THE EXPLANATION OF DISORDERED EATING: A METATHEORETICAL APPROACH

by

Hannah Hawkins-Elder

A thesis

submitted to the Victoria University of Wellington

in fulfilment of the requirements for the degree of

Doctor of Philosophy

In Psychology

Victoria University of Wellington

(2021)

Abstract

Eating disorders are an increasingly serious global health problem, resulting in substantial physical and mental health costs for sufferers and their families, and for society as a whole. There has been considerable research aimed at understanding and treating these conditions. Despite this, treatment for disordered eating remains relatively poor, leaving many with ongoing suffering. Although there are many reasons why this may be the case, a possible contributor not investigated by the literature is the quality of the theoretical explanations underpinning these treatments. In this thesis, therefore, I set out to examine the quality of the current dominant explanations for disordered eating upon which best-practice treatments are based. Upon evaluation, these explanations are found to be significantly lacking in quality, leading to subsequent weaknesses in their interventions. Underlying these specific problems, I identify three common meta-theoretical issues with these explanations: (1) the poor selection and description of explanatory targets, (2) the use of single explanatory models as complete explanations (i.e., the One Best Model perspective), and (3) the lack of systematic method informing their construction. In the remainder of the thesis, I propose solutions for each of these, arguing for (1) more appropriate selection and *compositional explanation* of explanatory foci, (2) transition away from a unified approach to explanation toward *model pluralism*, and (3) the development and application of a guiding methodological framework for explanation in psychopathology. Each of these solutions represents a significant shift in our approach to understanding and explaining psychopathological conditions. Their application to the specific problem area of disordered eating is likely to result in explanations for these conditions that are more valid, comprehensive, and fruitful, thereby providing superior platforms for clinical intervention and research.

Acknowledgements

I would like to begin by offering an especially large thank you to my supervisor, Tony Ward, whose intellectual and personal guidance have been invaluable throughout this process. With his encouragement and support, I have learned to have confidence in my own ideas and opinions, and to value my work. I could not have managed this feat without his help. Many thanks, Tony.

I would also like to thank the intellectual community I have been a part of here at Victoria – both the clinical psychology programme and the Explanation of Psychopathology and Crime (EPC) lab, as well as other stand-alone individuals. Each has provided its own source of intellectual stimulation and contributed in its own way to the production of this thesis. Special mention to Kristopher Nielsen, with whom I was able to share many philosophical discussions within the first two years of my PhD and whose work helped to inspire my own in many ways, as well as to the various postgraduate friends and colleagues who aided in the proofreading of this thesis and supporting me during its composition. I am very grateful.

Finally, the biggest thank you must go to the many friends and loved ones who have supported me throughout this process. Thank you to my close friends and flatmates, Adam and Alex, who have been understanding and consistent throughout my thesis journey. Thank you also to the Wellington Brazilian Zouk community, who have helped to bring a lot of joy to my life outside of university that has been vital to getting me through these last three years. To my partner, Connor Kennedy: thank you for all your love and support. Although you have only been with me for half of this process, you have been an important part of its completion. And last but definitely most, to my family – especially my parents, Lesley and Ron, and my sister Jamie – thank you for always believing in me and my abilities and supporting me at every turn. Not every person is as lucky, and I am grateful for it every day. Thank you for your love and encouragement, always.

I would like to dedicate this thesis to my grandmother, Valerie, who died in the year prior to its submission. Her selflessness, optimism, and compassion for people have been and will remain a constant inspiration to me. She was always my greatest cheerleader and I miss her every day. I hope this would have made you proud, Granny Val.

Publications Included in the Text

Hawkins-Elder, H. & Ward, T. (2020). Theory construction in the psychopathology domain: A multiphase approach. *Theory & Psychology, 30*(1), 77–98. Copyright © 2020, SAGE Publications. doi:10.1177/0959354319893026

Hawkins-Elder, H., & Ward, T. (2020). The explanation of eating disorders: A critical analysis. *Behaviour Change, 37*(2), 93-110. Copyright © 2020 The Authors, Published by Cambridge University Press. doi:10.1017/bec.2020.6

Hawkins-Elder, H. & Ward, T. (2021). Describing disorder: The importance and advancement of compositional descriptions in psychopathology. *Theory & Psychology, OnlineFirst,* 1-25. Copyright © 2021, SAGE publications. doi:10.1177/09593543211021157

Hawkins-Elder, H. & Ward, T. (2021). From competition to cooperation: Shifting the 'One Best Model' perspective in psychopathology. *Theory & Psychology, OnlineFirst*, 1-21. Copyright © 2021, SAGE Publications. doi:10.1177/0959354321995900

Permissions

SAGE publications Green Open Access policy (https://us.sagepub.com/en-us/nam/journal-author-archiving-policies-and-re-use) states that published journal article content may be re-used by the author in a dissertation or thesis, including where it may be published in an online repository or database provided that the journal containing the published version of the material is appropriately credited by including the full citation and reference information for the article and if access to the document is provided at no charge. This applies to the three articles published in *Theory & Psychology*.

Cambridge University Press guidelines state: “Permissions requests are waived if: the author of the work wishes to reproduce a single chapter (not exceeding 20 per cent of their work), journal article or shorter extract in a subsequent work (i.e. with a later publication date) of which he or she is to be the author, co-author or editor”. This applies to the article published in *Behavior Change*, which on its own does not constitute greater than 20% of this thesis.

Contents

List of Tables

List of Figures

List of Abbreviations

ED	*Eating Disorders*
AN	*Anorexia Nervosa*
BN	*Bulimia Nervosa*
BED	*Binge Eating Disorder*
OSFED	*Other Specified Feeding or Eating Disorder*
UFED	*Unspecified Feeding or Eating Disorder*
ARFID	*Avoidant Restrictive Food Intake Disorder*
NZ	*New Zealand*
B-P	*Binge-Purge*
BMI	*Body Mass Index*
FBT	*Family-Based Treatment*
IPT	*Interpersonal Theory of Disordered Eating*
IPT-ED	*Interpersonal Therapy for Eating Disorders*
CBT	*Cognitive-Behaviour Therapy*
TCBT	*Transdiagnostic Cognitive-Behavioural Theory*
CBT-E	*Cognitive-Behavioural Therapy for Eating Disorders*
ADT	*Affect Dysregulation Theory*
DBT	*Dialectical Behaviour Therapy*
DBT-ED	*Dialectical Behavioural Therapy for Eating Disorders*
ACT	*Acceptance and Commitment Therapy*
HPA axis	*Hypothalamic-Pituitary-Adrenal axis*
SSRI	*Selective Serotonin Reuptake Inhibitor*
DSM(-IV, -5)	*Diagnostic and Statistical Manual of Mental Disorders (4th or 5th edition, respectively)*
APA	*American Psychiatric Association*
ICD(-11)	*International Classification of Diseases (11th edition)*
WHO	*World Health Organisation*
SNM	*Symptom Network Model*
RDoC	*Research Domain Criteria*
RCT	*Randomised Controlled Trial*
fMRI	*functional Magnetic Resonance Imaging*
CSF	*Cerebrospinal fluid*

OBM	*One Best Model*
OBT	*One Best Treatment*
ATOM	*Abductive Theory of Method*
PDM	*Phenomena Detection Method*
PDM-TC	*Phenomena Detection Method of Theory Construction*

Chapter 1. Introduction

> I am forever engaged in a silent battle in my head over whether or not to lift the fork to my mouth, and when I talk myself into doing so, I taste only shame. I have an eating disorder.
>
> —Jena Morrow, *Hollow: An Unpolished Tale*

The "Eating Disorder Problem"

Eating disorders (EDs) like Anorexia Nervosa (AN) and Bulimia Nervosa (BN) represent a serious global health problem. Each year, an average of 2.2% of women and 0.7% of men worldwide will experience an ED – amounting to over 100 million individuals annually (Galmiche et al., 2019). Across the lifespan, this balloons to 8.4% and 2.2% respectively (Galmiche et al., 2019) – just shy of 500 million people; an alarming sum. The proportion of ED sufferers becomes even higher within at-risk populations such as young women, of whom it is estimated as many as 1 in 7 may be experiencing ED symptoms (Schmidt et al., 2016). This is significantly higher than in previous decades. ED prevalence shows a notable rising trend, particularly within high-income countries (Galmiche et al., 2019; Schmidt et al., 2016). In New Zealand (NZ), specifically, EDs were last estimated to affect around 24,000 people annually – most of these, again, are young women (Browne et al., 2006). Considering both the rising trend in EDs globally as well as the drastic increases in the utilisation of ED services documented by the NZ Ministry of Health (40% from 2012-2016; 2016, 2018), today this figure is likely to be much higher.

The significant and growing prevalence of EDs is extremely concerning given the detrimental effects these conditions have, both on sufferers and on wider society. EDs carry considerable economic burden. The annual healthcare costs of an individual diagnosed with an ED are 48% greater than that of the average person (Striegel Weissman & Rosselli, 2017; Stuhldreher et al., 2012; van Hoeken & Hoek, 2020) and approximately 97% of cases report significant financial hardship (Gatt et al., 2014). Diagnosis of an ED is also associated with lower earning and employment rates than the general population (Samnaliev et al., 2015), reducing the ability of sufferers to fund their treatment and/or to financially recuperate afterwards.

More importantly, however, EDs are responsible for significant physical and mental disability. In 2017, it was estimated that 3.3 million healthy life years were lost as a direct

result of EDs (an age-standardised rate of 43.1 years per 100000 population; van Hoeken & Hoek, 2020). Furthermore, in the 2013 Global Burden of Disease study, AN and BN together ranked as the 12th leading cause of disability adjusted life years (DALYs) in young women living in high-income countries (Hoek, 2016)[1] – a substantial ranking given the low prevalence of EDs relative to other mental and physical health conditions such as diabetes or depression. Furthermore, in contrast to most psychiatric diagnoses, EDs often directly result in severe medical consequences, such as damage to the cardiac and gastrointestinal systems, reduced bone density, and atrophy of the brain – effects that can persist even if the ED resolves (Mehler, 2018; Mehler & Brown, 2015). Eating disorders are thus particularly pernicious conditions, severely impacting both mental and physical wellbeing. Even in those who do eventually recover, typical illness duration is still several years (six on average; Schmidt et al., 2016), during which the individual is significantly impaired.

EDs are also notoriously difficult to treat. Many ED sufferers – particularly those with anorexic presentations – do not want to be treated at all. This typically occurs when the ED is *egosyntonic*: despite its potential for harm, the disordered eating behaviour is viewed by the individual as in harmony with their goals (e.g., weight-loss, emotion regulation) and an acceptable price to pay to achieve them (Giordano, 2005; Gregertsen et al., 2017). In these cases, motivation for treatment is often extremely low, leading to poor engagement and adherence and, thus, poor efficacy (Gregertsen et al., 2017). Treatment may be forced on these individuals for their own safety – for example, they may be nasogastrically force-fed to prevent imminent starvation. Although arguably necessary, such practices can create therapeutic ruptures and further intensify client resistance, making continued treatment even more challenging (Giordano, 2005).

When ED sufferers *are* motivated for treatment, there is still no guarantee that intervention will be successful. Even after a full course of treatment, a significant proportion of ED patients continue to exhibit symptoms, and many others remain chronic or develop a lifelong pattern of remission and relapse (Berkman et al., 2007; Eddy et al., 2017; Hay, 2020; Keel, 2018; Murray, 2020; Murray et al., 2018; Watson & Bulik, 2013). A recent long-term follow-up study of ED sufferers found that over two-thirds of AN sufferers and one-third of BN sufferers remained unwell at 9-year follow-up and roughly a third of clients still had not recovered after 22 years had passed (Eddy et al., 2017). It is

[1] Out of 306 mental and physical disorders (Hoek, 2016).

also not uncommon for ED sufferers to die during treatment, most often by suicide or starvation (Keel, 2018). AN, specifically, has the highest mortality rate of any mental disorder, with a risk of premature death 10 times that of healthy individuals (Arcelus et al., 2011; Hoek, 2016; Keel, 2018).

Given the gravity of the "ED problem", our ability to treat these disorders effectively is critical. Accordingly, the evaluation and improvement of existing interventions and the development of novel therapies have been, and remain, key foci of ED research. Despite this effort and attention, however, treatment for EDs still broadly lacks the degree of efficacy for which both clinicians and clients would hope. Treatment for AN, generally considered the most serious ED diagnosis, typically produces poor results (Murray et al., 2018; Watson & Bulik, 2013). No medications show reliable efficacy for the condition (Bulik et al., 2007; Davis & Attia, 2017; Flament et al., 2012; Hay, 2020; McElroy et al., 2018)[2] and only one psychotherapeutic intervention – Family-Based Therapy (FBT; Le Grange & Rienecke, 2018) – is considered reliably effective. [3] Even then, FBT typically leads to remission in only 25-33% of clients (Murray, 2020). These positive effects have only been demonstrated among youth and are typically poorly maintained, with a large number of patients relapsing or remaining chronic (Berkman et al., 2007; Keel, 2018; Khalsa et al., 2017; Murray et al., 2018; Watson & Bulik, 2013). For adult cases of AN, there is no form of treatment, either pharmacological or psychological, that is proven to be efficacious (Bulik et al., 2007; Hay, 2020; Watson & Bulik, 2013). So, for those individuals whose condition persists past adolescence – a not unlikely scenario given AN's high likelihood of chronicity and relapse – the picture becomes exceedingly bleak.

Treatments for BN and Binge-Eating Disorder (BED) show slightly more promise. Although evidence is still limited, Selective Serotonin Reuptake Inhibitors (SSRIs) and other antidepressant medications show moderate effectiveness for both disorders (Davis & Attia, 2017; Flament et al., 2012; Hay, 2020; McElroy et al., 2018; Shapiro et al., 2007; Svaldi et al., 2018). The d-amphetamine Lisdexamfetamine Dimesylate (LDX) – a

[2] Although there is preliminary evidence that the antipsychotic Olanzepine may have a small effect (Hay, 2020; McElroy et al., 2018).

[3] There is some evidence suggesting that Enhanced Cognitive Behaviour Therapy (CBT-E) may also be efficacious for AN, showing comparable short- and long-term results to FBT in some studies (e.g., Atwood & Friedman, 2020; Grave et al., 2021). However, CBT-E is not yet considered a frontline treatment for anorexic presentations in childhood and adolescence, but rather an alternative treatment option (Hay, 2020).

stimulant medication used in the treatment of Attention-Deficit Hyperactivity Disorder (ADHD) – also appears to reduce binge eating in cases of BED (Davis & Attia, 2017; Flament et al., 2012; Hay, 2020; McElroy et al., 2018; Reas & Grilo, 2015; Shapiro et al., 2007; Svaldi et al., 2018). In terms of psychotherapies, there is currently evidence for the efficacy of Enhanced Cognitive-Behaviour Therapy (CBT-E; Fairburn, 2008; Wilson, 2018; Shapiro, 2007), Interpersonal Therapy for EDs (IPT-ED; Burke et al., 2018), and Dialectical Behaviour Therapy for EDs (DBT-E; Brown et al., 2020; Chen et al., 2018; Peat et al., 2017; Svaldi et al., 2018) in treating both conditions. However, even with these treatment options, evidence suggests that a substantial proportion of individuals (approximately 30-50%) remain symptomatic following treatment or relapse in the future (Berkman et al., 2007; Keel & Brown, 2010; Linardon & Wade, 2018).

Reasons for the "Eating Disorder Problem"

So why are current ED treatments failing our clients? There are many possible reasons, including prescription errors (e.g., inappropriate treatment allocation, insufficient duration or intensity of intervention), delivery problems (e.g., poor treatment fidelity, insufficient clinician training), client factors (e.g., low motivation, poor adherence, coexisting physical or mental health conditions, inadequate support network), therapeutic obstacles (e.g., poor therapeutic alliance, transference/countertransference issues), institutional restrictions (e.g., lack of funding/staff/resources, limited sessions per client), and the broader sociocultural context in which treatment is situated (e.g., persistence of cultural values that promote disordered eating, such as the Western "thin-ideal" or the objectification of women's bodies; Ata et al., 2014; Fredrickson & Roberts, 1997; Moradi & Huang, 2008; Schaefer et al., 2018). For the most part, each of these possibilities has been, and continues to be, investigated by empirical researchers.

A possible contributor yet to be investigated is the quality of the theoretical explanations underpinning current ED treatments. Therapeutic interventions are typically constructed based on *explanations*, which describe how the disorder began (i.e., *aetiology*) and/or why it has persisted over time (i.e., *maintenance*). In much of medicine, this can be achieved through identification of a disease entity or process, such as infection with a microorganism, growth of a tumour, or the under or over-production of a particular hormone, that is believed to have elicited and be maintaining the patient's signs and

symptoms.[4] However, for mental disorders and some more nebulous medical problems, such as chronic fatigue or fibromyalgia, explaining the onset and maintenance of symptoms is more complicated, as we have thus far been unable to identify a central disease process behind the symptoms of these disorders. Explanations for these conditions therefore typically involve a complex of biological (e.g., brain chemistry, physiological systems), psychological (e.g., belief systems and schemas, personality traits, cognitive and emotional styles), and social (e.g., family dynamics, socio-cultural influences) factors or processes that come together to prompt and maintain the client's problems.

Explanations are important for many reasons, such as facilitating accurate classification, stimulating novel research, and guiding clinical assessment. However, arguably their greatest value lies in their ability to inform the development of intervention strategies. In theory, if we know what factors or processes are causing a specific problem, we can construct an intervention to modify or eliminate them and thereby reduce or resolve the client's undesirable symptoms. However, on the other hand, if our explanations are flawed, then it is likely that the interventions we construct based on them will be as well. Without accurate understanding of what factors and processes are involved, or exactly how they are involved, interventions are likely to be misguided in their aims, targeting either the wrong factors and processes altogether, or the right ones in the wrong ways. Interventions therefore become less efficacious than intended or, in some cases, totally ineffective.

It is this possibility that I examine in this thesis – namely, whether current ED explanations provide a suitable foundation for interventions, or whether they are crippled by fundamental flaws that misguide treatment. However, before I begin to address this question, I must first define what "disordered eating" and "eating disorder" encompass in the context of this thesis.

Defining "Disordered Eating"

At present, both in clinical practice and in research, the definition of "disordered eating" is dominated by the *Diagnostic and Statistical Manual of Mental Disorders 5th*

[4] It should be noted that aetiological explanations for the development of cancers or infections necessarily include a broad spectrum of factors – e.g., environmental stressors, genes, lifestyle, diet etc. It is also acknowledged that such factors may exacerbate or improve the patient's symptoms via their effects on the central disease process. Nevertheless, unlike psychopathological disorders, the symptoms of these medical conditions can be directly linked to a single underlying disease entity or process, such as a tumour or microbial infection.

Edition (DSM-5; American Psychiatric Association (APA), 2013)[5], which outlines a number of disordered eating classifications within the section *Feeding and Eating Disorders* (pp. 329-354). The three defined categories that are representative of the typical conceptualisation of an "eating disorder" are:

1. Anorexia Nervosa (AN), characterised by restriction of energy intake, fear of weight gain, undue influence of weight and shape on self-evaluation, and disturbance in the way body weight and shape is experienced.
2. Bulimia Nervosa (BN), characterised by binge eating and purging, fear of weight gain, and undue influence of weight and shape on self-evaluation.
3. Binge Eating Disorder (BED), characterised by recurrent binge eating that causes distress and impairment.

For similar disordered eating presentations that do not fit into these three categories there are two others. The Other Specified Feeding or Eating Disorder (OSFED) category, which includes identifiable cases of disordered eating that are similar to AN, BN, and BED but do not meet full criteria for diagnosis, such as subclinical presentations or atypical AN, and the Unspecified Feeding or Eating Disorder (UFED) category, which includes presentations of disordered eating that do not fit a recognised ED pattern but are deemed clinically significant (causing distress and/or impairment; APA, 2013).

The DSM-5 also includes three other minor ED diagnoses that fall outside the scope of this thesis:

1. Avoidant and Restrictive Food Intake Disorder (ARFID), characterised by avoidance or disinterest in eating or the restriction of specific foods or food groups often due to their sensory characteristics to a degree where there is significant weight loss or nutritional deficiency.
2. Pica, characterised by the compulsive ingestion of non-nutritive, non-food substances, such as dirt, hair, paper, or soap.
3. Rumination Disorder, characterised by the repeated regurgitation of food (re-chewing, re-swallowing, or spitting out food; APA, 2013).

[5] Or, in Europe, the *International Classification of Diseases* (11th Edition (ICD-11); WHO, 2018).

These disorders are not dealt with by this thesis as they are arguably less related to the other EDs, showing distinct epidemiology, presentation, associated features and course, and seldom co-occurring with them (Delaney et al., 2015).

In contrast to the above EDs, ARFID is typically observed in younger children, particularly those diagnosed with Autism Spectrum Disorder (ASD; Dovey et al., 2020; Dickerson Mayes & Zickgraf, 2019). Indeed, in the previous version of the DSM (i.e., DSM-IV; APA, 2000), ARFID was called Feeding Disorder of Infancy or Early Childhood. The disturbance in eating behaviour observed in ARFID seems to revolve primarily around conditioned aversion to the sensory characteristics of specific foods or food groups – such as their taste, texture, or scent – rather than body image disturbance and/or desire for weight-loss (Zimmerman & Fisher, 2017). Accordingly, most cases of ARFID are successfully treated through behavioural and nutritional interventions that gradually reintroduce restricted foods into the individual's diet and address any nutritional deficiencies resulting from restriction (Zimmerman & Fisher, 2017).

Pica occurs in very different populations to AN, BN, or BED, most frequently observed in pregnant women, small children, and those with developmental disabilities (Fawcett et al., 2016; Stiegler, 2005). In some cases, Pica is caused by a particular nutrient deficiency, such as anaemia during pregnancy (Borgna-Pignatti & Zanella, 2016; Fawcett et al., 2016). In other cases, Pica represents a form of self-injurious behaviour in those with severe developmental disabilities (Williams & McAdam, 2012). Accordingly, Pica is typically treated by nutritional rehabilitation or a combination of differential reinforcement (introducing and solely reinforcing an alternative behaviour) and environmental enrichment, depending on the identified cause (Williams & McAdam, 2012). Hence, the overall clinical picture of ARFID and Pica are significantly different from the three core ED presentations outlined above.

Regarding Rumination Disorder, there is a notable lack of consensus on the diagnosis, theory, and treatment of this condition (Murray et al., 2019), which makes it difficult to locate conceptually within the ED space. However, in many cases the regurgitation of food seems to be an automatic response to food intake that is unintentional and often fought off by the sufferer (Murray et al., 2019). Furthermore, rumination behaviour seems to lack explicit goals related to the modulation of food intake (e.g., weight loss) or concerns about the consequences of eating (e.g., weight gain) commonly observed in the three central ED

diagnoses (AN, BN, BED). Such a presentation puts this condition in a very different light to AN, BN, or BED, which involve disruption to eating behaviour that is (at least initially) arguably more intentional and goal oriented.

Outside of whether an individual meets the DSM criteria, it is difficult to define exactly what does or does not constitute disordered eating. As humans, all of us must eat to survive. However, there is great variation in how we do this and, arguably, no single "right" way to do it. Eating behaviour varies hugely between individuals, families, societies, and cultures, and may fluctuate within individuals in response to changes in season, mood, location, food availability, hormone levels, cultural norms, values, goals, or social context. There is therefore a huge range of eating behaviour that might be considered "normal", making "abnormality" and "pathology" in this area difficult to define conclusively.

For example, if we consider purely statistical abnormality, the behaviour of deliberate food restriction would not qualify. A huge proportion of individuals today engage in some form of "dieting" (e.g., paleo, keto, low-carb) or food restriction (e.g., vegetarianism, veganism), in which they purposely restrict the amount or type of food they consume, and many cultures and religions engage in fasting (e.g., Ramadan, Lent) or abstinence from specific food groups (e.g., Halal). Most of us would not consider all these individuals "disordered". Similarly, many individuals might engage in binge or purge-style behaviours on occasion (e.g., overeating on holidays or special occasions, vomiting due to food poisoning or overindulgence, taking over-the-counter laxatives), but again we would not consider them mentally unwell.

So, the question remains: when does eating become "disordered"? I would argue that this occurs when an individual's eating patterns are no longer serving them, either physically or mentally. This perspective is influenced by Nielsen and Ward's (2018, 2020b) Embodied-Enactive approach to mental disorder, in which they propose that mental disorders occur when a particular set of behaviours begins to violate an individual's *functional norms*: behaviours of the whole system that serve the continued survival and adaptive functioning of the organism (e.g., seeking food and shelter, forming and maintaining social bonds; Nielsen & Ward, 2020b). These are not based on statistical normativity, nor on the theorised evolutionary process of one system or another, but rather are specific to the individual's biology, psychology, and sociocultural context (Nielsen & Ward, 2020b). In the case of EDs, therefore, an individual's eating behaviour could be classed as "disordered" when it is

impairing their ability to continue functioning adaptively in the context of their own biological, psychological, and sociocultural needs.

This definition helps to differentiate non-problematic patterns of eating, such as normative dieting, from those about which we would be clinically concerned, such as the life-threatening food restriction seen in AN. Regular dieting does not typically impair the individual's ability to function in an adaptive manner. Although dieting behaviour may revolve around eating and weight concerns, the degree is not such that it results in marked distress or interferes with engagement in life. Physical health is also typically minimally affected and may even improve in the case of those who are overweight. In contrast, those engaging in the kind of food restriction that would warrant clinical attention – such as in AN – are typically highly distressed by eating and experience significant anxiety or depression around food, such as at mealtimes. Food, eating, and related concerns about weight and shape have also typically taken on a central role for that person, to the degree that their ability to engage in day-to-day life is impaired (e.g., being distracted from or eschewing usual activities to engage in food-related tasks or disordered eating behaviours). Finally, because of the extreme degree of energy restriction taking place, the person is generally experiencing significant negative physical effects that prevent them from functioning optimally (e.g., fatigue, physical weakness, poor concentration) or endanger their life (e.g., cardiac damage and risk of Sudden Death Syndrome).[6] Similarly, in those engaged in binging or purging behaviour, there is typically significant psychological distress or physical harm taking place, such as clinical levels of anxiety and depression or damage to gastrointestinal, oesophageal, dental, and other structures.

Hence, in general one can conceive of "disordered eating" and "eating disorders" as patterns of eating behaviour that markedly impact the individual's ability to function adaptively within their world, causing psychological or physical distress to the extent where they are unable to live a fulfilling life. For the most part, it is this more encompassing functional definition with which this thesis is concerned. However, as mentioned previously, there is a cluster of ED presentations, including AN, BN, and BED, that seem to be more related to each other and conceptually distinct from other types like ARFID, Rica, and Rumination Disorder. Furthermore, the majority of ED research and theory to date has

[6] *Sudden death syndrome* is a rough umbrella term for a variety of cardiac syndromes that cause sudden cardiac arrest resulting in death (Koplan & Stevenson, 2007).

revolved around this cluster of disorders, providing a larger pool for exploration and analysis. Hence, it is on the explanation of this collection of EDs that this thesis will focus.

Structure of the Thesis

Within this thesis, I aim to explore whether our current explanations for disordered eating possess sufficient theoretical quality to provide robust foundations for clinical intervention and, if not, how they could be created to better serve this purpose. However, before assessing the quality of such explanations, it is critical to outline what an *explanation* is, how it differs from related theoretical concepts, such as *theories* and *models*, what its role is within scientific inquiry, and why it matters to the ED field in particular. It is also important to identify what qualities make for a "good" explanation, so that the criteria against which current explanations should be assessed are clear. Hence, in the next chapter (Chapter Two: *Explanation*), I discuss each of these topics in depth to orient the reader to this broader theoretical context.

In Chapter Three (*Examining Current Explanations for Eating Disorders*) I conduct a critical evaluation of current ED explanations using the criteria outlined at the end of Chapter Two. The explanations chosen for analysis are those underlying the most prominent psychotherapeutic interventions for EDs: Family-Based Therapy (FBT; Le Grange & Rienecke, 2018), Enhanced Cognitive-Behaviour Therapy (CBT-E; Fairburn, 2008), Interpersonal Therapy for Eating Disorders (IPT-ED; Burke et al., 2018), and Dialectical Behaviour Therapy for Eating Disorders (DBT-E; Chen et al., 2018). These explanations are, respectively, Minuchin et al.'s (1975, 1978) Psychosomatic Family Theory, Fairburn et al.'s (2003) Transdiagnostic Cognitive-Behavioural Theory of Eating Disorders, Rieger et al.'s (2010) Interpersonal Theory of Disordered Eating, and Linehan & Chen's (2005) application of Affect Dysregulation Theory to binge-purge ED presentations. In the process of evaluating these explanations, I identify three broader issues with their construction that I argue contribute to the specific problems of each: 1) the poor selection and description of explanatory targets, 2) the endorsement of the One Best Model perspective, and 3) the lack of a guiding methodological approach for explanations of psychopathological problems. The remainder of the thesis is then devoted to addressing each of these meta-theoretical issues.

Chapter Four (*Selecting and Describing Explanatory Targets Effectively in Psychopathology*) addresses the first issue identified: the poor selection and description of explanatory targets. This problem can be divided more specifically into 1) a focus on flawed

DSM syndromes, and 2) the failure to describe the composition of psychopathological phenomena sufficiently. This chapter deals with each issue in turn. First, I discuss the problems inherent in DSM syndromes that preclude them from serving as appropriate targets for explanation and propose a way forward in this area: shifting our investigative focus from diagnostic syndromes to *clinical phenomena*. Second, I discuss *compositional explanation* and how it contributes to understanding psychopathological phenomena in theory, research, and practice. I argue that the psychopathology field currently lacks sufficient compositional explanations of mental disorder by evaluating several types of descriptive account within the psychopathology domain against the criteria that would be expected of a "good" compositional explanation. In response to this critique, I outline what effective compositional explanations of psychopathological problems might look like and be constructed, using the ED symptom of "binge eating" as an example.

Chapter Five (*Engaging with a Pluralistic Approach to Psychopathological Explanation and Treatment*) is devoted to addressing the fact that isolated ED models – and models of mental disorder more generally – are often presented as *complete* explanations. This problem reflects a more systemic issue within the psychopathology domain: the pervasive belief that mental disorders can be explained or understood using a single model. I term this belief the One Best Model (OBM) perspective. In this chapter, I outline this perspective, discuss how it has come to dominate the explanation of mental disorder, and highlight the ways in which it is ultimately unhealthy for psychopathology research and practice. In response to this problem, I propose a significant shift from *explanatory unification* to *model pluralism* within psychopathology explanation and discuss how such a shift might be implemented when building explanations, translating explanations into interventions, and applying interventions in practice.

Chapter Six (*Developing a Methodological Framework for Constructing Explanations in Psychopathology*) addresses the lack of a clear, guiding methodological process for building explanations of EDs and mental disorder more broadly. Within this chapter, I outline the role of method in science and argue that explanation is a scientific task that should be subject to the same methodological rigor expected of empirical research. In particular, I highlight the current lack of methods and methodologies that can inform the generation of psychopathology explanations and the impact this has on existing explanations of mental disorder. To remedy this, I present the Phenomena Detection Method of Theory Construction (PDM-TC) – a preliminary methodological framework for constructing explanations of

mental disorder – in order to demonstrate a systematic approach to theory construction and provide a starting point for the continued development of explanatory methods and methodologies in psychopathology.

Finally, in Chapter Seven (*Conclusions and Future Directions*) I reflect on and summarise the content of this thesis and discuss the implications it has for research and practice surrounding EDs. I also discuss some of the obstacles that stand in the way of these ideas being applied and how such barriers might be overcome. Finally, I provide some ideas about how future research could build on this thesis, and the steps I see as necessary to achieve greater efficacy in ED explanation and treatment.

Chapter 2. Explanation

There is some disagreement in the scientific literature over what exactly constitutes an explanation and how explanations relate to scientific models and theories. These terms also carry lay meanings that may significantly differ from their scientific definitions. It is therefore important to outline how these terms are being used in this thesis, as different readers will likely have different understandings of the concepts. Also important to the reading of this thesis is the fact that explanations can exist at a variety of conceptual levels (e.g., biological, psychological, sociocultural) and be framed in different ways (e.g., in terms of mechanisms, intentions, or patterns). The first half of this chapter is therefore devoted to unpacking these various concepts for the reader.

I begin by discussing the concept of an "explanation" in science, and how it is believed to relate to "theories" and "models". I also give a brief account of the two core components of an explanation: the *explanandum* (or "explanatory target") and the *explanans* (or "explanatory account"). Following this, I provide a brief outline of explanatory levels (i.e., levels of analysis) and several different types of explanation that have been distinguished by the literature and are relevant to explanatory discussions surrounding mental disorder: nomological explanation, functional explanation, agential explanation, mechanistic explanation, and aetiological and compositional explanation.

The second half of this chapter discusses the relevance of explanation to psychopathological problems, such as eating disorders (EDs). Specifically, I describe how explanation (as a general task) and explanations (as specific accounts) can inform a variety of research and practice tasks, such as clinical research, diagnosis/classification, assessment, formulation, and intervention. Finally, I discuss the qualities that a "good" explanation should possess and outline a set of criteria against which the selected explanations of EDs will be evaluated in Chapter Three.

Explanations, Theories, and Models

Explanations are constructed in a variety of scientific disciplines to account for the occurrence of specific states and outcomes, such as diseases, mental disorders, chemical reactions, species evolution and survival and so on. It is not easy to define what does and does not constitute an explanation. Indeed, there is much dispute among philosophers of science as to what exactly is encompassed by the term. At the most basic level, explanations

can be defined as conceptual structures formulated to explain why certain phenomena exist and persist. Trout (2016) provides a simple and helpful definition of explanation as "the description of underlying causal factors that bring about an effect" (p. 18). More specifically, an explanation is generally thought of as an account that provides an understanding of a phenomenon's causes, composition, context, or consequences (Faye, 2014; Ruphy, 2016). However, for the purposes of this thesis, an *explanation* is considered *the most complete explanatory account possible* of the phenomenon in question, consisting of the entire bank of explanatory knowledge related to it (Craver & Kaplan, 2020). Under this expanded definition, models and theories may play a critical role in generating explanations but are seldom (if ever) equivalent to them. Instead, models and theories can be considered *tools* that can be put to the task of explanation (Mantzavinos, 2016; Savulescu et al., 2020).

A *theory* is an integrated system of concepts and ideas that can be used to construct an explanation, although it does not provide an explanation in and of itself. Theories are generally less specific in their content than models, usually detailing more abstract phenomena (e.g., "human behaviour" in general versus specific types or instances of behaviour). Any one theory can be used to explain different subsets of phenomena (Bailer-Jones, 2003). For example, Operant Conditioning Theory (Staddon & Cerutti, 2003) has been applied to the occurrence and persistence of a variety of different psychopathological problems, such as child conduct problems (e.g., *the coercion cycle*; Dishion & Patterson, 2015; Granic & Patterson, 2006), self-harm (e.g., Chapman et al., 2006; Gratz, 2003), eating disorders (e.g., Haedt-Matt & Keel, 2011; Pearson et al., 2015), and childhood anxiety (e.g., *parental overprotection*; Hudson & Rapee, 2004; McLeod et al., 2011). In contrast, *explanations* are more specific to their phenomenon of interest (e.g., heart disease, depressed mood, sexual offending) and should provide a complete account of the phenomenon rather than a partial one. Operant Conditioning Theory (Staddon & Cerutti, 2003) is but a partial explanation for human behaviour as it is unable to account for all events – for example, it struggles to provide an effective explanation for the emergence of novel behaviours, about which the individual has no operant information. Hence, although theories may have explanatory properties, may be used to inform the construction of models, and may contribute to the generation of explanations, they are not themselves explanations in toto.

A *model* is a representation of a more concrete empirical phenomenon – an identified factor, system, or process, such as "clinical depression" or "binge eating". Typically, this is an idealised and simplified representation (Bailer-Jones, 2003; Haig, 2014). For example, a

model may depict the phenomenon's constitution or aetiology at only some *levels of analysis* (see section below: *Levels of Explanation*) or may exclusively detail the contribution of a single factor or process involved in comprising or causing the phenomenon of interest (e.g., the role of distorted cognitions in depression). Hence, models typically represent only *some* aspects of a phenomenon, rather than its whole. In some cases, models may represent localised applications of theories (Bailer-Jones, 2003) – for example, as mentioned earlier, the coercion cycle model represents an application of operant conditioning theory (Staddon & Cerutti, 2003) to the specific phenomenon of child conduct problems (Dishion & Patterson, 2015; Granic & Patterson, 2006).

Models can be thought of as *tools* for explanation, as explanations generally require the contribution of explanatory models. However, not all models have explanatory properties. Take, for example, *phenomenal models*, which are descriptive in nature and aim to depict patterns, like co-occurring events or symptom clusters (Craver & Kaplan, 2020; Ward & Fischer, 2019). Even when models *are* explanatory, they seldom, if ever, constitute complete explanations on their own. A model may equate to an explanation if it represents the entire bank of explanatory knowledge about the phenomenon in question – when the model alone can explain the phenomenon adequately without being overly complicated or sacrificing critical detail. However, this is rarely, if ever, the case. More often, models represent *partial* explanations (Bailer-Jones, 2003). When the phenomenon of interest is more complex, such as psychopathological problems, multiple models will likely be needed to fully explain it (Kendler et al., 2020).

Explanatory Components

Any explanatory account, whether partial or whole, will necessarily involve an *explanandum* and an *explanans* (Hempel & Oppenheim, 1948). The *explanandum* (also known as the "explanatory target") refers to the phenomenon to be explained, whereas the *explanans* refers to the set of facts about the world that are used to explain it (Hempel & Oppenheim, 1948). A simple example might be the question "why is there smoke?", for which a basic explanation might be "because there is a fire". In this case, the fire (explanans) is providing an explanation for the existence of the smoke (explanandum). However, typically, multiple facts will comprise the explanans. These facts must necessarily be related in a way that makes sense and that makes the existence of the explanandum more understandable. For example, if the explanandum were "why did the first child hit the

second?", a multi-statement explanans could be (1) the second child launched a verbal insult, (2) the first child became angry, and (3) the first child wanted to punish the second child. These statements relate logically to one another and allow us to understand why the explanandum occurred.

However, these days in science, it is seldom the case that an explanandum can be so easily accounted for, as most of the phenomena we are interested in are routinely more complex, such as mental disorder. We require explanations that can account for multiple factors and processes acting and interacting at a variety of levels of analysis. The explanandum and explanans therefore necessarily become much more complicated, typically involving multiple phenomena. Nevertheless, it is critical to distinguish between the target of an explanation and the factors or processes being invoked to explain it, so that it is clear as to *what* is being explained and *how*.

Levels of Explanation

There are a variety of *levels of analysis* at which explanation may occur. *Levels of analysis* (also sometimes called "levels of explanation") refer to the hierarchical division of material composition into distinct conceptual strata. These strata are not broadly agreed upon, and different divisions may be made for different purposes. However, each is considered to be comprised of "classes of structures and processes of relevant similarity" (Brooks & Eronen, 2018, p. 34). For example, one may make a distinction between phenomena that exist at the "biological" level (e.g., neurotransmitters, cells, nerves) and phenomena that exist at the "psychological" level (e.g., beliefs, emotions, subjective experience). However, one could also further divide each of those levels depending on the purpose for which they are being used – for example, dividing "biological" into "molecular", "cellular", and "neural", and "psychological" into "cognitive", "emotional", and "phenomenological/experiential".

These levels are typically considered to be hierarchically organised, with those concerning smaller or more local phenomena, such as the molecular or biological, to those that are larger or operate more broadly, such as the environmental or cultural (Brooks & Eronen, 2018). However, there is also notable conceptual overlap between these levels such that their boundaries are necessarily blurred. For example, what exactly separates the biological from the psychological or the psychological from the social? Should thoughts be considered biological patterns of neural activation or abstract cognitive constructs? Can the psychological construct of emotion ever be separated from its social or biological function?

Hence, some philosophers of science have argued that levels of analysis are better conceptualised as *scales*, as this allows for a more dimensional understanding and for consideration of both temporal and spatial relations (Potochnik & McGill, 2012). However, such a debate is beyond the scope of this thesis. I therefore refer to scales and levels of analysis interchangeably throughout, as both denote the same broad idea.

Types of Explanation

Explanations may be framed in different ways depending on their purpose. For example, an explanation in psychopathology might seek to identify a person's reasons for acting as they did, what general pattern of thought, emotion, and behaviour they are demonstrating, what factors and processes underlie or maintain their difficulties, or what causal pathway led to the development of their problems. Each of these questions is likely to require a different form of explanation. In psychopathology, several types of explanation are particularly relevant: nomological explanation, functional explanation, agential explanation, mechanistic explanation, and aetiological and compositional explanation. Many of these types of explanation are not mutually exclusive. For example, an explanation may be both mechanistic *and* appeal to agential factors. Nevertheless, they all represent a unique form of explanation. I will now provide a brief account of each.

Nomological Explanation

Nomological explanation – sometimes also known as the "covering law model" of explanation – refers to explanation that is based on rules or "patterns" and describe outcomes in terms of certain fundamental laws (Wilsch, 2016). Such laws are common in sciences like physics, such as the laws of thermodynamics or motion. These laws can be applied to specific phenomena to explain their occurrence. For example, Newton's law of gravity, which states that all objects are acted on by gravitational force (see Verlinde, 2011), could be used to explain why a glass falls to the ground when knocked off a table. Behaviourist theory (Skinner, 1976) represents an attempt to apply this kind of explanation within psychology, stating that all behaviours are learned via the same *laws of conditioning*, such as operant conditioning and classical conditioning.

