PBHE, illusions turned out to be particularly common, with 39% of the widowed reporting a sense of presence of the dead spouse; visual and auditory hallucinations have also been reported by 14% of the sample. Most of these phenomena appear to last many years, and are particularly common during the first ten years of widowhood (p. 37). However, passing of time usually marks a decrease in frequency: for example, visual hallucinations were reported by 4.5% of those widowed for more than 20 years, and by 17% of those widowed for less than 20 years. Notably, most of the widowed subjects interviewed by Rees presented PBHE in a positive light, as a "helpful accompaniment of widowhood" (p. 37). The numbers in this sense are particularly striking: "Altogether 78% of the visually hallucinated,

66% of the auditorily hallucinated, 73% with illusions of the deceased's presence, and 82% of those who spoke to the dead person found it helpful" (p. 40). More recent studies on PBHE confirmed the high incidence of the phenomenon among widowed subjects: a comprehensive meta-analysis carried out by Castelnovo and her collaborators (2015) has uncovered numbers ranging from 30% to 60%.

At first sight, it seems difficult to pin down the differences between cases of PBHE and the experience reported by patients affected by Fregoli delusion. Indeed, some of the case studies discussed in Chapter Two also involve widows who misidentify hospital staff with their dead husbands (Moriyama *et al.* 2007; Turkiewicz *et al.* 2009). In one sense, this experience closely resembles the phenomenon of felt presence: while perception remains somewhat intact – i.e. "I don't see him or hear him" (PBHE); "He looks different" (Fregoli) – the resulting judgment is significantly altered – i.e. "I know he's physically present" (PBHE); "I know they are the same person" (Fregoli). Moreover – in PBHE as well as in Fregoli – hyperfamiliarity appears to arise from hypersalience: due to the dead person's affective importance, a subject might start "seeking" his presence in the environment and this may give rise to "finding" behavior even in absence of the object sought (see Parkes 1970). More generally, both in PBHE and Fregoli the degree of emotional attachment to the person who gets misidentified plays an important role. Not only she is a family member or friend, but often someone who has acquired particular importance because of a traumatic event (e.g. death, break-up) or due to a particular obsession (e.g. recall the woman who was seeing her favorite actress everywhere – Courban & Fail 1927). In terms of risk factors involved, the two situations also appear strikingly similar: in fact, *duration* may be even longer in PBHE, since the process of bereavement tends to last several years as opposed to acute psychotic episodes that often disappear in a few weeks or months (see Moriyama *et al.* 2007). *Frequency* also does not help to distinguish PBHE from the misidentification errors occurring in

Fregoli. Indeed, the studies discussed above indicate that PBHE episodes may occur at "variable times throughout the day", with some people being "continually hallucinated" and others experiencing these phenomena more often at a particular time of the day – e.g. morning or night (Rees 1971, p. 40). The degree of *intensity* is probably very high in both situations, with one crucial difference: patients affected by Fregoli tend to feel persecuted, haunted or followed by their loved ones, whereas people who experience PBHE perceive these phenomena as helpful and even pleasant. In this sense, intensity could be high in both situations but exhibiting a positive valence in the non-pathological case and a negative one in its pathological counterpart. Such a difference in valence might indicate that people experiencing PBHE benefit from the action of several protective factors. For instance, some reports indicate that these individuals are able to maintain a certain degree of *control* over the situation, by telling themselves that they are just imagining, or that what they have seen must not be real. For example, one of the subjects interviewed by Parkes reports: "I hear him and see him although I know it's only imagination" (1972, pp. 61-62). Another indicator that PBHE subjects may exhibit a higher level of control with respect to Fregoli patients is that the former are able – at least most of the time – to regard these experiences as private and as something that can be integrated in their everyday life without disrupting it. Unlike psychotic patients, widowed subjects are "able to integrate the experience and keep it secret" (Parkes 1972, p. 41). In this sense, PBHE experiences probably work as coping strategies themselves, as they help fulfilling a frustrated desire – e.g. to see the loved one again – and function as ways of abreacting or processing some otherwise stressful content – e.g. by talking with the deceased spouse.

Hypo. At the low end of the familiarity dimension there are mild and moderate experiences of *dissociation*. These phenomena take a variety of forms but they usually involve a sense of detachment from people and situations that would be otherwise familiar to the subject. In

more severe cases, dissociation is characterized by a state of fragmented consciousness involving amnesia and feelings of being disconnected from oneself and one's environment (see Steinberg 2001). Although dissociative experiences often go undetected because of their subtle, fleeting and elusive character, they are in fact strikingly common (see Goleman 1991). Frequent instances include gaps in awareness, such as realizing afterwards that we have not followed part of a conversation or that we have become so engrossed in a book to lose track of time. Depersonalization symptoms are also common: in one of the first systematic studies on the topic, 29% of the participants reported that they occasionally see themselves as characters in a movie, and 14% had the experience of not recognizing themselves in the mirror (Ross, Norton & Anderson 1988). This sense of detachment from one's own body and mind becomes particularly salient when people are involved in life-threatening events. For example, a woman who fell from a third-story balcony describes her accident this way: "I experienced it as if I were standing on another balcony watching a pink cloud float down to the ground" (Goleman 1991, p. 1). Similarly, people who have been involved in natural disasters or terrorist attacks often report a sense of being strikingly detached from their environment, feeling like they were dreaming or "going through the motions" without any emotional reaction (Steinberg 2001, pp. 68-72). These examples suggest that dissociative experiences may be part of an adaptive response to traumas and extremely stressful situations. Depersonalization may act as a survival mechanism that allows people to react efficiently in danger situations by shutting down emotions and looking for an immediate solution (e.g. finding a way out of a building on fire).

Besides cases in which dissociative symptoms are automatically triggered, there are also practices where states of altered awareness are actively promoted as intense spiritual moments. In her ethnographic work on American spirituality, Luhrmann (2005) draws a parallel between dissociation and the capacity of *absorption* promoted by Christian groups such as Pentecostals

and Southern Baptists. These communities put a lot of emphasis on the intimate relationship that the believer establishes with God through prayer. Through a series of exercises, believers learn to become absorbed in inner sensory stimuli (e.g. finding God's voice within one's thoughts) while gradually losing awareness of external ones (p. 148). This is how one of the subjects interviewed by Luhrmann describes her experience: "You're just sort of removed from what's going on around you. Very often, the words that are coming out of your mouth aren't your own words, or the picture that you're describing isn't something you're conjuring up out of your own brain. It's like, you know, an image is put into your mind and words are put into your mouth that you're just meant to speak out. So there is that sense of being *disconnected from yourself* in some capacity" (p. 152. Italics mine). This wide range of cases suggests that alterations along the dimension of familiarity should be seen as lying on a continuum, with fleeting episodes of absorption on the one hand (e.g. becoming engrossed in a book or movie) and more frequent or intense experiences on the other hand (e.g. losing touch with one's environment in order to contact God through prayer).

There are important similarities between these cases of dissociation and disorders of hypofamiliarity such as Capgras delusion. During episodes of depersonalization, subjects experience a puzzling lack of familiarity with respect to objects, people, or situations that they in fact know well. For example, Steinberg reports the case of a bank manager who has unexpectedly been fired a few years before retirement: "I looked at my boss, and suddenly I didn't know him anymore. He didn't look real. The room looked strange and unreal too – almost like being out of focus" (2001, p. 73). In particularly stressful situations, subjects experiencing depersonalization might even have trouble recognizing close family members. One of the subjects interviewed by Steinberg fails to recognize her mother during a heated argument: "She looks like a stranger to me, and I hear myself thinking: 'Who is this person? She is not my mother. *This person is not my*

mother'" (p. 74). With respect to the thought content expressed, these cases look strikingly similar to the Capgras patients who are unable to recognize their spouses, sons or daughters (see Christodoulou 1977, Frazer & Roberts 1994).[48] Another element of similarity between dissociation and Capgras is the gap between perceptual and emotional recognition. While subjects are aware that the person in front of them looks exactly like their significant other, they fail to feel the affective response connected to this perception. As Steinberg puts it: "There's an awareness that a familiar emotion is lacking in the normal experience of Mother. The person recognizes Mother intellectually, but the emotion is withdrawn. The person is aware that the usual emotions associated with Mother – feelings of love, closeness, comfort, and security – are missing" (2001, p. 71).

Despite these similarities, the distinction between pathological and non-pathological cases hinges on quantitative factors such as the frequency, duration, and intensity of the relevant episodes. This is exactly what Steinberg had in mind when she started developing her structured interview for dissociative disorders, now widely known as SCID-D (*Steinberg Clinical Interview for DSM-IV Dissociative Disorders*). On her view, the core distinction between mild and severe forms of dissociation turns on the idea that non-pathological experiences would be "brief", "rare" and with a "minimal effect" on people's ability to function socially or professionally. By contrast, disordered cases would exhibit episodes that are "persistent, recurrent and disruptive to social relationships or job performance" (2001, p. 12). On the model outlined here, the relevant risk factors that would help us distinguish between normal and pathological cases would thus be frequency, duration, and intensity. In terms of protective factors, one important difference might have to do with the degree of *control* that a subject is able to exercise on different aspects of her personality. While we all play different roles in different situations (e.g. at home and at work)

[48] See Chapter Two, §3.1. for more details on the Capgras cases.

patients affected by a dissociative disorder are not able to predict in advance which "alter" would make her appearance. For example, one of Steinberg's patients painfully describes her inability to control a particularly aggressive and demeaning alter (i.e. The Mean) who comes out in situations of emotional conflict (p. 148). A great part of Steinberg's therapy thus consists in helping the patient learn to discuss and negotiate with her alters, by actively listening to them and regulating their intervention. This form of dialogical therapy would also provide the patient with a higher degree of *abreaction*, because she would eventually learn to work through identity conflicts by talking to different parts of her personality as if they were different people.

§2.4. Agency: False Confessions and Mind Wandering

Hyper. One intermediate case lying on the agency dimension is the phenomenon of false confessions, where an individual confesses to a crime that he or she has in fact not committed (see Gudjonsson 2003). These cases straightforwardly qualify as instances of hyperagency because someone who falsely confesses to a crime takes responsibility for an action that extends beyond her control. The idea of a non-mentally disordered person willing to face legal charges for something she has not done appears very counterintuitive. Yet, studies in forensic psychiatry show that false confessions are relatively frequent, although their exact number is obviously difficult to determine. For example, in the early Eighties 10% of the defendants assessed in Birmingham and 24% of those in the London pleaded "not guilty" at their trial after having provided the police with a written confession (see Gudjonsson 2003, p. 184). In his extensive work on the topic, Gudjonsson shows that false confessions are not confined to the mentally ill and that "the view that apparently normal individuals would never seriously incriminate themselves when interrogated by the police is wrong" (p. 243). A famous case of false confession is the one involving Peter Reilly, an 18-year-old from Connecticut who found his mother dead

upon returning home and – following police interrogation – came to believe to have murdered her. One important fact preceding Peter's confession was his "failure" to pass the polygraph test, after which both his confidence and memory began to weaken: "Now there is doubt in my mind. Maybe I did do it"; "The polygraph thing didn't come out right. It looked like I've done it" (Connery 1977, p. 66). Peter exhibited no evidence of cognitive or psychiatric disorder, and his IQ placed him in the "bright-normal" range. During the trial he was finally found to be innocent, and all charges against him were dropped (Gudjonsson 2003, p. 236). Why would someone like Peter confess to a crime that he has not committed? The answer to this question is quite complex. First, the coercive methods used by the police play an important role, as well as the conditions in which the custodial confinement occurs – e.g. sleep-deprivation, under- or over-stimulation, inadequate diet and physical discomfort (see Hinkle & Wolff 1956). Some interrogation techniques even appear to elicit memory distrust and distortion when combined with situations of emotional shock or extreme stress. Second, false confessors usually exhibit a set of personality traits that make them particularly vulnerable to suggestion: lack of self-confidence, exaggerated trust in authority, eagerness to help and difficulty to detect discrepancy between what is recalled and what is suggested (see Ofshe 1989). Forensic psychologists group false confessions into three categories: a) *voluntary*, where one spontaneously confesses without being interrogated, either to protect someone else or for pathological reasons – e.g. self-punishment, reality distortion; b) *coerced-compliant*, where one confesses as the result of an interrogation to obtain some immediate gain – e.g. escape from an intolerable situation, having one's sentence reduced; c) *coerced-internalized*, where one confesses as the result of an interrogation because he comes to believe to have committed the crime (Gudjonsson 2003, pp. 192-195). C) cases are the most

interesting for our purposes because they constitute a non-pathological example of hyperagency while at the same time resembling the instances of pathological guilt discussed in Chapter Two.[49]

What makes false confessions importantly different from the voluntary confessions of psychotic patients? Again, there are some striking similarities among the two situations: in both cases, a subject falsely – although sincerely – comes to believe to have committed an action that falls beyond his control, and takes moral as well as legal responsibility for it. A strong feeling of guilt features in both kinds of confessions. Psychotic and depressed subjects often feel the need to expiate real or alleged transgressions (p. 195), whereas others feel guilty for not having been in control when the crime was committed (e.g. because of alcohol or drug intoxication), or for not being able to trust their memory in recalling events without confusion (p. 238). Despite these similarities, mentally disordered subjects appear to exhibit a pre-existing feeling of guilt that makes some actions particularly salient (e.g. Emmy von N), while false confessors experience guilt after having lost confidence about their ability to recollect what happened. As a consequence, the degree of internalization with respect to the confession also differs: while voluntary confessions are spontaneous and rarely retracted, coerced confessions are almost invariably taken back by the subject even if the timing of retraction varies from a few hours to several years (p. 182). In this sense, the factor of *duration* can be taken as a reliable indicator to distinguish between pathological and non-pathological cases: the least pressured and the hardest to retract the confession, the higher its pathological import. This point also allows us to describe a number of borderline cases: some false confessions may be characterized as transitory mental disorders from which people recover soon after the stressful situation has ended, while other – much longer – processes may indicate that the person has crossed the threshold to pathology. Pathological and non-pathological cases appear to markedly differ also in terms of *urgency* and

[49] See Chapter Two, §4.1.

intensity. For instance, psychotic subjects voluntarily contact the police and appear distressed for having committed the crime in question ("I did it"; "It was me"), whereas false confessors initially proclaim their innocence and then come to confess in a tentative fashion ("I must have done it"; "I think I did"). Protective factors such as *strength* and *control* play an important role as subjects often appear to confess after a prolonged period of physical discomfort and psychological stress. Gudjonsson describes the process as follows: "The forces pushing people towards confessing are strengthened (e.g. persuading people that it is in their own interest to confess, that there is substantial evidence to link them to the crime) whilst forces maintaining resistance are weakened (e.g. by tiredness, lack of sleep, exhaustion, emotional distress)" (2003, p. 189). In this sense, another difference between pathological and non-pathological cases may lie in the degree of effort required by the subject to regain a sufficient level of strength and control over the situation. In some cases, the state of confusion and memory distortion leading to the false confession would fade quite easily, while in others the recovery process may take longer or fail to occur at all.