One could argue that psychiatric diagnosis using the DSM forms a kind of nomological explanation – or, at least, is often used as such. The diagnostic categories of the DSM represent clusters of symptoms and signs that "hang together" statistically – they are patterns of dysfunctional thought, feeling, or behaviour. There is sometimes a tendency to use

these categories as explanations for individuals' disordered behaviour – for example, "Cindy starves herself because she has an eating disorder" or "depression is a condition that causes low mood, anhedonia, and changes to sleep and appetite". This explanatory use of diagnostic categories stems from the medical discipline. In medicine, many diagnoses do indeed have explanatory properties, as they signal the presence of a known common disease process, such as the growth of a tumour, infection with a pathogen, or a breakdown in the functioning of a physiological system. However, in psychopathology, we have been unable to identify the precise pathogenesis of disorders, and so our diagnoses specify no such disease entity. Instead they are mere descriptions of sets of signs and symptoms that are statistically associated. Their use as explanatory is therefore tautological – "person X presents with certain signs and symptoms because they present with those signs and symptoms". Hence, although they may convey some explanatory information – such as suggesting possible causal processes based on research and other clinical cases (Muang, 2016) – they do relatively little explanatory work themselves.

Functional Explanation

Functional explanation (also sometimes called *teleological explanation*) accounts for the existence of a phenomenon in terms of the function it serves, rather than the causes by which it arises (Lombrozo & Carey, 2006). For example, a functionalist would answer the question "why do people have hearts?" with the response "to pump the blood around their bodies". In other words, functional explanations assert that phenomena exist because of the function they serve. A psychological example could be that the function of the attachment system is to promote survival (Ainsworth, 1989). Functional explanations may therefore be clinically useful in psychopathology, as they provide the purposes for which a particular problem behaviour might be engaged in. For example, self-harm could be explained as functioning to regulate an individual's emotions. Another example would be compulsions in Obsessive-Compulsive Disorder (OCD), which could be explained as serving the function of reducing the anxiety associated with obsessions. Such an explanation allows us to understand the function of the client's problem behaviour, even if they themselves are unaware of it (as in "automatic" or habitual behaviours). We may then be able to help them replace their behaviour with more adaptive alternatives that serve a similar function.

Agential Explanation

Agential explanations (also sometimes referred to as *intentional* or *mental-state explanations*) are particular to human behaviour and describe behavioural events as a function of an individual's agency (free-will, desire, beliefs, attitudes; Hacker, 2009). In other words, it explains behaviour in terms of the choices or intentions behind it. For example, the act of someone making a cup of tea could be agentially explained in a variety of ways – for example, to comfort themselves/someone else, to be social, to quench thirst, to improve health, to keep warm – as there may be a variety of reasons why they chose to engage in that behaviour. Agential explanations of psychopathology may ask why individuals choose to engage in certain behaviours or respond to their thoughts or feelings in the way that they do, typically with reference to beliefs, attitudes, or goals. For example, an agential explanation of self-starving behaviour might cite goals such as losing weight, communicating distress, numbing emotion, or delaying pubertal development, and the individual's beliefs and attitudes related to these. Like functional explanations, agential explanations have notable value for clinical practice in that they allow clinicians to understand the reasons why their client is engaging in particular behaviours, which provides the data needed to generate alternatives. This information may also be especially useful in egosyntonic conditions like Anorexia Nervosa (AN), in which the individual's values and goals are often aligned with the disordered behaviour.

Mechanistic Explanation

A *mechanistic explanation* "describes, represents, or provides information about the *mechanism* producing, underlying, or maintaining the phenomenon to be explained" (Povich, forthcoming, p. 1). There is some debate in science over what exactly constitutes a mechanism (see Andersen, 2014; Illari & Williamson, 2012). However, a mechanism is broadly considered to be an organised collection of entities and activities that produces, underlies, or maintains a phenomenon (Povich, forthcoming, p. 1). In other words, a mechanism comprises a set of parts and the relations between them that combine to produce a particular outcome, which may be different under different circumstances. An example of a mechanistic explanation from biology is *the Krebs cycle* (also known as the "citric acid cycle"). The Krebs cycle explains how aerobic organisms release stored energy via a process of chemical reactions within the cell. As this process is organised (non-random) and involves

the contribution of multiple entities (molecules) and their interrelations (chemical reactions), it constitutes a mechanism.

From a mechanistic perspective, explanations of disease in medicine attempt to discover the impact of an aetiological agent, such as exposure to a toxin or bacterial agent, on the production of a suite of symptoms through the operation of underlying biological causal mechanisms (*pathogenesis*; Kendler et al., 2011). In psychopathology however, there is little agreement on the nature of the causal mechanisms generating mental disorders (Murphy, 2006). It also presents a particular challenge in that some mechanisms may not be directly observable and therefore largely inferred, such as cognitive mechanisms (Bechtel, 2008). However, mechanistic explanations are of great use in psychopathology, as they are likely to best represent the complex nature of mental disorder, which is caused and constituted by a variety of different factors interacting in a multitude of ways. Despite this, no mechanistic explanations of psychopathological phenomena currently exist, although there have been preliminary attempts (for example, Ward & Clack's (2020) mechanistic "sketch" of anhedonia in clinical depression; see Chapter Five).

Aetiological versus Compositional Explanation

Philosophers of science make a distinction between *aetiological* and *compositional* explanations of scientific phenomena (e.g., Craver, 2007; Kaiser & Krickel, 2016). An *aetiological explanation* depicts the causes that result in the emergence of a subsequent phenomenon, such as an effect, event, or structure – for example, heating water (a cause) until it boils (an effect) or stimulating the dendrites of a neuron (a cause) to the level required for an electrical impulse to be transmitted down an axon (an effect). Causes are independent of effects and can take place without them necessarily occurring (Illari & Russo, 2014). A psychopathological example of an aetiological explanation is the proposal that traumatic experience (a cause) can result in the experience of intrusive memories, elevated arousal, and avoidance behaviour (a *post-traumatic stress* syndrome). Another would be Beck's (2008) cognitive model of depression, which implicates the combination of "cognitive vulnerabilities" developed as a result of adverse early developmental experiences, such as negative self-schemas or information processing biases, and precipitating stressful events in the onset of a major depressive episode.

In contrast, a *compositional explanation* of a phenomenon describes the factors and processes that comprise it. For example, the symptom of low mood is likely to be composed

of processes at the phenomenological, cognitive, neurobiological, and physiological levels (Ward & Clack, 2019). Thus, the goal of a compositional explanation is to describe the underlying structures and interactions that constitute a phenomenon, which are viewed as part of it rather than causing it (Craver, 2007; Gillett, 2021; Kaiser, 2015). However, compositional explanations may include causal relationships. In psychopathology, for example, a compositional account of a disorder state may include the causal processes that *maintain* it, such as mutual reinforcement between symptoms (e.g., insomnia and low mood; Konjarski et al., 2018) or behavioural reinforcement cycles (e.g., the coercion cycle; Dishion & Patterson, 2015).

An explanation cannot be both aetiological and compositional at the same time. However, the same explanans may represent an aetiological *or* a compositional explanation depending on the nature of the relevant explanandum. For example, if our explanandum were the symptom “low mood”, then an aetiological explanation might include insomnia as a causal factor, given that literature shows that insomnia shares a bidirectional relationship with depressed mood (Konjarski et al., 2018). However, if our explanandum were the syndrome “clinical depression”, then the causal relationship between low mood and insomnia would instead form part of a compositional explanation, as this relationship partially constitutes the experience of depression. Hence, factors or relations that may be aetiological in one explanatory context could be compositional in another.

Explanation in this Thesis

To summarise, within this thesis, an *explanation* is defined as being the complete explanatory account of a phenomenon; that which encompasses all relevant explanatory knowledge. Given the nature of psychopathological problems as multifactorial and multidimensional, an adequate explanation under the definition of this thesis would need to refer to multiple explanatory levels (e.g., molecular, neural, physiological, cognitive, emotional, interpersonal, cultural). In terms of the nature of the explanatory account itself, it is likely that multiple types of explanatory reasoning (e.g., agential, mechanistic, compositional, aetiological) will be needed to fully understand psychopathological problems, and each may hold unique value for aspects of clinical practice. This will be further explored as this thesis progresses. For now, I move on to discussing the practical value of explanation – both as a process and as specific accounts – for research and practice in psychopathology.

Why Does Explanation Matter to Psychopathology?

As discussed above, explanations seek to identify the causes, composition, context, or consequences of a specific phenomenon. They are therefore valuable sources of information that allow us to understand phenomena of interest. In psychopathology, explanations allow us to comprehend the psychological problems experienced by clients in depth, including disordered eating patterns, such as self-starvation, binging, and/or purging. Explanations can tell us how these phenomena are constituted – what factors, processes, or experiences comprise them at various levels (compositional explanation) – and how they may have begun and been sustained over time – what mechanisms may be responsible for the emergence of the problem (aetiological explanation). This kind of detailed understanding is useful for a variety of scientific and practice tasks including clinical research, diagnosis/classification, assessment, formulation, and intervention.

Clinical Research

Clinical research can be considered as any empirical research relevant to the clinical context of psychopathology. Examples include studies investigating prevalence and onset, comorbidity, disorder features and correlates, risk factors and aetiological mechanisms, maintenance cycles or processes, prognosis, course and outcomes, treatment efficacy, treatment mediators and moderators, and therapist or therapeutic process factors. Such research plays a role in the generation of explanations, as this process typically draws on available empirical data to identify phenomena and mechanisms that might be relevant for understanding the problem in question. However, explanation also contributes meaningfully to the production of empirical research. Explanations conceptualise and contextualise empirical findings, determining what various findings mean in the context of a particular problem, knitting them together in meaningful ways, and situating them within the broader psychopathological context. Explanations therefore help empirical research to proceed fruitfully, as the theoretical conceptualisation of existing findings may provide new hypotheses to test, identify research gaps, or indicate promising avenues for future research. Furthermore, in cases where empirical research does not yet exist for a particular phenomenon, the generation of preliminary explanations (through inference from models of related phenomena or the application of broader psychological theories) can help to make decisions about where investigation should begin. In other words, explanation helps to organise and direct empirical research, which is then used to revise the original explanation

or generate new ones. Explanation and empirical research share a symbiotic relationship, each informing the other in valuable ways.

Diagnosis/Classification

Classification refers to the scientific categorisation of phenomena into groups or classes according to shared characteristics (features, properties, or functions; Mattu & Sullivan, 2020). Executing this task requires a degree of knowledge of the phenomenon one seeks to classify: to determine how an object, event, experience, or organism may best be categorised within a proposed taxonomy, one must be sufficiently aware of the properties or attributes it possesses relevant to that system of classification. Without sufficient information, classification schemes are more likely to categorise phenomena inappropriately based on a faulty understanding of their nature.

Take, as an extended example, the current classification of EDs under the DSM (APA, 2013). When this classification scheme was first proposed, scientific knowledge of EDs was relatively sparse. Accordingly, organisation was based on incomplete and superficial understandings of these problems. Researchers have since recognised that there is substantial overlap across these categories, such that a case of Anorexia Nervosa (AN) and a case of Bulimia Nervosa (BN) might be distinguished by only one or two arbitrary features such as Body Mass Index (BMI), which is widely considered to be a flawed method of measurement (Burkhauser & Cawley, 2008; Prentice & Jebb, 2001). ED cases also often migrate between different diagnostic categories over time – it is not uncommon for an individual diagnosed with AN to be reclassified under the BN diagnosis several years later (Eddy et al., 2008; Fichter & Quadflieg, 2007; Tozzi et al., 2005) – suggesting poor reliability of these diagnoses. Furthermore, the OSFED/UFED diagnostic categories – intended as "catch-alls" for atypical cases – are far more common than would be expected, accounting for around 50% of ED cases (Agras et al., 2009; Forbush et al., 2018; Keel et al., 2011; Machado et al., 2013). Hence, it has become clear that the diagnostic categories imposed on EDs by the DSM are an inaccurate representation of these conditions. This can be viewed, at least in part, as a reflection of the current poor understanding of these problems. If the psychopathology field possessed better understandings of these conditions – detailed explanations of their constitution and causes – we might have a better chance of grouping them in meaningful and constructive ways that more accurately reflect how they exist in the world.

At present, psychiatric classification is purely descriptive and relies on *syndromes* – sets of symptoms and signs that have been observed to consistently cooccur – such as AN or Major Depressive Disorder (MDD). Hence, psychological disorders are currently classified solely according to their external features (symptoms and signs), rather than any identified causal mechanism(s). This contrasts with the classification of many medical conditions, such as cancers, diabetes, or bacterial and viral infections, the detailed study of which has allowed for classification based around causes. For instance, viral and bacterial tonsillitis are recognised as different diagnoses based on their distinct causes of viral versus bacterial infection despite their often-identical symptom presentation (swollen tonsils, sore throat, fever, etc.). These classifications have *explanatory power* – they provide information about the cause of the client's symptoms rather than merely describing them – and therefore are much more clinically useful (e.g., able to identify targets for treatment). Such classifications are impossible without a comprehensive understanding of the factors and processes responsible for causing the problem – in other words, an explanation. Without this, researchers and clinicians cannot hope to provide diagnoses that are conceptually valid and meaningfully account for client problems. The explanation of psychopathological conditions like EDs is therefore critical for the development of classification schemes that can usefully inform diagnosis and treatment.

Assessment

During clinical assessment, explanations for disorders provide background information that can suggest to a clinician what areas they may need to investigate, both initially and as the assessment proceeds and new information arises. For example, a clinician who receives a referral for a client showing signs of depressed mood would know, based on existing explanatory information, to inquire about factors such as depressive cognitions (e.g., the *negative cognitive triad*; e.g., Beck et al., 1987), recent stressful events (e.g., job loss, relationship breakup, bereavement), and family history of depression or other psychiatric illness, all of which have been implicated in theoretical explanations for depressed mood (Beck & Bredemeier, 2018). During assessment, however, the client may additionally describe engaging in binging and purging behaviour, in response to which the clinician might adjust their line of questioning to investigate factors implicated in explanations for this phenomenon, such as body dissatisfaction, impulsive and perfectionistic personality traits, high levels of anxiety and depressed mood, and emotional lability or poor emotion regulation (Berg et al., 2013; Jacobi et al., 2018; Fairburn et al., 2003). Explanation therefore plays a

crucial role in guiding clinical inquiry, indicating how the clinician might direct their investigation to obtain support for or evidence against particular clinical hypotheses and gather relevant information for *formulation*.

Formulation

Formulation is, by definition, an explanatory task: it involves constructing a theoretical conceptualisation of the client's problems that provides an understanding of why they have developed and persisted (Johnstone & Dallos, 2014). Explanation is therefore integral to successfully formulating a client. To do so requires both good theoretical grounding in explanatory meta-theory and methodology, as well as skill at using explanatory reasoning in practice. Without this, formulation is more likely to be flawed; inaccurate, illogical, or incomplete.

However, explanation is also relevant to the content of formulations. Formulation typically draws on broader explanations of psychopathological phenomena and applies relevant models to the client's unique situation, considering their specific biology, psychology, and sociocultural context. For example, to explain why a child presents with a pattern of disobedient, defiant, and aggressive behaviour, a clinician could apply *the coercion cycle* model (Granic & Patterson, 2006) to the specific patterns of thought and behaviour demonstrated by that child and their family. However, this model alone is unlikely to fully explain the problem. A clinician will almost always need to draw on multiple explanatory theories and models to inform their formulation. As such, formulation is directly affected by the quality of explanations. If explanations of mental disorder (and their constituent theories and models) are flawed – lacking accuracy, depth, or validity – then the formulations they are used to inform will be also.

Intervention

Intervention is arguably the most important practical use clinicians have for explanations. Interventions are typically sought to either prevent a problem occurring in the first place or treat it once it has emerged. In both cases, they are usually designed based on a hypothesis or proposed understanding of what causes or maintains a specific problem (an explanation), as in theory this tells us where to intervene to prevent or eliminate the resulting problems. For example, after much research linking the development of cancer with cigarette smoking (e.g., Gandini et al., 2007), explanations were developed postulating that the harmful chemicals in tobacco smoke damage a person's DNA in ways that, over time,

promote the development of cancer (US Department of Health and Human Services, 2010). In order to target these mechanisms, public health initiatives for the prevention of lung cancer began to focus on stopping smoking or encouraging a transition to (arguably) less harmful vaporisers or nicotine replacements, such as chewing gum and patches (US Department of Health and Human Services, 2010).

Similarly, the treatment of "Strep throat" with antibiotics is based on the explanation that the symptoms result from infection of the throat with Streptococci bacteria. Antibiotics are introduced, killing the offending bacteria, and the infection and associated symptoms are subsequently resolved (Mustafa & Ghaffari, 2020). An example more specific to psychopathology is the treatment of depression with SSRIs, which augment the amount of serotonin present in the synapse, based on the explanation that persistent depressed mood results, at least in part, from reduced serotonin activity in the brain ("the serotonin hypothesis"; Albert et al., 2012; Cowen & Browning, 2015). SSRIs decrease the amount of serotonin being recycled from the synapse back into the axon terminal, leaving a greater amount of serotonin present in the synaptic cleft, thereby increasing overall serotonergic activity and, in time, alleviating depressive symptoms (Nutt et al., 1999).

In fact, it would be fair to say that the development of an intervention plan critically depends on the acceptance of a causal explanation, however rudimentary and conceptually thin this may be. Without some idea of what is causing a particular problem, it is difficult to justify a proposed intervention plan. Explanations identify the causal factors and mechanisms responsible for initiating and maintaining the client's symptoms, allowing them to then be targeted in treatment. It is worth noting that without an accurate understanding of what was causing the problem it would be very difficult to intervene effectively and the condition would be less likely to resolve. This is now being observed in SRRIs, having recently been revealed as far less effective than once believed and potentially possessing additional or alternative mechanisms of action to those originally proposed (Galecki et al., 2018; Kirsch, 2014; Wolkowitz et al., 2011).

This is often the case in the psychopathology field, where most explanations are insufficient (e.g., impoverished, incomplete, lacking validity; see Chapter Three). Accordingly, many psychotherapeutic treatments show poor efficacy. For example, cognitive behaviour therapy (CBT), the current gold-standard intervention for most psychiatric disorders, demonstrates only low to moderate effectiveness, both in remission rates and effect

sizes, across many mental disorders, such as MDD (Craighead et al., 2015), anxiety disorders (Springer et al., 2018), and EDs (Wilson, 2018). Treatments for many of the more serious psychiatric conditions, such as bipolar disorder and schizophrenia, typically fail to restore premorbid functioning and frequently amount to mere management of symptoms, often with gradual deterioration across the lifespan (Abbas & Lieberman, 2015; APA, 2010; Geddes & Miklowitz, 2013). Hence, our ability to effectively attend to clients' problems fundamentally depends upon the quality of the scientific explanations we possess to account for them. If these explanations are false or otherwise flawed, treatments are likely to have poor or insufficient efficacy, as is currently the case.

What Makes a "Good" Explanation?

The theory of Inference to the Best Explanation (IBE; Haig, 2014) proposes that explanations be judged in relation to their rivals based on their *explanatory worth* (or "explanatory power"). In other words, explanatory accounts should be assessed relative to explanatory criteria and the explanation that best meets them should be deemed most likely to be correct (Haig, 2014). There are various different ways in which explanatory accounts may be evaluated for their explanatory worth. Ultimately, what constitutes a "good" explanation will depend on the purpose to which it is put. For example, an explanation may be useful for prevention but not treatment, such as those involving *epigenetics* (see Feinberg, 2008). However, various *epistemic* criteria have been proposed to help evaluate explanatory accounts on a more theoretical level.[7]

Thagard's (1989, 1992) Theory of Explanatory Coherence (TEC) proposes that an explanatory account is accepted if it *coheres* better than its competitors. Explanations cohere or "hold together" based on the quality of their explanatory relations. An explanation may therefore be made more coherent by relations that make logical sense and correspond to background knowledge, and more incoherent by propositions that contradict each other or are incompatible with broader understandings. The TEC proposes seven principles by which the degree of coherence in an explanatory relation may be assessed: *symmetry*, *explanation*, *data priority*, *analogy*, *contradiction*, *competition*, and *acceptability*. At the level of the overall explanation, these principles can arguably be effectively captured by the following epistemic

[7] *Epistemic* describes something that relates to knowledge or the degree of its validation. It comes from the discipline of *epistemology*, which is the branch of philosophy that is concerned with knowledge, its nature, origin, and scope (Fumerton, 2006). *Epistemic criteria* therefore refer to theoretical values that increase the likelihood of an obtained piece of knowledge, such as a theory or model, being accurate or "true".

criteria (Hooker, 1987; Newton-Smith, 2002; Ward et al., 2006): *predictive accuracy*, *empirical adequacy*, *scope*, *internal coherence*, *external consistency*, *unifying power*, *heuristic value*, *simplicity*, and *explanatory depth*.

- *Predictive accuracy* refers to how well an explanatory account can predict outcomes. An example is Bowlby's Attachment Theory (Ainsworth, 1989), which predicted relations between insecure attachment, self-esteem, coping style, and emotional intelligence that have subsequently been supported by empirical research (e.g., Howard & Medway, 2004; Kafetsios, 2004; Laible et al., 2004).
- *Empirical adequacy* and *scope* assess both how well an account explains existing empirical findings, as well as the range of phenomena in need of explanation. Operant Conditioning Theory (Staddon & Cerutti, 2003) has excellent empirical adequacy and scope, having been able to account for numerous empirical findings over the years, as well as being successfully applied to a wide range of psychopathological phenomena, including child conduct problems, binging and purging behaviour, anxiety disorders, and self-harm (Chapman et al., 2006; Dishion & Patterson, 2015; Granic & Patterson, 2006; Gratz, 2003; Haedt-Matt & Keel, 2011; Hudson & Rapee, 2004; McLeod et al., 2011; Pearson et al., 2015).
- *Internal coherence* evaluates the degree to which an explanatory account is logically consistent, without gaps or contradictions in reasoning. One way in which explanatory accounts may lack internal coherence is when they cannot be *falsified*: they contain fallacies or tautologies of logic that prevent them from being scientifically tested. Many of Freud's theories fall into this camp (e.g., Psychoanalytic Theory, Dream Interpretation, Theory of Infantile Sexuality; Freud, 1964), as their claims relating to the "unconscious mind" cannot be empirically verified.
- *External consistency* reflects how well an explanatory account aligns with existing background theories. Bonta and Andrews' (2017) General Personality and Cognitive Social Learning Theory of criminal conduct, although arguably flawed in many other ways (see Fortune & Heffernan, 2019; Heffernan & Ward, 2017), has good external consistency as it draws together ideas from earlier criminological theories, such as Differential Association Theory and Strain Theory (Bonta & Andrews, 2017), as well as applying broader psychological theories about personality, social learning, and cognition.

- *Unifying power* assesses whether an explanatory account draws together phenomena from related domains, such as connecting research findings previously considered unrelated. Ward and Brown's (2006) Good Lives Model (GLM) of offender rehabilitation is an example of an account with good unifying power, as it connects the phenomenon of offending behaviour with other psychological, social, and lifestyle problems – such as mental disorder, social isolation, poverty, and unemployment – by conceptualising each as a form of "breakdown" in the process of an individual seeking "primary human goods" – such as *mastery* or *relatedness* – necessary for a "good life".
- *Heuristic value* represents the ability of an explanatory account to generate new and fruitful avenues of inquiry or, in clinical practice, interventions. Behaviourism (Skinner, 1976) is one example of a theory that has stimulated a particularly large body of research, as well as forming the basis of many different intervention strategies for psychopathological problems, such as *exposure and response prevention, behavioural activation, extinction, contingency management,* and *differential reinforcement.*
- *Simplicity*, or "parsimony", is a feature of explanatory accounts that make the fewest special assumptions, rather than requiring multiple specific conditions for their reasoning to make sense. Skinner's (1976) Behaviourism Theory is also a good example of this, as it explains a huge variety of phenomena using very simple concepts based on observable behaviour.
- Finally, *explanatory depth* assesses how well an explanatory account details underlying mechanisms or processes. Clack and Ward's (2020) compositional "sketch" of the symptom "anhedonia" within the context of depression is a good example of explanatory depth in action, as it models the factors and processes that comprise the symptom at a variety of progressively deeper levels (see Chapter Five).

These criteria are somewhat idealistic: it is unlikely that an explanation will perfectly exemplify them all. In some cases, values may contradict each other, such that the achievement of some may create an obstacle to the achievement of others. For example, the ability of an explanatory account to explain a wide range of empirical findings (empirical adequacy) and clinical phenomena (scope) is also likely to make it more convoluted, thereby reducing its simplicity. Similarly, an explanatory account built on previous theories may have good external coherence, but could lack internal coherence, explanatory depth, or heuristic

value if those theories are themselves poor quality. Some values may also be less important for a particular explanatory account to achieve based on the surrounding scientific context and the purposes for which it is intended. For example, external coherence becomes a less meaningful evaluative criterion at the point of a paradigm shift, during which background theories and models are being discarded. However, these criteria represent a useful guide for the comparison of rival scientific explanations, as well as a tool by which explanations may be evaluated for further development.

A "good" explanation will be one that balances these epistemic values effectively in its account of the phenomenon in question. In other words, one that provides a logical, elegant, and sufficiently detailed account of a phenomenon and its related empirical findings, integrates diverse information, assimilates existing theories, and signals fruitful avenues for investigation and clinical practice. This is not to say that explanations that do not achieve this standard should be dismissed or viewed as "bad". To attain such a level of explanatory calibre is extremely challenging (if not impossible) and may not currently be possible within some domains of scientific study, such as psychopathology, that are still in their infancy. However, these values still provide a useful standard against which explanatory accounts may be measured.

Summary

This chapter comprehensively discussed explanation. In this thesis, an *explanation* refers to the most complete explanatory account of a phenomenon; one that sufficiently details the factors and processes involved across levels of analysis. In the case of more complex phenomena, such as those in psychopathology, this is likely to involve the generation of multiple models and/or the application of diverse theories. Overall, this explanation should, as comprehensively as possible, account for either the causes, composition, context, or consequences of the phenomenon depending on the type of explanation in question (e.g., aetiological versus compositional). Such explanations can then be used to inform the classification of psychopathological conditions, as well as the assessment, formulation, and treatment of clients' difficulties. They may also, in some cases, suggest new and fruitful avenues of clinical research.

Explanations may be of varying quality, however, and therefore require careful theoretical appraisal to determine their likelihood of providing an accurate and useful explanatory account. This chapter concluded by presenting a set of epistemic criteria that may

be used to appraise explanations for their explanatory worth – namely, predictive accuracy, empirical adequacy, scope, internal coherence, external consistency, unifying power, heuristic value, simplicity, and explanatory depth. A “good” explanation will be that which most effectively balances these epistemic criteria – a tall order. In the next chapter, these criteria are used in a critical evaluation of explanations for disordered eating that underlie best-practice treatments.

Chapter 3. Examining Current Explanations for Eating Disorders

In this chapter, I review and critically evaluate the four explanations that are associated with current "best-practice" interventions for EDs: Minuchin et al.'s (1975, 1978) Psychosomatic Family (foundation of Family-Based Treatment (FBT); Le Grange & Rienecke, 2018)[8], Fairburn et al.'s (2003) Transdiagnostic Cognitive Behavioural Theory of Eating Disorders (TCBT; foundation of Enhanced Cognitive Behaviour Therapy (CBT-E); Fairburn, 2008), Rieger et al.'s (2010) Interpersonal Theory of Disordered Eating (IPT; foundation of Interpersonal Therapy for Eating Disorders (IPT-E); Burke et al., 2018), and Linehan and Chen's (2005) Affect Dysregulation Theory (ADT; foundation of Dialectical Behaviour Therapy for Eating Disorders (DBT-E); Chen et al., 2018).[9]

At the close of the previous chapter, I presented a set of epistemic criteria that can be used to evaluate the quality of explanations – predictive accuracy, empirical adequacy, scope, internal coherence, external consistency, unifying power, heuristic value, simplicity, explanatory depth. I now apply these to the above explanations to examine whether these accounts provide a sound basis for the clinical interventions they underlie. For each, I briefly outline the content of the explanation, critically analyse it, and discuss the effect that any conceptual shortcomings may have on the efficacy of its associated intervention. In the course of my analysis, I identify three broader issues that span ED explanations, and psychopathological explanations in general: 1) the poor selection and description of explanatory targets, 2) the equation of models with explanations, and 3) the lack of a guiding methodological approach for constructing explanations. These issues are discussed briefly at the close of this chapter. The subsequent chapters of this thesis then address each in turn, elaborating on the nature of the problem and proposing how it could be resolved.

Critical Analysis of Current ED Explanations

As discussed in the previous chapters, one of the most important uses for explanation is as a guide for clinical intervention. As such, I chose to focus my analysis on explanations associated with the current best-practice psychotherapeutic treatments for

[8] Also known as the Maudsley Method of Family Therapy, due to the first study being conducted at the Maudlsey Hospital in London (Le Grange & Rienecke, 2018).

[9] It should be noted that although several of these explanations are labelled by the authors as "theories" or "models", they are presented as *explanations* – as though they represent the entire bank of explanatory knowledge relative to the disordered eating phenomenon in question. Hence, they are evaluated herein as complete *explanations*, rather than theories or models.

EDs, as these are likely to represent the best current understandings of EDs, or at least those that are most widely accepted. These are, Minuchin et al.'s (1975) Theory of the Psychosomatic Family, Fairburn et al.'s (2003) Transdiagnostic Cognitive-Behavioural Theory of Eating Disorders, Rieger's (2010) Interpersonal Theory of Disordered Eating, and Linehan and Chen's (2005) Affect Dysregulation Theory. Each forms the foundation for one of the current best-practice psychological interventions in this area – FBT, CBT-E, IPT-E, and DBT-E respectively – and is frequently used to direct ED research. In my analysis of each, I provide a brief account of the content of the explanation, analyse it in relation to the evaluative criteria from the previous chapter, and discuss how the problems identified may impact the efficacy of its corresponding intervention approach.

The Theory of the Psychosomatic Family

Minuchin et al.'s (1975, 1978) account of the Psychosomatic Family is the explanation upon which FBT (Russell et al., 1987) – currently the most widely used treatment for Anorexia Nervosa (AN) in adolescence – was originally constructed (Le Grange & Rienecke, 2018). The explanation proposes that certain qualities of a family system lead to the development and maintenance of psychosomatic problems in children, such as AN.

First, these families are thought to be highly *enmeshed*: there is a lack of appropriate interpersonal boundaries, members are overly involved in each other's lives and children often inappropriately adopt parenting roles or become enlisted as "allies" by one parent against the other (Minuchin et al., 1975, 1978). Second, it is theorised that these families are extremely *overprotective*: family members show a high level of concern for each other's welfare and are hypersensitive to signs of distress, which are met with excessively nurturant and protective responses. Third, the structure of these families is conceived of as highly *rigid*: families are extremely fixed in their behaviour and experience great difficulty managing the necessary structural changes that come as children mature, such as negotiations of rules and boundaries to facilitate age-appropriate autonomy. Finally, it is proposed that the family lacks the ability to resolve conflict effectively: in some families, problems are avoided or denied, while in others they are openly confronted but resolution is impeded by obstacles like interruptions or subject changes.

The child's psychosomatic symptoms – in this case, disordered eating behaviour – are thought to be prompted or encouraged by this "dysfunctional" family environment

(Minuchin et al., 1975, 1978). Primarily, the child's illness is thought to act as a sort of conduit for familial conflict, especially that between parents. The presence of a sick and vulnerable child within the family is thought to help to unite warring parents in care and concern or to facilitate avoidance of deeper relationship conflict by channelling this into superficial disputes about how the sick child should be managed. The illness behaviour is therefore thought to be reinforced due to its ability to suspend or redirect familial conflict, thereby regulating the internal stability of the family.

Upon evaluation, several problems emerge with Minuchin et al.'s (1975, 1978) explanation. Empirically, the Psychosomatic Family account lacks evidentiary support – in fact, evidence now actively refutes it (Eisler, 2005). The account makes bold (and rather judgmental) claims about the nature of families with a child suffering from AN – that they are "overprotective", "rigid", and broadly "dysfunctional". Research has attempted to validate these claims by studying families of anorexic individuals to see if these "dysfunctional" patterns emerge but have been unable to identify either this or any other set of characteristics that reliably typifies these families (Le Grange & Rienecke, 2018). Families of anorexic patients appear instead to be a very heterogeneous group, both in terms of their collective characteristics – such as their sociodemographic profile, patterns of interaction, and relational dynamics – and the way they respond to the presence of the disorder (Eisler, 2005; Larsen et al., 2015; Le Grange & Rienecke, 2018; Le Grange et al., 2010; Schaumberg et al., 2017). Hence, the empirical adequacy of the Psychosomatic Family Theory is very poor, as it has now been largely disproven by scientific research. It also demonstrates problems with scope, as it fails to explain why anorexic behaviour can commence or continue in adulthood when an individual is no longer embedded in this "disordered" family system.

The Psychosomatic Family Theory also lacks unifying power, in that it implicates the family system as the sole causal and maintaining process in AN – specifically, that the primary function of the individual's disordered eating behaviour is to regulate the family's internal conflict. This seems an overly simplistic interpretation of a highly complex behaviour, which could be serving a variety of different functions across individuals, such as emotion regulation, social acceptance, or the management of physical sensations.[10] Even

[10] Indeed, research has begun to identify a variety of motives that might contribute to anorexic behaviour, such as seeking identity or security, avoiding distressing emotions, communicating with or accessing care from others, or achieving a sense of mastery (Lavis, 2018; Norbø et al., 2006; Stockford et al., 2018).

if regarded as a *model* of how the family system specifically influences disordered eating behaviour rather than a complete explanation, it is relatively impoverished, ignoring how factors and processes at levels other than the familial – such as genes, neurotransmitters, appetitive biology, cognitive distortion, emotion regulation, or cultural norms and values – could interact with family dynamics.

For example, there is no consideration of individual difference factors that could impact why individuals who develop an ED may be differentially affected by the family processes proposed, such as temperament or neurotransmitter differences. This is particularly important given that families of anorexic individuals often have other children, who typically do not present with disordered eating problems (Maon et al., 2020).[11] If the proposed "dysfunctional" family dynamics were the sole causal factor for anorexic behaviour, one would expect to see all children in a family being similarly affected – yet they are not. There is also no mention of how the effects of starvation, such as physical emaciation or alterations in mood and cognition, could initiate or escalate the familial patterns proposed – for example, it is understandable that a family's protectiveness might increase in *response* to an anorexic child's excessive weight loss and increased physical vulnerability, rather than being a causal contributor. Granted, much of the evidence for the contributions of these other domains may not have been available at the time the Psychosomatic Family account was constructed. Nevertheless, to place such responsibility on the family system alone in my view represents an extremely narrow perspective.

Despite the lack of empirical support for the Psychosomatic Family Theory, the intervention based on it – FBT – is reliably shown to be the most effective current treatment for AN (Le Grange & Rienecke, 2018). In response to this unusual situation, FBT has since progressed to a place of *theoretical agnosticism*: its developers have chosen not to align it with *any* theoretical account and claim that it is "atheoretical" (e.g., Strategic Family Therapy; Lock & Le Grange, 2005).

While it is commendable that FBT has renounced the debunked Psychosomatic Family account, the choice to forgo an association with *any* explanatory account is problematic. All interventions should arguably be informed by theory – it is not good enough to simply accept that a treatment "works" without trying to understand *why* (Del

[11] However, they are shown to demonstrate poorer overall quality of life and have an increased risk of developing psychopathology and ED-related symptoms (Maon et al., 2020).

Giudice, 2018; Kendler et al., 2011). This leaves clinicians without a coherent understanding of *what* mechanisms an intervention is targeting and *how*, which makes building on it effectively or prescribing it appropriately challenging. For example, FBT does not show efficacy for all cases of AN – such as chronic, treatment-resistant, or adult presentations (Berkman et al., 2007; Keel, 2018; Murray, 2020; Murray et al., 2018; Watson & Bulik, 2013) – and there is currently no conclusive or satisfactory explanation for this. It is therefore difficult to determine which clients are likely to benefit from a course of FBT and which might need additional or alternative therapies, as well as how FBT could be adjusted or supplemented to make it more effective. Furthermore, without an idea of how a treatment is working, there is always the risk that it may inadvertently cause harm – for example, due to its interaction with individual or contextual factors or other treatment elements – akin to prescribing a drug without knowledge of its pharmacological action (Holmes et al., 2018).

Although its developers have disassociated it from the Psychosomatic Family account, FBT's origin in an invalid explanation cannot be fully escaped – especially as no alternative explanatory account has been generated to replace it. Since dissociating itself from the Psychosomatic Family theory, FBT no longer focuses on modifying family dynamics, but revolves largely around weight restoration, with the family viewed as a resource for achieving this. Although an important clinical goal, this approach would seem to suggest that simply restoring weight should be enough to prevent relapse, and thereby fails to consider or address factors that may have prompted the anorexic behaviours responsible for the weight loss in the first place. This may partially account for the high relapse rates seen in anorexic patients post-treatment (Khalsa et al., 2017; Murray et al., 2018) – as the factors that led them to initially engage in their disordered eating behaviours are not theoretically formulated nor clinically addressed as a matter of course.

The Transdiagnostic Cognitive-Behavioural Theory of EDs

Fairburn et al.'s (2003) Transdiagnostic Cognitive Behavioural Theory of EDs (TCBT) is the explanation behind CBT-E, the current frontline treatment for both BN and BED (Wilson, 2018). It proposes that central to all EDs is the same "core psychopathology" – the "over-evaluation of eating, shape and weight and their control" (Fairburn et al., 2003, p. 510) – which is thought to inform the strict dieting and weight-control behaviour observed in ED sufferers. In individuals with an anorexic presentation, these behaviours are

proposed to lead to significantly low weight, which prompts a "starvation syndrome" (Fairburn et al., 2003, p. 521) characterised by intense hunger, impaired concentration, cognitive narrowing, and elevated satiety. This "starvation syndrome" is then proposed to reinforce the "core psychopathology" and disordered eating behaviour in a variety of ways – for example, intense hunger and impaired concentration increasing the individual's anxiety that they won't be able to maintain their strict dieting, which provokes more intense restriction to compensate (see Fairburn et al., 1998).

In those with a bulimic presentation, failed attempts to stick to excessively strict dietary rules are hypothesised to trigger binge eating, as breaches provoke strong negative emotions that cause restrictive efforts to be temporarily abandoned.[12] These binge episodes are thought to reinforce the "core psychopathology" by amplifying the concern about being able to control eating, shape, and weight, which then leads to redoubled restrictive efforts. For some individuals, binge episodes also prompt purging or compensatory behaviours, such as self-induced vomiting, laxative abuse, and excessive exercise, with the aim of preventing weight gain that might otherwise result from binge eating. However, this also causes the key deterrent for binge eating – weight gain – to be undermined, making binge eating more likely to occur in future.

As well as these core disorder mechanisms, the TCBT theorises that four additional maintenance processes are present in some cases: 1) *clinical perfectionism,* 2) *core low self-esteem,* 3) *mood intolerance,* and 4) *interpersonal difficulties* (see Figure 1). *Clinical perfectionism* is defined by the authors as "the over-evaluation of the striving for, and achievement of, personally demanding standards, despite adverse consequences" (Fairburn et al., 2003, p. 515). This trait is hypothesised to amplify or contribute to the individual's "core psychopathology" as these individuals are likely to adopt particularly high standards for their physical appearance and restrictive eating behaviour, to be extremely self-critical of their performance, and to have an intense fear of failure. All of this is thought to encourage even greater efforts at controlling eating, weight, and shape. *Core low self-esteem* is used to refer to individuals who have an underlying "global negative view of themselves" (Fairburn et al., 2003, p. 516). This is thought to prompt an even more determined pursuit of achievement in their valued domain of eating, weight, and shape

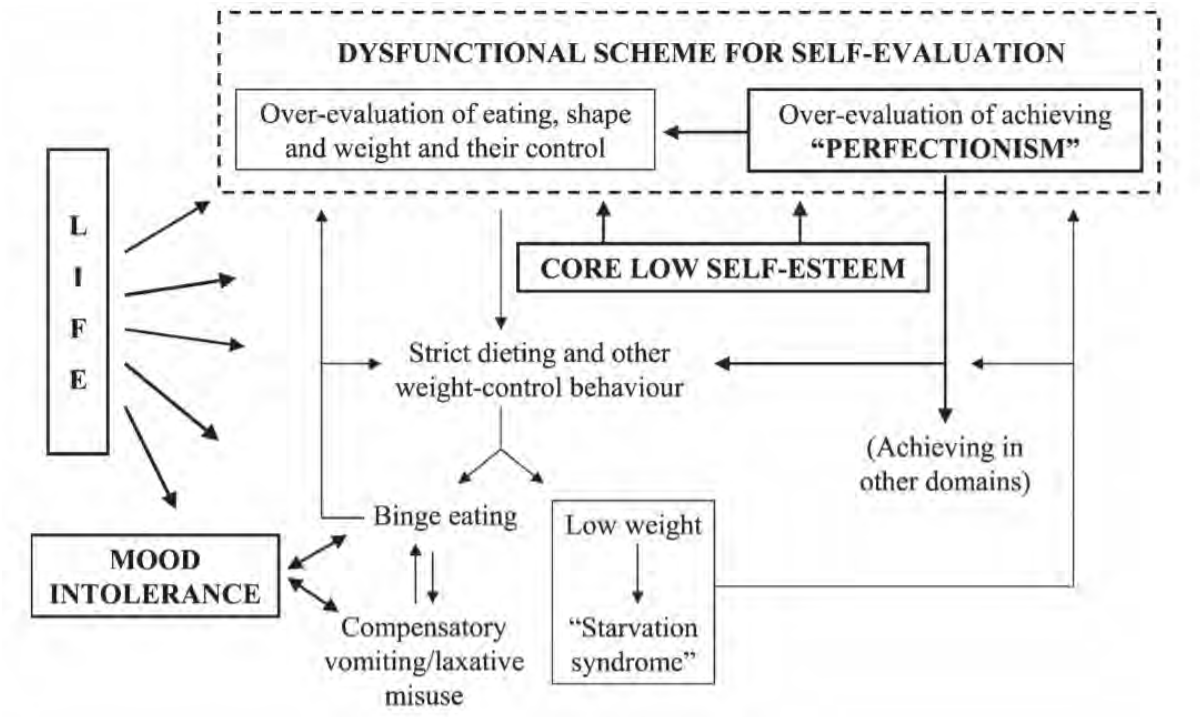
[12] This is often described as the *abstinence violation effect* (Ward et al., 1993).

control in order to achieve a sense of self-worth, as well as causing them to interpret any failures in that area as further evidence of their low value.

Mood intolerance is hypothesised to account for those individuals who have an inability to cope with intense emotional arousal – most often negative affect, such as anger, sadness, and shame. Instead of managing emotions adaptively, these individuals are thought to engage in various forms of "dysfunctional mood modulatory behaviour" (Fairburn et al., 2003, p. 517), including disordered eating behaviours, such as binging, excessive exercise, and self-induced vomiting, as well as other dysfunctional behaviours like self-harm or substance abuse. Finally, *interpersonal difficulties* refers to a range of interpersonal stressors that may be present for an individual and contribute to their disordered eating, either directly (such as family members criticising weight and shape or modelling behaviours and cognitions that promote disordered eating) or indirectly (such as lack of control experienced as a result of chaotic or conflictual familial environments with which disordered eating may be used to cope).

Overall, TCBT represents a better explanation than the previous one evaluated here. Unlike the Psychosomatic Family account, TCBT shows good empirical adequacy, being developed based on empirical literature and supporting each theoretical assertion with specific research findings. It also incorporates several widely accepted psychological theories in its account, such as Operant Conditioning Theory (Staddon & Cerutti, 2003) and the *abstinence violation effect* (Ward et al., 1993), and thus boasts good external consistency. Furthermore, the model has very broad scope, as it takes a transdiagnostic approach to the explanation of EDs, and therefore seeks to explain multiple clinical phenomena by combining explanation of the AN and BN diagnostic categories. In many ways, the TCBT would seem to present a rather elegant solution to the ED problem. At its core, however, the explanation encompasses but a single model, which arguably tries to explain too much. Although the transdiagnostic approach of the TCBT increases the explanation's scope by accounting for multiple ED phenomena, it sacrifices several other epistemic values along the way.

Figure 1. *A Schematic Representation of the Transdiagnostic Cognitive-Behavioural Theory of Eating Disorders*



Note. Originally captioned "A schematic representation of the 'transdiagnostic' theory of the maintenance of eating disorders". Reprinted from "Cognitive behaviour therapy for eating disorders: A "transdiagnostic" theory and treatment" by C. G. Fairburn, Z. Cooper, & R. Shafran, 2003, *Behaviour Research and Therapy, 41*(5), p. 523. Copyright © 2003 Elsevier Science Ltd. All rights reserved.

First, by combining syndrome clusters, the explanation attempts to account for an extremely large and heterogeneous collection of phenomena within a single model and, in the process, greatly forfeits explanatory depth. As one example, the "mood intolerance" construct – invoked to explain why some individuals appear to use ED behaviours to help regulate their emotions – is only vaguely defined. One might enquire, what are the mechanisms that cause an individual to have difficulty tolerating their emotional states? Is it because their emotions are experienced more intensely than other peoples', such that the associated physiological arousal is more pronounced and unpleasant? Is it because they struggle to detect and identify emotional states (*alexithymia*; Nowakowski et al., 2013), such that the experience of intense emotion is confusing and, thereby, distressing? Or is it due to an inability (or perceived inability) to tolerate and manage emotional states, such as

a lack of coping skills or low self-efficacy? Each of these possibilities presents a very different understanding of "mood intolerance" and would suggest different strategies for intervening in it.

It is likely that "mood intolerance" is, in fact, a composite construct, which may result from one or more of these mechanisms. It therefore follows that the clients who present with this trait may have developed and sustained it via different processes and may therefore require different interventions to address it. For example, a client presenting with alexithymia might be helped by mindfulness – to improve their awareness and recognition of physiological emotion sensations – and psychoeducation – to improve their emotional vocabulary and ability to identify which emotions are associated with which physical sensations. In contrast, a client who lacks effective strategies for managing their emotional distress might find coping skills training more beneficial. Alternatively, a client's difficulty tolerating intense mood states may result from both mechanisms, in which case both approaches are likely to be useful. However, without these mechanisms specified within the explanation, interventions risk targeting this construct inappropriately in all or some clients. Although some clinicians may seek to understand the exact nature of their client's "mood intolerance" and intervene in ways specific to their case, they could easily formulate this inappropriately due a lack of theoretical guidance. Other clinicians may fail to recognise that further specification is necessary, and thereby problematically reify and rely on this construct in clinical practice.

Second, although the TCBT attempts to create a unified account of the cognitive and behavioural aspects of EDs – both of which are incorporated effectively – it often fails to consider empirical research that suggests that factors in other domains – most notably, biological factors, such as neurotransmitters and appetitive hormones – might underlie or influence these in significant ways. For example, it may be that, at least in some individuals, dysfunction in neurotransmitters like serotonin may underlie or contribute to the mood and appetite problems invoked by the explanation. Several research studies indicate significant links between serotonin activity and both these and other psychological and physiological processes, such as impulsivity, anxiety, and the regulation of various neurochemical systems (Brewerton, 1995; Leibowitz, 1990; Steiger, 2004). Indeed, the treatment generally considered most effective for BN and BED is typically a combination approach involving both CBT-E and antidepressant medication, such as SSRIs (McElroy et al., 2012; Van den Eynde & Schmidt, 2008). This implies that the modulation of

neurotransmitters like serotonin either directly influences BN symptoms or helps to facilitate the therapeutic effects of CBT-E. There is scant consideration of this within the TCBT account.

As the model within this explanation already attempts to provide an account of the cognitive-behavioural processes involved in initiating and sustaining ED behaviour, it would be inappropriate to expand it to include a comprehensive account of neurotransmitter function as well. This would over-complicate the account and could be better handled via the inclusion of a separate model, or models, detailing neurological mechanisms. However, it is worth making some reference within the existing account to how these processes might underlie or interact with the cognitive-behavioural factors invoked, in the same way that reference is made to environmental contributing factors such as interpersonal difficulties, which one can concede may be expanded on by other explanatory models. Making these kinds of theoretical connections will arguably help with treatment integration down the line, for example providing clearer understandings of how treatment with CBT-E and SSRIs might interact or for which clients it may be most appropriate.

Finally, there are a couple of issues with the internal coherence of the TCBT, largely in terms of some of the theoretical constructs invoked. Most notable is the “core psychopathology” posited: “the over-evaluation of eating, shape, weight, and their control” (Fairburn et al., 2003, p. 510). The key problem with this construct is that it is both highly normative and vaguely tautological. To suggest an “over-evaluation” implies there is some “normal” level to which one should value these factors and some point at which the pursuit of these qualities becomes “abnormal”. Indeed, there may be a statistical norm that could be identified by surveying the general population – however, being statistically outside of population norms is not necessarily a problem (e.g., high achievers in academic subjects). Thus, the prima facie labelling of this construct as pathological seems to be an overly bold claim.

Although an individual may value these factors to a greater degree than the majority of the population, it does not automatically follow that this is inherently “maladaptive” or “abnormal”. One can conceive of non-pathological scenarios in which these factors may be similarly prized, such as among competitive athletes, health and nutrition enthusiasts, and general fitness fanatics. It therefore seems overly presumptuous to assert that the possession of these goals is so uniquely central to the development of disordered eating behaviour as to

be characterised as a "core psychopathology". It seems much more likely that this suite of values may confer differential *risk* depending on the co-existence of various other factors, both cognitive (e.g., cognitive biases, personality variables) and behavioural (e.g., observational or reinforcement learning).

This distinction is important as currently the explanation suggests that this "core psychopathology" is doing the bulk of the explanatory work and, accordingly, should be the main target of intervention. This is problematic as it prematurely assumes that high valuing of eating, shape, weight, and one's ability to control these is inherently problematic, and downplays the importance of other contributing factors. It may not necessarily follow that this goal needs to be rejected, as it is not pathological in and of itself – indeed, the value placed on physical appearance by society rationalises and recommends its pursuit, and there are many occupations that align with it, such as personal training, dietetics, or appearance medicine. Instead, it is likely that the goal is being sought maladaptively or excessively due to additional factors, such as a rigid cognitive style, low self-worth, or anxiety. These factors are therefore doing important explanatory work but are downplayed by the explanation's claim of a "core psychopathology". It may actually be possible to help clients reformulate or redirect the achievement of this goal in a more appropriate and adaptive way, without citing it as inherently "bad". For one individual it could be that it is their rigid and perfectionistic pursuit of these values that is creating the problem, and by managing these tendencies appropriately they could pursue their valued goals in a non-problematic way, such as through a career in health and fitness. For another, it may be that improvement of their self-esteem eliminates the need for them to pursue the goal, as the goal has no inherent value to them other than its ability to achieve a sense of personal worth. In both cases, the explanatory work is being done by these additional factors, rather than the proposed "core psychopathology".

The Interpersonal Theory of Disordered Eating

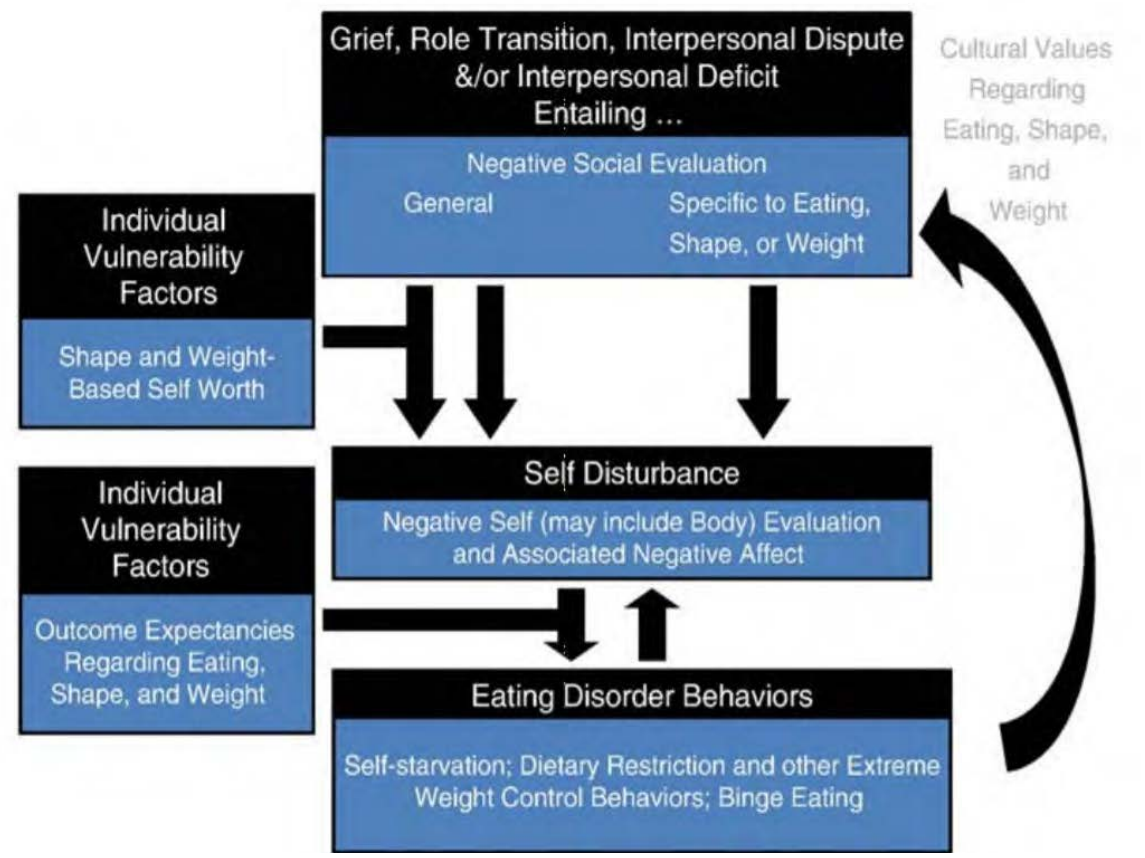
General interpersonal theories of psychopathology perceive interpersonal functioning as "a critical component of psychological adjustment and well-being" (Burke et al., 2018, p. 288). Rieger et al.'s (2010) Interpersonal Theory of Disordered Eating (IPT) applies this assumption to the context of EDs, proposing that problems with interpersonal functioning are at the heart of these conditions (see Figure 2). These interpersonal problems may take one or more forms – namely, *grief* (a recent or past loss), *interpersonal role*

disputes (conflicts with significant others that arise from differing expectations about their relationship to the client), *role transitions* (difficulties around changes in life status, such as puberty, moving house, or job loss), or *interpersonal deficits* (social isolation or involvement in unfulfilling relationships; Burke et al., 2018). These problems are proposed to result in *negative social evaluation* – actual or perceived negative evaluation by a valued social group, such as critical comments/bullying, unfavourable social comparison, exclusion, or insufficient social reinforcement – which then leads to a *disturbance of self*, involving negative self-evaluation and associated affect (Rieger et al., 2010).

According to IPT, disordered eating behaviours, such as self-starvation, binge eating, dietary restriction and other weight-control strategies, represent the individual's attempt to cope with this disturbance, primarily by avoiding the associated negative affect. For instance, binge eating might allow an individual a temporary escape from aversive emotions by distracting them or allowing them to transform confusing emotional sensations into more familiar physiological ones, such as hunger and fullness. Self-starvation may have a "numbing" effect, allowing affective experience to be avoided entirely.

It is important to note, however, that the causal links from negative social evaluation to disturbance of self, and from disturbance of self to disordered eating, rely on additional "individual vulnerability factors", such as "shape- and weight-based self worth[sic]" and "outcome expectancies (i.e., beliefs regarding the consequences of engaging in eating, shape, and weight related behaviors[sic] such as the belief that eating is helpful in managing negative affect or that social acceptance and self-esteem can be enhanced through attaining thinness)" (Rieger et al., 2010, p. 404). Over time, it is proposed that disordered eating behaviour comes to replace the function of healthy social interaction – namely, to maintain healthy self-esteem and positive affect. It is also thought to exacerbate existing interpersonal difficulties – for example, via familial conflict over disordered eating practices or withdrawal from social situations involving food – which in turn promote further disordered eating behaviour.

Figure 2. *Schematic of the Interpersonal Theory of Disordered Eating*



Note. Originally captioned: "The expanded IPT of disordered eating: The role of individual vulnerability factors". Reprinted from "An eating disorder-specific model of interpersonal psychotherapy (IPT-ED): Causal pathways and treatment implications" by E. Rieger et al., 2010, *Clinical Psychology Review, 30*(4), p. 404. Copyright © 2010 Elsevier Ltd. All rights reserved.

As with TCBT, IPT takes a transdiagnostic approach to EDs, making it subject to the same tension between scope and explanatory depth. Although the explanation provides an account of multiple ED behaviours, it does so at the expense of providing sufficient detail about the factors and mechanisms it proposes. One example is the theorised relationship between body dissatisfaction and negative affect. Body dissatisfaction is stated to provoke negative affect because, in Western society, appearance is a highly valued attribute that is attached to various desirable consequences such as social acceptance, occupational success, and finding a partner. However, there is a lack of information about how body dissatisfaction in the context of this cultural valuation specifically elicits negative affect – in other words, what *mechanisms* are at play. "Negative affect" is itself a composite construct, encompassing a range of experientially diverse emotions. For

example, being dissatisfied with one's body could conceivably cause an individual to feel sad (if one considers their body as central to their self-worth), guilty (if one believes their body reflects moral attributes like self-control or laziness), angry (with themselves, if they perceive themselves as failing societal standards, or with society, if they perceive such standards as unjustly imposed), ashamed (if they perceive others as judging them negatively based on their body), or fearful and anxious (if they believe their appearance may preclude love and acceptance). Each of these are distinct emotional states, accompanied by different cognitions, physical sensations, and behavioural responses, and brought about via different causal pathways. Each is therefore likely to be best addressed by a different intervention approach, which is not facilitated by an explanation that theoretically homogenises them.

Another example is the lack of detail provided about the "individual vulnerability factors" proposed to moderate the relationships between negative social evaluation, "disturbance of self" (negative affect and self-evaluation), and disordered eating behaviours (see Figure 2). These factors appear to do a lot of explanatory work – for instance, "disturbance of self" *only* leads to disordered eating behaviours when these vulnerability factors are present – but are thinly characterised. There is little detail provided about what they are or how they exert their effects. What *are* all the outcome expectancies around eating, weight, and shape that would lead to engagement in disordered eating behaviours? What are the specific cognitive processes (beliefs, thoughts, perceptions) that comprise these expectations? What happens in the case of multiple conflicting expectancies – for example, "eating will provide me with comfort" against "being thin will give me social acceptance"? And how do these expectancies come to exist in the first place? A significant amount of detail is missing here. Without such information, it is difficult to comprehend these crucial mechanisms adequately and to address them effectively in intervention.

Despite these problems, it is worth noting that the explanation succeeds to a degree in its external coherence and unifying power. The explanation applies an existing accepted theory of psychopathology (Interpersonal Theory; Stack-Sullivan, 1955) to the specific problem context of EDs. In doing so, it highlights the role interpersonal difficulties may play in initiating and maintaining disordered eating behaviours; a connection that had not previously been emphasised. However, its central thesis, that the key function of ED behaviours is to provide goods that would otherwise be achieved through healthy interpersonal relationships, seems overly simplistic. As with Minuchin et al.'s (1975, 1978)

Psychosomatic Family theory, which claims a single essential function for AN behaviour, this explanation makes a very bold claim about the fundamental purpose of disordered eating. There is significant research suggesting that disordered eating behaviours can serve a variety of functions for an individual, including mastery, security, control, and emotion regulation (see Nowakowski et al., 2013), and it seems reductionist to assume all of these could be adequately provided solely by adaptive social interaction (although undoubtably an important contributor to good mental health). There is also evidence that disordered eating behaviours may be partially biologically self-sustained by their various physiological effects – for example, the effects of "starvation syndrome" mentioned in the TCBT above (Casper, 2006; Dwyer et al., 2011; Keys et al., 1950). To therefore claim the role of interpersonal distress as so central to the aetiology and maintenance of disordered eating seems unjustified.

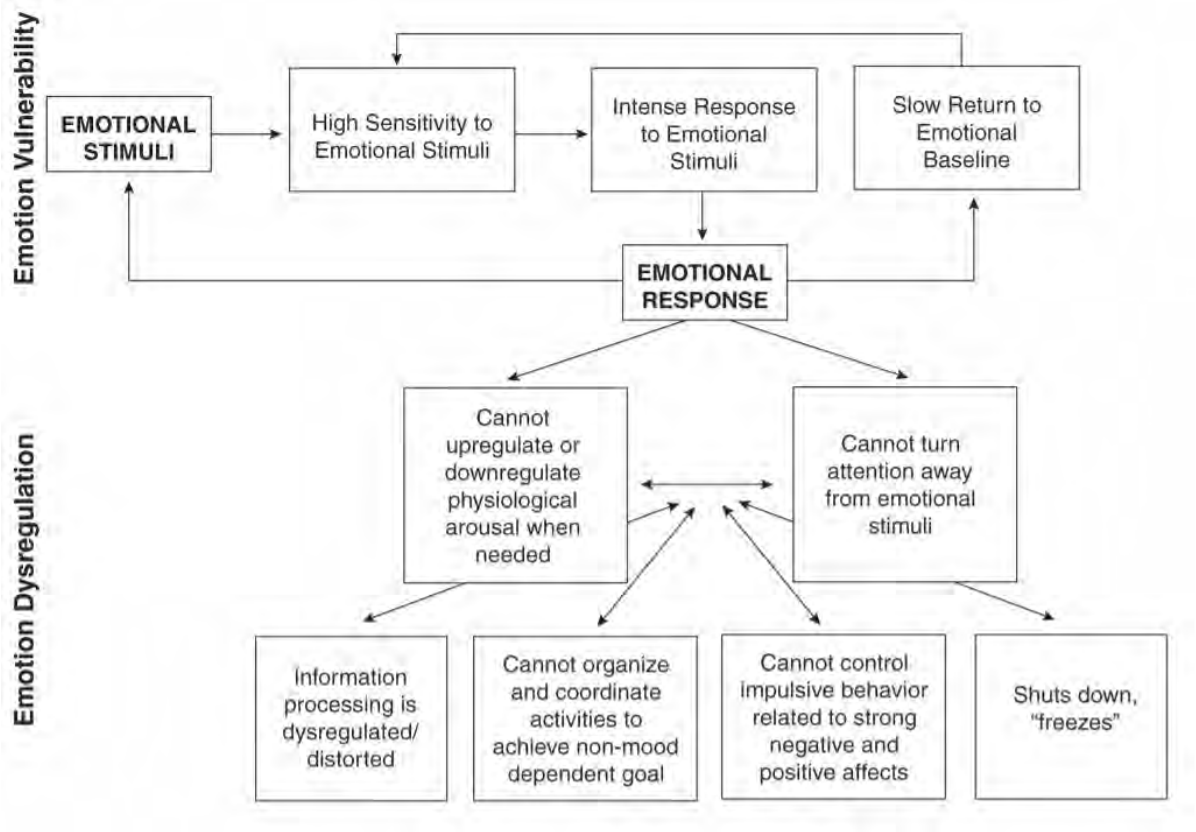
Affect Dysregulation Theory

The Affect Dysregulation Theory (ADT) was originally developed to explain borderline personality disorder (BPD) and underlies Dialectical Behaviour Therapy (DBT), the frontline treatment for this problem (Linehan, 1987). ADT proposes that certain problematic behaviours, such as self-harm and substance abuse, are engaged in to regulate aversive emotional states that individuals lack the skills to cope with adaptively (Chen et al., 2018). Individuals are proposed to struggle with emotion regulation for two reasons: 1) "emotional vulnerability" – which consists of an elevated sensitivity to emotional stimuli, intense emotional responding, and/or a slow return to baseline after emotional arousal – and 2) "difficulty modulating emotional reactions" – which involves difficulty inhibiting and organising behaviour in the context of intense emotion, self-regulating physiological arousal, diverting attention from emotional stimuli, and tolerating emotional experience (Bankoff et al., 2012). The maladaptive behaviours that these individuals use to cope with their aversive emotions become strongly reinforced because they provide immediate relief from distress (Chen et al., 2018). However, the physical, interpersonal, and occupational harm caused by these behaviours, as well as their social undesirability, contributes to additional negative affect like guilt and shame in the long-term, which the maladaptive behaviour is then further used to regulate (Chen et al., 2018). Hence, a vicious cycle of emotional distress and maladaptive regulation develops over time.

DBT is now also considered a promising treatment for ED sufferers who engage in binge-purge (B-P) behaviour, particularly cases that are difficult to treat and that other standard treatments – such as FBT, CBT-E, and IPT – have failed (Chen et al., 2018). Applied to this context, ADT suggests that B-P behaviour is used to block or distract from aversive emotional experience by providing the individual with an alternative, highly salient focus for their attention: the strong physical sensations produced by B-P behaviour (Linehan & Chen, 2005). This is thought to allow distraction from distressing thoughts and feelings and/or to translate the physical emotion sensations that the individual struggles to understand into the more comprehensible and less distressing sensations associated with eating (Linehan & Chen, 2005). As with the maladaptive strategies originally formulated by ADT, B-P behaviour is thought to be highly reinforcing, due to its ability to relieve emotional distress (Linehan & Chen, 2005). However, it also tends to provoke additional aversive emotions, such as guilt, shame, or disgust, that promote further B-P engagement and, over time, contribute to the development of a self-sustaining behavioural cycle (Linehan & Chen, 2005).

The ADT can be commended for its focus on a specific disordered eating phenomenon – B-P behaviour – rather than trying to account for all or multiple ED behaviours like the previous two approaches discussed. It also further refines its range of explanation exclusively to the emotional factors contributing to this phenomenon. Although this reduces its explanatory scope, it provides greater potential for explanatory depth and provides a relatively simple and elegant account of the role of emotional distress in B-P behaviour. Furthermore, although the explanation has been imported from another disorder (BPD), it has been appropriately and effectively applied to EDs by considering how the processes invoked by the original explanation account for a specific ED phenomenon: B-P behaviour.

Figure 3. *Schematic of the Affect Dysregulation Model*



Note. Originally captioned: "Schematic of the Affect Regulation Model of pervasive emotion dysregulation". Reprinted from "Dialectical behavior therapy for pervasive emotion dysregulation: Theoretical and practical underpinnings" by M. M. Linehan et al., in *The Handbook of Emotion Regulation*, 2006, New York, NY: The Guilford Press.

Despite its virtues, however, the model still falls short in several areas. First, it should be noted that the explanation does not account for *why* individuals engage in B-P behaviour rather than alternative strategies that might serve a similar function, including those which the explanation was originally developed to explain. B-P behaviour often occurs alongside other maladaptive regulatory strategies, such as self-harm or substance abuse, but not *always*. The explanation therefore lacks specificity, as it is ambiguous about what leads individuals to engage in B-P behaviour specifically versus these other strategies.

Second, although it focuses on the emotional factors and processes involved in B-P behaviour and provides a reasonable account of these, the explanation fails to provide any reference to factors at other levels that may underlie or interact with these or may play an

independent role in causing or maintaining B-P behaviour. For example, vulnerability to emotional stimuli could conceivably be affected by biological or cognitive factors such as an *overactive Hypothalamic-Pituitary-Adrenal (HPA) axis* (Chami et al., 2019),[13] distorted beliefs about emotional expression and coping (Meyer et al., 2009), or general sensory hypersensitivity (Merwin et al., 2013). There is also substantial evidence to suggest that individuals who engage in B-P behaviour have difficulty detecting, integrating, and interpreting bodily sensations (*impaired interoception*; Martin et al., 2019), which could contribute to difficulties in identifying and regulating both emotions (*alexithymia*; Panayiotou et al., 2021; Sfärlea et al., 2019) and appetitive states like hunger and satiety (Robinson et al., 2021). An intervention focusing solely on the emotional component of the problem, without acknowledging the need for complementary approaches at other levels, may therefore neglect the contribution of broader underlying mechanisms.

Third, although there is more room in this account than those previously discussed for the rich detailing of factors and processes, it is still somewhat lacking in explanatory depth, as there is insufficient description of the mechanisms underlying the two phenomena proposed to cause emotion regulation difficulties – "emotional vulnerability" and "difficulty modulating emotional reactions". "Emotional vulnerability" is proposed to result from several possible mechanisms – namely, greater sensitivity to emotional stimuli, more intense emotional responding, and extended time taken to return to emotional baseline – none of which are sufficiently detailed. For example, there is no explanation provided for why one individual may be more sensitive to emotional stimuli than others. Such a phenomenon could result from biological factors such as imbalances in hormones (e.g., oestrogen; Hiroi & Neumaier, 2011) or from cognitive factors such as attentional biases to threat or emotion-related beliefs (Ford & Gross, 2018), each of which would be best addressed by a different treatment approach. Similarly, a variety of causes are proposed for an individual's "difficulty modulating emotional reactions" – including lack of adaptive strategies, reduced executive function/self-control, beliefs about emotions and emotion management, and learned maladaptive responses – none of which are adequately described

[13] The *Hypothalamic-Pituitary-Adrenal* (HPA) axis is used to refer to the system of interaction between the hypothalamus, pituitary gland, and adrenal glands, the primary function of which is to stimulate and regulate the body's stress response (Lo Sauro et al., 2008). The HPA axis is activated in response to stress and causes the body to release stress hormones like *cortisol* that prepare the body for "fight or flight". In overactive HPA axes, there is persistent activation of this mechanism, such that the body is in a near-constant state of hyperarousal and alertness (Tafet & Nemeroff, 2020).

and explained. Associated intervention is therefore likely to lack specific guidance about how to address each of these problems.

General Issues with ED Explanations

In the process of my evaluation of ED explanations, I observed several broader, meta-theoretical problems: 1) the poor selection and description of explanatory targets, 2) endorsement of the One Best Model perspective, and 3) the lack of a guiding methodological framework for the explanation of psychopathological phenomena. As opposed to the specific epistemic deficits identified with the explanations above, these issues represent broader, meta-theoretical problems with how researchers currently approach the task of explaining psychopathological problems, including EDs. They critically influence how psychopathological explanations are structured and, thereby, may promote or discourage particular conceptual problems within them, such as the epistemic deficits identified in the explanations above. In Chapters Four, Five, and Six, I address each of these issues in turn; discussing the problem in depth and providing a detailed account of how I believe it could be resolved. However, I provide a brief account of this material here to introduce the reader to the discussion in the following chapters.

Issue 1: The Poor Selection and Description of Explanatory Targets

This issue relates to the nature of the explanatory targets typically used by psychopathological explanations and their description. As discussed in Chapter Two, an *explanatory target* is the phenomenon that an explanation, theory, or model claims to explain. For example, a model seeking to explain the onset of low mood, anhedonia, reduced appetite, insomnia, and suicidal ideation might have as its explanatory target the syndrome “clinical depression”. The problem with psychopathological explanations at present is that they typically focus on explaining either DSM syndromes like AN and BN – which lack conceptual validity – or transdiagnostic classes such as “eating disorders” – which collapse multiple syndromes into one large, heterogeneous category. Both types of target are problematic and lead to subsequent explanations that are inherently flawed. This point will be argued further in Chapter Four.

As well as being inappropriate in nature, current explanatory targets are also poorly defined. Descriptions of most psychopathological phenomena are “thin” and lack depth: many constructs are underspecified, and relevant factors, processes, or entire levels of analysis may be omitted. This poor description of explanatory targets means that there is

little information to guide intervention, classification, and aetiological explanation. The richer a description of a phenomenon is, the more content there is to provide clues about how it came to be and maintains its structure. For example, if the explanatory target were a cake, knowing that the cake is chocolate sponge strongly suggests that chocolate or cocoa was an ingredient used to make it, thereby narrowing down potential aetiological hypotheses and directing us to promising avenues for further investigation. It also gives us information about its structure, which helps to identify ways we could disrupt it (such as applying water or pressure) and how we might seek to classify it (for example, as a "sponge cake", "chocolate cake", or "baked cake"). Without this information, it is much harder to execute these tasks effectively, and much more likely that mistakes will be made.

In Chapter Four, I propose solutions for these problems. First, I propose that psychopathological explanation should shift focus from DSM syndromes and transdiagnostic classes to *clinical phenomena* (Ward & Clack, 2019). Clinical phenomena refer to specific patterns, experiences, or effects observed in or reported by individuals experiencing mental disorders. In some cases, they may be signs and symptoms as currently conceptualised, such as "anhedonia" or "binge eating". However, given the ambiguity and disagreement surrounding what exactly constitutes a "symptom" or a "sign" (see Wilshire et al., 2021),[14] these may in other cases be collections or "complexes" of these constructs that represent reliable patterns of behaviour, cognition, and emotion identified within psychopathology (such as the B-P cycle; Nielsen & Ward, 2020a). Such phenomena are more empirically valid and less heterogeneous than DSM syndromes. They comprise a more restricted number of features (e.g., binge eating is a less heterogeneous category than BN), the boundaries between them are less arbitrary (e.g., the symptoms "binge eating" and "self-starvation" are arguably more distinct from each other than are AN and BN), and transition in or out of them is more likely to reflect important psychological changes than, say, a transition from a diagnosis of AN to a diagnosis of BN. They therefore serve as more valid and reliable foundations for explanation than current alternatives.

Second, I argue that psychopathology research should devote more attention to the adequate description of psychopathological phenomena – in other words, *compositional explanation*. Compositional explanation refers to the comprehensive description of factors and processes involved in constituting a phenomenon rather than causing it. Compositional

[14] As such, throughout this thesis, I will use the terms "symptom", "sign", and "clinical phenomena" somewhat interchangeably, in accordance with the current nebulous definition of this space.

explanations have tremendous value for a variety of scientific and practice tasks in the field of psychopathology – such as aetiological explanation, classification, empirical research, assessment, treatment – and have the power to greatly aid our understanding of mental disorders. At present, few (if any) compositional explanations of psychopathological phenomena exist. I demonstrate this in Chapter Four by assessing whether current descriptive accounts of psychopathology – DSM/ICD syndromes, Symptom Network Models, transdiagnostic approaches, dimensional models, historical accounts, case narratives, the Research Domain Criteria (RDoC), existing "explanations" – meet criteria for a good compositional explanation. I then illustrate how effective compositional explanations of psychopathological phenomena *could* be constructed by grounding the explanation in data and constructing multiple interconnecting models of the phenomenon at different scales of analysis, using the ED symptom of "binge eating" as an example.

Issue 2: Endorsement of the One Best Model Perspective

Current psychopathological explanations typically subscribe to the notion that a single model should provide a complete explanation for a phenomenon, and therefore equate models with explanations (this argument is made in full in Chapter Five). Because individual models are perceived as complete explanations, there is seen to be no need for the contribution of additional models, and the relationship between alternative models of the phenomenon becomes, at best, detached and, at worst, antagonistic. The unspoken assumption is that (1) the model provided by the favoured explanation is correct or, in the unlikely event that it is not, (2) the correct explanation will be another single model. I term this philosophy the One Best Model (OBM) perspective.

There are several problems with the OBM. First, it is antithetical to the *ontology* of mental disorder.[15] Mental disorders are highly complex phenomena and therefore arguably impossible to explain with a single model. Such a model is likely to end up either extremely convoluted, and therefore unworkable, or impoverished, forgoing description of some factors and processes despite their relevance to the phenomenon of interest. As seen in the evaluation above, current ED explanations tend towards the latter, lacking explanatory depth. Second, the OBM approach contributes to *epistemic distance* between alternative models of mental disorder, in which models isolate themselves from one another at the

[15] *Ontology* refers to the branch of philosophy concerned with the nature of existence and reality (Smith, 2003).

theoretical level. This is a mistake, as it means there is little to no consideration of how different models could fit together, and this affects clinical practice down the line.

As a solution to the OBM problem, I propose that the psychopathology field adopt a *model pluralism* (Hochstein, 2017; Mitchell, 2002; Mitchell & Dietrich, 2006; Sullivan, 2017) approach to the explanation and treatment of mental disorder. This perspective advocates for the generation of a collection of models at different scales and levels of analysis that "hang together" to form an overarching explanation. Such an approach has the power to improve explanatory depth, as alternative models can be included that detail the processes, factors, or levels that other models might abstract away from. Indeed, if the explanations evaluated in this chapter were reframed as complementary models within an overarching explanatory network rather than rival explanations, our explanatory knowledge of EDs would arguably become much richer and more coherent.

Issue 3: Lack of A Guiding Methodological Approach to Explanation in Psychopathology

This issue is more meta-methodological in nature, relating to the process of building psychopathological explanations themselves. At present, there exists no appropriate methodological framework that can guide the generation of explanations of psychopathological problems. Methodological ambiguity like this lends itself to guesswork and, thus, a greater likelihood of error and oversight, which may result in epistemic deficits, such as those identified with the ED explanations above. Those involved in formulating such explanations are also not currently expected or required to have any training in explanatory theory or meta-theory, which might guide their reasoning in its place. Furthermore, because researchers are not expected to detail their reasoning process, there is no way for others to determine whether such error or bias may have occurred and critique or build on the theory in response. The validity of explanations thus becomes largely assumed, which is dangerous for the efficacy of the treatments built on them.

In Chapter Six, I argue that psychopathological explanation should not be methodologically exempt; that explanatory theories ought to be held to a similar standard of methodological rigor and transparency as empirical research. Systematic method should be used to guide the development of explanations of psychopathological problems, and published papers posing such explanations should, like empirical studies, be required to include method sections, in which the process of reasoning used to develop the explanatory account is explicitly laid out. A key barrier to this goal at present, however, is that no such

methodological approaches currently exist. Chapter Six therefore largely focuses on addressing this critical deficiency by applying two existing theoretical frameworks – Ward & Clack's (2019) Phenomena Detection Method (PDM) and Haig's (2014) Abductive Theory of Method (ATOM) – to the task of building psychopathological explanations. This gives rise to the Phenomena Detection Method for Theory Construction (PDM-TC) – a provisional methodological framework for the construction of explanatory accounts of psychopathological problems.

Summary

In this chapter, I critically evaluated four explanations for EDs that form the foundations of current best-practice ED treatment. Each was found to possess specific epistemic flaws, such as poor explanatory depth, low empirical adequacy, and a lack of internal coherence. As well as these specific issues, several broader problems were identified as being common to all the explanations, and to psychopathological explanations more generally: 1) the poor selection and description of explanatory targets, 2) the equation of models with explanations, and 3) the lack of a guiding methodological approach for constructing psychopathological explanations. In the chapters that follow, I address each of these issues in turn, proposing solutions for each.

In Chapter Four, I highlight the importance of choosing appropriate explanatory targets and describing their constitution comprehensively within compositional explanations. I then argue for a shift in focus within psychopathology explanations – from syndromes to symptoms – and illustrate how researchers could build effective compositional explanations of these phenomena, using the symptom of "binge eating" as an example.

Chapter Five addresses the problem of equating models with explanations in psychopathology – a problem labelled herein as the One Best Model (OBM) perspective. As a solution to this, I advocate for the widespread adoption of a *model pluralism* approach within psychopathology in which psychopathological phenomena are explained using a collection or network of compatible explanatory models. I then discuss how researchers and clinicians could begin to apply this philosophy to modelling mental disorder, translating models into interventions, and applying interventions in practice.

Finally, in Chapter Six, I discuss the importance of method to the construction of explanations in psychopathology. Then, due to the lack of existing methodological

approaches to psychopathological explanation, I construct and present a provisional methodological framework for this task: The Phenomena Detection Method of Theory Construction (PDM-TC).

Chapter 4. Selecting and Describing Explanatory Targets Effectively in Psychopathology

This chapter addresses the first issue identified by the analysis in the previous chapter: the poor selection and description of explanatory targets. There are two separate, but related problems encompassed by this: 1) that explanatory targets currently used by eating disorder (ED) explanations, and psychopathological explanations more generally, make for poor explanatory foci, and 2) that these targets, once selected, are poorly described. The first of these issues relates to the nature of the explanatory targets themselves – namely, what sorts of phenomena are being selected as the foundation for explanation within psychopathology and are they likely to yield productive explanatory results? The second relates to the description – or *compositional explanation* – of these targets – in other words, are the factors and processes that comprise these explanatory targets currently being sufficiently described within psychopathology?

These two issues are necessarily related, as the nature of an explanatory target affects how possible it is to construct a high-quality compositional explanation. Explanatory targets that are more heterogeneous will be more challenging to compositionally explain, as there are a huge number of phenomena that must be comprehensively described. Take the target "eating disorders" for example. This transdiagnostic category encompasses the entire range of phenomena observed in disordered eating presentations – a huge collection of diverse symptoms and signs. A good compositional explanation would be required to fully detail the constitution of all of these, which is a monumental task. However, it is not merely the size of the target that can be problematic. Targets with poor conceptual validity, such as DSM syndromes like Anorexia or Bulimia Nervosa (AN and BN), are also problematic, as the resultant compositional explanations will arguably describe the constitution of artificial phenomena ("pseudo-phenomena"). Such descriptions will therefore have limited utility to the broader field of interest, as they do not describe phenomena as they exist in the real world. Given this relationship between the selection and compositional explanation of explanatory targets, I address both within this chapter.

First, I discuss the selection of explanatory targets within psychopathology. At present, explanations of psychopathological problems, such as EDs, tend to focus on DSM syndromes – constructs that possess significant conceptual flaws and are now widely

agreed to be poor representations of psychopathological problems (Lilienfeld & Treadway, 2016; Nielsen & Ward, 2018, 2020a; Whooley, 2014; Zachar & Kendler, 2017). When this is not the case, the alternative focus is most often transdiagnostic categories, such as "internalising disorders", "anxiety disorders", or "eating disorders", which inadvertently make for explanatory targets that are far more heterogeneous and, thus, more complex to explain. In the first section of this chapter, I examine and critique the current use of these phenomena as explanatory targets in psychopathology. In response to the considerable problems identified with these targets, I then propose a shift in focus from *syndromes* to *symptoms*. In other words, I argue that psychopathology explanations should cease the use of DSM categories and transdiagnostic classes as explanatory targets and instead concentrate on *clinical phenomena*: clinically relevant signs and symptoms or appropriately scaled complexes of the two.

Second, I address the lack of adequate compositional explanation in the psychopathology discipline. At present, the phenomena that psychopathologists seek to understand are "thinly" described, most often consisting of sets of diagnostic criteria that invoke vaguely defined constructs, such as "body dissatisfaction". This lack of description has arguably contributed to our current poor understanding of psychopathological signs and symptoms (see Wilshire & Ward, 2021) as well as, to an extent, our poor aetiological explanations and clinical interventions. I elaborate on this argument in the second part of this chapter by critically evaluating the various descriptive approaches that have been proposed within psychopathology and could stake some claim of providing compositional accounts of psychopathological problems – namely, DSM syndromes, transdiagnostic categories and dimensions, Symptom Network Models (SNMs), historical accounts of mental disorders, clinical case narratives, the Research Domain Criteria (RDoC), and existing "explanations" of mental disorder (such as those evaluated in the previous chapter). Based on this evaluation, I argue that none of these approaches provide adequate compositional explanations of mental disorder. I respond to this by demonstrating how we could begin to build more sufficient compositional explanations of psychopathological problems; accounts that are grounded in empirical research and comprehensively detail the phenomenon at all levels of analysis. I do this by presenting a compositional "sketch" of the ED symptom "binge eating".