Hypo. Phenomena like distraction, daydreaming or mind wandering are extremely common and part of our everyday experience. We are working on an important project and we suddenly start thinking about the grocery list or our plans for the evening. We try to concentrate on a task, when memories pop up and absorb us for some time before we are able to resume our previous activity. In most cases these thoughts arise automatically and are very difficult to regulate: they can thus be seen as paradigmatic cases of *hypoagency*. Despite their pervasiveness in our ordinary life, phenomena of mind wandering have become the object of systematic scientific investigation only recently, mostly due to the growing number of neuroimaging results showing that the brain is active also in rest conditions. This neural pattern has come to be known as the Default Mode Network (DMN henceforth) and its discovery suggests that mind wandering

might constitute a psychological baseline from which people depart when engaging in demanding tasks and to which they return when their attention is not requested elsewhere (see Mason *et al*. 2007).[50] Although cases of excessive mind wandering have been at times granted pathological status (see Schupak & Rosenthal 2009), this phenomenon has also been associated with an increase in creativity and problem-solving abilities. Indeed, the neural profile of brains in DMN is similar to the one exhibited by subjects engaged in conceptual processing and problem-solving tasks (see Smallwood & Schooler 2006). In the past decade, researchers working in different fields – philosophy of mind, psychology and neuroscience in particular – have attempted to shed light on the nature of mind wandering while formulating hypotheses on its adaptive value. Metzinger (2013) characterizes mind wandering as a form of "mental autonomy loss" because of its spontaneous, automatic and task-unrelated nature. The notion of mental autonomy proposed by Metzinger partially overlaps with what I call agency, and comprises the ability to causally determine one's actions as well as the ability to control the conscious content of one's mind. Due to the ubiquitous interruptions caused by mind wandering, Metzinger suggests that mental autonomy should be regarded as "the exception rather than the rule" (p. 5). On his view, mind wandering would then be adaptive because it allows individuals to maintain a baseline arousal activity where past, present and future episodes hang together in a virtually unitary whole. Similarly, Smallwood & Andrews-Hanna (2013) argue that mind wandering has a number of positive effects on psychological functioning, such as consolidating memories by connecting past and present self, planning future events and delaying gratification. This activity thus grants the mind some freedom from the "here and now" and allows agents to perform mental actions that are not simply responses to the outside world (p. 4). If this is correct, it becomes easier to see how mind wandering might be connected to creative and problem-solving processes. Similarly to

[50] See also Andrews-Hanna 2012 for a comprehensive review of the neuroscientific studies on DMN up to date.

what happens in dreaming, our thoughts would drift away from external stimuli and focus on internal ones, thereby facilitating the integration of unassociated information and the formulation of original solutions to previously encountered problems. In a recent study on the topic, Baird *et al.* (2012) assigned the Unusual Uses Task (UUT) to 145 participants, asking them to generate as many uses as possible for a common object – e.g. a brick – in a given amount of time. After having read the list of objects, three groups of participants were subject to an incubation period during which some subjects were administered a demanding task, others an undemanding task and still others were allowed to rest; a fourth group proceeded to solve the problem without taking a break. The results indicate that participants engaging in the non-demanding task during the incubation period performed significantly better than the ones who were assigned a demanding task, no task at all or that did not have an incubation period (p. 5). According to Baird and colleagues, engaging in a simple task allowed participants to mind wander during the incubation period and this in turn helped them formulating more creative solutions to the UUT.

As I suggest above, mind wandering can be regarded as a case of hypoagency for a number of reasons. First, it typically starts out as an automatic and spontaneous mental phenomenon over which we have little control (e.g. when we are reading a text and some unrelated memories and thoughts catch our attention). Second, we often have a hard time accounting for the content of thoughts generated during mind wandering (e.g. when a song is stuck in our head and we have no idea where it came from).[51] What makes these cases different from pathological phenomena of hypoagency such as intrusive thoughts or auditory verbal

[51] One might argue that even in these milder cases agency is impaired (e.g. we can't get rid of the song stuck in our head even if we try). Even granting this point, there seem to be different degrees of severity at play. In the song case the functioning of the agency dimension is somewhat preserved: for instance, we normally perceive the tune as "popping up from nowhere" but *not* as externally generated or inserted by someone else in our mind. By contrast, in pathological cases – such as Auditory Verbal Hallucinations (AVH) – the agency dimension ceases to play its functional role and we completely lose the sense of what is self-generated and within our boundaries. Of course, in extreme cases this could also apply to songs, provided that the impairment goes so far that the song is then perceived as externally generated, inserted, implanted, etc. I would like to thank Peter Langland-Hassan for allowing me to clarify this point.

hallucinations (AVH)? Again, the two groups of phenomena appear very similar in terms of factors such as *duration* and *frequency*. On the one hand, intrusive thoughts can be pervasive, with subjects experiencing them up to hundreds of times per day (see Rachman & De Silva 1978; Purdon & Clark 1994). Similarly, patients affected by AVH report that the experience of voice hearing becomes particularly distressing when the voices grow in number and intensity, acting like a "running commentary" of one's life (Longden 2013). On the other hand, researchers studying mind wandering indicate that subjects "spend almost half a day engaged in the experience" (Smallwood & Andrews-Hanna 2013, p. 1) or even "roughly two thirds of their lifetime" (Metzinger 2013, p. 6). A crucial difference between the two cases may be the person's capacity to exercise a certain degree of *control* on the phenomenon. For instance, Smallwood & Andrews-Hanna (2013) suggest that task-context – i.e. how demanding is the activity one is engaging in – might heavily influence the nature of the mind wandering episode, making it adaptive or disruptive as a result. Roughly put: when we are engaging in a relatively non-demanding task, the experience of mind wandering is likely to be less disruptive and more conducive to positive outcomes (e.g. creative solutions) because our mental resources need not be fully absorbed in the completion of the task at hand. Conversely, when the current task requires our undivided attention an episode of mind wandering may qualify as an unwelcome and distressful interruption. Therefore, one's ability to control or regulate the context in which mind wandering episodes occur appears to play an important role in well-being. For example, one might learn to confine mind-wandering to non-demanding situations – e.g. washing dishes – while fending it off from demanding ones – e.g. work or study. This ability to control the context in which mind wandering episodes arise could also be improved through mindfulness and meditation techniques targeted to sharpen one's focus on particular thoughts (see Teasdale 1999; Naranjo & Schmidt 2012). Notably, this process of gaining control over internally generated

thoughts and speech acts is similar to the one described by recovering AVH patients. For instance, Longden (2013) learns to incorporate the voices in a larger autobiographical narrative and starts regarding them as neglected parts of her self. Similarly, one of the patients treated by Romme & Escher (2013) talks about setting boundaries and being able to push back the unwanted intrusions to a later time (p. 263). The ability to exercise a certain degree of control within a paradigmatically uncontrolled activity might thus be crucial to distinguish between ordinary – or adaptive – cases of mind wandering and their pathological counterparts. Studies on creativity have consistently shown that original solutions to problems are more likely to arise when people are given some unstructured time to mind wander (see Dijksterhuis & Meurs 2006). As the comedian John Cleese explains: "If you just keep your mind resting against the subject in a friendly but persistent way, sooner or later you will get a reward from your unconscious, probably in the shower later. Or at breakfast the next morning, but suddenly you are rewarded, out of the blue a new thought mysteriously appears" (1991).

Conclusion

In this chapter I set out to accomplish two goals. First, I rendered the notion of vulnerability to mental disorders more precise by distinguishing between three types of vulnerability. Second, I employed this distinction to explore a number of intermediate cases lying at various points of the dimensions outlined in Chapter Two. All the cases discussed are summarized in the table below (see Table 2). After presenting each case, I illustrated the core similarities and differences with respect to its pathological counterpart. By doing so, I contributed to disentangle the relationship between intermediate and disordered cases while focusing on the role played by risk and protective factors.

This discussion also served two more general purposes in the defense of a dimensional account of mental disorders. On the one hand, I showed that drawing the line between normality and pathology on the basis of gradual differences is a feasible enterprise. Indeed, for each of the cases discussed, I explained the transition to pathology in terms of a strengthening or weakening of risk and protective factors. Focusing on intermediate forms thus contributed to explain away the apparent discontinuity between normality and pathology that makes categorical views intuitively appealing. On the other hand, this detailed analysis of intermediate forms has substantiated the idea that borderline cases are both frequent and explanatorily relevant.[52] This shifts the burden of proof to the categorical opponent, who is now required to provide a principled way of distinguishing between intermediate and pathological cases without appealing to quantitative factors. Overall, with this chapter I contributed to make a dimensional approach to mental disorders more plausible.

[52] See Buckner 2016 for a similar argument about psychological kindhood.

Dimension	Intermediate	Pathological
Salience	Sadness	Depression
	Neuroticism	Interrogative attitude, obsessive disorders
Confidence	Impostor's syndrome	Self-reproach, depression
	Dunning-Kruger effect, pep-talk	Delusion of grandiosity
Familiarity	Dissociation, estrangement, depersonalization	Capgras delusion, Dissociative Identity Disorder (DID)
	Post-Bereavement Hallucinatory Experiences	Hallucinations, Fregoli delusion
Agency	False confessions	Pathological guilt
	Mind Wandering & Daydreaming	AVH, intrusive thoughts

Table 2: Synthetic table of intermediate cases and their pathological counterparts

Valentina Petrolini

Vulnerable Populations: Migration & Schizophrenia, Gender & Depression

"Stress research has emphasized the efforts that individuals can make to ward off distress or disturbance, but has paid much less attention to socially structured variations in exposure to stressors […] The importance of differential vulnerability has been overestimated to the degree that *differential exposure to stress* has been underestimated"

(Turner & Avison 2003, pp. 490 & 500. Italics mine).

Abstract

In this chapter I explore the relation between environment and psychopathology by introducing the notion of *vulnerable population*. The main goal of this discussion is to uncover various ways in which the environment can be pathogenic. In what follows I explore the reasons behind the higher incidence of pathology in some populations and I show that specific forms of disadvantage can be connected to specific disturbances. For example, the experience of social isolation that frequently accompanies *migration* may be connected with the development of positive symptoms of schizophrenia, such as paranoid delusions (see Janssen *et al.* 2003). Similarly, the repeated exposure to discrimination experienced by *women* in most societies may trigger feelings of hopelessness and frustration that are characteristic of depression (see Ussher 2010). After discussing these case studies, I expand the model outlined in Chapter Four by adding to it the notion of vulnerable population. While in the previous chapter I focus on psychological vulnerability, here I assess another group of factors that mark the distinction between healthy and pathological individuals. Specifically, I characterize environmental vulnerability as a different type of vulnerability arising from the higher exposure to social adversity. I also argue that situations of disadvantage such as the exposure to repeated discrimination may act as *triggers* of mental disorders, by causing or exacerbating psychiatric symptoms. Taken together, Chapters

Four and Five thus offer a more complete picture of what it means to be vulnerable to psychopathology.

Exploring the role played by *environmental factors* in the development of psychopathology is crucial to achieve the dissertation's overall goals. Indeed, the research discussed in this chapter indicates that looking at individual-level symptoms may not be enough to draw a satisfactory distinction between normality and pathology. Rather, important differences lie in conditions that are *external* to the patients, such as social pressures related to gender or discrimination due to minority status (Berg *et al.* 2014; Gutierrez-Lobos *et al.* 2000). In this sense, two individuals experiencing similar symptoms might greatly differ in terms of environmental circumstances and thus experience different degrees of vulnerability.[53] To render the picture more precise, I identify some environmental factors that have been found to be highly correlated with psychopathology. In doing so, I set out to go beyond general formulations (such as "social disadvantage") to uncover more specific mechanisms through which environmental factors affect psychological well-being.

The chapter is divided into three sections. In **§1** I offer an overview of empirical studies that discuss the role played by environmental factors in the development of psychopathology. In **§2** I explore two cases in more detail: first I discuss the recent studies on the high incidence of schizophrenia among migrants (**§2.1**) and then I report some older data on the higher rates of depression among women (**§2.2**). Both cases serve to illustrate situations in which the environment (as opposed to some sort of personal-level vulnerability) may be seen as pathogenic.

[53] Similar explanations have been offered to account for the differential vulnerability that people seem to have with respect to drug addiction. For example, many patients undergoing long-term hospital care become physiologically dependent on opioids (i.e. they experience withdrawal symptoms), but most of them stop taking the drugs when their therapy is over. Some patients instead become pathologically addicted, and some even get addicted after only one or few uses. Socioeconomic factors, as well as cultural ones (e.g. belonging to a social group that approves of recreational drug use) have been thought to be responsible for such a difference (see Szalavitz 2016; Brummett *et al.* 2017 for some recent reviews).

In **§3** I refine the model of vulnerability proposed in Chapter Four by fleshing out the notions of *vulnerable population* and *triggers*.

§1. Environment and Psychopathology

Within the history of Western psychiatry, Freud has probably been the strongest defender of the key role played by environmental factors in the onset of mental disorders. As I discuss in Chapter One, he offers a multi-factorial account of how mental disorders arise and distinguishes his view from others that regard mental illness as "organic inferiority" (Adler 1907) or "degeneracy" (Janet 1894). At various points throughout his career he talks about two kinds of pathogenic determinants: *dispositions* (or constitutional factors) on the one hand, and *experiences* (or accidental factors) on the other. The former are described as elements that "a person brings along with him into his life", whereas the latter are the ones that "life brings to him" (Freud 1913, p. 2623). The very idea that environmental factors would play an important role in the development of psychopathology has been introduced by Freud & Breuer in their *Preliminary Communication*: "[Our results] are valuable theoretically because they have taught us that external events determine the pathology of hysteria to an extent far greater that is known and recognized" (1893, pp. 3-4). A few years later, Freud insists that mental disorders should not be treated on a par with cases of "mental degeneracy" but that they should be seen as motivated responses to traumatic life events.