Explanatory Targets in Psychopathology

In this section, I discuss current explanatory targets in psychopathology. I begin by outlining the role that explanatory targets play in psychopathological explanation – namely, how they impact the structure and utility of the explanations produced. I then discuss current explanatory targets in psychopathology – typically, DSM syndromes and transdiagnostic categories. I argue that the conceptual and structural nature of these phenomena make them poor choices of target for explanation. In response to this, I introduce *clinical phenomena* as alternative explanatory targets and argue that explanations in psychopathology should shift their focus to these explananda. Specifically, I argue that a focus on these phenomena is likely to prove more valid and fruitful than the current orientation towards DSM syndromes.

Why Are Explanatory Targets Important?

As discussed in Chapter Two, an explanatory target – or *explanandum* – is the phenomenon that an explanation claims to explain. For example, the Psychosomatic Family" model, discussed in Chapter Three claims to explain the DSM syndrome of Anorexia Nervosa (AN) – this syndrome therefore serves as its explanatory target. The explanatory target plays a critical role in determining the nature of the explanation that is developed and, thereby, what type and amount of information it provides about the area of interest. Explanations represent attempts to solve a problem of sorts: they ask questions about phenomena – for example, "why do eating disorders occur?" – with the aim of identifying ways in which they might be predicted or controlled. The way that problem is defined therefore has substantial bearing on the nature of the explanation that will be developed – in other words, "stating the problem is half the solution" (Haig, 2014, p. 134).

Explanatory targets form the foundation of an explanation. Hence, the overall explanation will be fundamentally affected by the nature of its explanatory target. The appropriateness of an explanatory target will ultimately depend on its intended purpose. At the most basic level, if we are hoping to understand why individuals starve themselves, choosing the explanatory target "Bulimia Nervosa" (BN) is unlikely to be particularly helpful, as self-starvation is not central to this presentation, although it may occur at times. Similarly, if we hoped to understand why depression occurs, choosing to focus our explanation on the symptom of "insomnia" is likely to provide limited insight. Although often present in depressive presentations, insomnia is not central to the phenomenon of depression and frequently occurs independent of depressed mood. In contrast, explanatory

targets like the syndromes Major Depressive Disorder (MDD) and Dysthymia or symptoms characteristic of depression, such as “anhedonia” or “depressed mood”, are likely to produce more informative understandings.

Even when the explanatory target *is* appropriate to the area of interest, however, explanations can still be impacted by the epistemic properties of an explanatory target. The explanatory target selected has the power to significantly impact the explanatory value of the resulting explanation. For instance, the *scope* of an explanatory target can significantly affect the nature of the explanation generated to account for it. For example, the targets “eating disorders” and “binge eating” are notably different in scope: the first encompassing all ED presentations and their associated phenomena, the second only those presentations involving binge eating. Although both are able to provide information relevant to the ED field, an explanation focused on “eating disorders” will be very broad, as it must simultaneously account for the occurrence of all ED presentations, whereas an explanation centred around “binge eating” will be more focused and specific, accounting for a smaller and more homogenous subset of phenomena. Similarly, the *coherence* or *validity* of an explanatory target has significant bearing on an explanation. Explanatory targets with poor conceptual validity will produce explanations with similar epistemic failings, as they are arguably accounting for phenomena that do not genuinely exist as described. Explanations focused on DSM syndromes (see below) are prone to numerous epistemic flaws due to the dubious validity of these explanatory targets.

Explanatory targets selected are also important from a research and practice perspective. As discussed in Chapter Two, there is a symbiotic relationship between explanatory theory and empirical research, with each helping to direct what is studied by the other. Much empirical research is devoted to testing the hypotheses or predictions made by explanatory theories. In doing so, empirical studies typically adopt the explananda specified by these explanations. The Psychosomatic Family Theory critiqued in the previous chapter has subsequently been tested in a variety of empirical studies. As per the explanation, most of these studies focused on individuals meeting the AN diagnostic criteria and their families (e.g., Eisler, 2005; Garfinkel et al., 2009; Rowa et al., 2001; Wewetzer et al., 1996). The topics of empirical study also frequently inform the structure of explanatory theory. Hence, the use of the same explanatory targets can easily be perpetuated, resulting in a mutually reinforcing cycle such as that observed with the current entrenchment of DSM syndromes in psychopathology (e.g., Hyman, 2011; Tsou, 2016). Similarly, as research and theory form the

foundation of therapeutic treatment under the current *evidence-based practice* model (see Howick, 2011), their explanatory targets are also likely to filter through to clinical practice, potentially prescribing the focus of assessment and treatment (for instance, whether symptoms and signs or syndromes or "transdiagnostic" categories). The seemingly insignificant act of choosing a phenomenon to explain may therefore have broader systemic effects within the psychopathology discipline.

Current Explanatory Targets in Psychopathology

Current psychopathological explanations are typically built around DSM syndromes, such as AN, BN, or MDD. These categories are widely recognised as having extensive flaws (Lilienfeld & Treadway, 2016; Nielsen & Ward, 2018, 2020a; Whooley, 2014; Zachar & Kendler, 2017) and are highly resistant to changes recommended by new research – for example, routinely failing to adopt the recommendations of DSM working groups (Whooley, 2014). The DSM-5 ED syndromes, for example, lack conceptual validity due to the high frequency of diagnostic migration over time (Agras et al., 2009; Eddy et al., 2008; Fichter & Quadflieg, 2007; Forbush et al., 2018; Tozzi et al., 2005) and the similarities that exist between their clinical presentations – such as overlaps in diagnostic criteria and shared comorbidities (APA, 2013) – and associated empirical findings – such as shared genetic predisposition (Strober et al., 2000), neural network correlates (Bailer & Kaye, 2010; Broft et al., 2011), and personality variables (Culbert et al., 2015; Farstad et al., 2016). There is also evidence that EDs, both within specific diagnoses and generally, may serve a variety of different functions for the individual – such as avoidance, security/control, achieving care or communicating with others – that are likely to be critical to understand in treatment, yet not all of which are present in every case (Gagnon-Girouard et al., 2019; Nordbø et al., 2006). Furthermore, the fact that the Other Specified Feeding and Eating Disorder (OSFED) diagnosis – intended as a "catch-all" category for presentations not encompassed by the main diagnoses – remains the most common ED diagnosis, accounting for around 50% of cases (Keel et al., 2011; Machado et al., 2013) strongly suggests that these classifications are poor representations of disordered eating problems.

As theoretical constructs, DSM syndromes lack validity and reliability. The problem with their use as explanatory targets is therefore twofold. First, their conceptual deficits set researchers up to construct explanations with similar epistemic problems – an explanation is only as valid as its target. Hence, explanations based on DSM syndromes have poor

explanatory worth, as they are arguably providing accounts of pseudo-phenomena rather than psychopathological problems as they genuinely exist in the real world. Their clinical utility is therefore limited, as most client's difficulties will be insufficiently explained by these accounts – this is likely partly to blame for the low efficacy of current best-practice treatments. Second, the use of DSM syndromes as the focus of explanation lends these constructs undeserved conceptual support, by implying to academics and clinicians that they are indeed valid constructs. This practice contributes to the ongoing *reification* of these syndromes, in which these abstract theoretical constructs become falsely regarded as real material entities (Hyman, 2010)[16]. This further entrenches the DSM categories in research and practice and thereby hinders efforts to replace or modify them.

Some researchers have attempted to circumvent the problems associated with DSM syndromes by adopting a "transdiagnostic" approach that emphasises the similarities across diagnostic classes like "eating disorders" or "anxiety disorders", arguing that these should represent a single category. Although commendable in its intention to disavow the DSM taxonomy, this approach has its own problems. Despite trying to increase validity by eliminating the DSM's "fuzzy" boundaries, transdiagnostic categories actually make for *less* reliable targets, as they combine multiple syndromes into one broad and extremely heterogeneous category. Collapsing across different diagnostic presentations also problematically ignores the evidence that meaningful differences exist between these categories. For example, the DSM ED diagnoses, while entailing many taxonomic flaws, have been shown to differ significantly in age of onset, course, outcome, and level of impairment, as well as their associated psychological features and medical problems (Birmingham et al., 2009; Keel, 2018; Mehler, 2018; Swanson et al., 2011). All of this represents explanatorily relevant information, which transdiagnostic categories completely disregard.

Current explanations of mental disorder that evidence appropriate explanatory targets are those that focus in on a single symptom or feature of psychopathology, such as "binge eating" (for example, Self-awareness Theory, Heatherton & Baumeister, 1991; Dieting Theory, Polivy & Herman, 1985). This makes for a more valid and reliable explanatory target, as many of the problems associated with DSM syndromes are reduced if not eliminated by this reduction in scope – construct boundaries become less blurred, and

[16] *Reification* refers to the process of considering or representing an abstract concept or idea as a material or concrete entity (Hyman, 2010).

heterogeneity is reduced. For example, "binge eating" is more distinct from "restrictive eating" than AN is from BN and encompasses far fewer phenomena than either syndrome. The existence of these "feature-focused" explanations within psychopathology demonstrates that it is possible to construct explanations around alternate explanatory targets that are reduced in scope: psychopathological explanations need not revolve solely around DSM syndromes.[17]

Moving Forward: From Syndromes to Symptoms

Several theorists argue that, to move forward, psychological explanation should, at least for now, focus on symptoms and signs (*clinical phenomena*) rather than syndromes (Berrios, 2013; Borsboom, 2017; Costello, 1992; Fried & Nesse, 2015; Mojtabai & Rieder, 1998; Ward & Clack, 2019; Wilshire et al., 2021). For instance, instead of trying to describe and explain the syndrome of BN, which comprises a cluster of diverse symptoms, one might instead focus on a single symptom of ED pathology, such as "binge eating". At a conceptual level, this approach makes sense, as symptoms and signs have greater validity than diagnostic syndromes, arguably representing genuine phenomena as opposed to artificial categories. Compared to diagnostic syndromes, symptoms have more defined boundaries, less heterogeneity, and greater stability over time. For example, the symptom of "binge eating" is more clearly different from other symptoms like "self-starvation" or "purging" than the diagnostic category of BN is from other ED diagnoses such as AN. Furthermore, a client shifting away from this symptom to either alternative ED symptoms or a state of recovery is likely to be much more psychologically meaningful than a transition from one ED diagnosis to another, which can be accomplished by changes in less theoretically meaningful factors, such as BMI (see above). At a practical level, it is also useful to reduce the scope of our explanatory focus – it is much easier to recount in detail the composition of a single symptom or clinical phenomenon than a large and diverse collection of them (a *syndrome*).

At an ethical level, symptoms and signs also make for an appropriate focus of investigation, as they represent the actual concerns of clients – each is a valid and important aspect of the client's difficulties that clinicians and academics should aim to understand. At the coarser grain level of syndromes, although the accounts generated are concise and practical, they are likely to neglect the description and explanation of some symptoms in

[17] Although it should be noted that current "feature-focused" explanations typically suffer from other conceptual issues – namely, shallow description of the target and the explanatory processes proposed to account for it (i.e., poor explanatory depth).

favour of providing a brief and uncomplicated account of that syndrome. At finer grain levels, such as the neurobiological, although useful and detailed information may be generated that can be used to inform theoretical conceptualisation of clinical phenomena, we are no longer centring our theoretical accounts on the problems experienced and reported by clients – which arguably should be our paramount concern as clinicians – and risk decontextualising clients' mental distress (Whooley, 2014). Although to fully recount the nature of a symptom we no doubt need to investigate and describe phenomena and processes at smaller scales, such as the neurobiological or molecular, and consider their relationship to other symptoms (as do syndrome clusters and symptom networks), I argue that the appropriate starting point for the compositional explanation of disorder states should, at least for now, be psychopathological symptoms.

Compositional Explanations in Psychopathology

In this section, I discuss the lack of compositional explanations within the psychopathology discipline. First, I outline what role compositional explanations have in theory, research, and practice in psychopathology. Second, I discuss the current overwhelming lack of compositional explanations for psychopathological problems. I argue this point by evaluating a variety of accounts that appear to describe the composition of mental disorders to some degree – namely, DSM/ICD syndromes, transdiagnostic/dimensional approaches, Symptom Network Models (SNMs), historical accounts, case narratives, the Research Domain Criteria (RDoC), and existing "explanations" of mental disorder – and argue that none provide effective compositional explanations of psychopathological problems. Finally, I discuss how we might go about building effective compositional explanations of mental disorder, by grounding our descriptions in relevant data, modelling phenomena at multiple levels of analysis, and richly detailing the factors and processes involved.

What Is the Role of Compositional Explanation in Theory, Research, and Practice?

As outlined in Chapter Two, a *compositional explanation* describes how a given phenomenon is constituted or "made up" – it provides an account of the factors and processes that comprise the phenomenon rather than cause it (Craver, 2007; Gillett, 2021; Kaiser, 2015). However, compositional explanations may contain causal relations if those interactions play a role in constituting the phenomenon – as in the bidirectional relationship between insomnia and low mood in clinical depression (Konjarski et al., 2018; see Chapter

Two). Sometimes this may include *maintenance cycles*, such as the "binge-purge cycle" in BN, which arguably do not cause the disorder itself but serve to maintain its structure. Effective compositional explanations provide detailed information about how a particular phenomenon is constituted – that is, they comprehensively describe the factors and processes comprising it and how they interact. The criteria for a "good" compositional explanation are therefore the same as those outlined in Chapter Two for explanations in general. A "good" compositional explanation will be one that provides a logical, elegant, and sufficiently detailed account of a phenomenon and its related empirical findings that incorporates relevant information at all levels of analysis.

Compositional explanation plays a vital role in all aspects of clinical inquiry – theory, research, and practice – but are especially valuable from a theoretical perspective. Compositional explanations hold significant value for the development of aetiological explanations: the more you know about a phenomenon, the easier it is to hypothesise about its causes. The more detailed our compositional understanding of a phenomenon, the more "clues" we have about the causal processes involved in its aetiology. For example, if we were trying to explain how a cake came to exist and had no other information about it other than it being a cake, we could reason that its aetiology probably involved certain components and processes known to be common to most cakes, such as flour, sugar, or being baked. However, with further detail about how it is composed – whether it is chocolate sponge, multi-layered, or cream-filled – there is additional information that can help us refine our aetiological reasoning, strengthening existing ideas about the involvement of some factors and processes, like flour and being baked, and suggesting novel ones, such as whipping cream or cocoa.

When compositional understandings are "thin" (less detailed) it can promote errors in causal reasoning. Relevant causal factors and processes may be neglected or deliberately omitted, and others that are irrelevant or flawed may be included. For example, knowing only that the phenomenon is a cake, multiple aetiological theories could be developed hypothesising the occurrence of various processes one might expect to be involved based on prototypical background knowledge of cakes. However, all the while, the phenomenon in question was actually an ice cream cake – a type of cake indeed, but one involving none of the causal processes theorised. Possessing detailed characterisations of the phenomena to be explained is thus highly valuable for aetiological understanding and should not be overlooked.

Compositional explanations also hold value for clinical research and practice. Their primary value in research is for classification. Although compositional explanations describe disorder phenomena, they are *not* taxonomies: they do not necessarily claim to provide the best way of classifying the psychopathological phenomena with which they are concerned. Sometimes they may be constructed with a specific classificatory approach in mind. However, they are generally agnostic about how the phenomenon should be taxonomically organised. Nevertheless, because of their informational value, compositional explanations are highly useful for those aiming to develop taxonomies of mental disorder (Wilshire et al., 2021). Their rich description of the constitution of a presentation or specific symptom may help to signal connections between certain disorder states or psychopathological phenomena, and thereby suggest novel ways of considering and grouping them. Compositional explanations chart the psychopathological landscape in rigorous detail, which allows it to be effectively surveyed by those who wish to classify it.

In clinical practice, compositional explanations hold value for both assessment and treatment. For a start, the informational value that compositional explanation holds for aetiological explanation and classification has positive flow-on effects for assessment and treatment. For example, improved classification of mental disorder is likely to lead to more accurate and meaningful diagnoses in clinical practice. Likewise, better aetiological explanations for psychopathological problems are likely to improve the quality of prevention and treatment initiatives.

As well as these flow-on effects however, compositional explanations hold independent value for clinical practice. First, compositional explanations provide a wealth of information about the structure of psychopathological problems, and thus how they may present in practice. This makes them highly useful during the clinical assessment process, as clinicians have more information about relevant features of the condition that might signal its presence in a client or indicate that they are at risk for developing it. Second, and perhaps most critically, as compositional explanations typically describe maintenance processes, they are extremely useful for initial intervention. One of the first steps of intervention is disrupting the factors and processes considered to be perpetuating the client's problems, such as reinforcement processes and maintaining cycles. For example, treatment of EDs typically starts by addressing the processes perceived to be sustaining the individual's disordered eating behaviours – such as "starvation syndrome" or the "binge-purge cycle" – and then focuses on the processes that led to their initial development once eating behaviour has been

stabilised (e.g., FBT, CBT-E; Le Grange & Rienecke, 2018; Fairburn, 2008). Indeed, for psychopathological problems currently considered chronic, such as treatment resistant cases, chronic pain conditions, schizophrenia, or bipolar disorder – which are typically *managed* rather than treated – the disruption or reduction of constitutional processes is absolutely key to treatment in order to prevent symptoms from being exacerbated and the client's condition from deteriorating. Compositional explanations aim to describe these constitutional processes in a much higher level of detail and at multiple levels of analysis. They can therefore be an incredibly rich platform for the development of primary intervention strategies or management protocols.

Finally, compositional explanations may help to identify and pre-empt possible issues in the therapeutic process. The rich descriptions of phenomena that they provide offer clinicians a large amount of information about potential factors and processes affecting their client's mood, cognition, and behaviour. Some of these may have relevance to decisions about how therapy proceeds. For example, knowing that a particular disorder state tends to include cognitive flexibility or poor attentional control might cause a clinician to change how they present information to their client in session and/or contraindicate the use of interventions that require high levels of cognitive effort, at least initially. Likewise, understanding that a particular symptom, such as alexithymia, typically involves difficulty sensing physical sensations (*impaired interoception*) might influence how a clinician introduces sensorily-based exercises, such as mindfulness, to a client – for example, by devoting extra time to helping the client identify physical sensations.

The Lack of Compositional Explanations in Psychopathology

At present, there is a significant lack of attention to the task of compositional explanation within psychopathology. The comprehensive description of psychopathological states is not broadly recognised as being of practical or theoretical importance. Compositional explanation is not routinely recognised as part of the scientific inquiry process, such that most researchers are unaware of this task or the value that it holds for research and practice. As such, sufficient compositional accounts of psychopathological problems – or even attempts to provide them – are few and far between. However, as discussed above, there are substantial advantages to possessing comprehensive understandings of how psychopathological problems are structured and maintained. Compositional explanations have great value for developing theoretical understandings and taxonomies of mental disorder, conducting clinical

assessments, and identifying primary intervention targets. Their lack within psychopathology is therefore a significant oversight.

Some might argue that, although accounts are not currently labelled as "compositional explanations", such descriptions *do* currently exist within psychopathology. However, although there are various descriptive accounts of psychopathological problems, in my view none of the options equate to a sufficient compositional explanation. In this section, I will argue this point by evaluating the current approaches available that could be considered as "candidates" for compositional explanations within psychopathology. These are 1) classificatory approaches (e.g., DSM, ICD, transdiagnostic and dimensional approaches, symptom network models), 2) historical accounts of emerging syndromes, and 3) clinical case narratives, 4) the RDoC framework, and 5) existing "explanations" of psychopathological problems. I now critically evaluate these approaches relative to the criteria for a "good" compositional explanation, arguing why each is unable to fulfil this role within the scientific inquiry process.

Classificatory Approaches.

Classificatory approaches herein refer to those approaches within psychopathology that propose a particular taxonomy of mental disorder. Perhaps the most well-known examples within psychology and psychiatry are the DSM (APA, 2013) and ICD (WHO, 2018). However, some researchers have taken issue with the DSM/ICD's categorical approach, and have developed alternative taxonomic philosophies, such as the transdiagnostic and dimensional approaches to mental disorder classification (e.g., Carragher et al., 2015; Caspi & Moffitt, 2018; Kotov et al., 2018; Krueger & Eaton, 2015; Wright & Simms, 2015), as well as the Symptom Network Theory (Borsboom, 2017; Fried et al., 2017). These approaches are evaluated here because, although they each propose a specific taxonomic approach, they each provide descriptions of psychopathological problems, whether as a set of defined criteria, a position on one or more spectra, or a network of causally related symptoms. However, in my view, none of these approaches generate accounts that provide effective compositional explanations of psychopathological phenomena.

Diagnostic Syndromes: DSM-5 and ICD-11.

The DSM (APA, 2013) and ICD (WHO, 2018) are perhaps the most prominent attempts to conceptualise and describe mental disorders. Both group mental disorders into discrete syndromes (collections of symptoms and signs) comprised of a set of diagnostic

criteria, such as Borderline Personality Disorder (BPD) or AN. These syndromes are frequently used in place of compositional explanations in research and theory – both aetiological explanations and empirical studies often revolve around DSM categories. For example, the Psychosomatic Family Theory critiqued in Chapter Three has a DSM-5 ED category as the explanatory focus – namely, AN. Despite their use as such, however, DSM/ICD syndromes fail to provide effective compositional explanations in several important ways.

First, DSM/ICD syndromes lack *validity*. As opposed to identifying genuine divisions that exist within the psychopathology space ("carving nature at its joints"; see Slater & Borghini, 2013), DSM/ICD syndromes impose categories on mental disorder that poorly represent the nature of these phenomena. DSM/ICD syndromes have been consistently criticised in recent years for their significant conceptual flaws, such as symptomatic heterogeneity, diagnostic migration, and artificial comorbidity (see Lilienfeld & Treadway, 2016; Nielsen & Ward, 2018, 2020a; Whooley, 2014). Such problems do not necessarily make these diagnostic categories worthless, as they have some use in clinical contexts, such as communication between health professionals (e.g., *practical kinds*; see Kendler et al., 2011). However, they are arguably not accurate representations of the psychopathological landscape. This challenges their ability to serve as compositional explanations of mental disorder, in that they fail to describe these conditions as they actually exist in the world. They therefore arguably provide accounts of pseudo-phenomena rather than genuine psychopathological problems.

Second, DSM/ICD syndromes lack sufficient explanatory scope. Each is characterised by a relatively small number of descriptively "thin" criteria spanning but a few levels of analysis (e.g., behavioural, cognitive, emotional). For example, the criteria for AN (see Table 1) describe relatively few features of the condition, despite the fact that empirical research has identified many other features common to this type of ED presentation, such as alexithymia (Nowakowski et al., 2013; Westwood et al., 2017), impaired interoception (Stinson, 2019), cognitive deficits (Hedges et al., 2019), and autistic traits (Westwood et al., 2016). Furthermore, both the DSM and ICD taxonomies tend to outline only those features that are most salient at a clinical level – that is, that are most readily observable in practice or considered most central to the disorder's pathology. Although appropriate, and often useful, in a practice context, as compositional explanations these accounts omit relevant features that may be less clinically salient – those that are harder to identify (e.g., emotional

comprehension, executive functioning, interoceptive ability) or with less perceived relevance (e.g., attentional biases, *central coherence*)[18] – but nonetheless characteristic of the disorder state.

Table 1. *DSM Criteria for Anorexia and Bulimia Nervosa*

DSM Criteria: *Anorexia Nervosa*
A. Restriction of energy intake relative to requirements, leading to a significantly low body weight in the context of age, sex, developmental trajectory, and physical health. Significantly low weight is defined as a weight that is less than minimally normal or, for children and adolescents, less than that minimally expected.
B. Intense fear of gaining weight or of becoming fat, or persistent behaviour that interferes with weight gain, even though at a significantly low weight.
C. Disturbance in the way in which one's body weight or shape is experienced, undue influence of body weight or shape on self-evaluation, or persistent lack of recognition of the seriousness of the current low body weight.
Specifiers:
Restricting type: During the last 3 months, the individual has not engaged in recurrent episodes of binge eating or purging behaviour (i.e., self-induced vomiting or the misuse of laxatives, diuretics, or enemas). This subtype describes presentations in which weight loss is accomplished primarily through dieting, fasting, and/or excessive exercise.
Binge-eating/purging type: During the last 3 months, the individual has engaged in recurrent episodes of binge eating or purging behaviour (i.e., self-induced vomiting or the misuse of laxatives, diuretics, or enemas).
DSM Criteria: *Bulimia Nervosa*
A. Recurrent episodes of binge eating. An episode of binge eating is characterized by both of the following:
1. Eating, in a discrete period of time (e.g., within any 2-hour period), an amount of food that is definitely larger than what most individuals would eat in a similar period of time under similar circumstances.
2. A sense of lack of control overeating during the episode (e.g., a feeling that one cannot stop eating or control what or how much one is eating).
B. Recurrent inappropriate compensatory behaviours in order to prevent weight gain, such as self-induced vomiting; misuse of laxatives, diuretics, or other medications; fasting; or excessive exercise.

[18] *Central coherence* refers to an individual's ability to derive overall or "global" meaning from a set of stimuli – i.e., the "big picture" (Frith, 2003).

C. The binge eating and inappropriate compensatory behaviours both occur, on average, at least once a week for 3 months.

D. Self-evaluation is unduly influenced by body shape and weight.

E. E. The disturbance does not occur exclusively during episodes of anorexia nervosa.

Finally, DSM/ICD syndromes lack sufficient explanatory depth. The features or symptom criteria listed as comprising these syndromes are typically described in a low level of detail. For instance, "disturbance in the way in which one's body weight or shape is experienced" (APA, 2013, p. 339) is listed as a necessary criterion for AN (see Table 1) but there is no detail about the exact nature of this "disturbance". Is it a distortion in sensory perception or cognitive evaluation (Mölbert et al., 2017)? Does it encompass the body in general or does it tend to be focused on specific areas (Cash & Deagle, 1997)? Body image is recognised by researchers to be a "multi-faceted construct consisting of a variety of measured dimensions" (Thompson, 2004, p. 8), including perceptual, conceptual, and emotional (Stinson, 2019). There are therefore many factors and processes involved in constituting body image, and an effective compositional explanation of the AN syndrome would need to provide a more comprehensive account of the exact nature of its "disturbance" in individuals presenting with this problem. Most other features included in the DSM/ICD criteria for EDs are similarly ill-defined: there is insufficient information provided about exactly how these constructs are structured or organised within the phenomenon. Hence, DSM/ICD syndromes lack the requisite detail that a successful compositional explanation would be expected to provide.

Transdiagnostic Approaches and Dimensional Approaches.

Transdiagnostic approaches refer to classificatory perspectives that advocate dispensing with the existing problematic diagnostic syndromes in favour of broader classifications based on shared characteristics. In some cases, this involves collapsing diagnostic syndromes into a broader disorder category like "anxiety disorders" or "eating disorders", in others using a different common factor as the means of classification, such as in the "internalising/externalising" model *(*Achenbach et al., 2016; Krueger & Eaton, 2015).[19]

[19] The internalizing/externalizing model classes psychopathological problems as either "internalising" or "externalising" depending on whether the problematic behaviours, cognitions, and emotions are directed inwardly or outwardly. Examples of "internalising" problems are anxiety and depression; examples of externalising problems are aggression and impulsivity (Achenbach et al., 2016; Krueger & Eaton, 2015).

Dimensional approaches refer to classificatory approaches based around spectra or scales rather than discrete categories, such as the Five Factor Model (FFM) of personality disorders (Trull & Widiger, 2013; Widiger & Costa, 2013). Some approaches present both a dimensional *and* transdiagnostic perspective on mental disorder, such as the Hierarchical Taxonomy of Psychopathology (HiTOP) model, which proposes a hierarchical organisation of mental disorder, consisting of transdiagnostic spectra at the top (e.g., general psychopathology, internalising/externalising) and syndromal subfactors (e.g., eating problems), shared symptoms/signs, and traits at progressively lower levels (see Kotov et al., 2017; Kotov et al., 2018).

Although transdiagnostic and dimensional perspectives on mental disorder may provide useful ways of conceptualising and organising psychopathological problems from a *nosological* perspective,[20] they do not necessarily describe psychopathological phenomena any more fully than do DSM/ICD syndromes. In some cases, they may even provide weaker descriptions. For example, the “internalising/externalising” model, although highlighting links between diagnostic categories and to some degree overcoming the arbitrary boundaries between them, provides even less information about the psychopathological phenomena it classifies. Identifying a problem as an "internalising disorder”, although useful for some purposes, gives very little information about the precise nature of that problem (e.g., whether it involves anxiety, dysphoria, or social withdrawal) or the minutiae of its presentation – the factors and mechanisms that comprise the problem.

The HiTOP model provides somewhat more information thanks to its hierarchical structure, which conceptualises psychopathological problems at both more general (broad spectra) and more specific levels (symptoms, signs, and traits). However, even then, this model lacks the richness of information that a compositional explanation should aim to achieve. Symptoms, signs, and traits are not broken down into their constituent factors or processes, nor are any relevant relationships between them modelled. Furthermore, although specific problems are conceptualised at coarser, transdiagnostic grain sizes, they are not described contextually at higher levels of analysis, such as the sociocultural, interpersonal, or political – layers of meaning that are arguably necessary to fully comprehend the makeup of any psychopathological problem.

[20] *Nosology* typically refers to the branch of medicine concerned with the classification of diseases. However, more broadly, it encompasses “the underlying theory about the nature of the conditions that are being classified and the principles and rules of the classifying process.” (Jablensky, 2012, p. 77).

From the perspective of compositional explanation, transdiagnostic and dimensional perspectives represent a lateral move from DSM/ICD syndromes. These perspectives do not elaborate much further on the constitution of the phenomenon but rather change how it is being conceived of nosologically – trading categories for dimensions or moving from a finer to coarser grain size or vice versa. Sometimes these changes provide greater detail about a disorder's presentation, sometimes less; sometimes they produce categories with greater conceptual validity, sometimes not. However, never do they detail disorder phenomena with sufficient depth at all the relevant levels of analysis – this is not their intended purpose. Hence, although important and useful to conversations about the classification of mental disorder, they are no more effective at providing compositional explanations than DSM/ICD syndromes.

Symptom Network Models.

The Network Theory of mental disorder proposes that psychopathological symptoms should be conceptualised as directly causing each other (for example, persecutory delusions resulting in paranoia, which subsequently leads to social withdrawal) rather than as caused by an underlying disease process (for example, delusions, paranoia, and social withdrawal as all arising from a common cause, such as a particular neurobiological dysfunction or genetic mutation; Borsboom, 2017). Symptom Network Models (SNMs) represent the application of this theory to specific syndrome clusters – often (but not always) DSM/ICD syndromes. A network structure is generated by depicting the direct causal links that exist between symptoms of that condition, including their strength and direction (Borsboom, 2017). SNMs can also model the relationships between symptoms across disorders – for example, between eating disorders and depression or anxiety (Smith K. E. et al., 2018) – which makes them particularly useful in accounting for comorbidity (Fried et al., 2017). However, although SNMs provide a useful and interesting description of the relationships between psychopathological symptoms, they fail to provide complete compositional explanations of mental disorders.

SNMs often retain DSM/ICD syndromes as their basis, and simply elaborate on the symptom-symptom relationships within these (e.g., Forrest et al., 2018; Gay et al., 2020; Smith T. E. et al., 2017; Solmi et al., 2019; Weintraub et al., 2020). Although this provides a richer description of the presentation of these syndromes and how their symptoms may be

connected, from a conceptual point of view they are subject to the same validity problems as DSM/ICD syndromes as they retain the problems inherent in these categories (see above).

More importantly, however, SNMs lack the depth and scope necessary to serve as compositional explanations of mental disorders. Although the relationships between symptoms are somewhat depicted within these models – namely, their strength and direction – neither the mechanisms responsible for these connections nor the nature of the symptoms themselves are fully explained. Each symptom is represented largely at the phenomenological level, with brief labels like "dysphoria", rather than described at each level of analysis. The meaning and structure of these constructs are therefore not fully unpacked. There are many levels at which any psychopathological symptom can be described. For example, anhedonia, a key symptom of depression, can be represented at the phenomenological level as involving both decreased "liking" and decreased "wanting", at the cognitive level as a reduced hedonic capacity, reduced reward motivation, and errors in reward learning, at the neural level as dysfunction in the "hedonic network" and mesolimbic pathways, and at the molecular level as reductions in opioid and dopaminergic activity (see Clack & Ward, 2020). Compared to a full analysis like this, SNM accounts of symptoms are significantly underpowered.

It is worth noting that SNMs could usefully serve as *part* of compositional explanations. They draw attention to symptom relations as an important part of mental disorders and provide a novel and useful way of conceptualising these. They also typically include more symptoms than DSM/ICD syndromes, thereby enriching the clinical picture and providing valuable insight into the presentation of psychopathological problems. However, on their own, they are simply too one-dimensional to serve as compositional explanations.

Historical Accounts.

Historical accounts refer here to descriptions of disorder states that either accompanied the first identification of a psychiatric syndrome or were developed around the time that the DSM-III (APA, 1980) was published, which represented a paradigm shift towards the conceptualisation of mental disorders used by psychiatry today (Mayes & Horwitz, 2005). Examples include Russell's (1979) initial characterisation of BN and Bruch's (1973, 1982) early descriptive accounts of AN,[21] considered to be the first modern description of the disorder (Marks, 2019). These sorts of accounts typically consist of a set of

[21] The most well-known of these being her book *The Golden Cage: The Enigma of Anorexia Nervosa* (Bruch, 1982), first published in 1978.

clinical case studies from which the author then draws broader conclusions. For example, Russell's (1979) initial characterisation of BN involved thirty patients, three of whom were presented as illustrative case studies, from which he drew conclusions about the disorder's typical features, such as demographics, symptomology, medical complications, and psychopathological correlates.

Historical accounts are often more descriptively comprehensive than the classificatory approaches above. However, they are nevertheless unsuitable as compositional explanations of mental disorder. Most problematic is that their explanatory scope extends beyond that of the disorder state in question. Although they do typically describe the unwell state of the disorder – as would a compositional explanation – they often also aim to represent the broader clinical picture, including premorbid features and correlates. For example, Bruch's (1982) account of AN makes claims about the aetiology of the disorder as well as describing its clinical presentation – for example, "the child's inability for constructive self-assertion and the associated deficits in personality development are the outcome of interactional patterns that began early in life" (p. 37) – with several chapters highlighting precipitating factors and speculating on the causal role of family dynamics (chapters *The Perfect Childhood* and *How It Starts*). Although this information is relevant and useful in a clinical context, providing clinicians with additional features to assess in eating disordered clients that can inform their formulation, theoretically it creates confusion.

First, it conflates the theoretical tasks of compositional explanation – describing a phenomenon's constitution – and aetiological explanation – hypothesising the potential causal factors and processes that led to it. Although these tasks are related, each informing the execution of the other, they are conceptually distinct, requiring different modes of theoretical reasoning: aetiological versus compositional (see Chapter Two). Attempting to achieve both within the same account is likely to create convolution and promote logical errors, thus impairing the integrity of each task.[22]

Second, even if such an account were presenting an impartial description of premorbid constitution and not implying causal relationships, this would problematically combine the compositional explanation of two different explanatory targets: the premorbid and unwell states. These are conceptually distinct phenomena, both in nature – one a state of illness dysfunction, the other a state of wellness or comparative functionality – and in time –

[22] As Confucius said, "The man who chases two rabbits, catches neither".

each unable to occur alongside the other. Although the premorbid state is certainly relevant to comprehending a disorder – for example, able to help suggest aetiological pathways – detailing its composition is a separate task. Historical accounts typically describe both within the same explanation, which may be highly useful where the purpose of the account is to guide clinical assessment. However, in doing so, they impair the richness of the description of each: the descriptive workload is essentially doubled (from the description of one phenomenon to two) and detail will likely be sacrificed for the sake of simplicity.

Historical accounts also tend to lack an empirical foundation. Being either the first characterisation of a disorder or one of its earliest formal descriptions, these accounts had little to no empirical data available at the time to inform their construction. As such, they are typically based on a small number of clinical or scientific case studies that, although potentially the best option available at the time of their development, fall short of the standards expected by modern scientific research and evidence-based practice (see Howick, 2011). As well as possessing unacceptably low sample sizes, these approaches often suffer from extensive sampling bias. In most cases, the sampling pool is typically confined to the authoring clinician's own practice, making the participants more likely to be culturally homogenous and of higher than average socioeconomic status due to the population's reduced access to mental health services at that time (Bell, 1985; Mechanic, 2007). Furthermore, cases are sometimes aggregated in pseudoscientific or anecdotal ways to illustrate the author's points, rather than analysed via scientific method, making them less likely to be reliable.

There has generally been little structured effort to update or expand such accounts in line with contemporary research. Hence, although there may be some aspects of historical accounts that have since been empirically verified, they may also make claims for which there is insufficient evidence, or that research now actively refutes. For example, Bruch's (1982) account of AN describes the disorder as affecting "the daughters of well-to-do, educated, and successful families, not only in the United States but in many other affluent countries" (p. vii), implying that AN exclusively or predominantly affects wealthy individuals in first-world, westernised countries (in other words, Western Educated Industrialised Rich Democratic (WEIRD) populations; Henrich et al., 2010). However, research now indicates that AN is not a culture-bound syndrome but is instead observed cross-culturally (Keel & Klump, 2003; Pike et al., 2014), and shows no reliable association with either ethnicity or socioeconomic status (Hadassah Cheng et al., 2018; Marques et al.,

2011; Schaumberg et al., 2017). Other historical disorder descriptions are likely to include similarly unfounded claims, and thus misrepresent the disorder they claim to describe.

Finally, as with DSM/ICD syndromes, historical accounts tend to refer only to the most salient features of a disorder. This neglects the description of features at other levels that are less easily observed, though none the less relevant to the disorder's presentation, such as neural network dysfunction, alterations in hormone or neurotransmitter systems, or attentional biases. This omission is understandable given that these accounts are largely based on client observations and some of the technology or methods required to detect these deeper features may have not yet been available. However, the lack of comprehensive description of all the relevant levels comprising the phenomenon means that, within the context of compositional explanation, historical accounts are notably impoverished.

Clinical Case Narratives.

Clinical case narratives are often provided to students and clinicians in textbooks or treatment manuals to provide an idea of how a certain disorder typically presents. This is generally done by providing a description of a prototypical case, which usually includes information about the characteristic symptoms and signs, demographics, relevant history, and triggers for that disorder. Take, for example, this extract from *Adult Psychopathology & Diagnosis* (7th Edition):

> Wendy was a competitive runner with hopes of qualifying for the Olympics. Two years ago, while on a training run on trails, she was attacked by a man who tried to rape her. She screamed and fought and managed to get away by running out of the forest as fast as she could, but she continued to have flashbacks to this attack every day when she was training on the streets or trails. Prior to the attack, all her attention was on training. She was a healthy eater, focused on performance rather than appearance, and adhered strictly to the recommendations of her trainer and sports dietitian. After the attack, she became increasingly anxious and had difficulty keeping up her training schedule. Her race times were increasing, as was her weight. She lost two races to one of her main competitors, and when she saw a picture of herself coming in second in the paper, she became fixated on her weight. She was up to 125 pounds, which was higher than she had ever been before. She started cutting back on calories, which she found easy to do. Wendy stated, "I was training hard and pushing through pain to get through marathons, so dealing with hunger is a piece of cake in

> comparison." One evening she went out for pizza and beer with her husband and some friends. She went to the restroom and was overcome by anxiety about what she had eaten. She decided to vomit and, much to her surprise, she felt less anxious afterward. At that point, she rationalised that this would be a way to keep people off her case about not eating enough; she could eat, but then get rid of it whenever she wanted. Soon the urge to vomit became overwhelming, and it seemed to be the only way she could control her anxiety. She started a few times a week, but she was soon up to 5 to 10 times a day. Her weight continued to drop and her times kept getting worse. Her coach and trainer were worried about her health, but she denied any problems. She started having difficulty concentrating at work, was sleeping poorly, and withdrew almost completely from her family and friends. She also withdrew from her husband. Even though they had been talking about starting a family, she became less and less interested in having sex and started even disliking being touched. She spent hours in the bathroom scrutinising her body—checking to see if her shape had changed, pinching the skin on her waist to make sure her shape wasn't changing, and weighing herself—sometimes 10 times per day. She became convinced that the only way to start winning again was to get down to 90 pounds. She was restricting her intake to about 800 calories per day, restricting fluids, and vomiting several times per day. Wendy made it through the half marathon on sheer will, but passing out at the finish line was the final event that brought her into treatment (Bulik et al., 2014, pp. 499-500).

Because they provide a description of a specific disorder, case narratives like this could conceivably serve as compositional explanations. However, they still fall short in similar ways to the options already discussed.

First, as with both DSM/ICD syndromes and historical accounts, case narratives tend to refer only to those features of the disorder that are easily observed or assessed in clinical practice. For example, the exemplar above refers largely to the phenomenological experience of the client, Wendy, and neglects to include information about any physiological, neural, or molecular processes at work. This makes sense given that such accounts are intended as illustrative examples of how the disorder presents clinically, and in clinical practice one would not routinely engage in the methods of investigation required to identify more in-depth structural phenomena, such as functional Magnetic Resonance Imaging (fMRI) or

Cerebrospinal Fluid (CSF) sampling. However, as previously discussed, it significantly limits their explanatory value.

Second, like historical accounts, case examples typically include information about client history like the information about the sexual assault provided in the example above, thereby conflating the tasks of describing the premorbid and morbid states, as well as implying the causal relevance of these factors (thereby alluding to an aetiological explanation). They therefore lack the appropriate theoretical organisation to serve as compositional explanations.