The idea that environmental factors could be importantly pathogenic has been further developed over the past forty years by researchers working within the *stress-vulnerability* paradigm (Brown & Harris 1978; Ormel & Neelman 2001). The hypothesis underlying this research program is that vulnerability factors (both individual and environmental) interact with mediating variables to give rise to pathological manifestations. For example, someone who

exhibits a tendency toward rumination (individual vulnerability) and has experienced childhood abuse (environmental vulnerability) would be more likely to cross the diagnostic threshold if confronted with negative life events (mediating variable). By studying environmental vulnerability, these researchers attempt to identify specific life circumstances that would increase or limit the exposure to stressors. Among these events, some have come to be known as "turning points" because of their relevance with respect to wellbeing. For instance, one's employment, marital and socio-economic status (SES henceforth) tend to influence the availability of personal resources and the exposure to stressors to a significant extent (see Turner & Lloyd 1999).

Acknowledging this point means acknowledging that *stress exposure* is not uniformly distributed across groups, and that people of different gender, ethnicity and SES may be differently exposed to stressful life events. An increased awareness to this aspect has recently paved the way to studies investigating the "socially structured variation in exposure to stress" (Turner & Avison 2003). Some of these experiments adopt multidimensional measures that take into account factors such as recent life events, chronic stressors, lifetime major events, and discrimination-related stress. When these potential stressors are considered together, patterns of structural vulnerability begin to emerge. For example, gender seems to determine a differential exposure to stressors that is compatible with the "cost of caring" hypothesis (see Henz 2010; Roxburgh 2005). Since women are usually expected to be more involved in the lives of others, they are more likely to develop a heightened vulnerability to events concerning other people's health and well-being (Nazroo *et al.* 1998; Das, Das & Das 2012). Ethnicity and SES are even more reliable indicators of differential stress exposure. Generally speaking, the relationship between SES and stress is monotonic (i.e. "the lower the SES, the higher the stress"). With respect to ethnicity, different exposure to stressors seems to account for a significant portion of

the mental health gap between African Americans and European Americans (Turner & Avison 2003, p. 497).

In the past twenty years the research on environmental vulnerability has become more specific, with studies focusing on the correlation between various situations of disadvantage and specific disorders (depression and schizophrenia in particular). Pioneering work in this sense has been conducted by the Dutch psychiatrist Jim van Os (1998), whose research on the *eco-genetics* of schizophrenia explores a complex system of interactions between genetic and environmental factors. His work uncovers a significant interplay between genetic elements (e.g. a family member affected by the disorder) and environmental risk-increasing elements, such as stressful life experiences. Besides gene-environment interactions, van Os and colleagues also explored other interactions that are significant for psychopathology without strictly depending on genetic factors. These have come to be known as "*environment-environment interactions*" (van Os 2003, p. 292). An example of such interactions would be the fact that people who grew up in rural areas have been found to be more resistant to the "psychotogenic effect" of urban settings (Kirkbride *et al.* 2007; van Os 2004). Another example would be that minority populations are not at higher risk of developing a psychotic disorder in situations where they become majority populations – e.g. when the ethnic density of the neighborhood changes (Veling *et al.* 2008). This distinction between gene-environment and environment-environment interactions is crucial because it allows us to see that the environment may act upon individuals at multiple levels. On the one hand, environmental elements combine with genetic predispositions and liabilities (e.g. family member affected by schizophrenia + sexual abuse during childhood). On the other hand, they interact with one another creating different patterns of vulnerability (e.g. urban upbringing + repeated exposure to discrimination).

More recently, van Os and colleagues further refined the study of environment-environment interactions by introducing the notion of *sensitization* (Collip, Myin-Germeys & van Os 2008). This notion draws on results coming from animal studies where rats that have been exposed to chronic social adversity develop greater responses over time – e.g. they increase self-administered cocaine consumption (Covington & Miczek 2001; Tidey & Miczek 1996). In other words, repeated experiences of defeat cause rats to develop a heightened vulnerability to defeat situations. Moreover, rats that experienced social adversity show lasting changes in response amplitude (i.e. they overreact to small stressors) to the point that alterations in dopaminergic concentration have been observed. van Os and collaborators hypothesize that prolonged social adversity, such as chronic exposure to isolation or discrimination, would have a similarly pathogenic effect on humans. Specifically, these experiences may contribute to strengthen cognitive biases (e.g. paranoid attribution) or alter dopamine neurotransmission and cortisol production thus facilitating the transition to psychosis (Kapur 2003; Jones & Fernyhough 2007). If this is correct, some environmental factors could be seen as *sensitizing* because they increase risk within a particular group. For example, many women live in contexts that are less likely to engender confidence and mastery (e.g. lower-paying jobs, less power in relationships) and these experiences of loss and entrapment have been reliably correlated with clinical depression (see Craig & Pathare 2001).

The sensitization hypothesis has been corroborated by a number of studies in the past few years. In a comprehensive meta-analysis, Beards *et al.* (2013) stress the impact that experiences of social adversity – e.g. defeat, isolation, repeated discrimination – have on the development of mental disorders (i.e. two-fold to seven-fold increased risk). In the next section I discuss two examples of environment-environment interaction in more detail. First, I explore the pathogenic effect that arises from the combination of “being a migrant” and “being subject to discrimination”

(**§2.1**). Indeed, migrant populations have been found to be at increased risk of *schizophrenia* only under specific conditions, such as being part of a visible minority (Berg *et al.* 2014) or being subject to high levels of discrimination (Janssen *et al.* 2003; Veling *et al.* 2008). Second, I discuss the well-known gender gap in the development of *depression*, where women's risk of being affected is twofold with respect to men (**§2.2**). Again, I argue that environment-environment interactions play a key role in explaining this asymmetry. For example, studies investigating employment and marital status uncover patterns of structural disadvantage that significantly affect mental health (see Gutierrez-Lobos 2000; Byrne, Carr & Clark 2004).

§2. Environment-environment interactions

§2.1 Migration & Schizophrenia

The idea that *migration* could constitute a significant risk factor for schizophrenia has been initially proposed by Ødegaard (1932) in his seminal study on Norwegian immigrants in the United States. Comparing immigrants coming from Norway with native-born Americans and Norwegians, Ødegaard found a two-fold incidence in first admission rates for schizophrenia in the migrant population. In the past ten years, studies investigating the relation between migrant status and schizophrenia have flourished following the recent waves of migration towards Europe. In the United Kingdom, migrants of Afro-Caribbean origin have been found to be especially vulnerable to schizophrenia, with first-admission rates up to three times higher (see Fearon & Morgan 2006; Fearon, Kirkbride & Morgan 2006). In the Netherlands, studies have repeatedly identified Moroccan and Surinamese immigrants as the groups at higher risk for psychosis (see Veling *et al.* 2008). It seems thus clear that a personal or family history of migration constitutes an important risk factor for schizophrenia. Indeed, after genetic markers,

migration and urbanization score higher than all the other factors commonly associated with this disorder (e.g. obstetric complications, paternal age).

Some initial hypotheses put forward to explain this phenomenon have already shown their limitations. For instance, rates of schizophrenia do not seem higher in the countries of origin (Bhugra 2004), and selective migration cannot apply to countries such as Suriname where almost half of the population migrated over time (Selten *et al.* 2002). More recent explanations tend to focus on migratory stress and on experiences of social adversity specifically connected with migrant status. A promising proposal in this sense is the one put forward by Cantor-Graae & Selten, which has come to be known as the "*Social Defeat Hypothesis*" (2005a & 2005b). On this view, the greater exposure to social adversity and the stress caused by the chronic experience of "outsider status" may be at the root of the problem (Cantor-Graae & Selten 2005b, p. 18). The higher rates of psychosis among migrants may thus be connected with phenomena such as pervasive social exclusion, repeated discrimination, socio-cultural isolation, and institutionalized racism. Indeed, migration experiences have been reliably associated with childhood and adult disadvantage – e.g. long-term separation from parents, unemployment, social isolation, etc. (see Morgan & Hutchinson 2010). The Social Defeat Hypothesis would also contribute to explain the results coming from the US, where the high incidence of schizophrenia among African Americans correlates with low family SES, segregated neighborhoods and exposure to discrimination (Bresnahan 2007, p. 756). As I mention above, situations of social defeat have been also found to affect the dopaminergic system in non-human animals (Jin *et al.* 2015), and alterations in dopamine release appear to be responsible for the experiences of "aberrant salience" typical of psychosis (Kapur 2003 & 2004). Increased cortisol activation has also been associated with symptoms of schizophrenia and appears to occur in situations of social threat, especially when people are "ignored or ostracized" (Jones & Fernyhough 2007, p. 1173). Post-migration

stress and related experiences of social defeat may thus be at the core of the mental health gap between migrants and residents.[54] Specifically, there seems to be something pathogenic in the experience of being "disadvantaged and different from the norm" (van Os 2012).

What are the specific mechanisms at work in the pathway from migration to schizophrenia? Important suggestions in this sense come from recent studies focusing on different aspects of the migrant condition, such as *second-generation* immigrants (SGI henceforth), level of *perceived discrimination*, and *ethnic density*. First, although the increase in risk extends to SGI, this transmission does not affect all migrant groups uniformly (Bourque *et al.* 2011). Second, the incidence of psychosis becomes increasingly higher in groups with visible minority status – e.g. different skin color (Berg *et al.* 2014). Third, the risk usually decreases in areas or neighborhoods where there is high own-group density (Das-Munshi *et al.* 2012).

According to the most extensive meta-analysis to date on SGI and psychosis, migrant groups that are highly discriminated exhibit a three-fold risk for schizophrenia with respect to host populations (Bourque *et al.* 2010). Notably, in these cases the risk *increases* in SGI and varies by host country: the highest incidence appears to be among Afro-Caribbean immigrants in the UK and the lowest among Ethiopians in Israel. These results suggest that post-migration experiences such as the exposure to persistent discrimination in the host country may play a crucial role in the development of psychotic symptoms (p. 906). Indeed, the perception of outsider status becomes particularly striking when people move from a non-discriminated majority to a discriminated minority, as it is the case for Africans coming to Europe (Cooper *et al.* 2008). In these cases risk also accumulates over time, suggesting that the duration of exposure

[54] For a more general hypothesis about the relation between social defeat and schizophrenia, see Luhrmann's work on homeless women in the Chicago area (Luhrmann 2007). On her view, the way in which psychiatric institutions are organized in the United States is partially responsible for perpetuating conditions of social defeat in low-income people affected by schizophrenia. Luhrmann describes this "institutional circuit" as a cycle of homelessness, supported housing, hospital and jail that patients have to endure (p. 148).

to a hostile host country may work as a risk factor.[55] This may also apply to groups that – despite a more ancient history of migration – have been consistently discriminated and segregated over time. For instance, one of the few studies conducted on a US cohort reveals a three-fold increased risk for African Americans with respect to European Americans (Bresnahan *et al.* 2007). On the contrary, risk appears to decrease when people migrate from a discriminated minority to a non-discriminated majority (e.g. Ethiopian Jews moving to Israel). A promising hypothesis would thus be that, among disadvantaged groups, the risk extends to SGI because the exposure to discrimination starts at a very young age and affects crucial developmental phases. While many people who migrate as adults come from a place where they belong to a non-discriminated majority, SGI are often born and grow up in a hostile environment. Moreover, the protection coming from the socio-cultural community of reference may be limited or absent, especially if the migrant group is small or fragmented in the host society.

The connection between perceived discrimination and schizophrenia has also been supported by research investigating the impact of *visible minority status*. For instance, a recent Norwegian study reports that immigrants from outside Europe score higher in terms of symptom severity than immigrants coming from other European countries. In particular, non-European immigrants exhibit more positive symptoms – i.e. delusions and hallucinations – than their European counterparts (Berg *et al.* 2011, p. 5). These results are compatible with the sensitization hypothesis, according to which repeated experiences of social defeat render individuals more vulnerable to future adversities. Such heightened sensitivity translates into physiological long-term changes (e.g. higher blood pressure, dopaminergic hyperactivity) that are reliably connected with the onset and exacerbation of psychotic symptoms. From a cognitive viewpoint, sensitivity may be expressed through forms of hypervigilance or cognitive bias (e.g. paranoid ideation,

[55] See Vega *et al.* (1998) for an older study with similar results focused on Mexican immigrants in the United States.

jumping to conclusions) that originate from chronic exposure to discrimination. In his paper "Elevators, social spaces and racism" (2008) Yancy offers a striking example of this phenomenon. Here Yancy describes an elevator encounter with a white woman whose body language, gestures and facial expressions clearly betray apprehension and anxiety in his presence. On his view, it is through the everyday experience of these non-verbal movements, as well as through more explicit forms of discrimination, that Blacks develop a distorted perception of their own self. This profound form of self-alienation derives from others systematically seeing you as different with respect to how you see yourself (e.g. as a threat, as a robber, as a rapist). As Yancy puts it: "[In the elevator] I am suddenly aware of how I am being perceived" (p. 858). These recognizable occurrences of racism then produce forms of hypervigilance and heightened sensitivity that become shared by entire communities. When interpreting the woman's behavior in the elevator, Yancy stresses this point forcefully: "Her gestures cohere with my knowledge of white racism, her gestures cohere with other experiences that I have had vis-à-vis whites performing racist gestures in the past and my experience is consistent with the shared experiences of other Blacks" (p. 849).

In this sense, the higher incidence of positive symptoms among discriminated groups – with auditory hallucinations and paranoid delusions being particularly frequent (Berg *et al.* 2011) – may reflect an extreme manifestation of the attitude that Yancy describes. To put it roughly: the more a group is exposed to systematic discrimination, the more sensitive its members become to racist or otherwise discriminatory occurrences. This in turn contributes to increase the degree of self-alienation, with people perceiving themselves more negatively – e.g. as more disadvantaged and less worthy overall (Cooper *et al.* 2008). A recent study from the Netherlands focused on different migrant groups supports this hypothesis (Veling *et al.* 2008). Similarly to other European countries, the groups found to be at higher risk in this study are the ones with the

lowest SES and experiencing the most severe discrimination. In particular, Moroccan males turn out to be the highest risk group for schizophrenia and also the one reporting the highest rates of racial as well as cultural discrimination. Specifically, the Moroccan immigrants interviewed reported more experiences of discrimination (42% versus 8% of Turkish immigrants) and filed more official complaints to the *Anti-Discrimination Bureau* than any other minority group (230/year versus 97/year from Turkish immigrants). Another interesting case is the one involving Surinamese immigrants, who exhibit high rates of schizophrenia despite their higher SES and degree of acculturation. Suriname is a former colony whose citizens speak Dutch fluently and tend to be culturally more uniform to people from the Netherlands. Yet, the high incidence of schizophrenia in this population might be explained in terms of cultural identity conflicts. That is, being similar to the residents with respect to language and culture while still being discriminated by them for purely racial factors (e.g. different skin color) might engender serious forms of self-alienation that make episodes of discrimination more painful to endure.

Finally, studies on *ethnic density* uncover yet another aspect of the complex pathway from migration to schizophrenia. Comparing neighborhoods or countries that differ in terms of ethnic or religious composition, researchers found that living in an area with high own-group density may act as a "buffer" against the environmental risk factors for psychosis (Das-Munshi *et al.* 2012). This may also explain the relatively low incidence of schizophrenia among migrant groups in Israel, despite the country's high immigration rates (i.e. 35% of the population). Indeed, as Weiser and colleagues (2008) explain, a great part of the migrants moving to Israel are religious minorities who flee discrimination and persecution to become part of the majority cultural identity group (p. 1116). Another buffering effect, similar to the one afforded by ethnic density, has been uncovered in multicultural but non-segregated neighborhoods (Kirkbride *et al.* 2007).

This finding therefore suggests that social capital itself (more than support coming from own-group members) may act as a protective factor against psychosis.

To sum up: disentangling the mechanisms underlying the high incidence of schizophrenia among migrants is extremely complex. Yet, thanks to a host of recent studies it has become clearer that migration may not be pathogenic *per se*, but that the focus should be on how host societies interact with minority groups (see Tarricone *et al.* 2015). Indeed, the migrant groups at higher risk overlap with the ones with lowest SES and highest reported discrimination. This case study thus helps us to see that mental disorders could be – at least partially – a product of structural injustices that may be prevented through social reform. In the following section I turn to gender and depression to illustrate a similar case.

§2.2 Gender & Depression

In this section I discuss the case of gender and depression as a further example of environment-environment interaction. I start by presenting some empirical evidence about the well-known high incidence of depressive disorders among women. Then I show how these disturbances can be reliably connected with specific forms of disadvantage. For example, the repeated exposure to discrimination and violence experienced by women in most societies may trigger feelings of humiliation, frustration and entrapment that are typical of depression (see Ussher 2010). As in the case of migration and schizophrenia, this discussion takes us a step further from the observation that some populations are at higher risk of developing a mental disorder. Indeed, it allows us to sketch a more precise explanation of such vulnerability, one that focuses on the multiple ways in which the environment can be pathogenic.

The high incidence of depressive disorders among *women* has been reported by several clinical and epidemiological studies over the last few decades (see Bebbington 1996; Jenkins

1987). These studies uniformly stress that women experience depressive symptoms at higher rates than men (numbers agree on a 2:1 ratio) but also that women are more likely to seek medical attention and to receive pharmacological treatment for depression (see LaFrance 2007). Some researchers appeal to biological or constitutional differences to explain this gender gap while others see it as the result of a social construction. Among the biological and constitutional factors we find hormone levels (Seaman 1997) as well as personality traits such as the tendency towards rumination and a low sense of mastery (Nolen-Hoeksema, Larson & Grayson 1999). Among the constructivist accounts, many regard depression as the product of structurally sanctioned conceptions about gender that end up pathologizing women's experiences. For example, Emmons (2008) argues that the pervasive characterization of women as "the emotional sex" lends itself to the idea that depression would be an intrinsically feminine disorder. As she puts it: "While depression in men tends to be presented as a stark departure from 'normal' feelings or emotions, depression in women is more likely to be understood as an *outgrowth* of women's complex emotional lives" (p. 112. Italics mine). However, women appear to be significantly more exposed to negative events during childhood – sexual abuse in particular – and to depression or anxiety during adolescence (Avison & McAlpine 1992).

In what follows I suggest a more complex etiological picture of the high incidence of depression among women, one that focuses on environmental factors without fully embracing a constructivist account. Indeed, one can maintain that the disadvantaged condition of women in most societies makes them more vulnerable to developing depressive symptoms without thereby denying the reality of the disorder.[56] In order to see the relation between gender and depression as an example of environment-environment interaction, it is thus not sufficient to isolate one group

[56] Similarly, one may acknowledge the impact of socio-economic factors on the higher incidence of somatic ailments in certain populations (e.g. diabetes in African Americans) without thereby regarding these diseases as social constructions.

of individuals because of their heightened vulnerability. It is also crucial to understand how the development of the disorder connects with a host of risk factors within the relevant group. For instance, *relationship* and *employment* status, along with *unequal workload distribution* have a significant impact in determining which sub-groups of women are at risk of developing depressive symptoms.

In the past few years, a number of studies have explored the reasons behind the gender gap in depression by investigating some of these factors in more detail. For example, an Austrian study that takes into account the intersection between gender, relationship and employment status, found that *marriage* acts as a protective factor against depression for men but not for women. Indeed, among married subjects women exhibited an incidence rate that was more than twice as high with respect to men – i.e. 72 versus 31 patients out of 100.000 (Gutierrez-Lobos *et al.* 2000, p. 204). The data concerning single subjects are equally interesting: while both never-married men and never-married women exhibit a higher risk for depression, "the difference in the women group is much less pronounced" (p. 207). As the researchers conclude, married status thus seems to carry more significant benefits for men than for women.[57] The same study also shows that *employment* significantly affects depression rates across the board, with higher incidence among the unemployed (p. 206). However, when marital and employment status are considered together the gender gap resurfaces, with unemployed married women showing a two-fold incidence of depression with respect to their male counterpart – i.e. 88 versus 41 out of 100.000 (p. 204). The situation appears reversed among the unmarried, with unemployed single men displaying higher rates of depression than women in a similar situation – i.e. 140 versus 128. This study is particularly interesting because it goes beyond the "gross oversimplification" of

[57] Recently, the idea of marriage as protective has been called into question altogether by studies showing that married couples report *lower* scores of health, happiness, etc. than their non-married counterparts (see Kalmijn 2017).

women being more liable to depression, and explores specific environmental factors that could be responsible for such a gap (p. 207). At the same time, the data concerning gender and employment status vary *cross-culturally*, with studies reporting higher rates of depression among employed women in India (Das, Das & Das 2012) and in rural European areas (Vazquez-Barquero *et al.* 1992). These results deserve special attention because they uncover conflicting pressures to which women may be subject in different societies. Indeed, it is plausible to assume that unemployment would affect women more negatively in societies where they are expected to contribute to the workforce (e.g. Austria). The opposite may be the case in societies where women are expected to conform to strictly defined gender roles that often comprise housework and childcare (e.g. India). In fact, in such contexts working women may be even more exposed to discrimination and violence because their status could be seen as "threatening the authority of men" (Das, Das & Das 2012, p. 1661).

These results uncover more profound reasons that may explain women's heightened vulnerability to depression. First, a growing number of women seem to be unhappy with the restrictions and expectations associated with traditional gender roles. In a research measuring the emotional states experienced by a group of North-American mothers and fathers, Larson, Richards and Perry-Jenkins (1994) found that women who are "full-time in the homemaker role" report more negative emotional states and lower degrees of happiness across the board (p. 1044). Second, couples with a depressed female partner are more likely to be structured along asymmetrical power relations (Byrne & Carr 2000). This means that depressed women who are in a relationship also tend to be both financially and psychologically weaker than their partners, and to experience greater inequality in terms of decision-making, task and resources distribution (pp. 18-19). Finally, another striking result concerns the fact that women with no children exhibit depression rates similar to men's. Indeed, studies quantifying the impact of childbearing as a risk

factor report a three-fold incidence of depression in mothers as opposed to non-mothers (Williams & McCutcheon 2004; Craig & Pathare 2001). As researchers stress, these results point to a specific "exposure to environmental adversity that goes along with marriage and motherhood" (Craig & Pathare 2001, p. 212).

What are the specific environmental adversities encountered by women? A promising answer comes from the "cost of caring hypothesis", according to which one of the causes of women's vulnerability should be researched in the chronic stress deriving from *unequal workload distribution*. This phenomenon has come to be known in sociology as the "second shift" (Hochschild 1989). The expression conveys the idea that, while dual-earner households have become progressively more common, gender roles have not evolved accordingly thereby forcing many women to take on a significant amount of unpaid domestic work. Various forms of caregiving are a case in point, as considerably more women than men reduce their labor market participation because of care-related duties (Henz 2004). Moreover, paid employment appears to limit men's availability for caring but not women's. In other words, women are expected to perform caregiving duties – e.g. childrearing, looking after elderly parents – even when they are occupied with paid work, while the same expectation does not hold for men (Henz 2010). This underlying inequality creates a vicious cycle of disadvantage: since fewer women are employed in the workforce, more of them are expected to take up other kinds of duties. Because their wages are substantially lower than men's, they are more likely to give up paid work in favor of unpaid domestic labor. Since most women are socialized to care for others from a very young age, they feel obliged to attend to other's needs irrespectively of their own time and resources (Pinquart & Sörensen 2006). Childrearing deserves a special mention here because it is well-established that women's greater investment in parenting exposes them to higher chronic stress and harmful consequences for health and well-being. In a recent study comparing dual-earner couples with

ones in which one member was either working part-time or not working, Roxburgh (2005) concludes that mothers are more adversely affected than fathers by parenting strains. However, this gap appears less dramatic in dual-earner couples, where parenting duties are shared more equally (p. 1065). Notably, in these cases women's stress level does not decrease, but men's stress level increases (p. 1075). This speaks against the idea that women would be constitutionally more vulnerable, and suggests that different stress levels may rather be the product of social disadvantage.

Women are thus placed in a structurally disadvantaged position. On the one hand, if they adhere to traditional gender roles they risk becoming financially and psychologically dependent on their partners and thus more vulnerable to experiences of humiliation, frustration and entrapment (Brown 2002; Craig & Pathare 2001). On the other hand, if they do not strictly conform to these roles they are exposed to additional stress deriving from second-shift experiences or various forms of discrimination (Das, Das & Das 2012). The gender gap in depression seems thus to depend on dramatic differences in exposure to stress between men and women. Growing up, women are more likely to experience abuse, humiliation and entrapment: in particular, they are more likely to be stuck in unrewarding relationships because of external constraints (Brown, Bifulco & Harris 1987). As adults, they are expected to conform to traditional roles characterized by power asymmetries (Byrne & Carr 2000) and they are disproportionately more likely to encounter unequal division of labor, economic discrimination, and job inequality (see Piccinelli & Wilkinson 2000). Similarly to the case of migration and schizophrenia discussed in §2.1, women's vulnerability to depression may thus be partially amended through social and cultural reform.

§3. Vulnerable Populations

The cases discussed in **§2** uncover a new kind of vulnerability that has not been yet explored in this project. In Chapter Four I discuss a number of cases of personal-level vulnerability and I characterize them in terms of increase of risk factors (*type a*) or weakening of protective factors (*type b*). For example, a neurotic person may experience obsessive thoughts with a certain frequency and intensity while still being able to successfully integrate these symptoms in her daily life. However, if the relevant risk factors (i.e. frequency and intensity) increase sufficiently over time, she would be more likely to cross the diagnostic threshold than someone who did not exhibit the same vulnerability. Similarly, someone may be struggling with mild depressive symptoms while finding social and emotional support in a significant other (e.g. a spouse or friend). Yet, if this protective factor suddenly disappears (e.g. the significant other dies or moves far away), this person's transition to pathology would become more likely or even inevitable. These two kinds of vulnerability – i.e. type a) and b) – are encompassed by what Turner & Avison call "differential vulnerability" (2003, p. 500). To put it roughly, people are differently equipped to deal with stressors because of a variety of personal-level differences. For instance, they exhibit significant differences in their personality traits (e.g. a neurotic person would obsess more over small details) and in the way in which they experience feelings and emotions (e.g. one may be more or less anxious). People also greatly differ in terms of the coping resources at their disposal. For example, some rely on problem-focused strategies (e.g. changing an unpleasant situation), while others employ emotion-focused ones (e.g. detaching oneself from something painful).[58] In the previous chapter I also acknowledge that these types of vulnerability often interact with one another and give rise to more complex patterns (*type c*).

[58] For a detailed discussion on different coping strategies, see Lazarus 1993.

The evidence presented in this chapter suggests that a further kind of vulnerability should be taken into account. Besides vulnerable individuals, entire groups may be seen as vulnerable due to their potentially *higher exposure to stressors*. In other words, being a member of a group that is more exposed than others to stressful or negative events should also count as a risk factor for mental disorders. Given that this kind of vulnerability concerns the population level, it is not necessary that the individual experiences environmental adversity first-hand for the risk to be present. Yet – as I explain below – being subject to stressors does facilitate the transition from risk to actual development of a disorder.

The studies discussed in §2.1 and §2.2 help us understand how this process unfolds. For instance, people who belong to visible minorities are more likely to be discriminated in their everyday life than people who belong to the ethnic majority. More than migrant status *per se*, it is thus the experience of being "disadvantaged" and "different from the norm" that engenders a particular liability to schizophrenia (van Os 2012). This emerges from the studies reporting that the most discriminated groups (e.g. Afro-Caribbean in the UK, Moroccan males in the Netherlands) are also the ones at highest risk of psychosis (Veling *et al.* 2008). This heightened exposure to stressors translates into a number of health-related consequences encompassing neurological and cognitive changes. At the neurological level, repeated experiences of social defeat are connected with increased cortisol activation and hyper-dopaminergic activity in non-human animals (Covington & Miczek 2001; Jin *et al.* 2015). In humans, both cortisol and dopaminergic transmission are significantly increased in psychotic states (Jones & Fernyhough 2007; Kapur 2003 & 2004). At the cognitive level, exposure to discrimination facilitates the insurgence of biases such as external attribution (e.g. others are blamed rather than self) and jumping-to-conclusion (i.e. a conviction is reached too soon, without gathering all the relevant facts). These biases are at the root of pathological beliefs such as persecutory delusions and

delusions of reference, where patients become convinced that other people are following them, threatening them, or invading their personal space (Freeman *et al.* 2002). The pathway from chronic discrimination to psychotic symptoms thus becomes increasingly clear. Due to structural injustices, ethnic minorities are significantly more exposed to phenomena of discrimination, harassment and segregation. These repeated experiences of exclusion may then facilitate the insurgence of perceptual and cognitive biases, with people paying more attention to social and behavioral cues and thus being more prone to interpret a neutral episode as persecutory or threatening.