Third, although case narratives refer to significantly more features of the disorder than most classificatory approaches, they still lack depth in their description of these. Typically, they fail to go into any features in detail – relying on brief, superficial representations despite the fact that these constructs are often multidimensional (for example, "body image"; see above) – and often refer to the thinly-defined constructs that have become entrenched in the psychopathological research.

Finally, case examples are just that: examples. Each presents a specific presentation of a psychopathological condition. Hence, although many features of the disorder may be represented, it is unlikely that *all* the features relevant to that disorder will be described. Real life cases of a disorder seldom present with every feature associated with it but, rather, a subset of them. They therefore lack the explanatory breadth required to serve as effective compositional explanations.

The Research Domain Criteria (RDoC).

The RDoC is a clinically independent research framework intended to guide empirical investigation into psychological mechanisms (Cuthbert, 2014; Cuthbert & Insel, 2013; Insel et al., 2010). It was developed as a reaction to the publication of the DSM-5 (APA, 2013), which many perceived as being a conservative development on the previous edition (DSM-IV; APA, 2006) that retained many of the problems originally identified, such as reification (Cuthbert & Insel, 2013; Whooley, 2014). The RDoC presents an "alternative nosological framework" (Whooley, 2014, p. 100) that seeks to advance psychopathology research, specifically neurobiological investigation (Cuthbert & Insel, 2013; Whooley, 2014). The RDoC assumes that mental disorders are "brain disorders" borne out of dysfunctions in neural circuitry, and therefore aims to build a nosology of mental disorder from the "bottom-up" using current neuroscience research (Cuthbert, 2014; Cuthbert & Insel, 2013; Insel et al.,

2010; Whooley, 2014). The hope is that, by doing so, more valid diagnostic categories will be developed that are anchored in neurobiology (Lilienfeld, 2014; Whooley, 2014).

The RDoC framework provides a two-dimensional matrix to guide psychopathology research. This matrix consists of 6 psychological "domains" into which investigation may be conducted – *negative valence systems* (e.g., threat, loss), *positive valence systems* (e.g., approach motivation, reward learning), *cognitive systems* (e.g., attention, working memory), *systems for social processes* (e.g., attachment, social communication), *sensorimotor systems* (involved in the acquisition, control, and execution of simple and complex motor behaviours, such as action selection, initiation, execution, habit development) and *arousal/modulatory systems* (e.g., sleep-wake, arousal) – and 7 "units of analysis" at which each domain may be studied – *genes*, *molecules*, *cells*, *neural circuits*, *physiology*, *behaviour*, and *self-report* (Cuthbert, 2014; Cuthbert & Insel, 2013; Lilienfeld, 2014). The RDoC assumes that mental disorders result from disruptions in the normal-range functioning of these processes and thus applies basic understandings of psychology and neuroscience to psychopathological problems (Lilienfeld, 2014).

It is important that the role of the RDoC within the scientific inquiry process be accurately understood. Although the RDoC has the potential to generate a wealth of detailed information about mental disorder – and may therefore contribute usefully to the development of compositional explanations – it does not generate compositional explanations itself. The RDoC is a research framework intended to scaffold investigation into psychological processes – both their function *and* dysfunction – to obtain insight into psychopathology and thereby "inform future classification schemes" (Insel et al., 2010, p. 748). Hence, although able to generate significant data about psychological processes, the RDoC does not conceptualise these findings at a theoretical level to generate aetiological or compositional explanations.

Even if it did include such a synthesis, the RDoC matrix is not specifically directed towards the investigation of psychopathological problems and would therefore likely struggle to provide coherent understandings of specific conditions. As mentioned earlier, the RDoC directs investigation towards a set of broader psychological processes – such as "positive valence systems" or "cognitive systems" – how they function and malfunction, and how they may contribute to the development and maintenance of mental disorder – rather than specific psychopathological problems, such as low mood, disordered eating, or anxiety. Hence, the

picture that the RDoC framework is likely to facilitate is a comprehensive description of these systems – how they function and malfunction and may therefore play a role in causing or constituting various mental disorders – rather than explanations of specific psychopathological conditions. Furthermore, the RDoC framework is not solely directed towards compositional investigation: the framework also encompasses aetiological investigation, such as genetic research, and there is no clear definition between these two types of study built into the RDoC matrix. As previously discussed, on a theoretical level this approach risks conflating the two distinct tasks of aetiological versus compositional explanation and may lead to problems further along in the explanatory process (see above).

Another key concern is the RDoC's notable neurocentricity. The framework focuses predominantly on understanding psychopathology from a neurobiological perspective, explicitly proposing that mental disorders be considered "brain disorders" that are borne out of dysfunctions in neurocircuitry (Cuthbert, 2014; Cuthbert & Insel, 2013; Insel et al., 2010; Lilienfeld, 2014; Whooley, 2014). Of the seven units of analysis prescribed by the RDoC, five are biologically based – genes, molecules, cells, neurocircuitry, and physiology – which gives neurobiological factors significantly more weight than other levels of analysis (Lilienfeld, 2014; Lilienfeld & Treadway, 2016). Furthermore, even though one of the investigative domains specified concerns social or interpersonal phenomena – namely, "systems for social processes" – broader sociocultural structures and influences are not addressed within the framework. Cultural factors are well-evidenced as playing a significant role in all aspects of mental disorder – including aetiology, maintenance, diagnosis, classification, and symptom expression – such that most mental disorders arguably cannot be understood independent of their social and cultural context. For example, some of the symptoms associated with post-traumatic stress disorder can be viewed as adaptive when viewed within the context of the trauma event or specific circumstances – for example, hypervigilance, physical hyperarousal, and emotional detachment may be useful in combat situations. Hence, even though the RDoC might lead to a rich neurobiological account of a disorder phenomenon, on its own it risks decontextualising mental distress such that lower levels of analysis (e.g., cellular, neural) are not considered within the broader context of the problem and higher levels of analysis (e.g., psychological, phenomenological, sociocultural) end up significantly underspecified despite their explanatory relevance (Whooley, 2014).

All of this is not to say that the RDoC cannot contribute valuably to the development of compositional explanations of psychopathological problems. There are certainly some

benefits to its approach. First, the RDoC usefully highlights the role of neurobiological processes in mental disorder and provides a platform for their investigation. Second, the wealth of data it intends to generate would provide substantial information to inform the construction of compositional accounts of psychopathological problems. However, overall, the RDoC lacks the appropriate theoretical and methodological structure for generating compositional explanations of mental disorder. It is not oriented directly towards the study of psychopathological phenomena or their theoretical conceptualisation, it fails to distinguish between compositional and aetiological information, and it risks side-lining non-neurobiological levels of explanation despite their proven roles in mental distress.

Existing "Explanations" of Mental Disorder

By *existing "explanations"* of mental disorder, I mean accounts like those evaluated in Chapter Three – theoretical representations of specific psychopathological problems that claim to explain their occurrence. These accounts are eligible for consideration as compositional explanations because they often include descriptive propositions about how a disorder state is *maintained* alongside claims about how it initially emerged. For example, the Affect Dysregulation Theory (ADT) evaluated in Chapter Three proposes that binging and purging behaviours are maintained by a negative reinforcement cycle in which binge-purge behaviour provides relief from aversive emotional experience, but also contributes to the generation of additional negative affect with which further disordered eating behaviour is used to cope (Linehan & Chen, 2005). Similarly, the Interpersonal Theory of Disordered Eating (IPT) proposes that disordered eating behaviour exacerbates the existing interpersonal difficulties thought to have prompted it in the first place, thereby contributing to further engagement in disordered eating in response.

It is worth noting that in their description of maintenance processes these accounts provide some useful information about the composition of psychopathological problems that is not purely superficial but attempts to describe causal mechanisms. As discussed earlier in this chapter, understanding the processes maintaining a particular problem has significant value for primary intervention, as it clues clinicians in on how the disorder state can be disrupted. Despite this, however, these accounts fail to provide sufficient compositional explanations.

First, these explanatory accounts are primarily concerned with aetiology, seeking to account for how the disorder state emerged in the first place rather than its existing structure.

This often leads to them providing some compositional information as well, as maintenance processes are also important for understanding disorder aetiology, often helping to explain how an individual's unhelpful way of thinking, feeling, or behaving developed into a *pattern* of harmful dysfunction (and thus became a "problem"). For example, the proposal made by the ADT, that binge-purge behaviour is maintained by negative reinforcement, also helps us to understand how initial engagement in this behaviour could progress to a full-blown pattern of dysfunctional behaviour. However, the explanatory focus in these accounts remains primarily on aetiology rather than constitution. As previously discussed, these are different theoretical tasks, each requiring a different process of reasoning. Attempting to achieve both is, in my view, attempting to achieve too much within a single account – one or the other is likely to be neglected or mishandled in the attempt. Hence, like historical accounts, clinical case narratives, and the RDoC, existing explanatory accounts lack the appropriate theoretical orientation to fill the role of compositional explanation.

Second, likely because of their focus on aetiology, the compositional account provided by existing "explanations" is typically relatively "thin". As discussed in Chapter Three, these accounts generally fail to incorporate all the relevant explanatory information about their target phenomenon that exists at different levels of analysis. Most rely on cognitive, emotional, and interpersonal factors and processes, with little consideration given to the role of lower-level neural, chemical, or physiological processes or higher-level socio-political ones. They also frequently fail to sufficiently define the constructs they invoke in their accounts, such as the "mood intolerance" construct in Fairburn et al.'s (2003) Transdiagnostic Cognitive-Behavioural Theory of Eating Disorders (TCBT; see Chapter Three). Both criticisms apply equally to the compositional information provided by these accounts. Indeed, this is even more the case because *aetiology* is the primary focus of these accounts, not *constitution*. There is therefore even less detail provided about the compositional factors and processes described and far more information that is omitted.

As with the RDoC, this is not to say that existing "explanations" or explanatory models cannot contribute to the development of compositional explanations within psychopathology. With their descriptions of maintaining factors and processes, many already approximate compositional models that could contribute to the generation of an effective compositional explanation. For instance, the ADT's proposal that binge-purge behaviour is maintained by a negative reinforcement cycle involving avoidance or escape from aversive emotion provides a useful basis for an emotion-focused model of the composition of the

"binge-purge cycle" phenomenon. However, such an account, even when further developed, would still form but a partial compositional explanation of this symptom.

Building Better Descriptions

The types of descriptive accounts we currently possess within the psychopathology discipline are insufficient as compositional explanations of mental disorder, lacking either the appropriate theoretical orientation or sufficient validity, depth, and empirical adequacy. This section therefore discusses how more effective compositional explanations could be constructed – namely, by using available compositional data as a basis, modelling at multiple levels of analysis, and providing detail-rich descriptions of relevant factors and processes.

Such an explanation should also be based around an appropriate target. As discussed in the previous section on explanatory targets, this is frequently a problem in psychopathology: explanations frequently use syndromes or transdiagnostic classes for this purpose, which lack sufficient reliability and validity. In response to this, I proposed that we shift our explanatory focus, at least to start with, to *clinical phenomena*: individual signs and symptoms of mental disorder or reliable "complexes" of signs and symptoms that represent reliable patterns of thought and behaviour, such as "the binge-purge cycle". Hence, in my view, better compositional explanations should seek to describe *these* phenomena, rather than syndromes or transdiagnostic groups as many of the above descriptive approaches do. Once such a phenomenon has been selected, work can then begin on the subsequent steps. I will now demonstrate how a quality compositional explanation could be constructed, using the clinical phenomenon "binge eating" to illustrate each part of the process.

Starting with Data.

All explanations should arguably be constructed based on empirical evidence. This begins with the unbiased gathering of relevant data – cross-sectional research from a variety of disciplines, such as clinical psychology, neuroscience, medicine – that is of high methodological quality – Randomised Control Trials (RCTs), meta-analyses, systematic reviews, or methodologically rigorous single studies. These data can then be mined for patterns that might represent constructs or processes that are part of the symptom's constitution.

For instance, within the symptom of "binge eating", we might gather a range of quality data – including meta-analyses, systematic reviews, and RCTs – from psychology,

medicine, and neuroscience that focus on this symptom and analyse it for patterns. Through this, we might discern a pattern of "impaired inhibitory control" evidenced by studies showing that individuals who binge eat demonstrate poorer performance on tasks designed to assess planning (Farstad et al., 2016; Fischer et al., 2008), decision-making (Guillaume et al., 2015; Perpiñá et al., 2017; Wai Sze Chan et al., 2014; Wu et al., 2016), and set-shifting ability (Kanakam & Treasure, 2013; Roberts et al., 2007; Wu et al., 2016); higher self-reported impulsivity (Farstad et al., 2016; Steward et al., 2017; Vervaet et al., 2003; Waxman, 2009; Weinbach et al., 2018) and engagement in additional impulsive or reckless behaviours like self-harm and substance abuse (Peebles et al., 2011); and hypoactivity in frontostriatal brain areas responsible for self-regulation and inhibitory control, such as the prefrontal cortex (Balodis et al., 2015; Donnelly et al., 2018; Friederich et al., 2013; Lavagnino et al., 2016). The fact that this pattern is evidenced by a wide range of robust research across multiple disciplines makes it more likely that this process is genuinely present and not the false product of biased reasoning or methodological error. Hence, we can be reasonably confident that this process is a genuine component of "binge eating".

This example demonstrates how factors and processes contributing to the constitution of a clinical phenomenon can be reliably identified. As they are based on quality empirical data, these phenomena are also more likely to be valid, and thus genuine contributors to the clinical phenomenon in question. The eventual compositional explanation generated will therefore have greater empirical adequacy and be more likely to represent the phenomenon as it actually exists.

Modelling at Multiple Levels.

Complete compositional explanations should describe their phenomena at all relevant levels of analysis. Indeed, Zachar (2008) describes psychopathological phenomena as structures with many overlapping levels. Compositional explanations should therefore be multi-level to reflect this adequately. To facilitate this thinking, compositional explanations can be constructed in a loosely stratified or "stacked" format, beginning with the experiential or phenomenological level (how the symptom/sign is observed or reported) and moving outwards, considering each level of analysis in turn, to identify factors and processes that might be relevant to the symptom's constitution. This is useful because factors and processes at one level may partially constitute or be otherwise related to those at another and building outwards in this manner may help the researcher to make these theoretical connections.

Consider the symptom of "binge eating". At the phenomenological level, we can recount how this symptom is experienced by the client based on self-report data from empirical research: privacy is important ("secretive eating"; Lydecker & Grilo, 2019), emotion is typically central to the event ("emotional eating"; Leehr et al., 2015; Pinaquy et al., 2003; Ricca et al., 2012; Waters et al., 2001; Zeeck et al., 2011), individuals tend to feel as though they can't control their eating ("lack of control"; Colles et al., 2008; Greeno et al., 2000), and may experience strong metabolic "hunger" and hedonic "craving" for food (Ng & Davis, 2013; Van den Eynde et al., 2012; Witt & Lowe, 2013; see Figure 1). From this level, we can then move downwards, considering factors and processes suggested by empirical research at lower levels that may comprise this symptom. For example, at a cognitive level, the "hunger" and "craving" experienced in binge eating could be represented as a "heightened appetite"[23] (see Figure 1). This, in turn, may be partially constituted, at the physiological level, by an impaired appetite feedback system – involving imbalances in hunger and satiety hormones (Culbert et al., 2016) and altered vagal nerve transmission (Peschel et al., 2016) – and, at the neurological level, by an impaired interoceptive network (aberrant responding of the insula and medial frontal gyrus; Gasquoine, 2014; Kim et al., 2012; Klabunde et al., 2017) and alterations in reward pathways (Avena & Bocarsly, 2012; Berridge, 2009; Frank, 2013; Schienle et al., 2009; Wierenga et al., 2014; see Figure 1).

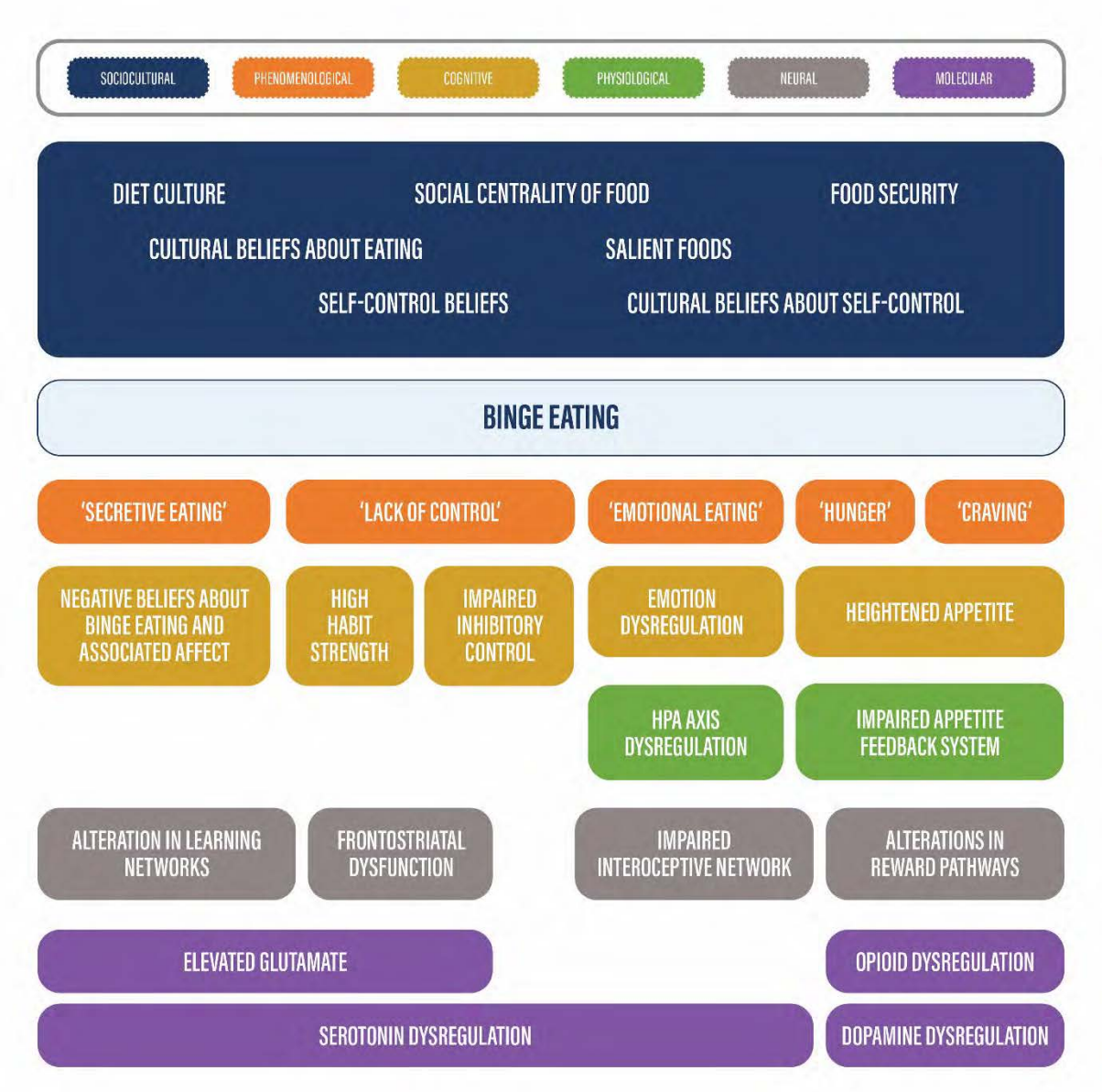
Further down, at the molecular level, serotonin dysregulation (Compan et al., 2012; Kaye et al., 2005; Klein & Walsh, 2004; Kuikka et al., 2001; Tiihonen et al., 2004) may be partially responsible for the experience of hunger surrounding a binge (due to its role in appetite regulation; Lam et al., 2010), and the experience of intense craving may be influenced by dysregulation of opioid and dopaminergic systems implicated in reward and addiction (Bello & Hajnal, 2010; Berridge, 2009; Broft et al., 2011; Majuri et al., 2017; Wang et al., 2011). At a higher level, it is also worth considering how sociocultural factors may influence how the symptom is experienced, enabling or inhibiting certain behaviours or altering symptom content. For instance, overeating, for either hunger or hedonic pleasure, is to some degree dependent on the socioeconomic availability of food ("food security"; Anderson-Fye, 2018), as food needs to be available in reasonable abundance to allow for its overconsumption. Similarly, the content of cravings is likely to be influenced by the

[23] *Appetite* refers to an individual's drive to eat, for either hedonic or metabolic purposes (APA, 2007; Booth, 2003).

individual's cultural environment, such as diet culture, the social centrality of food, and cultural beliefs about eating and self-control (Osman & Sobal, 2006).

We can conduct the same sort of process for each aspect of a symptom to build a multi-level explanation of the phenomenon's constitution (see Figure 1). Such an explanation provides a comprehensive account of the symptom, as it considers the relevant factors and processes at all levels and how they may comprise or influence each other. This prevents some levels being unfairly privileged over others and lends the overall explanation substantial explanatory depth.

Figure 4. *Illustrative Sketch of a Multi-level Approach to Compositional Explanation, Using the Symptom "Binge Eating"*



Detailing Domains.

As discussed above, a compositional explanation should aim to describe all aspects of the phenomenon in question. However, I would argue that a good compositional explanation should also aim to do so in a high level of detail. Factors and processes identified at each level of analysis should be richly described so that it is clear how they come together to constitute the phenomenon. Any abstract constructs referred to in the explanation should also be properly unpacked so that the overall account is as clear and unambiguous as possible. One way of doing this is to build smaller compositional models of the constructs that are "nested" within the larger account. This helps to avoid overcomplicating the broader compositional model above but ensures that all factors and processes are sufficiently outlined within the overall compositional explanation. Hence, the overall explanation becomes pluralistic in nature, comprised of multiple explanatory models at different scales (*model pluralism*; see Chapter 5).

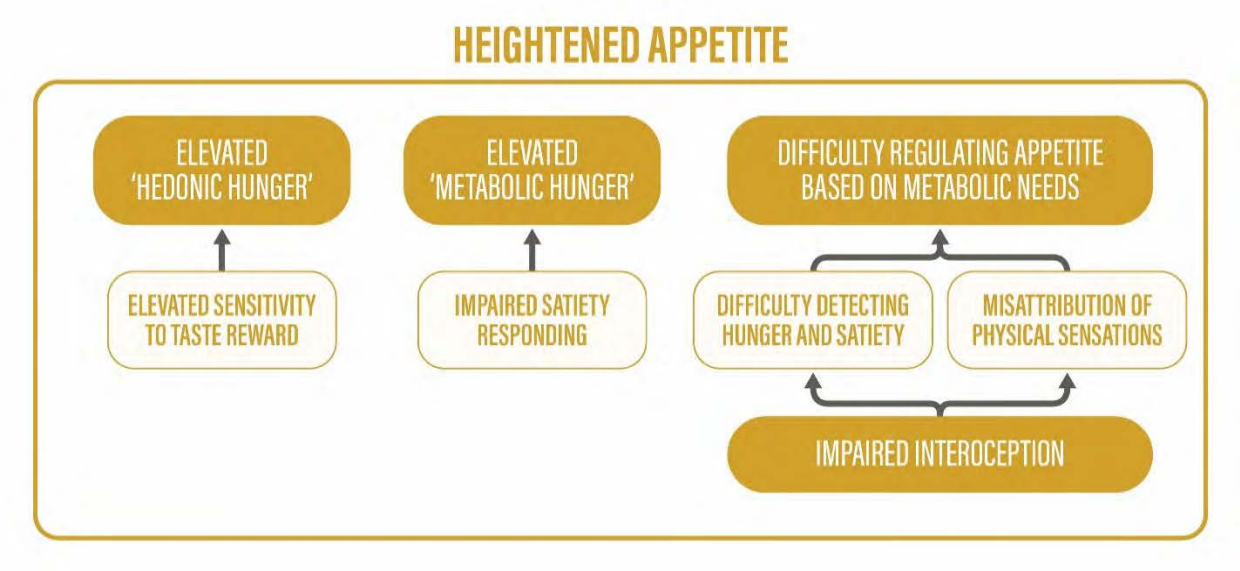
Consider the multi-level model of "binge eating" previously sketched out. Although this model provides a good overview of constitutional factors and processes at each level of analysis, the constructs referred to within each domain are still in need of further definition. By constructing nested sub-models of these phenomena, we can more clearly define the factors and processes invoked and thereby enrich the overall account. For example, consider the "heightened appetite" phenomenon identified at the cognitive level earlier. Based on the literature, we can construct a descriptive sub-model of this construct that elucidates its nature in more detail (see Figure 5):

Individuals who binge eat are more sensitive to the effects of reward (Harrison et al., 2010; Wierenga et al., 2014), including reward from food ("taste reward"; Schag et al., 2013). They are therefore likely to have greater hedonic hunger (Lowe & Butryn, 2007; Witt & Lowe, 2014) than non-binge eating individuals. Binge eaters also demonstrate impaired satiety responding (Klein & Walsh, 2004; Sysko et al., 2007). Their metabolic hunger is therefore likely to also be greater as they have less indication when they are full. Finally, binge eaters reliably demonstrate impaired *interoception* – the ability to sense and interpret internal sensations – than non-binge eaters (Herbert & Pollatos, 2018; Jenkinson et al., 2018; Klabunde et al., 2017). They may therefore struggle to detect physical cues, including those associated with appetite, and accurately interpret them, sometimes misattributing physical sensations as hunger or satiety. This may make it more difficult for them to modulate eating

behaviour appropriately according to their body's metabolic needs, as they are less able to accurately perceive the physical cues intended to communicate them (Herbert & Pollatos, 2018).

This model gives a much clearer and more comprehensive understanding of the "heightened appetite" phenomenon cited in the broader account.

Figure 5. *Example of a Compositional Sub-model Within the Symptom "Binge Eating", Detailing the Nested Cognitive Phenomenon "Heightened Appetite"*

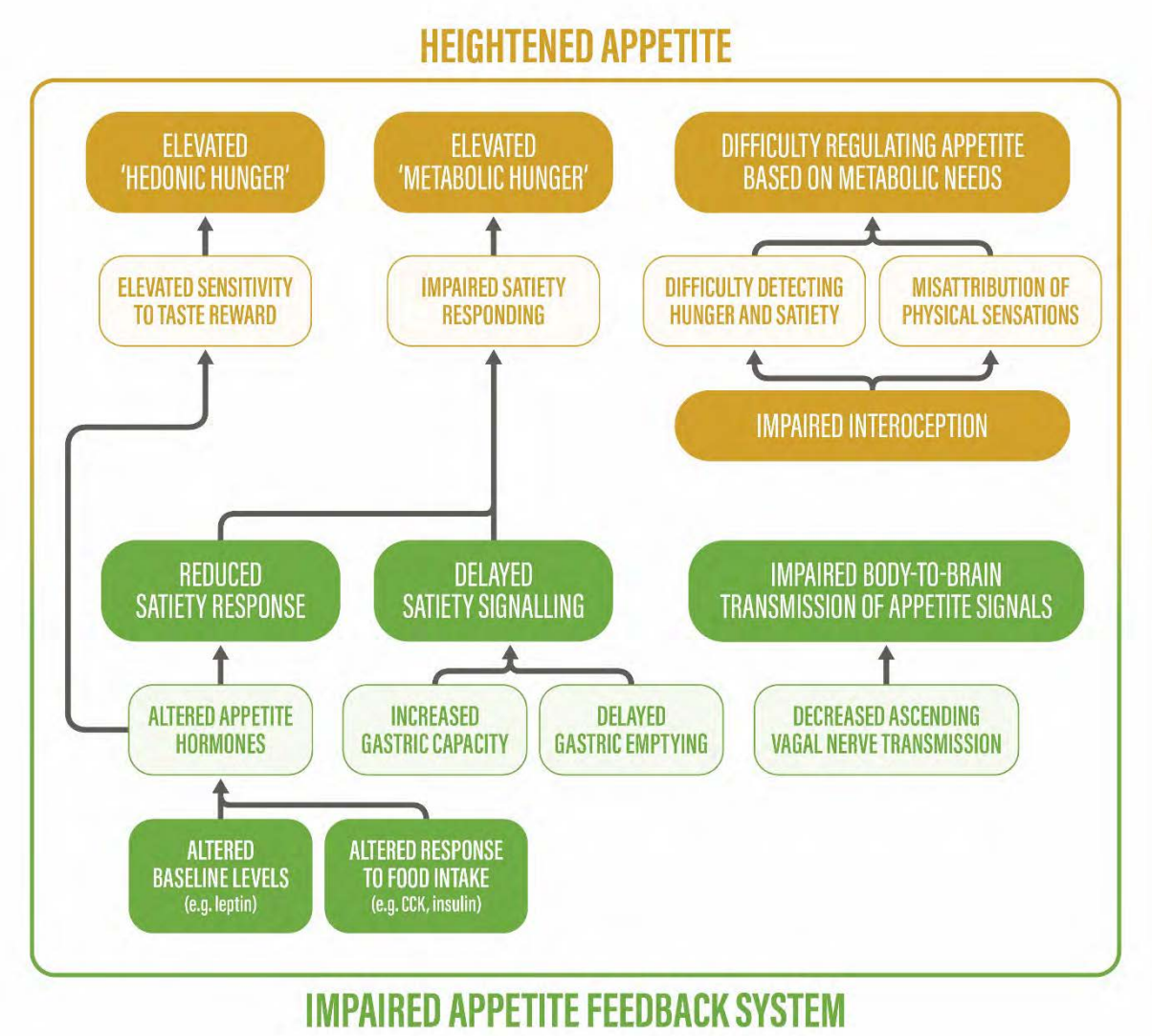


Sub-models at one level can also be linked to sub-models at other levels to further enrich the compositional account. For example, at the physiological level, we may construct a sub-model of the "impaired appetite feedback system", in which altered appetitive hormones, both at baseline and in response to food intake, contribute to a reduced satiety response (Culbert et al., 2016; Prince et al., 2009); increased gastric capacity and delayed gastric emptying result in additional delayed satiety signalling (Devlin et al., 2012; Klein & Walsh, 2004); and decreased ascending vagal nerve transmission contributes to impaired body-to-brain transmission of appetite signals (Faris et al., 2000; Murialdo et al., 2007; Simmons & DeVille, 2017; see Figure 6). This can then be linked to the cognitive model by relating the lower-level physiological processes to those at the higher cognitive level. For example, the impairments in satiety responding described at the cognitive level is likely to be at least partially constituted by the dysfunctions in appetitive hormones and gastric functioning at the physiological level (Berthoud, 2008; Zanchi et al., 2017; see Figure 6). Aberrant functioning of appetite hormones may also partially comprise altered perceptions of food-related reward

at the cognitive level due to the influence of some appetitive hormones, such as *leptin*, on dopaminergic networks (Cassioli et al., 2020). Physiological dysregulation of the vagus nerve may likewise play a part in the impairment of interoception, as the vagus nerve plays a key role in transmitting both internal and external sensory information from the body to the brain (Craig, 2002).

This additional modelling within and between levels greatly enriches the overall compositional explanation. There is far more clarity about how the factors and processes hypothesised at each level are comprised and connected, as well as how they relate to the constructs at other levels. The overall explanation therefore has far greater explanatory depth, as it richly details the deep underlying mechanisms contributing to the structure of the symptom.

Figure 6. *Example Model Linking the Compositional Sub-models of Cognitive ("heightened appetite") and Physiological ("Impaired Appetite Feedback System") Phenomena Within the Symptom "Binge Eating"*



Summary

Current explanations within psychopathology are typically built around targets that are conceptually flawed. This is problematic as it limits our ability to construct explanations of mental disorders that are valid and useful for understanding and treating these problems. Possessing appropriate targets for explanation is critical – the focus of explanation fundamentally affects the nature of the explanation generated, including what type and amount of information it provides about the area of interest. The current persistent focus on syndromes and transdiagnostic classes within psychopathology is therefore highly limiting, as the significant conceptual issues with these categories constrain the validity and depth of the

explanations they are used to construct. Until better classifications are developed, therefore, I argue that we should refine the focus of our explanations to *clinical phenomena*: individual signs and symptoms, such as "self-starvation" or "anhedonia", or "phenomena complexes", such as "the binge-purge cycle" (Nielsen & Ward, 2020a). These targets possess a higher degree of validity and reliability than the current alternatives and are likely to facilitate explanations with superior explanatory value.

As well as selecting more appropriate explanatory targets, we need to get better at describing them – that is, at compositional explanation. As discussed, good compositional explanations have significant value for theory, research, and practice in psychopathology – they help to scaffold better aetiological explanations and classification schemes, suggest worthwhile avenues of inquiry in clinical assessment, and indicate primary intervention strategies or potential therapeutic issues in treatment. However, none of the descriptive accounts we currently possess in psychopathology – DSM/ICD syndromes, transdiagnostic and dimensional classifications, symptom network models, historical accounts, clinical case narratives, the RDoC, and existing "explanations" – provide a sufficient compositional explanation.

In the second half of this chapter, I therefore suggested how better compositional explanations could be constructed, by using empirical data as a foundation and creating multiple models that richly detail all levels of analysis relevant to the phenomenon in question. This kind of nested modelling approach is a good example of how originally conceptually thin clinical phenomena can be elaborated into rich, multi-layered compositional accounts that could, ultimately, result in stronger aetiological theories, more accurate taxonomies, and more precisely targeted treatment. It also endorses a multiple model (*model pluralism*) approach to explanation in psychopathology, which relates to the second issue identified by the analysis in chapter three and forms the focus of the next chapter.

Chapter 5. Engaging with a Pluralistic Approach to Understanding Psychopathology

This chapter addresses the second issue identified by the analysis in Chapter Three: the endorsement of the One Best Model (OBM) perspective within psychopathology. Existing models of mental disorder are typically presented as autonomous, stand-alone explanations, with the implicit assumption that, eventually, One Best Model will rule them all. Accordingly, there is frequent and heated debate over which of the many available therapeutic interventions are likely to be superior, as each is associated with a specific single-model "explanation". Family-based approaches war with individualised therapy, cognitive-behavioural perspectives fight with interpersonal or sociocultural ones, and advocates of the medical model reject psychotherapeutic intervention in favour of psychoactive medications. Overall, the current prevailing attitude to the conceptualisation of mental distress is that there will emerge a single explanatory model that neatly accounts for each disorder presentation, or for a cluster of disorder presentations (the "transdiagnostic" approach; Fusar-Poli et al., 2019).

This assumption can be discerned throughout the history of psychopathology. Namely, in how the conceptualisation of mental disorder aetiology has progressed through various rival approaches: from early psychodynamic conceptualisations (mental disorder as the manifestation of unconscious desires; e.g., Freud, 1964), to behaviourism (mental disorder as learned patterns of maladaptive behaviour; e.g., Skinner, 1976), through social constructionism (mental disorder as normatively determined; e.g., Szasz, 1960), and cognitivism (mental disorder as errors in information processing; e.g., Beck & Haigh, 2014), to more recent neurobiological perspectives (mental disorders as brain disorders; e.g., the RDoC; Insel et al., 2010). All of these accounts assert that their conceptualisation of mental disorder is superior and that the level or scale of analysis they focus on – psychological, behavioural, sociocultural, cognitive, or neurobiological – is most important. It is seldom acknowledged that all of these models might hold value, each contributing uniquely to our understanding of mental distress.

Most psychopathology researchers and practitioners would now agree that mental disorders arise from a combination of biological, psychological, and social causes (the "biopsychosocial" model; Wade & Halligan, 2017), thereby acknowledging the contribution of multiple factors and levels to the aetiology of mental disorder. However, this perspective does not reflect the intention to simultaneously hold separate models specific to these

domains but, rather, to *integrate* them into a single overarching model. Current "explanations" of mental disorder often follow suit, aiming to unify the diverse contributors to mental disorder under a single integrative model – the Unified Model of Depression (Beck & Bredemeier, 2016) and the Transdiagnostic Cognitive-Behavioral Theory of Eating Disorders (Fairburn et al., 2003) are prime examples of this approach. Philosophers of psychiatry and psychology also frequently seek to develop theoretical frameworks for the integration of models or findings from different domains into a single account – for example, the RDoC (Insel et al., 2010) or Thomas and Sharp's (2019) "mechanistic science" account.

This notion that there exists "One Best Model" permeates psychopathology. At an empirical level, therapeutic interventions are repeatedly pitted against each other in comparative efficacy studies (e.g., Cuijpers et al., 2013; Hofmann et al., 2012; Khan et al., 2012; Tanner-Smith et al., 2013; Tolin, 2010), competing to be evaluated the most effective or declared "best-practice". Such rivalry translates into practice, encouraging clinicians to "pick therapeutic teams" – that is, to adhere strictly to a single approach for which they feel the evidence (either empirical or anecdotal) is most convincing. This automatic tendency to compare therapeutic approaches against each other frequently leads to amplification and dichotomisation of their differences. As a result, some theoretical camps are so at odds with each other that many practitioners will immediately consider them irreconcilable, without giving thought to how they might actually be similar or work together – for example, the antipathy often observed between new-wave Acceptance and Commitment Therapy (ACT; Hayes, 1999, 2006) and traditional Cognitive Therapy (see Gaudiano, 2011).

Even approaches that describe themselves as "trans-theoretical", such as the Stages of Change model (DiClemente & Prochaska, 1998) or the "transdiagnostic" approach (Fusar-Poli et al., 2019), do not offer comprehensive explanations for mental distress. Although each aims to provide a less insular perspective, both still advocate for the OBM viewpoint – namely, the achievement of a superior single model to explain mental disorder. One could even argue that they are more problematic, in that they inadvertently promote further reduction by implying that a single unified model should not only be able to account for one disorder presentation but multiple. Although such approaches have some value for treatment – for example, aiding resource efficiency and accounting for comorbidities – at an explanatory level they prevent in-depth exploration of the composition and aetiology of disorder symptoms. Arguably, a unified model simply cannot sufficiently incorporate all the relevant information while remaining comprehensible.

Overall, I believe the OBM perspective is broadly unhealthy for psychopathology research and practice. In its encouragement of model autonomy, it limits the flexibility of academics and clinicians, restricting them to the use of a single explanatory model rather allowing for multiple, each offering a unique and valuable perspective. The "messy" realities of clinical practice – such as comorbidity and treatment-resistance – mean that a single explanatory model or therapeutic approach is seldom sufficient to ameliorate a client's problems. Particularly with more chronic or complex psychological problems, such as personality disorders or complex Post-Traumatic Stress Disorder (PTSD; Brewin et al., 2017; Skodol, 2018), multiple therapeutic approaches are likely to be required for effective remediation (Cloitre et al., 2011; Ehrenthal & Benecke, 2019; Resick et al., 2012). However, the current theoretical segregation of explanatory models promoted by the OBM leaves clinicians without guidance as to how these different interventions may overlap and intertwine.

In this chapter, I therefore argue for a transition away from the OBM philosophy, toward a multiple model approach to psychopathology that is collaborative and pluralistic – in other words, a *model pluralism* approach. I begin the chapter by outlining the particulars of the OBM perspective and the problems it entails for psychopathology theory and practice. I then present my proposed solution of *model pluralism* and suggest how this may be applied to how the psychopathology discipline approaches the processes of modelling disorder, translating models into interventions, and applying therapeutic models in practice.

The One Best Model (OBM) Perspective in Psychopathology

The possibility of unifying different scientific domains (such as mathematics, physics, biology, or chemistry) or disparate phenomena with a single explanatory theory or model has long excited philosophers and scientists alike. This has most notably been observed in the ongoing pursuit in physics for a unified theory to explain all physical aspects of the universe (a "theory of everything"; Laughlin & Pines, 2000)[24]. However, some researchers argue that unification may not always produce the *best* explanation, depending on the nature of the phenomenon/a to be explained and the different goals an explanation may need to satisfy (e.g., Hochstein, 2017; Potochnik, 2010). As discussed in Chapter Two, an explanation *may* equate to a single model if that model alone represents the entire bank of explanatory knowledge about a phenomenon. However, most often, single models are unable to do this

[24] Often referred to as the "holy grail of physics" (Jones, 2008).

and therefore represent only partial explanations (Bailer-Jones, 2003). When the phenomenon of interest is more complex, as psychopathological problems are, a variety of models will likely be needed to explain it fully and effectively (Kendler et al., 2020). It is therefore perhaps unsurprising that mental disorder has thus far confounded all attempts at unified explanation (Kendler, 2005). As the problem of mental illness grows in public awareness, so too does the incentive to produce explanations for mental distress that are easily understood and communicated. The idea that there might be a single explanatory model that can characterise each, or several, mental disorder(s) is thus extremely enticing.

The OBM perspective refers to this implicit or explicit belief within psychopathology that a single model should be able to completely explain one or multiple presentation/s of mental distress. It claims that a model should equate to an explanation – either that this is currently possible (in that one of the existing models provides a sufficient account) or, more likely, that we can achieve this in the future with additional research.

The OBM perspective may manifest in a couple of ways within the psychopathology discipline. First, it may be expressed as the proposition that one level of explanation – for example, the neurological or cognitive-behavioural – is more important or central to the conceptualisation of mental distress, such that an explanatory model at that level is sufficient to fully explain a particular psychopathological problem. One example of this approach is the RDoC, which advocates for mental disorders being understood primarily as “brain disorders”, thereby prioritising the neurobiological level (Insel et al., 2010). Such assertions about a level’s “superiority” within psychopathology explanation likely play a role in the conflict and rivalry observed between different theoretical schools and psychopathological disciplines – no one likes to be told their perspective is less valued.

Alternatively, the OBM may appear as the belief that all the relevant levels of explanation can be sufficiently represented within a single model and should be integrated into a unified account. For example, Beck and Bredemeier (2016) argue that the somatic, behavioural, cognitive, affective, and physiological symptoms of depression can all be explained by a single Unified Model of Depression that is built around core dysfunctional cognitive schemas (*depressogenic beliefs*) and distorted information processing. Transdiagnostic approaches, recently popularised within psychopathology, are another particularly strong example of the OBM perspective, as they attempt to explain multiple

disorder presentations (that is, diverse phenomena) within a single, overarching model. The OBM is therefore strongly related to the notion of explanatory *unification*.