A similar mechanism may also explain the higher incidence of depression among women. Indeed, gender inequalities also render women more likely to be exposed to discrimination with respect to men. Most societies still expect women to care more about some aspects of life over others (e.g. family, reproductive sphere), and to sanction them when they do not conform to these expectations. This underlying inequality creates important health-related consequences. On the one hand, by conforming to traditional gender roles women find themselves in relationships characterized by asymmetries in terms of power bases (e.g. financial assets), outcomes (e.g. decision-making), and processes (e.g. bargaining power) – see Byrne & Carr 2000. On the other hand, those who do not conform to social norms risk being exposed to various forms of discrimination. These contradictory expectations constitute an exemplary case of "double bind" (Oakley 2000).[59] A specific example discussed by Oakley concerns the opposition between "competency" and "femininity". When it comes to leadership positions, women often feel they have to reach a balance between "speaking assertively but not too assertively" or "dressing 'like a woman' but not dressing 'too feminine'" (p. 325). This pressure on women leaders derives from underlying biases that associate femininity with incompetence and competence with

[59] Notably, these norms and expectations tend to be accepted and enforced by women themselves, due to a process known as "internalized oppression" (Bluhm 2011; Williams 2012).

stereotypically masculine traits (e.g. loud voice). As a result, many women in the workplace conclude that they must act, speak, or dress "un-feminine" in order to be perceived as competent or assertive. For all these reasons, one could see "being a woman" as a *de facto* vulnerability for depression. In other words, the sense of hopelessness and self-loathing typical of this disorder may be seen as an outgrowth of a social environment that systematically disadvantages women. Indeed, even personal-level traits such as mastery and self-esteem are largely influenced by how the environment responds to the individual's actions and efforts. As Turner & Lloyd (1999) point out, women tend to live in circumstances that are less likely to bolster their sense of confidence: for example, despite being more educated than men they work in lower-level jobs, experience lower wages, and have less power in relationships (p. 378).

Both case studies allow us to sketch a more refined characterization of what it means for a population to be vulnerable. I started from the epidemiological consideration that some groups exhibit a heightened incidence for specific mental disorders. For instance, a person of Afro-Caribbean origin living in the UK is three to six times more likely than a native Briton to develop psychotic symptoms, while women are twice as likely as men to experience depressive disorders. However, it is also possible to identify a host of environmental risk factors *within* the relevant groups. For example, the migrant groups at highest risk of psychosis overlap with those experiencing the highest levels of discrimination (Veling *et al.* 2008; Cooper *et al.* 2008). Similarly, while women appear to be more vulnerable to depression than men across the board, risk level increases in some circumstances – e.g. childrearing (Williamson & McClutcheon 2004). These finer-grained distinctions suggest that some groups should be regarded as vulnerable due to their potentially higher exposure to stressors. This idea of group-level vulnerability could be applied equally well to somatic medicine, where some risk factors affect a

segment of the population more or less uniformly: e.g. one is at higher risk of developing lung cancer is she lives in a particularly polluted area.

Yet, environmental vulnerability *per se* is clearly not sufficient for the development of psychopathology. After all, not all migrants and women are affected by mental disorders despite being members of a vulnerable group. Similarly, many people who live in polluted areas do not develop lung cancer. So, how does the transition from risk to actual danger occur? Given an existing vulnerability, some events may act as *triggers* that make the subject increasingly more exposed to negative events of the same kind – e.g. discriminatory actions, everyday humiliations. Triggers can thus be seen as actions, perceptions and experiences that extend over time and affect one's neurological and cognitive makeup. In other words, triggers tap onto a pre-existing vulnerability by increasing one's exposure to environmental adversity. This is for example what happens when a migrant moves from a non-discriminated majority to a discriminated minority, or when a woman decides to give up paid work and depend on her partner financially. In both cases, these life choices increase the individual's exposure to additional stressors or negative events. In the case of gender, these would be all the situations in which one experiences discrimination *qua woman* – e.g. being denied a raise, being sexually harassed, being subject to pressures concerning traditional gender roles, etc. In the case of migrants, stressors include episodes of social exclusion, discrimination and isolation to which minorities are routinely exposed – e.g. being searched by the police more frequently, being shunned by others on the bus, being excluded from housing or employment opportunities, etc.

The notion of vulnerable population outlined here goes hand in hand with the one of vulnerable individual discussed in Chapter Four. In both cases, a person or group exhibits a certain *vulnerability*, that may be seen as a "weak spot" at the personal level (e.g. being prone to obsessive thoughts) or at the environmental level (e.g. being a member of a disadvantaged

group).[60] Vulnerable people would then be more sensitive to *triggers*, namely events that strengthen risk factors or weaken protective factors. For example, an alcoholic who relapses by drinking a glass of wine (trigger) renders the thought of drinking again more frequent and intense (risk factors increase). Or again, the death or departure of a friend (trigger) provides fewer opportunities for venting to the person who is struggling with mild depression (protective factors weaken). When it comes to vulnerable populations, triggers would then be events that increase one's exposure to environmental adversity – e.g. by changing status, social setting, etc.

To sum up: the kind of vulnerability discussed here concerns groups that are (on average) more exposed to environmental adversity and stressors than others. In this sense, the members of these groups are at higher risk of developing a mental disorder even if – *qua* individuals – they might not experience stressful or negative events. The transition from risk to actual development of a disorder thus occurs when members of the vulnerable groups experience a sufficient number of *triggers*, namely events that *de facto* increase his or her exposure to adversity.[61] This is usually the moment when symptoms emerge and consolidate. For example, a woman may start developing feelings of anxiety or fear with respect to certain situations because of past experiences (e.g. traveling alone) and then choose – more or less consciously – to avoid similar circumstances. Notably, these coping strategies may be maladaptive because they deprive the person of future opportunities and contribute to the formation of self-fulfilling prophecies (Schmader, Johns & Forbes 2008). Similarly, a migrant belonging to a discriminated group may develop paranoid thoughts about other people's behavior as the result of repeated experiences of racism (see Yancy 2008). To defend himself from negative events, he may then choose to avoid

[60] Notably, these two levels of vulnerability should not be seen as mutually exclusive. In fact, vulnerable individuals who are also members of vulnerable populations would be particularly prone to developing mental disorders.

[61] At this stage, it is difficult to be more precise in determining the number of triggers required for a member of a vulnerable population to cross the diagnostic threshold. Similarly to the personal-level vulnerabilities discussed in Chapter Four, I believe that the model should allow for some degree of individual variability as well as for more specific considerations about the socio-cultural context at hand. I am planning on addressing this issue in more detail in future work.

social contact thereby exacerbating a sense of isolation and self-alienation that make psychotic symptoms more likely to occur.

There are at least three ways of crossing the diagnostic threshold and develop a clinically relevant manifestation. First, one may exhibit mild or attenuated symptoms (i.e. *type a)* vulnerability) and then experience one or more triggering events that strengthen the relevant risk factors. This would be the case of the recovering alcoholic who already thinks a lot about drinking and then develops a full-blown obsession after relapse. I characterize these situations as ones in which the threshold is crossed by *summation*. Second, one's existing vulnerability may be rendered more serious by the sudden or progressive weakening of the relevant protective factors (i.e. *type b)* vulnerability). This is what happens to the mildly depressed person after her friend has died or moved away. In these situations, I talk about crossing the threshold by *collapse*. Third, one may exacerbate an existing vulnerability by increasing the exposure to social adversity and stressors. This is the case of people who belong to a discriminated group and start developing progressively greater responses to negative events. As I mention above, in these cases the threshold is likely to be crossed through a process of *sensitization*.

Conclusion

Together with Chapter Four, the aim of this chapter is to offer a more precise characterization of the notion of *vulnerability* to psychopathology. In the previous chapter I focus on personal-level vulnerability and I distinguish between subjects who exhibit an attenuated version of a full-blown syndrome (*type a*) and subjects whose protective factors are weakening (*type b*). In this chapter I focus on environmental-level vulnerability and I discuss the complex interplay between environmental factors and psychopathology. I do so by introducing the notion

of *vulnerable population*, a group of individuals who are more likely to be exposed to social stressors due to structural disadvantage (Turner & Avison 2003; Turner & Lloyd 1999).

First, I reviewed a significant body of evidence suggesting that some groups of individuals are more vulnerable to specific disorders (i.e. schizophrenia and depression). Then I drew on the notion of environment-environment interaction (van Os 2003) to explore two cases where social factors play a key role in the development of psychopathology. These are the higher incidence of schizophrenia among migrants (**§2.1**) and the fact that women are more vulnerable to depression than men (**§2.2**). In discussing these cases in detail, I identified a number of environmental risk factors within each population. For example, migrants who are exposed to discrimination or women who experience power asymmetry in relationships have been found to be at higher risk of developing a mental disorder. This allowed me to sketch a characterization of vulnerability in terms of *higher exposure to adversity*, which complements the notions introduced in Chapter Four.

The focus on environmental triggers ties into a more general critique of the current state of psychiatric practice. Indeed, if social factors play a key role in the onset of mental disorders, therapeutic approaches centered primarily on the individual (e.g. pharmacological intervention, talk therapy) may fail to address the problem at the appropriate level. An important aspect of psychiatric work should thus focus on reducing stress exposure for groups at special risk. Given the impact that socially structured inequality has in the onset and exacerbation of mental disorders, the reduction of such inequality should become one of psychiatry's goals. The model proposed here contributes to this process, as it distinguishes among different kinds of vulnerabilities and among different ways of crossing the threshold to pathology. Being able to pry apart vulnerabilities and triggers may also help clinicians to assess the impact of risk and

protective factors in specific situations and therefore to better tailor therapeutic interventions to patients.

Appendix: Objections & Replies

A possible downside of the dimensional model outlined in Chapters Two, Four, and Five is that it may still seem too theoretical and *a priori*. Specifically, one might worry that the account would not be amenable to empirical falsification (or verification). After all, the preliminary evidence that I offer is mostly based on reports and autobiographical accounts of mental disorder, which – albeit suggestive – appear to lack the required bite.

In this short appendix I address this concern by sketching a few ways in which the account I propose may be falsified, along with some practical suggestions on how to test it.[62]

- **What would falsify the account?**

Generally speaking, I see this project as setting the necessary groundwork for empirical verification. The main goal of the dissertation is to outline a dimensional view of mental disorders that can come to be seen as a plausible alternative to its categorical counterpart. Yet, for the account not to be *a priori* true, I need to establish some conditions under which it would be falsified.

Here are a couple of suggestions:

a) *X is imbalanced on one or more dimensions to the same degree of Y, but Y exhibits a mental disorder while X does not.*

This result would falsify the account, provided that the relevant risk and protective factors are also sufficiently similar to make the two situations comparable. Going back to one of the examples that I use in Chapter Four: Mohammad Alì might have been imbalanced on the

[62] I would like to thank Peter Langland-Hassan for raising these issues and for encouraging me to think more thoroughly about possible ways to test my account.

confidence dimension *to the same degree* of someone who suffers from grandiosity delusions, but he arguably had some protective factors in place (e.g. humor, control) that prevented him from being *de facto* delusional.[63]

b) *X is diagnosed with a mental disorder (e.g. major depression) while displaying *normal* levels of confidence, salience, familiarity, and agency.*

This result would also falsify the account, although I want to stress that looking at one dimension alone may not be sufficient. For example, many cases of clinical depression result from disorders of confidence, but there might be situations where the hypoconfidence component is not too pronounced and other imbalances are more significant – e.g. anhedonia (hyposalience). Disorders should thus be assessed by looking at the four dimensions together: borrowing a Freudian term, I sometimes refer to this as a "pathological constellation". This point is relevant also because imbalances on the four dimensions tend to cluster in specific configurations. For example, a prototypical constellation for major depression would be characterized by low values across all the dimensions (i.e. low salience, low confidence, low familiarity, and low agency). This implies that it would be problematic for the account if someone were to be diagnosed with major depression while displaying normal (or high) levels of salience, confidence, familiarity, and agency.

- **How can we test whether someone is imbalanced on one or more dimensions? What would be the means to establish that?**

Another important issue concerns finding independent ways to determine where a person stands with respect to the four dimensions. In other words, we need reliable tools to establish

[63] This is of course debatable. Someone might instead argue that during the famous "Rumble in the Jungle" match against Foreman, Alì had in fact crossed the threshold to pathology (at least momentarily). I would like to thank Johannes Brandl for this observation.

both the sort of imbalance that a person exhibits – e.g. hyposalience – and its degree of severity – e.g. how close the person is to crossing the diagnostic threshold. I am not able to address this issue in detail here, but I still want to offer some suggestions on possible ways to test one's location on the dimensions.

One option would be to look at disorders that are already thought to be dimensional in nature (e.g. autism, psychopathy, or personality disorders) and formulate similar diagnostic methods employing the four dimensions. For example, the Hare PLC-R checklist for psychopathy assesses performance in three domains (i.e. interpersonal, affective, behavioral) with a specific cut-off score for each one. Similarly, rating scales and questionnaires for autism distinguish between mild, moderate, and severe forms depending on the degree of support required in a variety of tasks – e.g. social communication.

Generally speaking, dimensional disorders (and their severity) are assessed by looking at a variety of aspects:

- **Core domains**: e.g. social communication (autism); interpersonal behavior (psychopathy). The core domains in my model would be the four dimensions of *salience, confidence, familiarity,* and *agency.*

- Context*:* e.g. family history, social ties (autism; psychopathy). In my model this would be assessed by looking at *personal-level risk* and *protective factors,* and at the *environmental-level vulnerabilities* at play in the specific situation.

- **Degree of functioning or impairment**: e.g. requiring support (autism). This could be established by looking at whether the dimension's *functional role* is impaired or preserved, and to which degree (see Chapters Two and Three). For example: someone who fails to show up at work

because she is too busy rearranging her clothes by color would have lost sight of what is salient or important.

Another option would be to put together a battery of tests to directly assess where someone falls on each dimension (see Table 3 below).

Dimensions	**SALIENCE**	**CONFIDENCE**	**FAMILIARITY**	**AGENCY**
Tests	-Wisconsin Card Sorting Task (WCST) - Tower of London - Stroop Test - Iowa Gambling Task - California Verbal Learning Test (intrusions task)	- Rosenberg Self Esteem Scale (RSES) - Sorensen Self Esteem Test - Confidence ratings on metacognition tasks	- Structured Clinical Interview for Dissociative Disorders **(**SCID, especially depersonalization and derealization scales) - False memory task (Loftus) - Damasio (unconscious affect tasks) - Metamemory tasks - Benton Test of Facial Recognition	- Wegner's illusion of control experiments - Pronin's apparent mental causation experiments
Stipulated underlying mechanisms and/or constructs	Executive functions; selective attention (PFC; dopamine transmission)	Self-worth; self-love; self-preservation (amygdala)	Self-knowledge; other-knowledge; affective memory (unconscious memory; hippocampus)	Metacognition; self-other distinction; mental causation (sensory-motor control, PFC)

Table 3: Battery of tests to determine where one falls on a given dimension

A few caveats are in order. As you can see from the table, a crucial issue is to determine which underlying mechanisms these tests would be capturing. In other words: What are we measuring when we determine a person's location on a given dimension? A cognitive ability (e.g. selective attention), a neural correlate (e.g. PFC), or something else entirely? These are of course empirical questions, although – ideally – some degree of mapping between these levels would be desirable. For example, a person who turns out to be imbalanced at the cognitive level (e.g. hypersalience) should also exhibit a similar imbalance at the neurological level (e.g. hyper-dopaminergic activation).