Unification, in the context of explanation, may refer to several different processes, all of which aim to generate a single model or theory that fully explains one or multiple phenomena. In the above example from physics – the fabled "theory of everything" – and other early accounts of the subject (e.g., Friedman, 1974; Kitcher, 1989), *unification* refers to the explanation of multiple separate phenomena by a single, greater model or theory that claims these phenomena as related and explains them accordingly (Jones, 2008). Hence, in my view, the following decisions frequently made within psychopathology explanation reflect a unified approach to explanation, as each attempts to reduce the explanation of a phenomenon or collection of phenomena to a single model:[25] (1) the decision to collapse or combine multiple accounts of a phenomenon into a single explanatory model, (2) the choice to accept a single explanatory model as superior and disregard the information offered by others, and (3) the decision to reduce model that addresses multiple levels down to a single level of explanation.

It is unsurprising that academics would be drawn to the idea of unified explanations: using a single model to explain a phenomenon is far simpler and more parsimonious than having to rely on multiple. It can even be argued that the concept of unified explanation is inherently attractive. As humans, we are constantly trying to make sense of the world around us and, in many ways, are cognitively wired to seek out simplicity (Chater, 1999; Chater & Vitányi, 2003; Feldman, 2016). To be able to explain the aspects of life that fascinate and perplex us with a single model, rather than multiple, therefore has intuitive appeal. This is no less true of the psychological sciences. As it always has, the mystery of our own development, inner experience, and behaviour continues to fascinate us at a profound level – particularly in cases of mental disorder, in which we perceive these processes as having malfunctioned in some way. Given how these conditions currently vex us, it is understandable that we might be comforted by the idea that we can develop a single model to explain them. However, this belief entails several problems, both in theory and in practice.

[25] Unification therefore also often promotes *reductionism* – explanation or description of complex phenomena in terms of their most basic-level or fundamental constituents – which various philosophers of science and psychopathology theorists alike argue to be both epistemically and ontologically problematic (e.g., Borsboom et al., 2019; Potochnik & McGill, 2012).

Problems with the OBM Perspective

Theoretical Issues

Mental disorders are complex conditions that comprise a diverse array of phenomena. Indeed, recent theories of mental disorder conceptualise mental distress and the "mind" as including the entire body-brain system (*embodied*; Maiese, 2016), reliant on their surrounding physical and sociocultural environments (*embedded*; Fuchs, 2018), and comprising properties that emerge out of interaction between the individual's psychological capacities, needs, and values, and the outside world (*enacted*; Colombetti, 2013; de Haan, 2020). The aetiology of mental disorder is believed to be similarly multifactorial and multidimensional: an intricate interplay between biological and psychological predisposing factors and environmental/sociocultural experiences (Lehman et al., 2017; Wade & Halligan, 2017; Monroe & Cummins, 2015). To explain all or even one of these conditions with a single model is therefore likely an impossible task. In trying to describe all the factors and processes relevant to the phenomenon across all levels and in sufficient detail, a single model would become so convoluted as to be incomprehensible. Hence, what is typically observed when a single model seeks to unify all these explanatory levels is that valuable information is sacrificed for the sake of simplicity.

One way this occurs is through a reduction in the level of detail at which factors and processes are described. For example, Fairburn et al.'s (2003) Transdiagnostic Cognitive-Behavioural Theory of EDs (see Chapter Three) attempts to account for both the proximal aetiology and the ongoing maintenance of both AN and BN at both the cognitive and behavioural levels. This is an enormous range of phenomena to be explained. As such, the finished explanation is somewhat underspecified. For instance, the construct of "mood intolerance", invoked to explain why some individuals appear to employ disordered eating behaviours in response to intense emotion, is only vaguely defined within the explanation (see Chapter Three). The elaboration of other constructs invoked, such as "global low self-esteem" and "clinical perfectionism", is similarly lacking.

Another way information may be sacrificed is through the exclusion of entire levels of analysis. This may mean trying to represent factors and processes at one level at another (usually lower) one – for example, trying to represent cognitive and emotional experience purely at the biological level. Alternatively, it may involve assertions that the levels of analysis included in the model are the most important in comprehending the disorder

phenomenon and that therefore the inclusion of factors and processes at other levels is unnecessary – for example, arguing that cognitive-behavioural factors play a more important or useful causal role in mental distress than biological processes. Both approaches are problematic. First, it is not necessarily true that factors and processes at one level can be accurately represented at others. Features and processes at one level may possess properties that cannot be fully captured through representation at other scales ("emergent properties"; see Potochnik & McGill, 2012). For instance, the "mind" and its varying mental states, which exist conceptually at the psychological or phenomenological level, may be a product of and depend upon neural activity but cannot be fully represented at this level – they are separable phenomena that each have explanatory value (Voneida, 1998). The reduction of higher-level phenomena to lower-level factors and processes therefore often sacrifices considerable explanatory information.

Second, it is problematic to assume the superiority of any level of analysis in the explanation of mental disorder when we still know so little about the exact nature of these conditions. For example, we are only beginning to become aware of the role that biological systems outside of the brain play in mental disorder, such as the *gut-brain axis* (e.g., Foster & McVey Neufeld, 2013). The *gut-brain axis* is a bidirectional link between the central nervous system and the enteric nervous system, which controls gastrointestinal behaviour (Mayer et al., 2015). This includes direct and indirect pathways between cognitive and emotional neural areas and peripheral intestinal functioning (Mayer et al., 2015). On face value, we might not expect the gut to play a role in mental distress, as it has not been included in our historical understanding of psychology and the mind, which largely focuses on the role of the brain and the central nervous system. However, we are slowly discovering that the gut microbiome – the community of microorganisms, both commensal and pathologic, that reside in the gastrointestinal tract (Cresci & Izzo, 2019) – may have a vital role to play in a whole host of psychopathological problems, such as depression and anxiety (Foster & McVey Neufeld, 2013), disordered eating (Seitz et al., 2019), and schizophrenia (Nemani et al., 2015). It is likely that we will identify further diverse contributors to mental distress as time progresses.

Practice Issues

Clinical interventions are typically constructed based on explanations of mental disorder. The factors and processes hypothesised by the explanation to cause or maintain the disorder present potential treatment targets: modifying these phenomena should theoretically

reduce or eliminate the individual's psychopathological symptoms. For example, cognitive theories of depression, such as Beck and Bredemeier's (2016) Unified Model of Depression, implicate negative cognitions like beliefs and attributional biases in the onset and exacerbation of depressive symptoms. Based on this, cognitive therapies seek to target negative cognitions with therapeutic strategies such as *cognitive restructuring* (a set of techniques aimed at helping people identify and challenge their negative thought patterns) in the hope of improving depressed mood (Beck et al., 1987). The assumptions and premises of psychopathological explanations therefore have a fundamental impact on intervention approaches, affecting *what* they seek to target and *how*.

At present – thanks, in part, to the OBM perspective – there is often a 1-1 relationship between explanatory models and intervention strategies – from a single model springs a single associated therapeutic approach.[26] Because of this, interventions inherit the ethos of the models upon which they are based. For example, in the field of EDs, Enhanced Cognitive-Behaviour Therapy (CBT-E) is a direct product of Fairburn et al.'s (2003) Transdiagnostic Cognitive-Behavioural Theory (TCBT) and Rieger et al.'s (2010) Interpersonal Theory of Disordered Eating (IPT) is the sole basis for Interpersonal Therapy for EDs, and both interventions strongly incorporate the assumptions and principles of these founding "explanations". This causes the OBM perspective to filter through to treatment, engendering a similar rivalry and discord between therapeutic interventions, such as interventions vying to be accepted as the One Best Treatment (OBT) for a particular psychopathological problem. It also encourages clinicians to form allegiances with specific intervention approaches. For example, clinicians may identify as a "CBT practitioner" or "ACT practitioner" and use solely that approach rather than selecting from a variety of interventions based on their efficacy for specific types of symptoms or psychological difficulties. Such allegiances are concerning, as they have been shown to affect therapeutic outcomes negatively when clinicians are forced to use interventions that are outside of their preferred approach (Boccaccini et al., 2017).

Interventions are also often developed by the very same researchers responsible for their founding "explanations", which leaves something to be desired in terms of objectivity. Many also become branded and disseminated for profit (Holmes et al., 2018). Indeed, the

[26] Another key contributor to this is the entrenchment of and overreliance on DSM-5 diagnosis within clinical research and practice (Hyman 2011), as interventions are frequently developed to target diagnostic categories, rather than specific psychopathological mechanisms.

generation of therapeutic interventions has become a highly lucrative practice – developers charge high sums for their tools and trainings, which clinicians and organisations are forced to pay in order to stay up to date in their evidence-based practice. The high costs associated with these interventions may force services with limited funding, such as public health services, to choose between intervention approaches or to rely on outdated versions of them. Furthermore, innovative research that looks at modifying these interventions or combining them with rival approaches may be hindered due to intellectual property restrictions (Holmes et al., 2018). Each of these problems further contributes to an approach to treatment in which interventions are either used in isolation – which limits their adaptability to client needs – or integrated with secondary interventions poorly, without theoretical consideration of how they may interact or overlap.

A prime example of this issue is the combined treatment approach of antidepressant medications, such as SSRIs, and Cognitive-Behavioural Therapy (CBT), often used to treat anxiety or depression. These intervention strategies are typically used in tandem and have been reliably shown to be more effective together than alone in most cases (Hollon et al., 2014). However, they are still commonly perceived as independent approaches that address distinct facets of the problem (biological versus psychological), even though these domains are known to be intrinsically interconnected and the conceptual boundaries between them imprecise. Biological and psychological factors may be related constitutively (e.g., neurotransmitter dysfunction at the biological level partially representing the negative cognitive processes observed psychologically), causally (e.g., neurotransmitter dysfunction contributing to the development of negative cognitive biases or negative beliefs producing changes in neurotransmitter levels), or via a shared mechanism (e.g., neurotransmitter dysfunction and negative cognitive biases both independently increasing or intensifying negative mood). For example, some research indicates that antidepressants and CBT may have a shared therapeutic mechanism of action via the alteration of affective information processing (Pringle et al., 2011; Roiser et al., 2012).

Assumptions about the inherent separability of intervention approaches is further enhanced by the “split-treatment” model (Gabbard & Kay, 2001) – namely, the fact that different interventions are often administered by different clinicians. For example, a general practitioner or psychiatrist may administer antidepressant medications while a clinical psychologist or counsellor engages the client in psychotherapy. There is also often poor coordination between these different professionals (e.g., LoPiccolo et al., 2005; Meyer &

Simon, 1999a, 1999b), despite the fact that these conceptual levels (biological and psychological) and their interventions may be connected in a variety of ways (see above). In general, it is becoming more and more clear that psychotherapeutic interventions may have physiological or interpersonal effects, pharmacological drugs may affect psychological variables and social interactions, and interpersonal or community programmes may impact an individual's mood, beliefs, and physical health. Better theoretical understanding of therapeutic mechanisms is therefore critical.

It may be that, with further modelling of the relationships between the levels of analysis informing therapeutic interventions, these approaches can be more effectively combined. In the past, greater understanding of how interventions work has allowed for them to be improved via appropriate combination with other interventions. A simple example from the field of medicine is the augmentation of iron supplementation for anaemia with vitamin C, as vitamin C has been found to facilitate iron absorption and its utilisation from storage. Similarly, therapeutic interventions at one level may be less effective without concurrent or prior change occurring at others. Psychotherapeutic work to correct threat biases will likely have little effect before an individual has been safely removed from a trauma context (in which those biases are constantly reinforced), such as domestic abuse or a war zone, and better understanding of the neuropsychological effects of starvation – such as cognitive narrowing, set-shifting problems, and decision-making difficulties (Keys et al., 1950; Pender, et al., 2014) – suggests that severely anorexic individuals cannot engage effectively in psychotherapy until their malnutrition is addressed. Indeed, research shows that a variety of cognitively mediated psychological variables significantly improve upon weight restoration, allowing more productive engagement in subsequent cognitive therapy techniques (Bamford et al., 2014; Meehan et al., 2006; Pollice et al., 1998). Better understanding of factors and processes involved in causing or constituting mental disorder may therefore help to make treatment more time and resource efficient.

As in the earlier example of CBT and antidepressants, there is emerging evidence that interventions at different levels (such as the psychological and biological) may operate via shared mechanisms of action (Pringle et al., 2011; Roiser et al., 2012), and may therefore "represent different methods rather than different concepts" (Holmes et al., 2018, p. 252). Hence, it may be possible for treatment to be streamlined by eliminating one or other of these interventions – or at least the components within them responsible for the mechanistic

overlap. It may also provide clients with more choice of interventions, as several options may be able to target the same mechanisms in different ways.

Identifying interrelations between therapeutic models may also help to avoid potential harm that could come from combining treatments, in the same way that accurate knowledge of pharmacological interactions helps medical professionals to avoid prescribing drugs that will adversely react with their patient's other medications. There is already evidence for the existence of harmful or potentially harmful therapies within psychopathology, such as Scared Straight programmes and recovered memory techniques (Lilienfeld, 2007).[27] It is therefore reasonable to suggest that there might also exist potentially harmful therapeutic interactions. As a very basic example, there is significant evidence that a strong therapeutic relationship and specific therapist factors, such as empathy, positive regard, genuineness – which can arguably be thought of as interventions in their own right – have a significant enhancing effect on the efficacy of specific intervention strategies, such as CBT (Keijsers et al., 2000). Hence, the absence of these factors or the presence of a negative or judgemental therapeutic context would likely have a deleterious effect on the efficacy of a CBT intervention and may create psychological harm or discourage the client from seeking psychological support in the future. Even therapeutic interventions that appear to promote each other in the short-term could inhibit each other over time. For instance, the combined use of CBT and antidepressants or benzodiazepines for the treatment of anxiety disorders, although showing slight advantage over either treatment alone in short-term trials, has been shown to produce higher relapse rates over time relative to those who received CBT without drug treatment (Holmes et al., 2018). With clearer understanding of the factors and processes being targeted in treatment, we are more likely to be able to predict and prevent such interactions.

The use of single-model interventions also causes problems for two common clinical issues: psychiatric comorbidity and treatment resistance. Given the high rates of comorbidity across mental disorders (Merikangas & Kalaydjian, 2007), it is extremely common for clinicians to encounter clients with multiple forms of mental distress. In some cases, these problems interrelate in such a way that the recommended intervention for one of them works for them all – for example, if the problems share a common cause or one is producing the other/s. For instance, if a client's excoriation disorder is primarily a coping strategy for their social anxiety, then therapeutic treatment for anxiety using CBT may also, by extension,

[27] Other provisional examples include military style bootcamps for youth with conduct problems, hypnosis, and attachment therapy techniques such as *holding therapy* and *rebirthing* (Lilienfeld, 2007).

resolve their skin picking. However, in other cases, problems may be less directly related or involve self-reinforcing cycles that perpetuate them independently of comorbid problems, in which case multiple intervention strategies are likely to be required for effective treatment. For example, a client with comorbid social anxiety and disordered eating involving binging and purging is likely to require multiple interventions over the course of treatment that specifically address each set of problems as well as their relationship to each other, as both social anxiety and binge-purge behaviour are thought to involve self-reinforcing patterns – namely, *experiential avoidance* (Chapman et al., 2006; Howe-Martin et al., 2012) and "the binge-purge cycle" (Corstorphine et al., 2006; Haedt-Matt & Keel, 2011) – that are likely to maintain each problem independently of the other.

Furthermore, sometimes best-practice treatments may fail to treat a client effectively, such as in late-stage treatment-resistant cases of depression or AN (e.g., Fava, 2003; Souery et al., 2006; Strober, 2004). Such difficult-to-treat cases will often prompt clinicians to employ multiple interventions either in sequence or in tandem (Al-Harbi, 2012; Long et al., 2012; Shelton et al., 2010). The lack of theoretical communication or integration between intervention approaches promoted by the OBM perspective can cause problems in both cases. When interventions are applied sequentially – for example, chronic AN treated with an initial unsuccessful course of FBT, followed by IPT-E, then CBT-E, and so on – without theoretical consideration of the relations between them, each new intervention approach essentially starts from scratch. This could easily lead to repetition and redundancy that wastes resources or frustrates the client. If the second, we run into the same problem as combined CBT and antidepressant therapy – namely, that there may be more effective or efficient ways to combine the relevant intervention approaches that careful theoretical consideration would reveal. The current theoretical landscape for treatment provides little to no guidance for these common situations.

There is a disconnect between the reality of clinical practice, which frequently requires the application of multiple intervention strategies, and the field's general perspective towards the explanation and treatment of mental disorder. The OBM perspective has led to estrangement and rivalry between therapeutic interventions, which arguably may have contributed to the excessive dominance of some treatment approaches within the psychopathological sphere, such as CBT (see Dalal, 2018). It is extremely difficult to obtain traction for novel interventions when they are confined to the status of "adjunct therapies" rather than combined more equally and collaboratively with the dominant intervention

approaches. Clinicians may also be reluctant to employ such novel interventions, as there is little to no theoretical guidance on how they might integrate them into their existing practice (Stewart et al., 2012).

Moving Forward: From One Best Model to Model Pluralism

Ameliorating the problems associated with the OBM perspective requires a shift in how psychopathology, as a discipline, approaches the tasks of modelling mental disorders, translating models into intervention strategies, and applying interventions in practice. This demands broad, systemic change in the perspectives of both researchers and practitioners. In this section, I outline the changes that may prove helpful in each of these stages, specify how they may be achieved, and discuss how they might be able to address the problems associated with the OBM perspective.

Modelling Mental Disorder

There has long been debate around what "mental disorder" actually is (e.g., Bolton, 2008; Kendler et al., 2011; Lilienfeld & Marino, 1995; Stein et al., 2010; Szasz, 1960; Wakefield, 1992; Zachar, 2014). Early conceptualisations tended to rely on one or two factors or domains to explain mental disorder. However, over time we have become aware of the breadth of phenomena involved in the development and experience of mental distress. Research has revealed the contribution of a large range of factors – from genetic and molecular phenomena to contextual and cultural influences – that play varying spatial and temporal roles in constituting and causing mental illness. Recent approaches to the conceptualisation of mental disorder acknowledge this complexity and highlight the multifactorial nature of mental distress – namely, that mental disorders comprise and are generated by a complex interplay of multiple different factors at a range of different scales and levels of analysis (e.g., the *MPC kinds* perspective, Kendler et al., 2011; the Embodied-Enactive perspective, Nielsen & Ward, 2018). The OBM perspective therefore no longer adequately reflects our developing ontological conceptualisation of mental distress.

Moving forward, we need to view the theoretical representation of mental disorders through a more sophisticated lens that reflects their complexity. In my view, the OBM perspective should be set aside in favour of a more collaborative and pluralistic approach. *Model pluralism* involves the representation of a target phenomenon, such as a specific mental disorder or symptom, via a collection of models at different scales and levels of analysis that together form an overarching explanation or "explanatory network" (Mitchell &

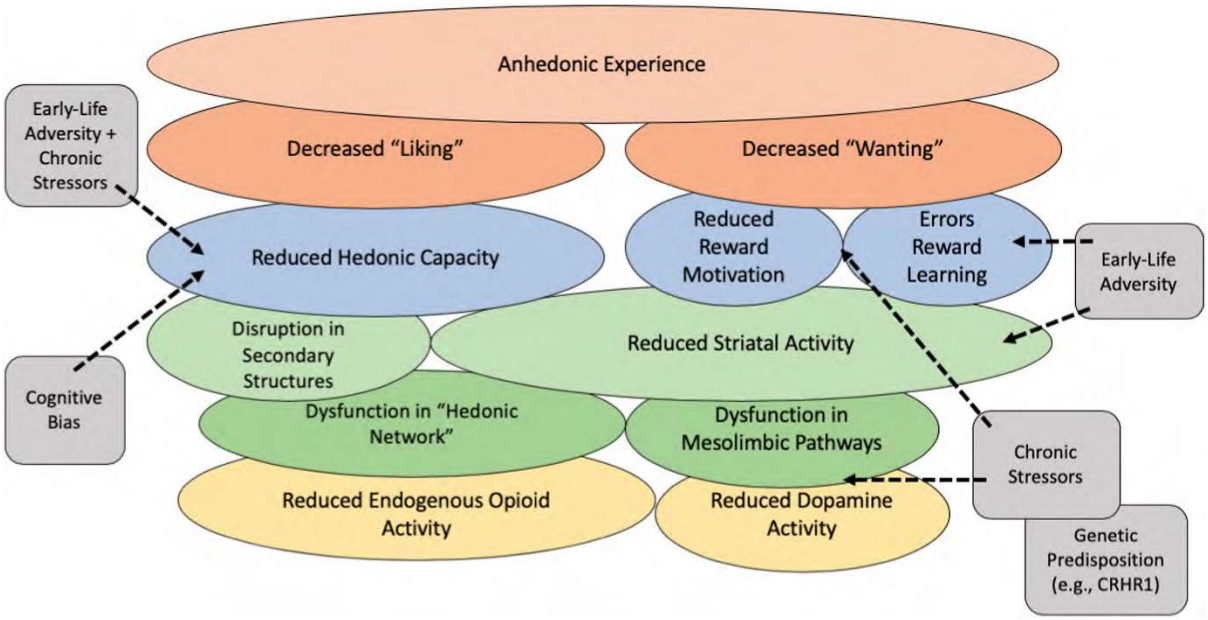
Dietrich, 2006). This overarching explanation may include neural models, cognitive models, emotional models, behavioural models, interpersonal models, and cultural models as well as models detailing specific processes in greater detail (models at smaller scales) or the interactions between these levels ("linking" models) – whatever is required to comprehensively represent the phenomenon in its entirety.

This approach to explanation is described by Hochstein (2017) as *explanatory holism.* Hochstein identifies a variety of explanatory goals that explanations are often expected to satisfy – for example, conveying understanding, allowing prediction and control, or elucidating causal mechanisms. However, "an individual scientific model cannot simultaneously satisfy *all* the scientific goals typically associated with explanation" (Hochstein, 2017, p. 1107). For instance, such goals are often at odds, such that "to satisfy some goals must come at the expense of satisfying others" (p. 1107). To provide an account that allows for reliable prediction and control, a model must identify general principles or patterns – a process known as *abstraction*. In doing so, however, it sacrifices the degree of detail (specificity) it provides about causal mechanisms. Theorists therefore frequently end up trying to decide which goal is most important or arguing that attainment of any one of the plethora of explanatory goals should be sufficient. However, both solutions require theorists to compromise their explanatory intentions – they must settle for being able to achieve but one or two of their potentially multiple goals. Instead, Hochstein argues that explanations can be created in ways that allow for the achievement of a variety of explanatory aims by distributing the explanatory work across multiple models, such that different models can be selected to satisfy different needs.

A good example of model pluralism in action is Clack and Ward's (2020) compositional "sketch" of the depressive symptom "anhedonia". They demonstrate how this symptom can be modelled independently at different levels of analysis (molecular, neural, cognitive, and phenomenological) and how these different models can be meaningfully related to one other and to aetiological explanations of anhedonia to form a more comprehensive overall explanation of this phenomenon (see Figure 7). This approach to the explanation of mental disorder allows for all relevant factors and processes to be equally represented and sufficiently detailed – thus providing specificity – without overcrowding a single model with information that makes it harder to comprehend – in other words, sacrificing abstraction. The explanation thus becomes much more flexible, as clinicians and

researchers can select out the component models relevant to their tasks yet still have the remaining information available if needed.

Figure 7. *Ward & Clack's (2020) Compositional "Sketch" of the Symptom "Anhedonia" Relating Compositional Sub-Models at Different Levels of Analysis*



Note. Originally captioned: "Example compositional explanation of anhedonia illustrating how central structures/processes are nested together and possible etiological factors.". From "Modeling the symptoms of psychopathology: A pluralistic approach" by T. Ward & S. Clack., 2020, *New Ideas in Psychology*, *59*, 100799.

psychopathological explanations would necessarily be more complex, as the relationships between explanatory models in psychopathology are obviously far more complicated. However, this analogy provides some idea of how such a "bank" could be structured.

I would also argue that models relating to specific scales or levels of analysis should be constructed by those with sufficient expertise and knowledge in that area, as they stand the best chance of accurately and effectively representing those phenomena (see Kendler et al., 2020). This is likely to require a great deal of communication and collaboration between different psychopathological disciplines – such as medicine, neurobiology, physiology, neuroscience, psychology, and sociology – in order to achieve the best representations of the factors and mechanisms involved, as well as their interrelations – in other words, a "coordinated pluralism" (Sullivan, 2017). Such cooperation is easier to achieve through a pluralistic lens, in which it is accepted that multiple models and perspectives hold value and are needed to represent the phenomenon of interest adequately. This more equivalent positioning of models has the potential to reduce model competition or rivalry, and thereby make space for greater connection and conversation between perspectives and disciplines. The generation of the aforementioned "explanatory bank" may also assist with this, as it would provide a platform for the sharing and relating of models, as well as an overarching framework within which these may be situated.

Translating Models into Intervention Strategies

As discussed above, the OBM perspective often promotes a similar approach to treatment, which results in practical problems. In brief, the perceived superiority of some therapeutic models can lead to the entrenched dominance of some interventions (e.g., CBT), rivalry between treatment approaches (e.g., CBT versus ACT or psychotherapy versus pharmacotherapy), and poorly informed integration of different intervention strategies (e.g., CBT and antidepressants). We know that mental disorders are highly complex, multifactorial conditions. Hence, their treatment is likely to be best achieved by a correspondingly broad and diverse approach.

The recommended change in theoretical perspective discussed above already contributes to changing the way that models of mental disorder are translated into clinical interventions. If multiple models are used to account for mental disorders, interventions must necessarily be informed by multiple models – thereby abandoning their previous 1-1 relation – or downsize to address fewer factors or processes but be applied alongside other

intervention strategies – retaining the 1-1 relationship, but eliminating the single-intervention approach. It is likely that a combination of both approaches will be necessary, and both have advantages.

The first may be more manageable and parsimonious – requiring application of just one intervention strategy – but may be less responsive to clients' diverse needs by proposing a single broad intervention across cases. The second, on the other hand, may be more complicated to design and manage – involving the application of multiple, smaller intervention strategies – but may allow for treatment to be more tailored to the clients' concerns and capacities, as well as their specific cultural needs (Holmes et al., 2018). It may also make it easier for interventions to be usefully evaluated, as they would be composed of one or a few elements designed to target a specific mechanism. As Holmes et al. (2018) state: "the more elements that are combined in a psychological treatment, the harder mechanistic specificity is to establish" and "improved precision is likely from assessing the mechanisms of particular procedural elements rather than combinations of elements" (p. 245). Overall, however, the most important thing here is that both approaches recognise and respond to the need for multiple, diverse representations of mental disorder.

It is the hope of this thesis that the shift towards a more pluralistic explanation of mental disorder will also change the focus of intervention strategies. At present, interventions are often directed at DSM/ICD syndromes, based on theoretical models that do the same. Given that these categories are fundamentally flawed, the generation of interventions to treat them is problematic, and may partially contribute to the poor efficacy of many approaches. One of the responses to this has been the advent of "transdiagnostic" interventions that aim to target processes thought to be shared by two or more disorders – such as CBT-E, which targets EDs using a transdiagnostic approach (Fairburn et al., 2003), or transdiagnostic treatments for depression and anxiety (Craske, 2012). Although a promising notion – and, indeed, an efficient one – this approach neglects disorder-*specific* processes and risks homogenising client experiences. For example, it has been estimated that there are at least 227 unique symptom profiles that could qualify for the DSM-5 diagnosis of MDD (Fried & Nesse, 2015). Hence, addressing all of these with the same intervention directed either at this syndrome or at "depressive disorders" more generally seems highly reductive and depersonalised. It is likely that interventions that are more specific to clinical phenomena (or "phenomena complexes", such as the binge-purge cycle) might be more effective, as they would be able to cater for both processes shared across disorders as well as those that are

more disorder-specific and allow greater responsivity to client needs by being tailored to their specific symptomology rather than heterogeneous syndrome profiles.

Applying Models and Interventions in Practice

Although it will take time for new, pluralistic explanations and treatments to be developed, it is possible for clinical practitioners to move towards model pluralism in their existing practice. At the most basic level, all that is required is an acknowledgement that mental disorders are complex, multifactorial problems that can be understood and approached from a variety of perspectives and that, similarly, may have multiple possible solutions. Many clinicians and services already frequently employ multiple interventions or perspectives out of need. However, I would argue that it is worth explicitly committing to pluralism as an underlying theoretical principle that drives the process of assessment and treatment, rather than for it to occur inadvertently as a product of the clinical process. The intention with this would be to foster, at the outset, an open-minded approach to the treatment of mental disorder and discourage therapeutic allegiances or "favouritism" that could impact clinical efficacy.

However, this open-mindedness should also be accompanied by critical evaluation and sound clinical judgement, such that therapeutic models and interventions are objectively appraised and appropriate decisions made about their use. At present, this critical appraisal is somewhat lacking. There is a general scarcity of theoretical literacy within the field of psychotherapy (Ward, 2019), seen in the widespread persistence of some therapeutic models or intervention strategies that are unsupported or theoretically flawed (see Chapters One and Three). The development of sound theoretical understanding and reasoning (*conceptual competence*; Aftab & Waterman, 2020) is greatly undervalued in many higher education programmes related to psychopathology, such as clinical psychology training and counselling certificates, despite it being essential to multiple clinical tasks, such as formulating clients, evaluating emerging research and interventions, and developing coherent treatment plans (Lilienfeld et al., 2013). It is partly this gap in education that promotes the adoption or rejection of therapeutic models and interventions solely based on personal preference or anecdotal reasoning (Ward, 2019) – a process not in line with the philosophy of evidence-based practice (Howick, 2011; Lilienfeld et al., 2013). Enhancing theoretical literacy among psychopathological practitioners is therefore of utmost importance to making sure that current and future therapeutic models are recognised for their strengths and limitations and applied appropriately.

In my view, it would be beneficial for psychopathology students and clinicians to be educated in the following topics:[28]

- Important theoretical concepts – such as *explanations*, *theories*, *models*, *phenomena*, and *data* – their roles, and how they relate to clinical practice (for example, formulation as a form of explanation, symptoms and signs as clinical phenomena; Ward & Clack, 2019).
- Accounts of scientific explanation and related concepts, such as *reductionism, unification, pluralism, mechanistic explanation*, and *explanatory modelling* (Faye, 2014).
- The nature and role of *values* (epistemic, normative, moral, etc.) in psychopathology science and practice (Johnson, 2014).
- Theories of scientific method relevant to the behavioural sciences – such as Haig's (2014) Abductive Theory of Method (ATOM) – and their application to the clinical process (e.g., Ward et al., 2016).
- Critical evaluation of explanatory models and empirical research within psychopathology – particularly psychotherapy research (Coyne & Kok, 2014; Lilienfeld et al., 2014; Mulder et al., 2017).
- The ontology and nosology of mental disorder – including Embodied-Enactivism (Nielsen & Ward, 2011), the RDoC (Insel et al., 2010), Symptom Network Models (Borsboom, 2017), and Mechanistic Property Clusters (Kendler et al., 2011).
- The nature of signs and symptoms (Wilshire et al., 2021).
- Ethical principles, such as *beneficence*, *non-maleficence*, and *justice* (Boyd, 2005; Gillon, 1994), and ethical dilemmas, such as the "dual-relationship problem" (Ward, 2013).
- The philosophies of evidence-based practice (Howick, 2011; Spring & Neville, 2011) and *science-informed practice* (Ward & Wilshire, 2021).
- The role of culture in constituting and explaining mental disorder – especially in countries where there has been colonisation of an indigenous population or with diverse ethnic or cultural composition; Strauss-Hughes et al., 2019) – and how culture can be handled safely in clinical practice (*cultural safety*; Curtis et al., 2019).

[28] A notable example of a clinical training programme that is working to incorporate such content is the Postgraduate Diploma in Clinical Psychology (PGDipClinPsyc) offered by Victoria University of Wellington (New Zealand), which devotes a significant proportion of its class and assessment content to the topics outlined here.

- Strength-based models and approaches to clinical practice, such as the Good Lives Model (GLM; Ward & Brown, 2006).

The combination of the dual mindsets of openness and critical objectivity, in my view, promotes a style of practice that is flexible and responsive to client needs and preferences while also maintaining a robust theoretical and empirical foundation. It is therefore a stronger embodiment of the notion of *evidence-based practice* (Howick, 2011; Spring & Neville, 2011). Such an approach should ideally:

- Be responsive to and informed by multiple theoretical models that consider all levels of analysis relevant to the case at hand, as well as their interaction – in other words, *model pluralism*.
- Employ multiple intervention strategies where feasible and appropriate to the case.
- Carefully consider how interventions may best be combined based on ideas about how they may overlap or interact with each other.
- View each intervention strategy employed as an equal contributor to an overarching, client-specific treatment plan.
- Be constructed based on individual case formulations rather than diagnostic syndromes – for example, emerging "personalised models" of mental health (see Holmes et al., 2018).
- Be mindful of client needs, resources, and preferences – for example, including consideration of client-treatment "match" (see Holmes et al., 2018).

Even within the current climate, where interventions are often perceived as "either/or" (such as ACT and CBT), it is possible to achieve this. Interventions based at different levels of analysis can usually be related in some way and do not have to compete or conflict. For example, as previously discussed, CBT and antidepressant medication – aimed at the psychological and biological levels respectively – can be theoretically related via a shared mechanism of action – namely, the modification of affective information processing (Pringle et al., 2011; Roiser et al., 2012). Even interventions that appear to have different philosophies, such as CBT and ACT, can search for common ground. Are any assumptions or ideas shared? Is there a way to reconcile or compromise where assumptions are different? Do the approaches have common goals or moral philosophies? For example, although CBT and ACT are often presented as alternative psychotherapeutic approaches, some research indicates they may be more similar than they are distinct and are, in fact, fully compatible

models (Arch & Craske, 2008; Hofmann & Asmundson, 2008). CBT and ACT could exert their therapeutic effects on anxiety symptoms via similar mechanisms, despite seeming to approach the client's problem from very different philosophical perspectives (Arch & Craske, 2008). As one example, Arch & Craske (2008) point out that the core goals of CBT and ACT – symptom reduction and "valued living" respectively – oft cited as a critical philosophical disparity between the two approaches, are not mutually exclusive and may actually promote each other:

> In ACT, the emphasis on values-driven behaviour may lead to behavioural exposures and, hence, to eventual anxiety reduction…In CBT, behavioural exposure to feared situations, which likely represent personally valued behaviours, leads to anxiety reduction, which, in turn, results in greater likelihood of engagement in previously avoided, valued behaviours. (Arch & Craske, 2008, p. 271).

This sort of perspective considers how two philosophically distinct models can be integrated or work collaboratively, rather than prima facie assuming their incompatibility.

Summary

Looking broadly at science, rarely is it the case that only one "best" explanation of a phenomenon is available. Nevertheless, for some reason this myth persists in several fields. This chapter examined the problem of the One Best Model (OBM) perspective within psychopathology. The OBM perspective refers to the seeking of unified explanations for mental disorder, in which explanations are equated with single theoretical models. This perspective was identified in Chapter Three as a core issue contributing to the conceptual flaws found within current explanations of disordered eating phenomena.

As this chapter discusses, the assumption that a single model can adequately explain a particular phenomenon often leads to explanations being overly convoluted, and thus unworkable, or (more likely) impoverished, neglecting or omitting relevant factors or processes for the sake of abstraction or simplicity. Furthermore, it contributes to estrangement and rivalry among interventions, as they are typically aligned with explanatory models that are similarly at odds (e.g., CBT versus ACT; Leahy, 2008). Hence, although the notion that a single model will be able to explain one or more mental disorders is inherently attractive, it is arguably a fool's errand, the continued pursuit of which is likely to be a theoretical and practical dead end.

Reflecting on the unity of science, Ruphy (2016) concludes that “theoretical unification, a long-standing quest and hallmark of scientific progress, is no longer seen as desirable across all disciplines” (p. xi). More often than not in science, a number of models are formulated at the same or different levels of analysis, each offering a distinct and valuable perspective on aspects of the world. The danger in assuming that only *one* model is likely to be of value is the creation of a culture of dogmatism and the premature rejection of promising ideas, as discussed above. Furthermore, it may result in a paradoxical situation in which there is, in fact, a multitude of different models each assuming that they are the superior – in other words, the OBM perspective.

Rather than settling for a chaotic marketplace where each candidate model refuses to acknowledge the value of alternative approaches, I argue that a vibrant explanatory pluralism needs to be coordinated and actively encouraged. An advantage of explicitly endorsing such an approach is that there is a better chance of providing a rich understanding of the causes and nature of psychopathology, and hopefully, arriving at more effective and flexible treatment strategies. It is also less likely that researchers will unjustifiably ignore other models and engage in the ultimately unfruitful and self-defeating game of the search for the One Best Model, and, relatedly, the One Best Treatment. In other words, shifting from the OBM perspective to a coordinated model pluralism approach to psychopathological explanation is likely to yield a richer, more comprehensive, more flexible, and ultimately more productive understanding of mental disorders.

Chapter 6. Developing a Methodological Framework for Explanation in Psychopathology

This chapter addresses the final issue identified by the analysis in Chapter Three: the lack of a coherent methodological framework for constructing psychopathological explanations. This issue, although specifically identified within the context of ED explanations, is common to psychopathology as a discipline. At present, there is no established methodological approach for constructing explanations of psychopathological problems. Although researchers may have their own internal or external processes by which they generate explanations, these are currently idiosyncratic and "off the record". Reasoning processes are therefore likely to be highly diverse, and there is no way to know if they are sound. This is a significant problem, as it means current explanations of mental disorder may be constructed using biased or logically flawed reasoning processes, and there is currently no way to hold researchers accountable for this. In other words, explanation in psychopathology sits within a sort of methodological "limbo" – the task of explanation is not widely perceived as being governed by scientific method and, as such, no methodology specific to the task of psychopathological explanation exists. However, in my view, there is no reason why explanation, being a scientific task, should be methodologically exempt.

In this chapter, I therefore discuss the role of method in science – specifically, how it relates to the task of scientific explanation. In doing so, I argue that there is significant need for a methodological framework for constructing explanations of psychopathological problems in order to ensure they are valid, detailed, and robust. As no such framework currently exists, I proceed, in the second half of this chapter, to present a preliminary "sketch" of such an approach – the Phenomena Detection Method of Theory Construction (PDM-TC) – based on existing theoretical tools developed by Ward and Clack (2019) and Haig (2014) for the formulation of psychopathological symptoms and the generation of psychological theory respectively.

The Role of Method in Science

Method is critical to science. Its importance is well articulated by Haig (2014) in his book *Investigating the Psychological World: Scientific Method in the Behavioural Sciences*:

> Although modern science is made up of many parts, scientific method is its centrepiece. The centrality of method to science stems from the fact that it provides scientists with the primary form of guidance in their quest to obtain knowledge about the world. As fallible inquirers, scientists face immense challenges in their efforts to learn about the complexities of nature. In good part, these challenges are met through the use of methods, which provide scientists with the cognitive assistance that they need to undertake successful inquiry. (Haig, 2014, p. ix)

Hence, any inquiry undertaken without the use of systematic method is arguably unscientific. Within science, the concept of *method* refers in general to the "systematic ordering of inquiry" (Haig, 2014, p. 12): procedures or sequences of actions that provide a process for investigating one or more research questions. Method therefore provides the backbone of science, ensuring that the study of a phenomenon is undertaken in a way most likely to yield accurate and enlightening results.

Methods may, of course, vary in quality: for example, case studies are considered less reliable sources of information than RCTs according to the *scientific hierarchy of evidence* (see Evans, 2003). However, whether methodologically good or bad, having an organised procedure for conducting inquiry is ultimately superior to investigation that is unstructured and chaotic. Method provides a common language by which the process of a piece of research can be communicated and justified to other academics. Method can provide information about the reliability and validity of findings, as well as the possibility that confounding variables were actually responsible for an identified effect. Furthermore, method provides a way of structuring the investigative process that may, depending on the quality of the method used, help to avoid or counteract errors or biases in reasoning. Hence, although many different theories of scientific method exist – for example, Bayesian, hypothetico-deductive, and inductive theories of method – and the exact nature of method in science is still debated, there at least seems to be a clear consensus that "method really matters to science" (Haig, 2014, p. 1).

Method and Explanation

To date, there have been few endeavours to develop methods for the generation of overall explanations of psychopathological problems. Indeed, the theoretical space within psychopathology broadly lacks the influence of scientific method. Although some methods exist that can help to evaluate or "test" psychopathological explanations – such as appraisal

relative to epistemic criteria (see Chapter Three) or hypothesis testing (Bayesian or frequentist) – and there are some methods available that can aid in detecting phenomena – for example, exploratory factor analysis (EFA) and symptom network modelling – there is no broader framework to guide how researchers should go about generating the overarching explanations that describe and account for these phenomena in a comprehensive and coherent manner. This represents a significant problem as, without being appropriately governed by method, explanation in psychopathology is arguably unscientific at present.

The lack of methodological attention to explanation is surprising given the emphasis on methods and methodology in other aspects of scientific inquiry. For instance, the hegemonic status of the *hypothetico-deductive* theory of method within science has yielded a large variety of methods geared towards hypothesis testing (Haig, 2014), as well as the expectation that these methods be explicitly outlined in research papers through the conventional inclusion of a Method section in empirical articles. These methods are rigorously designed, critiqued, and acknowledged as integral for obtaining valid results. The task of explanation, however, is not compatible with any of these methods.