A related issue with this table is that it groups together different kinds of tasks. Indeed, some of the tests listed here are explicitly cognitive (e.g. metacognition and executive function tests), whereas others are more similar to surveys or structured interviews (e.g. Rosenberg, SCID).[64] Although it looks like one's location on each dimension could be assessed both ways, it would be important to make sure that these scores reliably correlate. For example, salience levels assessed by executive function tests should come out sufficiently similar to the ones assessed through surveys or interviews. This is also part of the work that would be required to test the account empirically.

[64] Notably, structured interviews could be administered to the subject as well as to other people who know the person well (e.g. family members, close friends). This is in line with recent studies in personality psychology suggesting that others may know us better than we know ourselves (see Vazire & Carlson 2011; Vazire 2010).

References

- Adler, A. (1907). *Study of Organ Inferiority and Its Physical Compensation*. New York: Nervous and Mental Diseases Publishing.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (DSM-5®). American Psychiatric Publications.
- Andreasen, N. C. (2007). DSM and the death of phenomenology in America: an example of unintended consequences. *Schizophrenia Bulletin* 33(1): 108-112.
- Andrews-Hanna, J. R. (2012). The brain's default network and its adaptive role in internal mentation. *The Neuroscientist* 18(3): 251-270.
- Avison, W. & McAlpine, D. (1992). Gender differences in symptoms of depression among adolescents. *Journal of Health and Social Behavior* 33: 77-96.
- Ayduk, O., Downey, G., & Kim, M. (2001). Rejection sensitivity and depressive symptoms in women. *Personality and Social Psychology Bulletin* 27(7): 868-877.
- Baird, B. *et al.* (2012). Inspired by distraction mind wandering facilitates creative incubation. *Psychological Science* 23(10): 1-6.
- Banerjee *et al.* (2012). Is it light or dark? Recalling moral behavior changes perception of brightness. *Psychological Science* 23: 407-409.
- Bar-On, D (2013). Origins of meaning: must we 'go Gricean'? *Mind & Language* 28(3): 342-375.
- Barrett, L. F., & Bar, M. (2009). See it with feeling: affective predictions during object perception. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 364(1521): 1325-1334.
- Barrett, L. (2015). Conceptual Act Theory: a Roadmap. *The Psychological Construction of Emotions*, Barrett, L. & Russell J. (eds). Guilford Press.
- Beards, S. *et al.* (2013). Life events and psychosis: a review and meta-analysis. *Schizophrenia Bulletin* 39(4): 740-747.
- Bebbington, P.E. (1996). The origins of Sex Differences in Depressive Disorder: Bridging the gap. *International Review of Psychiatry* 8(4): 295–332.
- Bentall, R. P. (2004). *Madness explained: Psychosis and human nature*. Penguin UK.

- Bentall, R. P. et al. (2014). From adversity to psychosis: pathways and mechanisms from specific adversities to specific symptoms. *Social psychiatry and psychiatric epidemiology 49* (7): 1011-1022.

- Beran, M. J., Brandl, J., Perner, J., & Proust, J. (Eds.). (2012). *Foundations of metacognition.* Oxford University Press.

- Berg, A. O. *et al.* (2014). The impact of immigration and visible minority status on psychosis symptom profile. *Social Psychiatry and Psychiatric Epidemiology* 49(11):1747-1757.
- Berg, A. O. *et al.* (2011). Perceived discrimination is associated with severity of positive and depression-anxiety symptoms in immigrants with psychosis: a cross-sectional study. *BioMedCentral psychiatry* 11(1): 77-86.

- Berninger, A. (2016). Thinking sadly: In favor of an adverbial theory of emotions. *Philosophical Psychology* 29(6): 799-812.

- Bhugra, D. (2004). Migration and mental health. *Acta Psychiatrica Scandinavica* 109(4): 243-258.

- Billon, A. (2015). Making sense of Cotard's syndrome: insights from the study of depersonalization.

- Bluhm, R. (2011). Gender differences in depression: Explanations from feminist ethics. *International Journal of Feminist Approaches to Bioethics* 4(1): 69-88.

- Bonanno, G. A., & Kaltman, S. (2001). The varieties of grief experience. *Clinical Psychology Review* 21(5): 705-734.

- Boorse, C.

  (1977). Health as a Theoretical Concept. *Philosophy of Science* 44: 542-573.

  (1975). On the Distinction between Health and Illness. *Philosophy and Public Affairs* 5: 49-68.

- Bortolotti, L. (2010). *Delusions and Other Irrational Beliefs*. Oxford University Press.

- Bourque, F., van der Ven, E. & Malla, A. (2011) A Meta-analysis of the risk for psychotic disorders among first and second generation immigrants. *Psychological Medicine* 41: 897-910.

- Bresnahan, M. *et al.* (2007). Race and risk of schizophrenia in a US birth cohort: another example of health disparity? *International Journal of Epidemiology* 36(4): 751-758.

- Brewin, C. R., Christodoulides, J. & Hutchinson, G. (1996). Intrusive Thoughts and Intrusive Memories in a Nonclinical Sample. *Cognition & Emotion* 10(1): 107-112.

- Briñol, P., DeMarree, K. & Petty, R. (2010). Processes by which confidence (vs doubt) influences the self. R.M. Arkin, K.C. Oleson & P.J. Carroll (eds.). *Handbook of the uncertain self*. New York: Psychology Press.

- Brown, G. W. (2002). Social roles, context and evolution in the origins of depression. *Journal of Health and Social behavior* 43(3): 255-276.

- Brown, G. W., Bifulco, A., & Harris, T. O. (1987). Life events, vulnerability and onset of depression: some refinements. *The British Journal of Psychiatry* 150(1): 30-42.

- Brown, G. W., & Harris, T. (1978). Social origins of depression: a reply. *Psychological Medicine* 8(4): 577-588.

- Bubl, E. *et al.* (2010) Seeing gray when feeling blue? Depression can be measured in the eye of the diseased. *Biological Psychiatry* 68: 205-208.
- Buckner, C. (2016). Transitional Gradation in the Mind. Rethinking Psychological Kindhood. *British Journal of Philosophy of Science* 67(4): 1091-1115.

- Brummett *et al.* (2017). New persistent opioid use after minor and major surgical procedures in US adults. *JAMA Surgery* 152(6). Online only.

- Byrne, M., & Carr, A. (2000). Depression and power in marriage. *Journal of Family Therapy* 22(4): 408-427.

- Byrne, M., Carr, A., & Clark, M. (2004). Power in relationships of women with depression. *Journal of Family Therapy* 26(4): 407-429.

- Campbell, M. L., & Morrison, A. P. (2007). The subjective experience of paranoia: Comparing the experiences of patients with psychosis and individuals with no psychiatric history. *Clinical Psychology & Psychotherapy* 14(1): 63-77.

- Cantor-Graae, E. & Selten, J. P.

  (2005a). Schizophrenia and migration: a meta-analysis and review. *American Journal of Psychiatry* 162 (1): 12-24.

  (2005b) Social defeat: risk factor for schizophrenia? *The British Journal of Psychiatry* 187(2): 101-102.

- Carruthers, P. (2016). Are epistemic emotions metacognitive? *Philosophical Psychology* 29(8): 1-15.

- Castelnovo, A. *et al.* (2015). Post-bereavement hallucinatory experiences: A critical overview of population and clinical studies. *Journal of Affective disorders* 186: 266-274.

- Christodoulou, G. N. (1977). The syndrome of Capgras. *The British Journal of Psychiatry*, 130(6): 556-564.

- Cleese, J. (1991). Lecture on creativity. (Available online: https://genius.com/John-cleese-lecture-on-creativity-annotated)

- Collip, D., Myin-Germeys, I., & Van Os, J. (2008). Does the concept of "sensitization" provide a plausible mechanism for the putative link between the environment and schizophrenia? *Schizophrenia Bulletin* 34(2): 220-225.

- Connery, D.S. (1977). *Guilty until proven innocent.* New York: Putnam.

- Cooper, C. *et al.* (2008). Perceptions of disadvantage, ethnicity and psychosis. *The British Journal of Psychiatry* 192(3): 185-190.

- Cooper, R.

  (2005). *Classifying Madness.* Dordrecht: Springer.

  (2004). What is wrong with the DSM? *History of Psychiatry* 15(1): 5-25.

- Costello, E. J. *et al.* (2003). Relationships between poverty and psychopathology: A natural experiment. *JAMA* 290(15): 2023-2029.

- Courban P. & Fail G. (1927) Syndrome d'illusion de Fregoli et schizophrénie. *Bulletin Social Clinique Médication Mental* 15: 121-124.

- Covington H. & Miczek, K. (2001). Repeated social-defeat stress, cocaine or morphine. Effects on behavioral sensitization and intravenous cocaine self-administration binges. *Psychopharmacology* 158(4): 388-398.

- Craig, T. & Pathare, S. (2001). Gender differences in the experience and response to adversity. Harris (ed.) *When Inner and Outer Worlds Meet*. Routledge.

- Cross, P. (1977). Not Can But Will College Teachers Be Improved? *New Directions for Higher Education* (17): 1–15.

- Damasio, A.

  (1999). *The feeling of what happens: body and emotion in the making of consciousness.* Houghton Mifflin.
  (1996). The Somatic Marker Hypothesis and the Possible Functions of the Prefrontal Cortex. *Philosophical Transactions of the Royal Society: Biological Sciences* 351: 1413-1420.

- Darwin, C. (1859). *On the Origins of Species by Means of Natural Selection*. London: Murray.

- Das, J., Das, RK, Das, V. (2012). The Mental Health Gender Gap in Urban India: Patterns and Narratives. *Social Science and Medicine* 75: 1660-1672.

- Das Munshi, J. et al. (2012). Ethnic Density as a Buffer for Psychotic Experiences: Findings from a National Survey (EMPIRIC). *The British Journal of Psychiatry* 201: 282-290.

- Descartes, R., (2006/1641). *Meditations, objections, and replies.* Hackett Publishing.

- Dienes (2012). Is hypnotic responding the strategic relinquishment of metacognition? *The Foundations of Metacognition*, Beran, M., Brandl, J. L., Perner, J., & Proust, J. (eds.). Oxford University Press.

- Dijksterhuis, A., & Meurs, T. (2006). Where creativity resides: The generative power of unconscious thought. *Consciousness and Cognition* 15(1): 135-146.

- Dub, R. (2015). Delusions, Acceptances, and Cognitive Feelings. *Philosophy and Phenomenological Research*: 1-34.

- Dunlosky, J., & Metcalfe, J. (2008). *Metacognition*. Sage Publications.

- Elgin, C.Z. (2007). Emotion and Understanding. In G. Brun, U. Dogouglu and D. Kunzle (eds.) *Epistemology and Emotions*, London: Ashgate.

- Ellis, R. D., & Faw, B. (2012). Concluding observations Comparisons, contrasts. In *Categorical versus Dimensional Models of Affect: A seminar on the theories of Panksepp and Russell* (Vol. 7, p. 321). John Benjamins Publishing.

- Elpidorou, A. (2013). Moods and Appraisals: How the phenomenology and science of emotions can come together. *Human Studies* 36(4): 565-591.

- Emmons, K. (2008). Narrating the Emotional Woman: Uptake and Gender in Discourses on Depression. *Depression and Narrative: Telling the Dark.*

- Estes, Z. & Felker, S. (2012). Confidence mediates the sex difference in mental rotation performance. *Archives of Sexual Behavior* 41(3): 557-570.

- Fairbrother, N., & Woody, S. R. (2008). New mothers' thoughts of harm related to the newborn. *Archives of women's mental health* 11(3): 221-229.

- Falkenberg, I. *et al.* (2007). Sense of humor in patients with schizophrenia. *Schizophrenia Research* 95(1): 259-261.

- Fearon, P., Kirkbride, J.B. & Morgan, C. (2006). Incidence of schizophrenia and other psychoses in ethnic minority groups: results from the MRC AESOP Study. *Psychological Medicine* 36(11): 1541-1550.

- Fearon, P. & Morgan, C. (2006). Environmental factors in schizophrenia: the role of migrant studies. *Schizophrenia Bulletin 32*(3): 405-408.

- Fernandez, A. (forthcoming). Phenomenology and Dimensional Approaches to Psychiatric Research and Classification. *Philosophy, Psychiatry, & Psychology*.

- Fine, C. (2006). *A mind of its own. How our brain distorts and deceives.* Thriplow, England: Icon Books.

- Firestone, C. & Scholl, B. (2016). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *Behavioral and Brain Sciences* 39: 1-77.

- Frazer, S. J., & Roberts, J. M. (1994). Three cases of Capgras' syndrome. *The British Journal of Psychiatry* 164(4): 557-559.

- Freeman, D. *et al.* (2002). A Cognitive model of persecutory delusions. *British Journal of Clinical Psychology* 41: 331–347.

- French, L., Garry, M., & Loftus, E. (2009). False memories: A kind of confabulation in non-clinical. *Confabulation: Views from Neuroscience, Psychiatry, Psychology, and Philosophy.* Oxford University Press.

- Freud, S., & Breuer, J. (1893). On the psychical mechanism of hysterical phenomena: Preliminary communication. *Standard Edition* 2. London: Hogarth.

- Freud, S.

  (1894). The Neuro-psychoses of Defence. *Standard Edition* 3. London: Hogarth.

  (1895a). *Studies on Hysteria*. *Standard Edition* 2. London: Hogarth.

  (1895b). Obsessions and Phobias. *Standard Edition* 3. London: Hogarth.

  (1896). Further Remarks on the Neuro-psychoses of Defence. *Standard Edition* 3. London: Hogarth.
  (1901). *The Psychopathology of Everyday Life*. WW Norton & Company.

  (1905). *Jokes and their relation to the unconscious*. Standard Edition 8. Trans. and ed. J. Strachey. London: Hogarth Press.

  (1907). Obsessive Actions and Religious Practices. *Standard Edition* 9. London: Hogarth.

  (1909). Analysis of a phobia in a five-year-old boy. *Standard Edition* 10 (3). London: Hogarth.

  (1910). *Five Lectures on Psycho-analysis*. WW Norton & Company.

  (1911a). Formulations on the Two Principles of Mental Functioning.

  (1911b). *Psychoanalytic Notes on an Autobiographical Account of a Case of Paranoia. Standard Edition* 12. London: Hogarth.

  (1912). Types of Onset of Neurosis. *Standard Edition* 12. London: Hogarth.

(1913a). Observations and Examples from Analytic Practice. *Standard Edition* 12. London.

(1913b). The Disposition to Obsessional Neurosis. *Standard Edition* 12. London: Hogarth.

(1914). Remembering, Repeating and Working Through. *Standard Edition* 12. London: Hogarth.

(1915a). Instincts and their Vicissitudes. *Standard Edition* 14. London: Hogarth.

(1915b). Repression. *Standard Edition* 14. London: Hogarth

(1915c). The Unconscious. *Standard Edition* 14. London: Hogarth.

(1916-17). *Introductory Lectures on Psychoanalysis*. *Standard Edition* 19. London: Hogarth.

(1917). Mourning and Melancholia. *Standard Edition* 14. London: Hogarth.

(1923). The Ego and the Id. *Standard Edition* 19. London: Hogarth.

(1924a). Neurosis and Psychosis. *Standard Edition* 19. London: Hogarth.

(1924b). The Loss of Reality in Neurosis and Psychosis. *Standard Edition* 19. London: Hogarth.

(1926). *Inhibitions, Symptoms and Anxiety*. WW Norton & Company.

(1933). *New Introductory Lectures on Psycho-analysis*. WW Norton & Company.

(1937). Analysis Terminable and Interminable. *Standard Edition* 23. London: Hogarth.

(1938). *An Outline of Psychoanalysis. Standard Edition*. London: Hogarth Press.

(1950). *Project for a Scientific Psychology*. *Standard Edition* 1. London: Hogarth.

- Frijda, N. H. (1986). *The emotions: Studies in emotion and social interaction.* Cambridge University Press.
- Fuchs, T. (2003). The phenomenology of shame, guilt and the body in body dysmorphic disorder and depression. *Journal of Phenomenological Psychology*, 33(2): 223-243.
- Fusar-Poli, P. *et al.* (2013). The psychosis high-risk state: a comprehensive state-of-the-art review. *JAMA Psychiatry* 70(1): 107-120.
- Garety, P. A., & Freeman, D. (1999). Cognitive approaches to delusions: a critical review of theories and evidence. *British Journal of Clinical Psychology* 38 (2): 113-154.
- Gast, L. (1996). *When we were kings.*

- Gerrans, P. (2000). Refining the explanation of Cotard's delusion. *Mind & Language* 15(1): 111-122.

- Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond "heuristics and biases". *European Review of social Psychology* 2(1): 83-115.

- Goldberg, L. R. (1990). An alternative description of personality: the big-five factor structure. *Journal of personality and social psychology* 59(6): 1216-1229.

- Goldie, P. (2009). Getting feelings into emotional experience in the right way. *Emotion Review* 1(3): 232-239.

- Goleman, D. (1991). Feeling unreal? Many others feel the same. *New York Times*.

- Graham, G.

  (2015). *The Abraham Dilemma: A Divine Delusion*. Oxford University Press.
  (2013). *The Disordered mind: An Introduction to Philosophy of Mind and Mental Illness*. Routledge.

- Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology* 39(3): 281-291.

- Gudjonsson, G. H. (2003). *The psychology of interrogations and confessions: A handbook*. John Wiley & Sons.

- Gutiérrez-Lobos, K. *et al.* (2000). The gender gap in depression reconsidered: the influence of marital and employment status on the female/male ratio of treated incidence rates. *Social Psychiatry and Psychiatric Epidemiology* 35(5): 202-210.

- Haslam, N. (2002). Kinds of kinds: A conceptual taxonomy of psychiatric categories. *Philosophy, Psychiatry, & Psychology* 9(3): 203-217.

- Hawkes, E. (2012). Making Meaning. *Schizophrenia bulletin* 38(6): 1109-1110.

- Heidegger, M. (1927/1962). *Being and Time*. SUNY Press.

- Henz, U.

  (2010). Parent care as unpaid family labor: How do spouses share? *Journal of Marriage and Family*, *72*(1): 148-164.

  (2004). The effects of informal care on paid-work participation in Great Britain: a lifecourse perspective. *Ageing and Society 24*: 851-880.

- Hinkle, L.E. & Wolff, H.G. (1956). Communist interrogation and indoctrination of 'enemies of the states'. *American Medical Association Archives of Neurology and Psychiatry* 76: 115–174.

- Hirstein, W. (2009). *Confabulation. Views from Neuroscience, Psychiatry, Psychology and Philosophy.* Oxford University Press.
- Hobson, J. A. (2001). *The dream drugstore: Chemically altered states of consciousness.* MIT Press.
- Horwitz, A. V. & Wakefield, J. C. (2007). *The loss of sadness: How psychiatry transformed normal sorrow into depressive disorder*. Oxford University Press.
- Hochschild, A. (1989). *The second shift: Working parents and the revolution at home.* Penguin Books.
- Humpston, C. S., & Broome, M. R. (2016). Perplexity. In *An Experiential Approach to Psychopathology* (245-264). Springer International Publishing.
- Ivanova, A. *et al.* (2014). Sense of humor disorders in patients with schizophrenia and affective disorders. *Psychology in Russia: State of the art* 7(1): 146-157.
- Jamison, K. R. (1996). *An Unquiet Mind.* Vintage Books Edition.
- Janet, P. (1894). *État mental des hystériques*. Rueff.
- Janssen, I. et al. (2003). Discrimination and delusional ideation. *The British Journal of Psychiatry* 182(1): 71-76.
- Jaspers, K. (1913/1963). *General Psychopathology.* Chicago University Press.
- Jenkins, R. (1987) Sex differences in depression. *British Journal Hospital Medicine* 38: 485-486.
- Jin, H.M. *et al*. (2015). The effects of social defeat on behavior and dopaminergic markers in mice. *Neuroscience* 288: 167-177.
- Johns, L. C., & van Os, J. (2001). The continuity of psychotic experiences in the general population. *Clinical psychology review* 21(8): 1125-1141.
- Jones, S. R., & Fernyhough, C. (2007). A new look at the neural diathesis-stress model of schizophrenia: the primacy of social-evaluative and uncontrollable situations. *Schizophrenia Bulletin* 33(5): 1171-1177.
- Jordan, H. W., Lockert, E. W., Johnson-Warren, M., Cabell, C., Cooke, T., Greer, W., & Howe, G. (2006). Erotomania revisited: thirty-four years later. *Journal of the National Medical Association* 98(5): 787-793.
- Kalmijn, M. (2017). The Ambiguous Link between Marriage and Health: A Dynamic Reanalysis of Loss and Gain Effects. *Social Forces* 95(4): 1607-1636.

- Kapur, S.

  (2004). Psychosis as a State of Aberrant Salience: A Framework Linking Biology, Phenomenology, and Pharmacology in Schizophrenia. *American Journal of Psychiatry* 160: 13-23.

  (2003). How Antipsychotics Become Anti-'Psychotic' – from Dopamine to Salience to Psychosis. *Trends in Pharmacological Sciences* 25: 402-406.

- Kay, K. & Shipman, C. (2014). The confidence gap. *The Atlantic* 14: 1-18.

- Kendell, R. & Jablensky, A. (2003). Distinguishing between the validity and utility of psychiatric diagnoses. *American Journal of Psychiatry* 160: 4-12.

- Kendell, R. (1975). The Concept of Disease and its Implications for Psychiatry. *British Journal of Psychiatry* 127: 305-315.

- Kessler, R.C. (2003). Epidemiology of women and depression. *Journal of Affective Disorders* 74(1): 5-13.

- Killmister, S. (2015). The Woody Allen Puzzle: How 'Authentic Alienation'Complicates Autonomy. *Noûs* 49(4): 729-747.

- Kirkbride, J. B. *et al.* (2007). Neighbourhood-level effects on psychoses: re-examining the role of context. *Psychological Medicine* 37(10): 1413-1425.

- Kitcher, P. (1995). *Freud's Dream: A Complete Interdisciplinary Science of Mind*. MIT Press.

- Kobasa, S. C. (1979). Stressful life events, personality, and health: an inquiry into hardiness. *Journal of personality and social psychology* 37(1): 1-11.

- Kochuparackal, T., & Simon, A. E. (2012). A "Contemporary" Case of Frégoli Syndrome. *The Primary Care Companion to CNS Disorders* 14(1).

- Koriat, A. (2000). The feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and Cognition* 9(2): 149-171.

- Kruger, J. & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology* 77(6): 1121-1134.

- LaFrance, M. N. (2007). A Bitter Pill A Discursive Analysis of Women's Medicalized Accounts of Depression. *Journal of Health Psychology* 12(1): 127-140.

- Langdon, R. & Coltheart, M. (2014). The Fregoli delusion: a disorder of person identification and tracking. *Topics in cognitive science* 6(4): 615-631.

- Langland-Hassan, P. (2008). Fractured phenomenologies: Thought insertion, inner speech, and the puzzle of extraneity. *Mind & Language* 23(4): 369-401.

- Larson, R. W., Richards, M. H., & Perry-Jenkins, M. (1994). Divergent worlds: the daily emotional experience of mothers and fathers in the domestic and public spheres. *Journal of Personality and Social Psychology* 67(6): 1034-1046.

- Lazarus, R.

  (1993). From psychological stress to the emotions: A history of changing outlooks. *Annual Review of Psychology* 44: 1-21.
  (1991). *Emotion and Adaptation.* Oxford University Press.

- Lazarus, R. S., & Folkman, S. (1984). Stress, appraisal, and coping. *Springer*.

- Longden, E. (2013). Learning from the voices in my head. In Monterey, CA: TED Conferences. http://www.ted.com/talks/eleanor_longden_the_voices_in_my_head?language=en#t-203777

- López-Silva, P. (2015). Schizophrenia and the Place of Ego-dystonic States in the Aetiology of Thought Insertion. *Review of Philosophy and Psychology*: 1-18.

- Luhrmann, T.

  (2007). Social defeat and the culture of chronicity: or, why schizophrenia does so well over there and so badly here. *Culture, Medicine and Psychiatry* 31(2): 135-172.

  (2005). The art of hearing God: Absorption, dissociation, and contemporary American spirituality. *Spiritus: A Journal of Christian Spirituality* 5(2): 133-157.

- Maibom, H. L. (2016). Failure to Forecast and the empathic Imagination. *Knowledge Through Imagination*, Kind, A. & Kung, P. (eds.). Oxford University Press.

- Maiese (2015). Auditory Verbal Hallucinations and Sense of Ownership. Paper presented at the *Southern Society of Philosophy and Psychology*: New Orleans.

- Makari, G. (2008). *Revolution in Mind: The Creation of Psychoanalysis*. New York: Harper Collins.

- Marraffa, M., & Viola, M. (2017). Quale mappa per il dominio delle emozioni? *Sistemi Intelligenti* 29(1).

- Mason, M. F. *et al.* (2007). Wandering minds: the default network and stimulus-independent thought. *Science* 315(5810): 393-395.

- Matuszak, J., & Parra, M. (2011). That's not my child: a case of Capgras syndrome. *Psychiatric Times* 28(2): 1-4.

- Meehl, P. E. (1992). Factors and taxa, traits and types, differences of degree and differences in kind. *Journal of Personality* 60(1): 117-174.
- Meier, B. P. *et al.* (2007) When "light" and "dark" thoughts become light and dark responses: Affect biases brightness judgments. *Emotion* 7(2): 366-376.
- Merleau-Ponty, M. (1945/1958). *Phenomenology of Perception*. Routledge.
- Merton, R. K. (1948). The self-fulfilling prophecy. *The Antioch Review*, 8(2): 193-210.
- Metcalfe, J. (2010). Metacognition of agency across the lifespan. *Cognition* 116(2): 267-282.
- Metzinger, T. K. (2013). The myth of cognitive agency: subpersonal thinking as a cyclically recurring loss of mental autonomy. *Frontiers in Psychology* 4: 1-19.
- Minkowski, E.(1923). Contribution l'étude de l'autisme, l'attitude interrogative. In *Annales médico-psychologiques* 81(1): 421-435.
- Mitchell, J. (1974). *Psychoanalysis and feminism*. Pantheon.
- Morgan, C. & Hutchinson, G. (2010). The social determinants of psychosis in migrant and ethnic minority populations: a public health tragedy. *Psychological Medicine* 1(1): 705-7
- Moriyama, Y. *et al.* (2007). Frégoli syndrome accompanied with prosopagnosia in a woman with a 40-year history of schizophrenia. *The Keio journal of medicine* 56(4): 130-134.
- Morrison, A. P., French, P., & Wells, A. (2007). Metacognitive beliefs across the continuum of psychosis: Comparisons between patients with psychotic disorders, patients at ultra-high risk and non-patients. *Behaviour research and therapy* 45(9): 2241-2246.
- Morrison, A. P. *et al.* (2006). Psychological factors in people at ultra-high risk of psychosis: comparisons with non-patients and associations with symptoms. *Psychological Medicine* 36(10): 1395-1404.
- Muris, P., Merckelbach, H., & Clavan, M. (1997). Abnormal and normal compulsions. *Behaviour Research and Therapy* 35(3): 249-252.
- Murphy, D. (2006). *Psychiatry in the scientific image*. Cambridge, MA: MIT Press.
- Naranjo, J. R. & Schmidt, S. (2012). Is it me or not me? Modulation of perceptual-motor awareness and visuomotor performance by mindfulness meditation. *BMC Neuroscience* 13(1): 1-17.
- Nazroo, J. Y., Edwards, A. C., & Brown, G. W. (1998). Gender differences in the prevalence of depression: artefact, alternative disorders, biology or roles? *Sociology of Health & Illness* 20(3): 312-330.

- Nolen-Hoeksema, S., Larson, J., & Grayson, C. (1999). Explaining the gender difference in depressive symptoms. *Journal of Personality and Social Psychology* 77(5): 1061-1072.

- Oakley, J. G. (2000). Gender-based barriers to senior management positions: Understanding the scarcity of female CEOs. *Journal of Business Ethics* 27(4): 321-334.

- O'Connor, A. R., & Moulin, C. J. (2010). Recognition without identification, erroneous familiarity, and déjà vu. *Current Psychiatry Reports* 12(3): 165-173.

- Ødegaard, O. (1932). Emigration and insanity. *Acta Psychiatrica et Neurologica* 4(1): 1-206.