Explanation is arguably an *abductive* process of repeatedly reasoning "up" from data to phenomena and connecting them in plausible ways, rather than examining pre-formed hypotheses (Haig, 2014). *Phenomena* are general features of the world that are recurrent and relatively stable (see Bogen & Woodward, 1988; Haig, 2014). These features may come in a variety of forms, such as objects, states, processes, or events; the most well-established often being referred to as "effects" (Haig, 2014). Each represents a reliable pattern that has been observed within a collection of *data*, which, in contrast to phenomena, are "recordings or reports that are perceptually accessible; they are observable and open to public inspection" (Haig, 2014, p. 34). An explanation links the relevant phenomena together to construct a likely account of the causes, consequences, constitution, or context of a particular target phenomenon (the "explanatory target").

Explanation is as important to science as hypothesis testing. Both have critical roles within scientific inquiry: one to generate conceptual understandings of phenomena based on the available evidence, the other to test their legitimacy. However due to the predomination of the hypothetico-deductivist methodology in science, explanation has yet to receive the independent methodological attention it deserves. No explicit methodological frameworks for constructing explanations of psychopathology exist, nor has there been consistent focus

on developing any. Explanation in psychopathology therefore currently exists within a sort of methodological "vacuum", as if it were somehow exempt from scientific method, despite being a critical aspect of scientific inquiry.

The problem with this is that it takes explanation from a scientific task to an exercise in guesswork, in which there is a greater likelihood of bias, error, and oversight. A good example of this is the numerous conceptual issues identified in the ED explanations evaluated in Chapter Three – such as the reliance on problematic DSM syndromes, lack of richness and specificity, and the neglect or omission of relevant factors or levels of analysis. These problems can be inferred as reflecting the mismanagement of cognitive challenges that arose during the explanatory process, challenges that a systematic method could potentially counteract. For example, cognitive biases on the part of the researcher (e.g., attentional, confirmation, availability) could easily lead to relevant factors and processes being neglected or omitted within an explanation. This oversight could be counteracted by application of a systematic methodological approach that advocates a broad and thorough approach to the consideration of data, if only such an approach were available.

A Systematic Methodology for Psychopathological Explanation

At present, a key barrier to explanation in psychopathology being appropriately governed by scientific method is the fact that there is currently no broadly agreed on methodological process for their construction – that is to say, a *methodological framework*. Methodological frameworks are broadly considered to be structures of rules and principles that can be used to guide the execution of a particular scientific task. In other words, "a methodological framework provides structured practical guidance or a tool to guide the user through a process, using stages or a step-by-step approach" (McMeekin et al., 2020, p. 2). Such an overarching structure is highly useful, as it provides clear guidance as to how one should proceed in organising a reasoning process or investigation. Hence, in my view, a critical first step to ameliorating the methodological neglect of explanation within psychopathology is to develop an appropriate methodological framework that can inform the execution of this task. It is this problem that the rest of this chapter addresses.

In this section, I present the Phenomena Detection Method of Theory Construction (PDM-TC), a preliminary methodological framework for the generation of explanations of psychopathological problems. The PDM-TC represents a specific application of Ward and

Clack's (2019) Phenomena Detection Method (PDM) framework to the task of constructing explanations in psychopathology. As such, this chapter first briefly outlines the PDM, as well as Haig's (Haig, 2014) Abductive Theory of Method (ATOM), an important theoretical approach underlying it.

The Phenomena Detection Method (PDM)

The PDM (Ward & Clack, 2019) is a meta-theoretical framework for the detection and modelling of clinical phenomena (symptoms and signs) that is not dependent on problematic classification systems such as the DSM or ICD (discussed in Chapter Four). It has four phases: 1) formulating client complaints and/or accompanying signs, 2) discerning and analysing patterns in data related to these symptoms (detecting *phenomena*), 3) constructing multiple descriptive models of the phenomenon using different levels or units of analysis, and 4) linking in aetiological factors to develop aetiological explanations. It therefore emphasises the importance of detecting valid phenomena for explanation ("explanatory targets") and developing compositional explanations of those phenomena. The PDM also promotes and adheres to *model pluralism* as a core principle. It therefore aligns with the solutions proposed in Chapters Four and Five for the respective problems of poor explanatory target selection and description and the One Best Model (OBM) perspective. With the PDM as its key theoretical foundation, the PDM-TC therefore naturally incorporates these ideas, emphasising the tasks of careful explanatory target selection and comprehensive compositional explanation, as well as promoting the pluralistic explanation of psychopathological phenomena.

The Abductive Theory of Method (ATOM)

The PDM – and, thus, the PDM-TC – draw heavily from Haig's (2014) Abductive Theory of Method (ATOM). ATOM is a methodological framework for the scientific inquiry process more generally within psychology and related behavioural sciences. It takes a problem-focused approach to scientific inquiry, a perspective well-suited to psychopathology within which the primary concerns are the problems of clients (Ward et al., 2016). Crucially, ATOM stresses the importance of theory construction within science and makes a critical distinction between *data* and *phenomena* (see *Method and Explanation* above), which are often incorrectly conflated within scientific inquiry. ATOM comprises two broad phases – *phenomena detection* and *theory construction* – the second of which can be further divided into *theory generation* (the supposition of rudimentary, plausible

explanations), *theory development* (the elaboration of initial theories through analogical modelling), and *theory appraisal* (the evaluation of developed theories in relation to their rivals). The PDM and PDM-TC can be understood as both nested within the ATOM framework – as frameworks for the construction of theories or explanations – and influenced by its principles at each stage – as the processes of phenomena detection, theory generation, theory development, and theory appraisal occur throughout the PDM and PDM-TC.

Introducing the Phenomena Detection Method for Theory Construction (PDM-TC)

The PDM is a meta-theoretical framework for modelling psychopathological symptoms and is therefore somewhat agnostic about how it might be applied to specific tasks within psychopathology, such as formulation, explanation, or classification. The PDM-TC therefore represents the application of this framework to the specific task of explaining psychopathological phenomena. This results in phases 1 and 2 of the PDM being collapsed into a single phase – "selecting an appropriate explanatory target" – and phases 3 and 4 being revised respectively as phase 2 – "developing a rich compositional explanation" – and phase 3 – "developing a quality aetiological explanation". These phases are roughly sequential, the tasks of each essential to successfully executing those that follow. However, it may be necessary to return to and revise the work of the earlier phases in response to insights gleaned during the latter – none should at any point be considered finalised, but rather always open to ongoing revision and improvement. Indeed, the phases may at times be worked on concurrently, developing symbiotically as each provides insight into the others.

This section presents the phases of the PDM-TC in sequence, outlining the reasoning process advocated by each phase and how it can ameliorate the issues with that aspect of explanation in psychopathology (in other words, the issues identified in Chapter Three). A summary of the tasks pertaining to each phase and their advantages for the explanation of psychopathology is presented in Table 2. It should be noted that, although the PDM-TC is presented as a guiding methodological framework for psychopathological explanation, it is still in its infancy and requires further development to be considered a fully-fledged methodological approach. Nevertheless, in its current form the PDM-TC still arguably has value for researchers as it may begin to help guide their theoretical reasoning, as well as drawing attention to the current methodological deficits within

psychopathological explanation and thereby perhaps inspiring the development of additional methods.

Table 2. *Phases of the PDM-TC, Including Their Essential Tasks and Advantages*

Phase	Tasks	Advantages
Selecting an appropriate explanatory target	1. Identification of an area of interest 2. Broad gathering of data relevant to the area 3. Detection of multiple, valid phenomena 4. Selection of a single phenomenon to serve as an appropriate explanatory target	• Not reliant on DSM syndromes • Prevents relevant phenomena being overlooked and mistaken assumption of false phenomena • Increases the robustness of detected phenomena through grounding in empirical evidence • Increases attention to core features by advocating for the importance of centrality as a selection criterion • Selects more robust, reliable, and central targets for explanation
Developing a rich compositional description	1. Broad gathering of data relevant to the constitution of the target phenomenon 2. Detection of multiple, valid, "compositional phenomena" within the target 3. Modelling of temporal, causal, and spatial relations between compositional phenomena using multiple models 4. Evaluation and revision of individual models and the collective "model space"	• Prevents relevant phenomena being overlooked and mistaken assumption of false phenomena • Increases the robustness of detected phenomena through grounding in empirical evidence • Emphasises the value of compositional descriptions • Increases richness and specificity and prevents the privileging or neglect of levels of analysis through the construction of multiple models • Encourages evaluation and improvement of the models created
Developing a quality causal explanation	1. Broad gathering data relevant to the aetiology of the target phenomenon 2. Detection of multiple, valid, causally relevant processes and structures (i.e. aetiological phenomena) 3. Modelling of aetiological pathways to the target phenomenon, informed by the identified aetiological phenomena and using multiple models 4. Evaluation and revision of individual models and the collective "model space"	• Prevents relevant phenomena being overlooked and mistaken assumption of false phenomena • Increases the robustness of the detected phenomena through grounding in empirical evidence • Increases richness and specificity and prevents the privileging or neglect of levels of analysis through the construction of multiple models • Encourages evaluation and improvement of the models created

Phase 1: Selecting an Appropriate Explanatory Target.

Effective structuring of the problem space begins with the selection of an appropriate explanatory target. As outlined in Chapter Four, an explanatory target is the phenomenon that a theory, model, or explanation claims to explain. All explanations, by design, have an explanatory target. The issue with current explanations in psychopathology, is not that they lack explanatory targets in toto, but that their explanatory targets are inappropriate, generally lacking validity and/or reliability. As discussed in Chapter Four, current psychopathological explanations are typically built around DSM syndromes, such as AN and BN – which are widely recognised to have extensive conceptual flaws (Lilienfeld & Treadway, 2016; Nielsen & Ward, 2018, 2020a; Whooley, 2014; Zachar & Kendler, 2017) – or transdiagnostic categories, such as "eating disorders" – which represent extremely heterogeneous and unreliable targets. Both approaches set any subsequent explanation up for failure, either causing it to have poor validity or to be significantly underspecified. The ongoing use of these constructs in research and practice also contributes to their reification and further entrenchment within the psychopathology field. I therefore suggested in Chapter Four that psychopathological explanations shift their focus to clinical phenomena (clinical signs and symptoms or symptom "complexes"), at least as a starting point, as these provide more reliable and valid foundations for explanation. The PDM-TC outlines a logical process for the selection of these targets.

First, the researcher should identify their broad area of interest – for example, "problems with eating behaviour" or "problems with mood".

Second, they should engage in systematic data gathering to gain a general sense of the evidential landscape for that area. Data relevant to this stage may include clinical observations, case descriptions, and epidemiological studies. The intention here is to cast a wide net and, thereby, prevent a premature narrowing of focus that may cause relevant information to be overlooked. Furthermore, broadly reviewing the data in an area of interest means that the target selected for explanation will be firmly rooted in existing evidence, which makes it more likely to be valid.

Third, clinical phenomena should be detected – in other words, the available data should be mined for robust patterns that represent real-world states, effects, or processes (Bogen & Woodward, 1988; Haig, 2014). As discussed in Chapter Four, these may be anything from common clinical signs or symptoms (e.g., "binge eating", "purging"), to

components of these symptoms (e.g., “loss of control over eating”, “excessive caloric intake”, “self-induced vomiting”, “laxative abuse”), to collections of symptoms and signs that are considered to represent a single complex phenomenon (e.g., self-sustaining processes, such as “the binge-purge cycle”; Nielsen & Ward, 2020a). The phenomena most relevant to this initial stage are those representing core features of the area of interest (Ward & Clack, 2018). For example, if the area of interest were disordered eating, “the binge-purge cycle” would represent a more relevant explanatory target than, say, the symptom “anhedonia”.

Fourth, the researcher should evaluate the identified phenomena and select the most constructive explanatory target. This should be that which represents the best balance of *robustness*, *reliability*, and *centrality/relevance*.

Robustness refers to how well the phenomenon is evidenced by the data – for example, are there multiple pieces of data suggesting the same phenomenon (*triangulation*)? Is the data behind the phenomenon of good methodological quality (does it demonstrate a sufficient sample size, valid measures, and appropriate statistical analysis)?

Reliability refers to how consistent the phenomenon is – for example, is the phenomenon, coherent as a construct? Is the phenomenon relatively stable or transient? How consistently does the phenomenon present across individuals or contexts?

Centrality/relevance refers to how characteristic or important the phenomenon is considered to be to the relevant problem area – for example, how consistently do cases in the area of interest present with the phenomenon? How problematic is the phenomenon within that context? Does it occur outside of the area of interest/how specific is it to the problem?

Arguably, the most appropriate explanatory target will be the phenomenon which represents the best balance of these criteria, which requires theoretical judgement by the researcher. In eating disorders, for example, “excessive caloric intake” is arguably a less heterogeneous (and therefore more reliable) phenomenon than the symptom “purging”, which comprises multiple behavioural variants, such as vomiting and laxative/diuretic abuse. However, it is also less specific (and therefore less central) to the eating disorders area, as it commonly occurs among the general population, both in pathological presentations, such as obesity, and non-pathological presentations, such as on celebratory

occasions. Hence, in this case one could argue that the high centrality of purging to the eating disorder area would outweigh its poorer reliability and make it the preferred choice.

It is possible that the most appropriate explanatory targets in psychopathology will be the aforementioned "symptom complexes" – such as the "binge-purge cycle" in disordered eating or the *experiential avoidance* cycle in NSSI – as they have been robustly evidenced by the literature, represent reliable patterns of thought, mood, and behaviour, and are highly characteristic of and specific to the relevant problem areas. It may also be the case that several phenomena emerge as balancing the evaluative criteria to a relatively equal degree, in which case the researcher may select one at will to tackle first. In my view, by selecting the phenomenon that is, on balance, the most robust, reliable, and central to the problem, the researcher is better placed to construct an explanation of high theoretical quality that holds value for the area of interest.

Phase 2: Developing a Rich Compositional Explanation.

Once an appropriate explanatory target has been selected, the next step is to explain the composition of that phenomenon – in other words, to describe how it is constituted. The nature and role of compositional explanations, as well as their current lack within psychopathology and the problems that entails, have already been comprehensively outlined in Chapter Four. In brief, compositional explanations are highly valuable as they provide information about how psychopathological phenomena are structured, which includes information about their constituent components and how those are maintained. They therefore have great utility for both the generation of aetiological explanation and clinical assessment and intervention for mental disorder. To emphasise the important role of these explanations, therefore, the PDM-TC designates an entire phase for the construction of a sufficient compositional explanation. It also incorporates the pluralistic approach to explanation proposed in Chapter Five, as it prescribes that compositional explanations be generated by the construction of multiple models.

To build a rich compositional description of their chosen explanatory target phenomenon, the PDM-TC designates that the researcher should first gather data about the explanatory target by examining as much of the relevant research evidence as possible across different levels such as the biological, cognitive, emotional, behavioural, interpersonal, sociocultural, and political. From this information, the researcher can then identify smaller-scale phenomena that exist within the explanatory target at each level of

analysis ("compositional phenomena"). It is again important here to consider the quality of the evidence available to ensure the phenomena identified are robust.[29] As discussed previously, phenomena evidenced only by poor quality research evidence – such as that involving low sample sizes or comparison groups that are inappropriate or absent – cannot be reliably attested as phenomena at all, as they may reflect flawed scientific practice rather than genuine features of the world.

Once these relevant "compositional phenomena" have been isolated, the task becomes that of modelling the temporal, spatial and causal relations between them, both at and between levels of analysis, that can be reliably inferred from the literature.[30] First, initial, rudimentary plausible models should be generated, with influence from current empirical research. These models should then be comprehensively elaborated by exploring and richly detailing the possible structures and processes involved (Ward & Clack, 2018). This may involve specific theory development strategies such as analogical modelling (Haig, 2014). The end goal is the creation of a collection of models at different levels of abstraction that "hang together" to comprehensively describe the structure of the explanatory target (Hochstein, 2016; Ward & Clack, 2018). For example, a comprehensive compositional account of "binge eating" may involve physiological models detailing altered appetite or satiety, neural models involving dysfunctions in self-control, mood, and anxiety/stress networks, molecular models depicting changes in neurotransmitter functioning, cognitive models describing dysfunctional beliefs, perceptive and attentional biases, or poor emotion regulation/management, and socio-political models citing cultural values or industrial influences (see Chapter Four).

As discussed in Chapter Five, the model pluralism approach to explanation helps to prevent the privileging or omission of levels of analysis, as each level can be attended to by one or more models, each of which is acknowledged as playing an important role in the target's overall constitution. The reduced descriptive load on each model also allows for greater specificity and richness in the explanation, as it creates space to elaborate on the relevant phenomena and processes. Furthermore, the overall explanation becomes more

[29] A good guide for the evaluation of evidence is the generally recognised scientific hierarchy of evidence – namely, meta-analyses and systematic reviews, followed by randomised control trials, all the way down to single case studies (Evans, 2003).
[30] This involves the ATOM stages of *theory generation* and *theory development* (Haig, 2014; see above).

workable, as each piece of the whole can be considered separately and understood in detail before being situated within the larger theoretical picture.

Finally, the compositional models developed should be evaluated and revised both individually and collectively to ensure that they provide an account of the explanatory target that is rich, logical, and coherent.[31] One way that researchers might go about this is by applying evaluative models such as Thagard's (1989) Theory of Explanatory Coherence (TEC), a method by which researchers may appraise competing theories relative to seven evaluative criteria (see Chapter Two). Such evaluation is necessary to ensure that the description provides a sound foundation for the subsequent construction of aetiological models. If significant epistemic flaws are identified within models, the researcher is able to go back and revise them to improve their explanatory worth. However, no model should ever be considered "finished", as empirical research is ongoing: there will always be new ways that models can be adapted or modified in response to emerging evidence.

Phase 3: Developing a Quality Aetiological Explanation.

Once a rich compositional description of the explanatory target exists, work may begin on the development of aetiological explanations. In psychopathology, this is where much of our interest lies: in distilling the causes behind disordered states so we might understand how they came to be. Current aetiological explanations in psychopathology possess significant conceptual flaws, such as those identified with the ED explanations evaluated in Chapter Three. It is worth noting that the lack of effective compositional explanations within psychopathology is likely partly to blame for this, as compositional explanation helps to scaffold subsequent aetiological explanation (see Chapter Four). The absence of good compositional models therefore makes the generation of quality aetiological models very difficult: the target to be explained is less well understood and there are fewer identified constitutional factors and processes to aid in the detection and linking of relevant aetiological factors (Ward & Clack, 2019). Hence, the development of effective compositional explanations of psychopathological phenomena advocated by the previous phase of the PDM-TC is likely to help greatly in the construction of robust aetiological explanations.

[31] As per the ATOM stage of *theory appraisal* (Haig, 2014; see above).

The PDM-TC process for constructing aetiological explanations is similar to that outlined above for developing compositional models. First, a broad range of aetiologically relevant data – such as longitudinal or experimental research – should be gathered. This is, again, important to prevent both premature foreclosure in the researcher's reasoning about aetiological processes as well as the possibility of spurious causal claims being made. From this collection of data, aetiologically relevant factors and processes ("aetiological phenomena") should be detected. It may be helpful to cluster these roughly into those that are more distal to the instigation of the explanatory target and those that are more proximal, to begin establishing their temporal role.

Next, multiple aetiological models should be constructed at a range of spatial scales and levels of abstraction to create a rich understanding of the explanatory target's aetiology (Hochstein, 2016; Mitchell, 2002; Ward & Clack, 2018). For example, one may construct multiple models detailing the biological pathways or psychological pathways to the onset of a disorder, as well as models depicting how these pathways interact with each other over time. This process, again, will involve more rudimentary models being proposed initially, and then fleshed out over time using theory development strategies (such as *analogical modelling*; Haig 2014).[32] In the end, the aetiology of the explanatory target should be accounted for by a network of multiple detailed representational models, which, as with compositional models, should be continually evaluated, with revisions and improvements undertaken where necessary.[33]

As with compositional explanation, the PDM-TC has the potential to prevent the existing issues with aetiological explanations of psychopathology – namely, its model pluralism approach allows for greater specificity and works to prevent the privileging and/or omission of levels of analysis. As discussed in Chapter Five, many existing explanations may hold value here and should not be discarded prima facie. Some may be able to be revised and improved by retroactively applying the PDM-TC framework. For example, Rieger et al.'s (2010) Interpersonal Theory of Disordered Eating could be reworked as a model of disordered eating that specifically denotes the interpersonal level of analysis or, alternatively, as a model representing the interaction of interpersonal and psychological factors. It could even be subdivided into models detailing the interpersonal,

[32] i.e., *theory generation* and *theory development* stages (Haig, 2014; see above).
[33] Again, as per the ATOM stage of *theory appraisal* (Haig, 2014; see above).

cognitive, and emotional processes at play, along with models representing the interactions between them.

Summary

Our current processes of explaining psychopathological phenomena have thus far produced unsatisfactory results – namely, explanations that contain numerous conceptual issues (see Chapter Three) and associated treatments that demonstrate inadequate efficacy (see Chapter One). This can, at least in part, be traced back to the fact that there is currently no methodological process to guide the construction of explanations in psychopathology. It is therefore not unlikely that current explanations are being developed in ways that are unsound. In order to address this issue, the current chapter presented the Phenomena Detection Method of Theory Construction (PDM-TC) as a preliminary methodological framework for the construction of psychopathological explanations. The PDM-TC consists of a three-phase process – 1) the selection of an appropriate explanatory target, 2) the development of a rich compositional explanation, and 3) the development of a quality aetiological explanation – each of which involves specific tasks to aid in the achievement of its objective. Hence, the PDM-TC endorses and incorporates the solutions proposed for the issues discussed in Chapters Four and Five; advocating for careful selection of appropriate explanatory targets (namely, clinical phenomena), rich compositional description of phenomena, and explanation using multiple models (in other words, *model pluralism*).

As well as working to address these "bigger picture" issues, the PDM-TC also has the potential to ameliorate the specific conceptual issues identified with psychopathological explanations in Chapter Three, through its emphasis on broad data consideration, phenomena detection, model pluralism, and model appraisal and revision. Such principles seek to ensure that avenues of inquiry are not prematurely closed, factors and processes identified are valid and sufficiently detailed, that no relevant levels of analysis are omitted from the overall explanation, and that the explanation is modified in response to emerging research or theoretical critique. Although the specifics of the PDM-TC require further development, even in its draft stage this framework can begin to guide the reasoning of researchers seeking to develop psychopathological explanations. Applying systematic method approaches like the PDM-TC to such a task will, in my view, promote construction of explanations of disorder phenomena that are richer and more robust, as well as more accountable for their reasoning. Such explanations are likely to allow for greater scientific

understanding of psychopathological problems and, thereby, a better chance of treating them effectively in practice.

Chapter 7. Conclusion

Clinicians working in EDs have an ethical responsibility to represent clients' concerns accurately and to treat them as effectively as they can. Similarly, researchers, have both an ethical and epistemic responsibility to generate reliable knowledge using valid and rigorous methods. At present, however, these ethical and epistemic commitments are not being sufficiently met. Despite efforts to ameliorate these conditions, EDs remain a significant global problem, negatively affecting individuals, families, and communities. Classification, diagnosis, and treatment of EDs is inadequate, leaving many patients symptomatic and disillusioned. Research, both theoretical and empirical, suffers from substantial conceptual and methodological deficiencies that question the validity of their hypotheses and outputs, and generally impede our understanding of this area. This thesis set out to investigate a potential contributor to this situation: ED explanations. Namely, whether they are of sufficient theoretical quality to provide us with understandings of EDs that can facilitate valid and effective research and practice. As it turns out, they are not, and we can do better.

In this chapter, I will briefly review the content of this thesis. I will progress from the "eating disorder problem" initially presented, through my evaluation of current ED explanations and the solutions I proposed for the issues identified with them. I will then move to discussing some of the clinical implications that these findings have for classification and diagnosis, clinical research, and therapeutic intervention, both for EDs specifically, and for psychopathological problems more generally. Following this, I briefly touch on some of the obstacles that might stand in the way of the ideas in this thesis being disseminated or implemented in research and practice, such as the entrenchment of current meta-theoretical and nosological perspectives and the cognitive challenge of contending with a more complex theoretical approach. Finally, I conclude by discussing the steps that I believe need to be taken moving forward, both in terms of the concrete application of the content of this thesis to ED explanations, as well as the development of approaches in related domains such as clinical research, intervention development and implementation, and classification/diagnosis that can complement and facilitate the significant theoretical "shifts" proposed herein.

CONCLUSION

The Thesis in Brief

At the outset of this thesis, I outlined the "eating disorder problem": that considering the serious nature and growing prevalence of EDs, our available treatments are just not good enough, providing only modest relief and often failing to prevent relapse later in life (Berkman et al., 2007; Eddy et al., 2017; Hay, 2020; Keel & Brown, 2010; Linardon & Wade, 2018; Murray, 2020). Although there are likely a variety of factors responsible for this – such as poor treatment fidelity, insufficient treatment duration, poor client adherence – one contributor yet to be comprehensively studied is the quality of the theoretical explanations upon which ED interventions are based. As discussed in Chapter Two, these models play a critical role in a variety of tasks related to eating disorders, such as clinical research, diagnosis/classification, assessment, formulation, and of course intervention. Theoretical explanations have bearing on what aspects of mental disorders we choose to investigate, what ways we decide are best to organise them, what characteristics and processes we are vigilant for or curious about in assessment, which causal and maintaining processes we invoke in formulations, and how we seek to target these both before and after a pattern of disorder has emerged. Hence, if these explanations are of low theoretical quality – if they are lacking sufficient conceptual validity, empirical support, and detailing of factors and processes – they will provide poor platforms for these tasks.

To investigate whether inferior explanations of disordered eating may indeed contribute to the current state of treatment in this area, this thesis evaluated the four explanations behind the current best-practice ED treatments: Minuchin et al.'s (1975, 1978) Psychosomatic Family model (basis of FBT; Le Grange & Rienecke, 2018), Transdiagnostic Cognitive-Behavioural Theory (foundation of CBT-E; Fairburn et al., 2003), the Interpersonal Theory of Disordered Eating (behind IPT-E; Rieger et al., 2010), and Affect Dysregulation Theory (associated with DBT-E; Chen et al., 2018). Upon review, each of these explanations was found to possess significant theoretical flaws likely to contribute to the poor efficacy of ED treatment. Although each model had its own specific shortcomings, across the explanations three common meta-theoretical problems were identified: a tendency for models (1) to be built around conceptually problematic explanatory targets, such as DSM/ICD categories, which they fail to sufficiently describe, (2) to try to explain a phenomenon completely using a single model and typically sacrifice relevant description of relevant detail or entire levels of analysis in order to do so – in other words, the One Best Model (OBM) perspective – and (3) to lack a valid methodological approach behind their

construction, being built in ways that are idiosyncratic and undefined, such that there is no way to verify or critique them. The remainder of the thesis therefore focused on how these problems might be ameliorated.

In Chapter Four, the role and current nature of explanatory targets was discussed. At present, explanatory targets in psychopathology typically consist of DSM/ICD syndromes or transdiagnostic classes, both of which entail significant validity and reliability issues. It was argued that the explanatory focus in psychopathology should be reoriented (at least to begin with) towards the explanation of clinical phenomena – symptoms, signs or reliable complexes of both, such as behavioural cycles – which represent a more valid and finer-grained focus that suitably restrains the scope of explanation. Such targets are also likely to be easier to explain compositionally, as covered in the second half of this chapter. In the second half of Chapter Four, the importance of compositional explanation – the comprehensive description of the constitution of a phenomenon – was highlighted, and it was argued through evaluation of current descriptive approaches that this important explanatory task is being neglected within psychopathology. The exact nature of symptoms and signs is currently poorly understood (Wilshire et al., 2021), as research quickly leaps to explanation, rather than first seeking to comprehensively *describe* that which it seeks to explain (the "explanatory target"). At the close of this chapter, it was illustrated how rich accounts of psychopathological symptoms could be generated by comprehensively mining available research data to construct a pluralistic network of descriptive models, each detailing factors and processes at different levels and scales of abstraction. This was done by sketching out how compositional explanation of the ED symptom "binge eating" might proceed, by identifying factors and processes relevant at each level of analysis and then creating detailed sub-models of these phenomena and how they interrelate.

Chapter Five dealt with the concept of explanatory pluralism in more detail, arguing that adequate understanding of psychopathology requires application of a variety of theories and models that work together to detail all relevant aspects of a phenomenon's causes and composition at multiple levels or scales. In this chapter it was argued that the psychopathology field currently conforms to the idea that a single model should be able to provide a complete explanation of either one or more than one presentation/s of mental distress – a philosophy I termed the One Best Model (OBM) perspective. The OBM can be seen reflected in the rivalry often observed in research and practice between different schools of clinical intervention, such as ACT and CBT (Leahy, 2008). Due to the epistemic distance

between the theoretical models underpinning these interventions, there is little practical consideration of how they may overlap or complement each other, and some clinicians end up pledging themselves to particular "therapeutic camps" due to this theoretical enmity.

Although the idea of uniting the various aspects of a mental disorder within a single model is inherently attractive – being much more parsimonious and manageable than juggling multiple models – there is arguably no way to understand the complexity of mental distress sufficiently within the confines of a single model. To resolve this problem, a significant shift was proposed, from this fruitless pursuit of single-model unification to a model pluralism approach in which psychopathological phenomena are explained using multiple models, each of which represents the phenomenon at a different level or scale. This approach allows for more detailed understanding of all the factors and processes involved in causing and constituting mental disorder, and thus a greater comprehension of our clients' problems and how we might approach their amelioration. I indicated in the second half of this chapter how we might begin to apply model pluralism within psychopathology – namely, by generating explanatory networks of mental disorder composed of multiple models at different scales, shifting the focus of intervention strategies so that they are no longer syndrome-oriented or held captive behind large paywalls, and increasing the theoretical knowledge of clinical practitioners so that they can better evaluate intervention models and apply them judiciously in practice.

Finally, Chapter Six developed and presented a multi-phase methodological framework for constructing psychopathological explanations – the Phenomena Detection Method of Theory Construction (PDM-TC). The method, grounded in Haig's (2014) Abductive Theory of Method and Ward and Clack's (2019) Phenomena Detection Method, consists of three broad phases that guide the explanation of psychopathological phenomena: (1) selecting an appropriate explanatory target, (2) developing a rich compositional explanation, and (3) developing a quality aetiological explanation. Each phase comprises a series of smaller tasks that guide the researcher to its successful accomplishment. For example, the selection of an appropriate explanatory target in Phase 1 begins with the identification of an area of interest, such as "disordered eating", followed by broad gathering of data relevant to the area, such as clinical observations, case descriptions, or epidemiological studies. This data is then used to help the research detect a variety of valid phenomena – robust patterns in the data that may represent individual signs and symptoms such as "binge eating" or "low mood", or reliable "complexes" of these, such as the "binge-

purge cycle". Finally, the researcher must make a judgement about which of these clinical phenomena to focus on explaining first, based on which seems to be the most valid, reliable, and central to the problem/area of interest. Possessing a systematic process like this for the explanation of psychopathological problems is highly valuable, as it provides a structure that can help researchers to scaffold their theoretical reasoning and reduce logical error. Furthermore, the PDM-TC incorporates the solutions proposed in the previous chapters – shifting explanatory targets away from DSM syndromes, emphasising the role of compositional explanation, and promoting the construction of multiple compositional and aetiological models (in other words, *model pluralism*), making it a useful guide for applying these meta-theoretical points to the development of actual explanations.

Clinical Implications

So, how do these "solutions" specifically address the problem identified at the beginning of this thesis? In other words, how might they help to solve the "eating disorder problem"?

Explanation is the foundation for ED practice. Our theoretical understanding of disordered eating problems directly influences how we avoid these conditions occurring or restore those who have developed them to their relative state of health or, ideally, to an even better one. At present, lack of understanding of these conditions – their causes and structure – is arguably causing them to proliferate and persist. In other words, we do not know enough about what causes and maintains disordered eating problems to be able to help those who suffer from them effectively and prevent others from becoming unwell. We need better explanations.

Each of the meta-theoretical solutions discussed in the previous chapters advocates for a significant shift in the way we approach the explanation of psychopathological problems: from a DSM-driven, conceptually "thin", single-model perspective, in which explanation is seen as a methodologically exempt process, to a more descriptively rich pluralistic approach scaffolded by scientific method. Overall, this aims to facilitate explanations of EDs and other psychopathological problems that provide a richer and more detailed understanding of how these conditions are constituted, initiated, and maintained. The development of such an understanding is critical, as it is central to our ability to prevent and treat these conditions effectively, and to thereby solve the "ED problem".

Classification and Diagnosis

At present, ED research and practice largely revolves around DSM/ICD syndromes – AN, BN, and BED. This is problematic due to the conceptual flaws possessed by these categories, such as their criterial similarity, poor specificity, and instability over time. These categories have become so entrenched in research and practice that it is now difficult to avoid them, despite their recognised flaws. Transitioning to the explanation of individual symptoms – in other words, *clinical phenomena* – is likely to help break this problematic cycle by loosening the hold of DSM/ICD categories and allowing room for alternative taxonomies to be proposed. At present most explanatory models focus on DSM/ICD syndromes (see Chapter Three). This, in turn, prompts empirical studies to be constructed around these categories, which in turn are used to inform the development of future explanations and thus, the vicious cycle of entrenchment persists. In order for change to occur, this process has to be disrupted at some stage. This is something that symptom-focused explanation offers: the focus of explanatory models is redirected away from diagnostic categories, providing fresh hypotheses and predictions for empirical testing that do not revolve around DSM/ICD classifications. This can then be used to inform alternative theories and interventions – thereby beginning to slowly dismantle the DSM monopoly.

Explanation focused on clinical phenomena (symptoms and signs) is also likely to help facilitate the development of alternative classification schemes. When the majority of ED research and theory is concentrated around DSM/ICD categories, as it is at present, it is much harder to generate radically divergent taxonomies, as the very data one would use to inform their construction is so deeply embedded in these existing classifications. Furthermore, the lack of understanding of ED symptoms and signs engendered by focusing on the syndrome level makes it much more difficult to conceptualise alternative ways of organising them. Explanatory work on symptoms is therefore likely to lead to theoretical and empirical data that can provide fertile ground for novel taxonomies. For instance, if we have more detail about what comprises ED symptoms, we might be able to isolate sub-types or variants of symptoms, or symptoms that can be meaningfully collapsed, that allow for more precise and nuanced organisation. For example, ED behaviours can be viewed as serving different functions – such as security, emotional avoidance, or communication of distress (Nordbø et al., 2006) – that could theoretically be linked to specific dysfunctions in underlying biological and psychological systems, such as HPA axis overactivity, nervous system hypersensitivity, emotional dysregulation, or social skills deficits. A classification

scheme based on this may classify presentations of the same ED symptom under different functional sub-types – for example, "communication-oriented" or "emotion regulatory" – or cause superficially dissimilar ED behaviours to be filed under the same category in some cases – for example, classifying some cases of self-starvation and binge eating both as strategies for avoiding emotional experience. It may even be that by furthering our understanding of the makeup and causes of ED symptoms we will eventually be able to generate ED classifications that are pathogenetic rather than descriptive, and thus more diagnostically useful.

Finally, the theoretical commitment to model pluralism within explanation expounded by this thesis has the potential to promote a corresponding approach to classification: a *taxonomic pluralism* in which multiple methods of classifying EDs are recognised and incorporated, each dedicated to the achievement of particular epistemic or practical goals (Carter et al., 2020; Sullivan, 2017). Like explanation, classification is a task that can be undertaken with many different goals or interests in mind. It is very difficult to meet all of these within a single classification scheme without a loss of taxonomic utility, as the style of taxonomic organisation most useful for achieving one aim may hold little value for the realisation of another. For example, organisation of individuals according to symptom presentation, such as syndromes, may be most useful for communication amongst clinicians or agencies where the primary goal is the succinct transmission of client presentation. However, such classification is poorly suited for treatment allocation, for which grouping individuals by factors relevant to intervention, such as causal or compositional mechanisms, is likely to be more helpful. The recognition and acceptance of multiple methods of classification is more appropriate to this context, and likely to be more productive. An explanatory approach that thoroughly models ED signs and symptoms and encourages the construction of multiple models to explain them promotes this in several ways. First, by making room for alternative ED classifications besides the DSM, as discussed above. Second, by generally modelling the application of a pluralistic approach to understanding EDs. And, finally, by generating a wealth of ED models and related empirical data that can be used to inform multiple novel taxonomies of disordered eating.

Clinical Research

As discussed above, most ED research currently focuses on studying DSM/ICD syndromes: their similarities, differences, correlates, and causes. Given the problems

identified with these categories, this is highly limiting for our understanding of EDs: when we study these syndromes, we are arguably studying pseudo-phenomena rather than disordered eating problems as they genuinely exist in the world. As one example, the focus on researching the three core ED syndromes – AN, BN, and BED – has led to a significant lack of understanding of ED cases that fall into the Other Specified Feeding or Eating Disorder (OSFED) and Unspecified Feeding or Eating Disorder (UFED) categories. These categories are extremely heterogeneous, and therefore very difficult to study meaningfully. Hence, they are frequently omitted from ED research. This focus on DSM/ICD categories is partly a function of ED explanations being similarly oriented, as much empirical research aims to test or explore existing theoretical models. It therefore follows that if most explanatory theories are focused on syndromes, the majority of empirical studies will be too. Shifting our explanatory focus to clinical phenomena, as advocated by this thesis, therefore has the power to similarly shift our research focus, promoting more investigation of individual signs and symptoms.

This thesis also advocated for more detailed modelling of the composition and aetiology of clinical phenomena, such as ED signs and symptoms like “binge eating”. This has the power to be highly fruitful for the research sphere: having both a greater number of theoretical models and a greater level of detail within those models provides researchers with an abundance of hypotheses to test. These hypotheses will also be more precisely defined, making it easier to design studies to test them effectively. Additional exploratory research is also likely to be promoted: the high level of detail and specificity prescribed by this approach is likely to expose a plethora of topics in need of further investigation. For instance, identifying that many constructs within EDs (such as “body image disturbance” or “mood intolerance”; see Chapters Three and Four) are significantly underspecified promotes additional research into the nature of these phenomena.

Intervention

As discussed earlier in this thesis, intervention for EDs critically relies on the explanations we have for them, as interventions are generally designed (at least initially) based on some explanation of what is causing or maintaining the problem of interest (see Chapter Two). Hence, the more accurate and detailed our understanding of the causes and composition of EDs, the more likely we are to develop interventions that alleviate these problems effectively and efficiently. The meta-theoretical approach put forward in this thesis

promotes such an understanding. Explaining EDs via multiple models at a variety of scales – each prescribing a high level of specificity – has the power to generate a much richer understanding of ED aetiology and structure. We will therefore be much more likely to identify the processes that are genuinely responsible for the onset and persistence of these conditions, and to have a more precise idea of *how* these mechanisms operate, giving us a better chance of designing interventions that can successfully target them.

A pluralistic approach to explanation also better aligns with how ED treatment typically operates, seldom consisting of a single intervention, but rather a variety of strategies or programmes that may be delivered by multiple practitioners or organisations. Recognition of the need for multiple models has the power to reduce rivalry between therapeutic models and promote greater consideration of how they may interrelate, and thereby how their associated interventions might affect each other (see Chapter Five). This has the power to provide much needed guidance to clinicians about how alternative therapeutic strategies might best be combined in practice. For example, the research discussed in Chapter Five that proposed theoretical similarity and correspondence between Cognitive Behavioural Therapy (CBT) and Acceptance and Commitment Therapy (ACT; Arch & Craske, 2008; Hofmann & Asmundson, 2008) suggests that these two interventions may be able to be successfully applied in tandem or substituted for one another depending on client preference. Arch and Craske (2008) indicate that ACT's focus on maximising "valued living" and CBT's focus on reducing problematic symptoms may actually promote each other, maximising value-driven behaviour leading to symptom reduction via exposure, and symptom reduction allowing for increased engagement in valued behaviour. Hence, clinicians could select the therapeutic approach that best matches the client's goals and attitudes and still promote the goals and values of both therapeutic schools. Some clients may be most interested in maximising their engagement in valued activities, whereas others may be most focused on reducing their distressing symptoms, and some may personally resonate more with the philosophy of one approach over the other.

Challenges and Limitations

One of the advantages of adopting a meta-theoretical perspective is that one can be somewhat idealistic about how a particular process or field should look or operate. However, in reality things are less clear-cut and there can be many obstacles to applying these ideas in practice. Although offering many benefits for both the ED field and psychopathology in

general, the solutions generated by this thesis are by no means simple to achieve. Each requires a significant perspective shift within the psychopathology domain – in how we think or how we practise. They will therefore be multiple challenges to their implementation.

Changing Minds and Breaking Habits

Primary amongst these is the entrenchment of the current ways of thinking and practising around mental disorder. The DSM has reigned supreme in psychology and psychiatry for a long time, being the foundation that most explanations and interventions are built on. Although originally intended as a temporary placeholder for the development of a better classification system, the DSM has come to influence the fundamental way we think about and address mental disorders, both academically and clinically. The DSM, and the psychopathology discipline in general, were based on the medical model of disease. Hence, although alternative perspectives are beginning to filter through – such as the Embodied-Enactive perspective (Nielsen & Ward, 2018) – many practitioners and laypeople still conceive of mental disorders as akin to physical illnesses – in other words, as entities with discrete, identifiable causes that can be effectively treated with medication.