- Ofshe, R. (1989). Coerced confessions: the logic of seemingly irrational action. *Cultic Studies Journal* 6: 1–15.

- Ormel, J. & Neelman, J. (2001). Toward a dynamic stress-vulnerability model of depression: the role of neuroticism, life events and gender. Harris (ed.) *When Inner and Outer Worlds Meet*. Routledge.

- Ottesen, A. *et al.* (2014). The impact of immigration and visible minority status on psychosis symptom profile. *Social psychiatry and psychiatric epidemiology* 49(11): 1747-1757.

- Papageorgiou, C., Lykouras, L., Ventouras, E., Uzunoglu, N., & Christodoulou, G. N. (2002). Abnormal P300 in a case of delusional misidentification with coinciding Capgras and Fregoli symptoms. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 26(4): 805-810.

- Parkes, C.M.

  (1972). *Bereavement: Studies of grief in adult life.* London: Tavistock Publications.

  (1970). "Seeking" and "finding" a lost object: Evidence from recent studies of the reaction to bereavement. *Social Science & Medicine* 4(2): 187-201.

- Parnas, J. & Gallagher, S. (2015). Phenomenology and the interpretation of psychopathological experience. *Re-visioning psychiatry: Cultural phenomenology, critical neuroscience, and global mental health* (65-80).

- Pessoa, L. (2013). *The cognitive-emotional brain: From interactions to integration.* MIT Press.

- Piccinelli, M. & Wilkinson, G. (2000). Gender differences in depression. *The British Journal of Psychiatry* 177(6): 486-492.

- Pinquart, M., & Sörensen, S. (2006). Gender differences in caregiver stressors, social resources, and health: An updated meta-analysis. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 61(1): 33-45.

- Plath, S. (1963). *The Bell Jar.* Faber & Faber.

- Proust, J.

  (2013). *The Philosophy of Metacognition: Mental agency and self-awareness.* Oxford University Press.
  (2009). Overlooking metacognitive experience. *Behavioral and Brain Sciences* 32(02): 158-159.
  (2006). Agency in schizophrenia from a control theory viewpoint. *The disorders of volition* (87-118). MIT Press.

- Pulford, B. D. (1996). *Overconfidence in human judgment.* University of Leicester (Unpublished Doctoral Dissertation).

- Purdon, C., & Clark, D. A. (1994). Obsessive intrusive thoughts in nonclinical subjects. Part II. Cognitive appraisal, emotional response and thought control strategies. *Behaviour Research and Therapy* 32(4): 403-410.

- Rachman, S., & de Silva, P. (1978). Abnormal and normal obsessions. *Behaviour research and therapy* 16(4): 233-248.

- Radden, J. (2009). *Moody minds distempered: Essays on melancholy and depression.* Oxford University Press.

- Ratcliffe, M.

  (forthcoming). Depression, Self-regulation, and Intersubjectivity.

  (2013). Why mood matters. *The Cambridge companion to Heidegger's Being and Time* (157-176)

  (2015). *Experiences of Depression.* Oxford University Press.

  (2010). Depression, guilt and emotional depth. *Inquiry* 53(6): 602-626.

  (2008). *Feelings of being: Phenomenology, psychiatry and the sense of reality.* Oxford University Press.

  (2005). The feeling of being. *Journal of Consciousness Studies* 12(8-9): 43-60.

- Rees, W. D. (1971). The hallucinations of widowhood. *British Medical Journal* 4(5778): 37-41.

- Reina, A. (2009). The Spectrum of Sanity and Insanity. *Schizophrenia Bulletin* 36(1): 3-8.

- Rogers, C. R. (1959). A theory of therapy, personality, and interpersonal relationships: As developed in the client-centered framework. In *Psychology: A Study of Science* 1(3), Koch, S. (ed.). McGraw Hill.

- Romme, M., & Morris, M. (2013). The recovery process with hearing voices: accepting as well as exploring their emotional background through a supported process. *Psychosis*, 5(3): 259-269.

- Romme, M., & Escher, S.

  (2013). *Psychosis as a personal crisis: An experience-based approach*. Routledge.

  (1993). *Accepting Voices*. London: Mind Publications.

- Ross, C. A., Norton, G. R. & Anderson, G. (1988). The Dissociative Experiences Scale: A replication study. *Dissociation: Progress in the Dissociative Disorders*.

- Roth, P. (1969). *Portnoy's Complaint*. New York: Bantam.

- Roxburgh, S. (2005). Parenting strains, distress, and family paid labor: A modification of the cost-of-caring hypothesis. *Journal of Family Issues* 26(8): 1062-1081.

- Ruscio, J., & Ruscio, A. M. (2000). Informing the continuity controversy: a taxometric analysis of depression. *Journal of Abnormal Psychology* 109(3): 473-487.

- Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *Journal of personality and social psychology* 76(5), 805-819.

- Russell, J.

  (2003). Core affect and the psychological construction of emotion. *Psychological Review* 110(1): 145-172.
  (1980). A Circumplex Model of Affect. *Journal of Personality and Social Psychology* 39 (6): 1161-1178.

- Saks, E. (2007). *The Center Cannot Hold: My Journey Through Madness*. Hachette UK.

- Samuels, R. (2009). Delusion as a natural kind. *Psychiatry as cognitive neuroscience: Philosophical perspectives* (49-79). Oxford University Press.

- Sander, D., Grafman, J., & Zalla, T. (2003). The human amygdala: an evolved system for relevance detection. *Reviews in the Neurosciences* 14(4): 303-316.

- Sartre, J. P. (1939/1962). *The Emotions: Sketch of a Theory* (translated by B. Frechtman). Philosophical Library [GA]

- Sass, L. (2004). Affectivity in schizophrenia: a phenomenological view. *Journal of Consciousness Studies* 11(10-11): 127-147.

- Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review* 115(2): 336-356.

- Schneider, K. (1938/1959). *Clinical Psychopathology*. New York: Grune & Stratton.

- Schreber, D. P. (1903/1955). *Memoirs of my nervous illness*. New York Review of Books.

- Schupak, C. & Rosenthal, J. (2009). Excessive daydreaming: A case history and discussion of mind wandering and high fantasy proneness. *Consciousness and Cognition* 18(1): 290-292.

- Seaman, M.V. (1997). Psychopathology in Women and Men: Focus on Female Hormones. *American Journal of Psychiatry* 154(12): 1641–7.

- Selten, J. P. *et al.* (2002). Incidence of schizophrenia in Surinam. *Schizophrenia Research* 54(3): 219-221.

- Simon, R. W. (2002). Revisiting the Relationships among Gender, Marital Status, and Mental Health. *American Journal of Sociology* 107(4): 1065-1096.

- Sizer, L. (2006). What feelings can't do. *Mind & Language* 21(1): 108-135.

- Szalavitz, M. (2016). Opioid Addiction Is a Huge Problem, but Pain Prescriptions Are Not the Cause. *Scientific American*.
  Available online: https://blogs.scientificamerican.com/mind-guest-blog/opioid-addiction-is-a-huge-problem-but-pain-prescriptions-are-not-the-cause/

- Smallwood, J. & Andrews-Hanna, J. (2013). Not all minds that wander are lost: the importance of a balanced perspective on the mind-wandering state. *Frontiers in Psychology* 4: 1-6.

- Smallwood, J. & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin* 132(6): 946-958.
- Smith, J. D. (2005). Uncertainty monitoring and metacognition by animals. *Current Directions in Psychological Science* 14(1): 19-24.

- Stanghellini, G., & Ambrosini, A. (2011). Depressione postpartum e typus melancholicus. Uno studio pilota. *Giornale Italiano di Psicopatologia* 17: 22-30.

- Stanghellini, G.

  (2004). *Disembodied spirits and deanimated bodies: The psychopathology of common sense*. Oxford University Press.

  (2000). Vulnerability to schizophrenia and lack of common sense. *Schizophrenia Bulletin* 26(4): 775-787.

- Steinberg, M. (2001). *The Stranger in the Mirror. Dissociation – The Hidden Epidemic*. Harper-Collins

- Stephan, A. (2012). Emotions, Existential Feelings and their Regulation. *Emotion Review* 4(2): 157-162.

- Stewart, E. S. & Shapiro, L. (2011). Pathological guilt: A persistent yet overlooked treatment factor in obsessive-compulsive disorder. *Annals of Clinical Psychiatry* 23(1): 63-70.

- Stocker, M. (2004). Some considerations about intellectual desire and emotions. Solomon, R. (ed.) *Thinking About Feeling: Contemporary Philosophers on Emotions.* Oxford University Press.

- Styron, W. (1991). *Darkness Visible.* Random House.

- Sutherland, S. (2007). *Irrationality: The Enemy Within.* Pinter & Martin.

- Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica* 47(2): 143-148.

- Tamietto, M., Castelli, L., Vighetti, S., Perozzo, P., Geminiani, G., Weiskrantz, L., & de Gelder, B. (2009). Unseen facial and bodily expressions trigger fast emotional reactions. *Proceedings of the National Academy of Sciences* 106(42): 17661-17666.

- Tarricone, I. *et al.* (2015). Migration history, minorities status and risk of psychosis: an epidemiological explanation and a psychopathological insight. *Official Journal of the Italian Society of Psychopathology* 46: 424-430.

- Teasdale, J. D. (1999). Metacognition, mindfulness and the modification of mood disorders. *Clinical Psychology & Psychotherapy* 6(2): 146-155.

- Thayer, R. E. (1996). *The Origin of Everyday Moods*. Oxford University Press.

- Thompson, E. (2014). *Waking, dreaming, being: Self and consciousness in neuroscience, meditation, and philosophy*. Columbia University Press.

- Tidey, J. W. & Miczek, K. A. (1996). Social defeat stress selectively alters mesocorticolimbic dopamine release: an in vivo microdialysis study. *Brain Research* 721(1): 140-149.
- Tranel, D., & Damasio, A. (1993). The covert learning of affective valence does not require structures in hippocampal system or amygdala. *Journal of Cognitive Neuroscience* 5(1): 79-88.

- Turkiewicz, G., Zanetti, M. V., Zung, S., & Cordeiro, Q. (2009). Coexistence of Capgras and Frégoli syndromes associated to frontotemporal volume reduction and cerebral white matter hyperintensities. *Archives of Clinical Psychiatry* 36(6): 240-243.

- Turner, R. J., & Avison, W. R. (2003). Status variations in stress exposure: Implications for the interpretation of research on race, socioeconomic status, and gender. *Journal of Health and Social Behavior* 44(4): 488-505.

- Turner, R. J., & Lloyd, D. A. (1999). The stress process and the social distribution of depression. *Journal of Health and Social Behavior* 40(4): 374-404.

- Ussher, J. M. (2010). Are we medicalizing women's misery? A critical review of women's higher rates of reported depression. *Feminism & Psychology* 20(1): 9-35.

- van Os, J., Linscott, R. J., Myin-Germeys, I., Delespaul, P., & Krabbendam, L. (2008). A systematic review and meta-analysis of the psychosis continuum: evidence for a psychosis proneness-persistence-impairment model of psychotic disorder. *Psychological Medicine* 39(2): 179-195.

- van Os, J.

  (2012). Psychotic experiences: disadvantaged and different from the norm. *The British Journal of Psychiatry* 201: 258-259.

  (2004). Does the urban environment cause psychosis? *The British Journal of Psychiatry* 184(4): 287-288.

  (2003). Can the social environment cause schizophrenia? *The British Journal of Psychiatry* 182 (4): 291-292.

  (1998). The ecogenetics of schizophrenia: a review. *Schizophrenia Research* 32(2): 127-135.

- Vargas-Tonsing, T. M. (2009). An exploratory examination of the effects of coaches' pre-game speeches on athletes' perceptions of self-efficacy and emotion. *Journal of Sport Behavior* 32(1): 92-111.

- Vazire, S. & Carlson, E. N. (2011). Others sometimes know us better than we know ourselves. *Current Directions in Psychological Science* 20: 104-108.

- Vazire, S. (2010). Who knows what about a person? The Self-Other Knowledge Asymmetry (SOKA) model. *Journal of Personality and Social Psychology* 28: 281-300.

- Vázquez-Barquero, J. L. (1992). Sex differences in mental illness: a community study of the influence of physical health and sociodemographic factors. *Social psychiatry and psychiatric epidemiology* 27(2): 62-68.

- Vega, W. A. *et al.* (1998). Lifetime prevalence of DSM-III-R psychiatric disorders among urban and rural Mexican Americans in California. *Archives of General Psychiatry* 55(9): 771-778.

- Veling, W. *et al.* (2008). Ethnic density of neighborhoods and incidence of psychotic disorders among immigrants. *American Journal of Psychiatry* 165(1): 66-73.

- Weiser, M. *et al.* (2008). Elaboration on immigration and risk for schizophrenia. *Psychological Medicine* 38(08): 1113-1119.

- Weissman MM, Klerman GL (1977) Sex differences and the epidemiology of depression. *Archives of General Psychiatry* 3: 98-112.

- Wheatley, T. (2009). Everyday confabulation. *Confabulation: Views from neuroscience, psychiatry, psychology, and philosophy* (203-222). Oxford University Press.

- Wilder, B. (1955). *Seven Year Itch.*

- Williams, T.K. (2012). *Understanding internalized oppression: A theoretical conceptualization of internalized subordination.* Unpublished doctoral dissertation.

- Williamson, V. & McCutcheon, H. (2004). Postnatal depression: a review of current literature. *Australian Midwifery* 17(4): 11-16.

- Wilson, T. D. (2002). *Strangers to ourselves: Self-insight and the adaptive unconscious.* Harvard University Press.

- Wittgenstein, L. (1972/1953). *On Certainty.* Harper & Row.

- Wong, M. Y. (2016). Towards a theory of mood function. *Philosophical Psychology* 29(2): 179-197.

- Yancy, G. (2008). Elevators, social spaces and racism. A philosophical analysis. *Philosophy & Social Criticism* 34(8): 843-876.

- Young, V. (2011). *The secret thoughts of successful women: Why capable people suffer from the impostor syndrome and how to thrive in spite of it*. Crown Pub.

- Zachar, P. (2000). Psychiatric disorders are not natural kinds. *Philosophy, Psychiatry, & Psychology* 7(3):167-182.

- Zahavi, D. (2014). *Self and other: Exploring Subjectivity, Empathy, and Shame*. Oxford University Press.

- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35(2): 151-175.

- Zuckerman, E. W. & Jost, J. T. (2001). What makes you think you're so popular? Self-evaluation maintenance and the subjective side of the 'friendship paradox'. *Social Psychology Quarterly*: 207-223.