Although there is significant awareness within the psychopathology discipline of the shortcomings of the DSM and medical model, as well as general agreement that change is necessary, there are still many practitioners and institutions that are heavily DSM-oriented – whether by choice or by necessity. A good example of this is the fact that government or insurance agencies in many countries require a DSM diagnosis before they will release funding to a client for treatment or extra assistance, making such a diagnosis a necessary evil for many clinicians. Although it may be incongruous with the treating clinician's theoretical methodology or formulation of their client, refusing to assign a DSM diagnosis may limit the client's access to the resources they need to recover. Furthermore, depending on their personal beliefs about the nature of mental disorder, some clients may strongly desire a DSM diagnosis and may be confused or put off by practice that does not provide one. Similarly, clinicians primarily trained in the medical model, or who have practised under this framework for many years, may be very attached to the DSM and highly resistant to adopting an alternate approach to their formulation and treatment, especially if that approach is more complex and challenging than the current paradigm. Hence, as radically different an approach as the one proposed in this thesis is certain to come up against significant resistance.

DSM entrenchment also creates barriers at the theoretical level. Arguably, the majority of psychopathology research is based around DSM/ICD categories. For example,

most studies of EDs will select their participant groups based on diagnoses like AN, BN, or BED rather than symptomology, such as the presence or absence of "binge eating", "body dissatisfaction", or "excessive exercise". This can make it difficult to obtain valid data on individual symptoms, due to the poor specificity of these categories as well as their conceptual overlap. One can try to approximate symptom-specific data from DSM-oriented studies – for example, by using findings focused on the AN syndrome as a proxy for the symptom of "self-starvation". However, this is always going to represent a limited subset of the symptomology of interest. Although cases of AN do reliably feature "self-starvation", this symptom may also occur in BN, OSFED, or UFED cases as well as sub-clinically. Similarly, given the dominance of both the neurobiological and cognitive-behavioural approaches to mental health, there is much more research investigating factors and processes at these levels. Hence, challenges may arise when trying to formulate symptoms at other scales – such as the interpersonal or sociocultural – due to a shortage of data.

It is simply a matter of course that the approach outlined in this thesis will need to be applied alongside the existing understandings of mental disorder that operate at different levels – including the DSM, medical model, and neurobiological/cognitive-behavioural perspectives – both in research and in practice. It is neither realistic nor helpful to think we can immediately reject these understandings when they form the backbone of most research, practice, and cultural understanding of psychopathology – it is unreasonable to assume researchers or clinicians can operate independent of their institutional, political, or cultural context. However, it is likewise mistaken to believe that these approaches are incompatible. It is possible for the comprehensive, symptom-oriented, pluralistic explanatory approach discussed here to operate alongside or within existing conceptual structures. It would be incorrect to say that these understandings have no value to us – there is much in our current understanding of mental distress that is useful to a symptom-oriented, pluralistic approach. For example, as discussed in previous chapters, diagnostic syndromes are highly useful for concise communication of a client's clinical presentation, and neurocentric frameworks like the RDoC have the power to generate a wealth of useful research data on neurobiological processes involved in mental distress (see Chapter Four). Hence, the approach presented in this thesis does not advocate a full rejection of these understandings, but rather encourages their reformulation and elaboration in new ways that are more productive for future development.

As discussed in relation to competing therapeutic models in Chapter Five (such as CBT and ACT), these different perspectives can be reconciled by searching for common

theoretical or practical ground, as well as by identifying the tasks to which each is most appropriate. For example, the approach proposed in this thesis holds superior value for the formulation and explanation of psychopathological problems, but is unable to provide a useful method of classifying these phenomena.[34] Hence, in a practice context, a clinician might use a symptom-oriented, pluralistic approach to formulate and treat their client's difficulties – for example, by explaining the relevant clinical phenomena using multiple theoretical models at a variety of levels or scales and using this to inform treatment by selecting intervention strategies to target relevant mechanisms (see Ward et al., 2017) – but use diagnosis to communicate their client's case succinctly to other professionals or to access funding as necessary.

Contending with Greater Complexity

In general, it is difficult both practically and psychologically to change one's epistemological perspective. Transforming how one thinks about a topic entails significant mental gymnastics: one must fully embrace and comprehend the new perspective and its methods of operating, and attempt to put these into practice, while simultaneously unlearning the habits of the old. This is even more difficult when the new perspective is more complicated and cognitively challenging than its predecessor. Identifying that a problem is much larger and more complex than originally thought, and that the solution is therefore likely to be similarly intricate, is extremely daunting. Ideally, we would like problems to be simple to formulate and have parsimonious solutions. Hence, both the decision to engage with the approach to explaining psychopathology advocated in this thesis and the resolve required to continue applying it instead of current easier (although arguably less productive) alternatives are likely to be highly challenging.

The approach advocated in this thesis does not represent a "quick fix" for explaining and treating EDs. Rather, it requires researchers and practitioners to contend with multiple models of a phenomenon, which operate at a variety of conceptual levels, in order to fully comprehend it, rather than being able to rely on a single unified explanation. It also stresses the need for a large amount of additional theoretical research to be conducted: it calls for the development of a plethora of both compositional and aetiological explanatory models, each requiring broad and thorough surveying of available data, careful conceptualisation, and ongoing evaluation. Each of these tasks is likely to be both time-consuming and cognitively

[34] Although it may help to inform the development of fruitful classification schemes down the line.

challenging. However, just because a task is difficult does not mean it is not worth attempting. Despite the significant research attention paid to this task, the efficacy of ED treatment has shown limited improvement over the past decade: we should arguably be making much greater progress. The stagnation of research and treatment in this area strongly advocates for the consideration of alternative perspectives.

Our current, more linear, and deterministic approach to psychopathology – in which we search for a few common causes at one or two levels of analysis that lead to the "effect" of a mental illness – is arguably not working. Mental disorders are inherently complicated, comprised of multiple phenomena at a variety of levels, which interact in indirect and multifactorial ways to produce diverse and often unexpected effects at other scales, many of which feed back to perpetuate the problem. Indeed, many of the features common to psychopathological problems would support their conceptualisation as a *complex system* (Fried & Robinaugh, 2020; Rzevski & Skobelev, 2014).[35] It therefore follows that an accurate understanding of these phenomena will necessarily entail similar complexity (Fried & Robinaugh, 2020). This is not something to be shied away from, but actively embraced: better to make slower and more hard-won progress on a path that is likely to be more productive, than to labour at a journey that is ultimately leading us nowhere. Helpfully, there exists a significant developing body of interdisciplinary literature on investigating and managing complex systems that could be applied to psychopathology to help us manage the increasingly recognised complexity of this discipline, such as "complexity thinking" (Fried & Robinaugh, 2020; Rogers et al., 2013).

Future Directions

Although this thesis represents a start to solving the "ED problem", the changes it advocates are at the more abstract meta-theoretical level. Hence, there is a need for these ideas to be concretely realised through the actual implementation of the framework proposed by this thesis – in other words, through the generation of tangible explanations guided by this perspective. There is also need for complementary changes in related domains, such as

[35] A *complex system* (as opposed to a simple one) is typically characterised as possessing such characteristics as 1) a large number of components or agents, 2) nonlinear relations (there is not a straight-line or direct relationship between independent variables and dependent variables), 3) feedback loops (a portion of the system's output is fed back into the system as a new input), 4) a lack of central control (the system is not produced by a central cause, but distributed across multiple such that eliminating a single causal element does not destroy the system), and 5) emergence (they involve self-organising behaviour in which large-scale effects emerge out of connections between smaller scale components, which are often surprising or difficult to anticipate; Boccara, 2010; Ladyman et al., 2013).

empirical research, intervention development and implementation, and classification, that work alongside the theoretical.

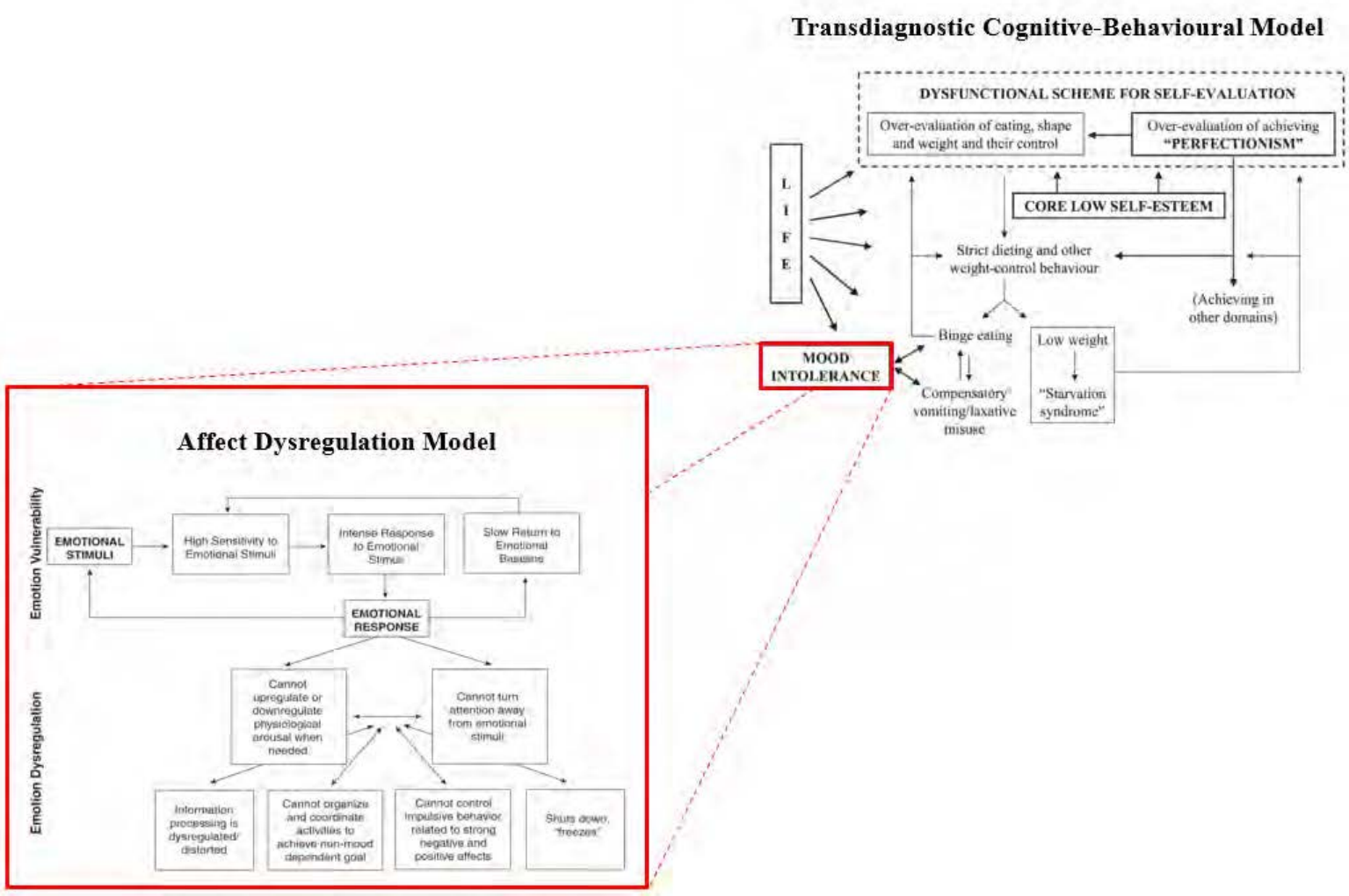
Generating Tangible Explanations

The obvious continuation of this thesis is its practical application to the explanation of EDs – the concrete generation of explanations of ED phenomena according to the perspective proposed. This is a monumental task, however, as there is a plethora of clinical phenomena associated with ED problems, the constitution and causes of which each need to be comprehensively modelled at all the relevant scales. This is why we describe the solutions proposed by this thesis as significant conceptual and methodological shifts: none represent "quick fixes", but rather a global reorientation in how we currently approach the explanation of mental distress, and EDs more specifically. They present a new foundation upon which understandings of psychopathology problems, such as EDs, can be built: a plan for the future of explanation in psychopathology. However, the execution of this plan will take considerable time, effort, and coordination.

In my view, a good place to begin is with the detailed conceptualisation and compositional explanation of ED symptoms, so we begin to develop more valid and specific constructs for empirical investigation and causal explanation. The current "thin" conceptualisation of ED symptoms is highly limiting for the field as a whole: it is difficult to generate valid research and theory if there is ambiguity and disagreement around what is being studied. Creating comprehensive compositional of constructs such as "body image disturbance", "fear of weight gain", "restrictive eating", or "purging", will likely begin to clarify this. A preliminary example of this kind of compositional modelling was sketched out in Chapter Four for the symptom "binge eating". This goal is also more immediately achievable than the construction of aetiological explanations, which are necessarily more complex – entailing a greater diversity of spatial and temporal contributors – and more difficult to achieve without robust understandings of the phenomena we seek to explain (i.e., ED symptoms and signs). It may also reveal and clarify links to other mental disorders by identifying common symptoms as well as shared compositional processes and structures. Some symptom network models have already done some work on this topic by identifying "bridge symptoms" between EDs and anxiety (Forrest et al., 2019; Levinson et al., 2018, Meier et al., 2020), post-traumatic stress (Vanzhula et al., 2019), and depressive presentations (Solmi et al., 2018).

It may also be useful to relate existing explanations of EDs to each other theoretically. These explanations possess various shortcomings (see Chapter Three) and will ideally be replaced eventually by an explanatory network of richer, symptom-oriented models. However, in the interim we can make the best of what we have by augmenting these existing models, interconnecting them and fleshing them out where necessary. For example, the Affect Dysregulation Model (ADM) of eating disorders could be nested within the Transdiagnostic Cognitive-Behavioural Theory (TCBT) in order to unpack more comprehensively the theorised construct of “mood intolerance”, which is poorly defined by the TCBT (see Chapter Three). The ADM provides a far more detailed explanation of this construct than does the TCBT, recognising the distinct contributions made by “emotional vulnerability” (greater sensitivity to emotional stimuli, greater response to emotional stimuli, slower return to baseline from emotional arousal) and “emotional dysregulation” (problems up- and down-regulating arousal or disengaging attention from emotional stimuli when needed). Hence, the ADM can provide a useful counterpart to the TCBT that enriches the broader explanation (see Figure 8).

Figure 8. *Example Relation of the Affect Dysregulation and Transdiagnostic Cognitive-Behavioural Models of Eating Disorders*



Complementary Changes to Research and Practice

Although improving our explanation of EDs and psychopathology more broadly is an excellent start to ameliorating the "ED problem", this alone represents only part of the solution. While the generation of better explanations of EDs *promotes* improvements in related domains, such as research and practice, by providing a theoretical platform for their achievement (see the section on Clinical Implications above), these areas need to choose to adopt this new theoretical orientation and apply it appropriately. Just because a particular theoretical orientation is adopted does not necessarily mean it will be applied in suitable or productive ways. For instance, empirical research may choose to investigate novel theoretical claims focused on symptoms but may do so poorly if its methodology is flawed. Similarly, novel interventions may be generated based on new symptom-oriented theoretical models but be poorly structured or implemented depending on the theoretical frameworks and methodologies used to translate the model from theory to practice. Hence, for the benefits of this thesis' explanatory approach to successfully flow-on to ED research and practice, complementary changes need to occur in the meta-theory and meta-methodology of these tasks.

Changes to Research.

As previously discussed, much of clinical and general psychological research is based around DSM categories, and this can create significant difficulties for researchers wanting to develop non-DSM-centred theories or classification schemes. Hence, a change that is needed to assist and better facilitate the implementation of the meta-theoretical approach outlined in this thesis is the complementary reorientation of empirical research into both EDs and psychopathology more broadly. This is likely to be stimulated partially by the development of new, symptom-oriented, theories, as theory is often a significant driver behind what we choose to study empirically and what methodologies we use to do so. However, it is unnecessary to wait for this "trickle-down" effect to emerge when we can begin to redirect empirical research now and generate data that can inform the development of novel explanatory models.

There are specific ways in which the empirical research process can be adapted to better align with the meta-theoretical approach proposed by this thesis. The first and most obvious of these is the redirection of empirical research away from DSM syndromes or transdiagnostic categories, which lack conceptual validity, towards specific symptoms or

clinical phenomena. This shift will allow us to gain more clarity around how these constructs are composed and caused, which will then provide a richer foundation for their theoretical explanation and classification. The RDoC project in many ways represents a promising step in this direction, as it seeks to investigate psychological phenomena at finer grain sizes independent of DSM categories. However, as previously discussed (see Chapter Four), the RDoC does not actually direct research toward psychopathological disorders or symptoms but, rather, towards broader "psychological systems" (e.g., positive valence systems, systems for social processes), both how they function and malfunction. The objects of study within the RDoC are therefore general psychological processes – such as "reward learning", "memory", or "social communication" – rather than particular psychopathological symptoms. Hence, although it may usefully complement symptom-oriented research by providing data on how these broader psychological processes may malfunction and thereby *contribute* to the constitution or aetiology of psychopathological problems, the RDoC framework is not suitable to guide research into psychopathological symptoms.[36] A useful first step in directing research towards these phenomena, therefore, would be to develop a theoretical framework that can guide the study of psychopathological symptoms.

Second, in order to complement the model pluralism approach taken in this thesis – that explanations of mental disorder should be comprised of multiple models operating at a variety of levels of analysis – empirical research will need to cultivate a similar culture of cooperation. There is a need for improved interdisciplinary communication and collaboration in the psychopathological sciences. At present, it is common for different disciplines involved in the study of psychopathology – such as neuroscience, psychology, sociology, medicine, biochemistry – to be somewhat isolated from one another, and their empirical research highly insular. Within institutions, different disciplines are typically sequestered in separate departments, thus reducing the frequency with which they interact, and students are typically taught only the philosophies and methodologies associated with their own particular field, with little tuition about other disciplines or the metacognitive knowledge needed to work with them ("higher-order cognitive skills"; see Boon & Van Baalen, 2019). In academia more broadly, many peer-reviewed journals or research databases are discipline-specific, making it possible for researchers or entire disciplines to become insulated from the philosophies and

[36] Although it should be noted that the RDoC framework would very successfully complement a meta-theoretical approach to the explanation of *psychological systems* that is similar to that outlined in this thesis for psychopathology.

methodologies of other disciplines. Even when journals *are* interdisciplinary, researchers are likely to focus on articles within their own field, as lack of interdisciplinary education can make engaging with other-disciplinary research feel bewildering and unrewarding.

Given the complex, multifactorial nature of mental disorder, it is unlikely to be solved by any one discipline alone. Each discipline will logically be better positioned to address some questions than others. For example, microbiologists would likely be better at researching the role of the microbiome in mental disorder (the *gut-brain axis*; Foster & McVey Neufeld, 2013; see Chapter Five) than psychologists or sociologists, who in turn would be better placed to investigate the phenomenological experience of mental distress or its socio-cultural representation. However, disciplines may be able to make useful contributions to problems with which they are not directly concerned. For instance, microbiologists' understanding of microbiota and the gut-brain axis may help psychologists to better understand how clinical symptoms are constituted or caused, even though microbiology is not directly concerned with psychopathological conditions. In psychopathology, there are subject areas that require input from multiple disciplines to fully comprehend. For example, the role of the family in mental disorder would likely be best understood though joint collaborations between geneticists (e.g., genes, epigenetic processes), psychologists (e.g., familial modelling, parenting practices, behavioural reinforcement), and sociologists or cultural anthropologists (e.g., familial composition and operation in different societies, cultural norms and values relating to family). Hence, interdisciplinary research is something that needs to be actively fostered within psychopathology. A recent example of one way this may be facilitated comes from Boon and Van Baalen (2019), who advocate for the development and application of "metacognitive scaffolds" for interdisciplinary work, such as *disciplinary matrices,* which analyse and articulate the way knowledge is constructed within the different disciplines in order to facilitate collaboration.

Changes to Intervention Development.

Although possessing better explanations of mental disorder provides a platform for improved clinical intervention, this relies on new interventions being developed based on these models. This requires an additional theoretical step: from explanatory models to *intervention models*. Explanatory models, or explanations, of mental disorder, as discussed in this thesis, aim to explain or model all or part of how a particular disorder state originated and is maintained. Ideally, this should identify causal factors and processes, the removal or

reduction of which should, theoretically, result in amelioration of the disorder symptoms. However, although it may identify and describe these causal factors and processes, it does not outline precisely *how* they should be targeted by an intervention. This is the role of *intervention models*. These models propose ways of practically targeting one or more theorised causal factors or processes within a clinical context, via specific therapeutic strategies or techniques. A good example of this would be cognitive therapy for depression (see Beck & Dozois, 2011). This model proposes that the negative cognitions thought to promote depressed mood, as theorised by cognitive explanations of depression (e.g., Beck, 2008), can be modulated by helping clients to notice and challenge unhelpful thinking patterns (*cognitive restructuring*). Intervention models thus dictate how intervention proceeds: *which* causal factors and processes to target and *how*.

Although the advent of better explanations in psychopathology will provide much more information about causal factors and processes involved in mental disorder, it will not produce better interventions on its own. We therefore need complementary intervention models to be developed that effectively target the causal constructs identified within these explanations. This requires further theoretical and empirical research into *therapeutic mechanisms of change* – namely, *how* intervention strategies work to change the causal constructs they target. This is particularly important given that most interventions currently do not target just one causal construct; typically, they are intended to target an array of causal factors and processes, via either one or multiple techniques. For example, in cognitive therapy for depression, negative cognitions are targeted using thought records, behavioural experiments, and Socratic questioning in session. Although we can currently evaluate the overall efficacy of these treatments, we are unable to determine *which* of the intervention strategies employed affected *which* of the constructs targeted, as well as *how* those effects occurred. In other words, "although there are many EBTs (*evidence based therapies*) available, there is little understanding of the mechanisms of change (i.e., precisely how they work)" (Kazdin & Blase, 2011, p. 28). This is problematic as it leaves us blind as to which components of our intervention strategies may be helpful, harmful, or utterly redundant.

The evaluation of these therapies is also flawed. There are multiple flaws in the design and execution of clinical trials of psychological interventions that create problems, including the "standards for reporting clinical trials, specification of treatment protocols and inclusion/exclusion criteria, choice of outcome measures, measurement of adverse effects and preventing bias in trial design and analysis" (Holmes et al., 2018, p. 68). Hence, as well as being unable to identify *how* or *why* a treatment works, we may be being misled about how

well it works. For us to appraise the efficacy of psychotherapeutic interventions accurately, the process and standards for conducting clinical trials within psychology need to be reviewed and improved. Furthermore, there is need for research into broader aspects of the therapeutic process that are not tied to any specific intervention but shown to predict treatment outcome, such as client motivation, therapeutic alliance and rapport, and treatment adherence (Chatoor & Kurpnick, 2001).

There have been various proposals on how the construction and evaluation of psychotherapeutic treatments could be improved, as well as the ongoing challenges inherent in the process of psychological treatment research.[37] In my view, there is need at the theoretical level for the generation of a meta-theoretical framework for the construction of intervention models, similar to the PDM-TC proposed in this thesis but specific to interventions. On the empirical side, there is a clear need for a more robust and valid framework for conducting clinical trials to evaluate psychological treatments, as well as more specific study of therapeutic change mechanisms. Development in each of these areas would, in my view, allow clinical practice to take full advantage of the greater understanding offered by richer and more detailed explanations of mental disorder.

Changes to Intervention Implementation.

As well as improvement in how psychotherapeutic interventions are constructed, there needs to be improvement in how they are implemented. Having robust intervention models that show high efficacy in clinical trials is incredibly important. However, this often does not reflect how interventions are realistically employed at the ground level. Clinical trials represent a “best-case scenario”: they test the efficacy of an intervention in isolation and in optimum, highly controlled conditions. Although this is useful for evaluating the optimum efficacy of intervention models, there is need for additional investigation into how they might be effectively applied in real-world settings. Clinical practice is much less “clean-cut”. There are often multiple professional bodies involved and a variety of interests or goals at play, each of which needs to be carefully considered and incorporated. Cases are also often complex – clients may bring a variety of developmental, interpersonal, occupational and/or socioeconomic problems into the therapy room along with their specific mental health

[37] Perhaps the best and most comprehensive example is Holmes et al.’s (2018) report for the *Lancet Psychiatry Commission.*

concerns. At present there is a lack of guidance about how to navigate psychopathological intervention within these complex contexts.

Psychology and psychiatry have tended to operate under the *evidence-based practice* model, in which successful clinical practice is considered a "three-legged stool" comprising scientific research, clinical expertise, and client values and preferences (Howick, 2011; Lilienfeld et al., 2013). However, such a model, while a useful meta-methodology for clinical practice, is too vague to help practitioners successfully navigate the translation of intervention models into the realities of clinical practice. In a recent paper, Ward and Durrant (2021) identify the need for a unique type of theory known as a *practice framework*, that can "offer program designers a tool for constructing and delivering a range of interventions to individuals … and constitute an epistemic hub into which relevant features of explanatory and treatment theories can be 'plugged'" (p. 1). Practice frameworks consist of three linked "levels" or sets of ideas: 1) the core values and principles it endorses (specific ethical and prudential values), 2) its knowledge-related assumptions (philosophical or meta-theoretical ideas about the nature and causes of the phenomena of interest, such as psychopathology or criminal offending), 3) their intervention guidelines (how clinicians should aim to work with clients in light of their values and assumptions).

A good example of a practice framework in the correctional setting is the Good Lives Model (GLM; Ward & Brown, 2006), which at the first level endorses human rights and agency, at the second level conceptualises offence behaviour as goal-directed behaviour intended to achieve particular "primary human goods", and at the third level positions clients as "fellow travellers" who hold equivalent moral status to their therapists, thereby encouraging a collaborative, nonpunitive approach to practice. Hence, the GLM prescribes how clinicians should practice with their clients, from what kinds of interventions might be suitable, to how they should be applied, to what a successful outcome might look like. Similar practice frameworks are lacking within the psychopathology area. Hence, there is need for the development of such frameworks that can help guide clinical practice, both across and within psychopathological problems.

Changes to Classification.

As discussed earlier in this thesis, changes in the focus of empirical research and theoretical modelling is likely to prompt or facilitate the advent of alternative classification schemes. However, as with interventions (see above), it is not a given that these classification

schemes will be generated in the most productive way. Taxonomies based on sound theoretical explanations can easily be flawed, as classification is a distinct task with its own predominating theoretical assumptions and methodology: if these are flawed, the resulting taxonomy likely will be too. Classification organises phenomena into non-arbitrary groups, creating a common language by which ideas about them can be communicated. It therefore represents a theoretical judgement about the best organisation of a set of phenomena – namely, which features or properties determine whether cases belong together or apart. For example, DSM-5 classification of EDs places greatest importance on external symptomology ("signs") to distinguish between different categories of ED – AN, BN, and BED are primarily differentiated by weight (BMI) and/or the presence or absence of certain ED behaviours, such as "binge eating" or "purging" – as well as relying to a degree on epidemiological factors such as average age of onset. The features prioritised by classification schemes then often become used to distinguish between groups in empirical studies and explanatory models – like how DSM categories have come to dominate psychopathological theory and research. The development of further flawed taxonomies thus also has the potential to undermine the changes in psychopathological explanation advocated by this thesis. There is therefore a need for similar meta-theoretical change to occur in the classification of psychopathology.

At present, classification in psychopathology subscribes to the notion that there is a single correct or "best" way of categorising mental disorders, similar to the One Best Model perspective extant in psychopathological explanation. Different taxonomic conceptualisations compete, much in the same way that explanatory models currently do within psychopathology (see Chapter Five). However, mental disorders are highly complex phenomena: there are multiple scales at which they may be conceptualised and multiple purposes for which a taxonomy may be created, including allocating interventions, communicating cases amongst clinicians, providing a platform for explanation, directing empirical research, or determining access to community resources or support. As with explanatory models, it can be argued that it is foolish to expect a single classification scheme to achieve all these objectives. It is much more likely that each will be best served by a different classificatory approach. For example, communication across clinicians about client presentation is likely to be best served by categorisation into disorder syndromes in order to optimise the efficiency of information exchange without sacrificing understanding, whereas allocating appropriate treatment is likely to be better served by classification based around

causes or functions of the disorder that can then be mapped onto appropriate intervention strategies.

Hence, in my view, it is necessary for classification to adopt a similarly pluralistic approach to that advocated for explanation in this thesis, in which psychological phenomena are classified in a variety of ways to serve a variety of aims, all of which are compatible but not collapsible into a single scheme. Carter et al.'s (2021) Pluralistic Classification Framework (PCF), developed to guide pluralistic classification of criminal behaviour, provides a useful scaffold for the development of a plurality of "categories and classification schemes that are fit for purpose" (p. 9). The PCF outlines a four-step reasoning process for determining the most appropriate means of classification based on the task at hand. It asks researchers to consider who will be using the classification system (e.g., clinical practitioners, academic researchers, psychopathological theorists), what task(s) they will be using it for (e.g., predicting outcomes, assigning individuals to treatment, case presentation), what conceptual categories are best suited to achieving the task (e.g., descriptive, causal, predictive), and what kinds of inferences and knowledge these categories will be able to produce (e.g., predictive categories cannot be used to make causal inferences). This framework can also be used as an integrative "hub", by which the various classification schemes within an area, such as "psychopathology", can be connected based on their relation to the above questions. The PCF may therefore provide a useful jumping off point for the development of pluralistic classification of EDs, and psychopathology more broadly.

For the explanation and treatment of EDs, a specific classificatory approach that may be useful to consider moving forward is the recent Functional Offending Behavioural Classification Framework (FOBCF), developed by Ward and Carter (2019) for the classification of crime-related behaviour. The FOBCF proposes classifying problematic (in this case, criminal) behaviours based on the function they serve, which may vary across individuals and contexts. The FOBCF outlines a framework for doing this based on a set of shared human motivational systems.[38] This functional understanding then provides a more

[38] These are 1) *bodily regulation* (drive to attend to physiological needs like warmth, hydration, and nutrition), 2) *aggression* (drive to respond to threats of harm to the self or loved ones), 3) *fear and security* (drive to avoid or escape immediate threat or predict and prevent future threat), 4) *disgust* (drive to avoid contact with biological, moral, or social "pollution"), 5) *status* (drive to achieve and maintain higher social status), 6) *mating and pair bonding* (drive to achieve and sustain romantic or sexual relationships), 7) *attachment and caregiving* (drive to acquire support and care within a close relationship), 8) *affiliation* (drive to seek out and maintain non-familial social relationships), 9) *reciprocity and acquisition* (drive for "justice", such as monitoring compliance to shared norms and facilitating fair trade), 10) *play* (drive to develop physical, social, and cognitive skills in the

robust platform for treatment. Once the function of a behaviour – or *functions*, as sometimes a behaviour will be intended to achieve multiple goals – is understood, practitioners can intervene to help the individual achieve their goal(s) via prosocial means, such as helping the client build the necessary capacities and increasing their awareness of alternatives.

As EDs are also primarily characterised by problematic behaviours, the FOBCF may provide a valuable way of classifying them. For example, self-starvation could be understood differently in the context of different motivational systems – for example, undereating due to inaccurate perception of nutritional needs may be linked to dysfunction in the bodily regulation system such as abnormal responding of appetitive hormones; undereating due to increased aversion to foods perceived as "bad" or "unhealthy" and bodily characteristics perceived as socially undesirable, such as being overweight or "fat", could be linked to over-sensitisation of the disgust system; and given the high value placed on thinness by society, undereating to lose weight could be linked to any of the status, mating/pair bonding, attachment, or affiliation systems. Given that some research already suggests that ED behaviours may serve different functions for those affected (e.g., identity, security, avoidance, care, mastery; Lavis, 2018; Norbø et al., 2006; Stockford et al., 2018), the FOBCF approach may indeed hold significant value for this area, and I hope to explore this in future work.

Concluding Remarks

The current explanations we possess for EDs are not serving us well. Limited by conceptual deficiencies and poor specificity, existing models provide a shaky foundation for research and practice and, as a result, empirical investigation and clinical intervention surrounding EDs has stagnated. For many years, theory in EDs has relied on the same explanatory approach: DSM-oriented, unified, and disorganised. Hence, progress has typically been rather artificial; consisting of minor revisions of existing models, or the development of new models that look remarkably like their predecessors. As stated by David Fleming in his book *Lean Logic* (2016), "forward movement is not helpful if what is needed is a change of direction." (p 472). It is time we stopped reinventing the wheel and made some meaningful progress. What is needed is not simply the revision of existing explanatory models, nor the generation of additional ones based on the same paradigm, but a

absence of threat), and 11) *curiosity* (drive to acquire knowledge about others and the world; see Ward & Carter, 2019).

comprehensive overhaul of the way that we currently approach the conceptualisation and explanation of disordered eating problems.

This thesis proposed several significant shifts in the way we explain both disordered eating problems and psychopathology more generally. First, it was argued that our current conceptually "thin" explanatory targets – namely, DSM/ICD syndromes and transdiagnostic classifications – should be traded in for a rich, detailed approach based around signs and symptoms in order to generate compositional understandings of phenomena and set up a more robust platform for aetiological explanation. Second, the current notion of a "unified", single-model approach to psychopathological explanation was disavowed in favour of a pluralistic approach in which multiple models are used to understand psychopathological problems in a way that is richer and more flexible to the diversity of needs within the psychopathology sphere. Finally, a methodological framework for the construction of psychopathology explanations was proposed – the PDM-TC – to guide the generation of richer, more robust explanatory models.

Based on these meta-theoretical developments, our explanations are primed to become richer, deeper, and more conceptually valid, thereby providing a more fruitful foundation for the treatment of psychopathological problems. However, there is still much work to be done, both in terms of the generation of actual explanations based on this approach and the corresponding changes needed in related areas (e.g., classification, research, intervention), and it is likely to be an uphill battle given the entrenchment of current perspectives. Nevertheless, this thesis represents a first step; reorienting our approach to explaining psychopathology in a direction that is likely to be more productive in understanding and relieving mental distress.

References

Achenbach, T. M., Ivanova, M. Y., Rescorla, L. A., Turner, L. V., & Althoff, R. R. (2016). Internalizing/externalizing problems: Review and recommendations for clinical and research applications. *Journal of the American Academy of Child and Adolescent Psychiatry, 55*(8), 647–656. doi:10.1016/j.jaac.2016.05.012

Aftab, A., & Waterman, G. S. (2020). Conceptual competence in psychiatry: Recommendations for education and training. *Academic Psychiatry, 45,* 203-209. doi:10.1007/s40596-020-01183-3

Agras, W. S., Crow, S. J., Mitchell, J. E., Halmi, K. A., & Bryson, S. (2009). A 4-year prospective study of eating disorder NOS compared with full eating disorder syndromes. *International Journal of Eating Disorders, 42*(6), 565-570. doi:10.1002/eat.20708

Ainsworth, M. S. (1989). Attachments beyond infancy. *American Psychologist, 44*(4), 709-716. doi:10.1037/0003-066X.44.4.709

Albert, P. R., Benkelfat, C., & Descarries, L. (2012). The neurobiology of depression - revisiting the serotonin hypothesis. I. Cellular and molecular mechanisms. *Philosophical Transactions of the Royal Society B, 367*(1601), 2378-2381. doi:10.1098/rstb.2012.0190

Al-Harbi, K. S. (2012). Treatment-resistant depression: Therapeutic trends, challenges, and future directions. *Patient Preference and Adherence, 6,* 369-388. doi:10.2147/PPA.S29716

Allegrini, A. G., Cheesman, R., Rimfield, K., Selzam, S., Pingault, J-B., Eley, T. C., & Plomin, R. (2019). The p factor: Genetic analyses support a general dimension of psychopathology in childhood and adolescence. *The Journal of Child Psychology and Psychiatry, 61*(1), 30-39. doi:10.1111/jcpp.13113

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Arlington, VA: American Psychiatric Association.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Arlington, VA: American Psychiatric Association.

American Psychiatric Association (2006). Treatment of patients with eating disorders (3rd ed.). *American Journal of Psychiatry, 163*(Suppl 7), 4-54. https://pubmed.ncbi.nlm.nih.gov/16925191/

American Psychiatric Association. (2007). *APA dictionary of psychology*. Washington, DC, US: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Association.

Andersen, H. (2014). A field guide to mechanisms: Part I. *Philosophy Compass, 9*(4), 274-283. doi:10.1111/phc3.12119

Anderson-Fye, E. P. (2018). Cultural influences on body image and eating disorders. In W. S. Agras & A. Robinson (Eds.), *The Oxford handbook of eating disorders* (2nd ed., pp. 187-210). New York, NY: Oxford University Press.

Arcelus, J., Mitchell, A. J., & Wales, J. (2011). Mortality rates in patients with anorexia nervous and other eating disorders: A meta-analysis of 36 studies. *Archives of General Psychiatry, 68*(7), 724-731. doi:10.1001/archgenpsychiatry.2011.74

Arch, J. J., & Craske, M. G. (2008). Acceptance and commitment therapy and cognitive behavioral therapy for anxiety disorders: Different treatments, similar mechanisms? *Clinical Psychology Science and Practice, 15*(4), 263-279. doi:10.1111/j.1468-2850.2008.00137.x

Ata, R., Schaefer, L. M., & Thompson, J. K. (2014). Sociocultural theories of eating disorders. In M. Levine & L. Smolak (Eds.), *The Wiley handbook of eating disorders, assessment, prevention, treatment, policy, and future directions* (pp. 269-282). West Sussex, UK: Wiley.

Atwood, M. E., & Friedman, A. (2020). A systematic review of enhanced cognitive behavioral therapy (CBT-E) for eating disorders. *International Journal of Eating Disorders, 53*(3), 311-330. doi:10.1002/eat.23206

Avena, N. M., & Bocarsly, M. E. (2012). Dysregulation of brain reward systems in eating disorders: Neurochemical information from animal models of binge eating, bulimia nervosa, and anorexia nervosa. *Neuropharmacology, 63*(1), 87-96. doi:10.1016/j.neuropharm.2011.11.010

Bailer, U. F., & Kaye, W. H. (2011). Serotonin: imaging findings in eating disorders. *Current Topics in Behavioral Neurosciences, 6*, 59-79. doi:10.1007/7854_2010_78

Bailer-Jones, D. M. (2003). When scientific models represent. *International Studies in the Philosophy of Science, 17*(1), 59-74. doi:10.1080/02698590305238

Balodis, I. M., Grilo, C. M., & Potenza, M. N. (2015). Neurobiological features of binge eating disorder. *CNS Spectrums, 20*(6), 557-565. doi:10.1017/S1092852915000814

Bamford, B. H., Attoe, C., Mountford, V. A., Morgan, J. F., & Sly, R. (2014). Body checking and avoidance in low weight and weight restored individuals with anorexia nervosa and non-clinical females. *Eating Behaviors, 15*(1), 5-8. doi:10.1016/j.eatbeh.2013.10.011

Bankoff, S. M., Karpel, M. G., Forbes, H. E., & Pantalone, D. W. (2012). A systematic review of dialectical behavior therapy for the treatment of eating disorders. *Eating Disorders, 20*(3), 196-215. doi:10.1080/10640266.2012.668478

Bechtel, W. (2008). Mechanisms in cognitive psychology: What are the operations? *Philosophy of Science, 75*(5), 983-994. doi:10.1086/594540

Beck, A. T. (2008). The evolution of the cognitive model of depression and its neurobiological correlates. *American Journal of Psychiatry, 165*(8), 969–977. doi:10.1176/appi.ajp.2008.08050721

Beck, A. T., & Bredemeier, K. (2016). A unified model of depression: Integrating clinical, cognitive, biological, and evolutionary perspectives. *Clinical Psychological Science, 4*(4), 596-619. doi:10.1177/2167702616628523

Beck, A. T., & Dozois, D. J. A. (2011). Cognitive therapy: Current status and future directions. *Annual Review of Medicine, 62*, 397-409. doi:10.1146/annurev-med-052209-100032

Beck, A. T., & Haigh, E. A. (2014). Advances in cognitive theory and therapy: The generic cognitive model. *Annual Review of Clinical Psychology, 10*, 1–24. doi:10.1146/annurev-clinpsy-032813-153734

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1987). *Cognitive therapy of depression*. New York, NY: The Guilford Press.

Bell, J. P. (1985). Mental health in the 80s. *American Journal of Social Psychiatry, 5*(3), 50-53. https://psycnet.apa.org/record/1986-22645-001

Bello, N. T., & Hajnal, A. (2010). Dopamine and binge eating behaviors. *Pharmacology Biochemistry and Behavior, 97*(1), 25-33. doi:10.1016/j.pbb.2010.04.016

Berg, K. C., Crosby, R. D., Cao, L., Peterson, C. B., Engel, S. G., Mitchell, J. E., & Wonderlich, S. A. (2013). Facets of negative affect prior to and following binge-only, purge-only, and binge/purge events in women with bulimia nervosa. *Journal of Abnormal Psychology, 122*(1), 111-118. doi:10.1037/a0029703

Berkman, N. D., Lohr, K. N., & Bulik, C. M. (2007). Outcomes of eating disorders: A systematic review of the literature. *International Journal of Eating Disorders, 40*(4), 293-309. doi:10.1002/eat.20369

Berridge, K. C. (2009). 'Liking' and 'wanting' food rewards: Brain substrates and roles in eating disorders. *Physiology & Behavior, 97*(5), 537-550. doi:10.1016/j.physbeh.2009.02.044

Berrios, G. E. (2013). Formation and meaning of mental symptoms: History and epistemology. *Dialogues in Philosophy, Mental and Neuro Sciences, 6*(2), 39-48. https://philarchive.org/rec/BERFAM-2

Berthoud, H. (2008). Vagal and hormonal gut–brain communication: From satiation to satisfaction. *Neurogastroenterology & Motility, 20*(s1), 64-72. doi:10.1111/j.1365-2982.2008.01104.x

Birmingham, C. L., Touyz, S., & Harbottle, J. (2009). Are anorexia nervosa and bulimia nervosa separate disorders? Challenging the 'transdiagnostic' theory of eating disorders. *European Eating Disorders Review, 17*(1), 2-13. doi:10.1002/erv.896

Boccaccini, M. T., Marcus, D., & Murrie, D. C. (2017). Allegiance effects in clinical psychology research and practice. In S. O. Lilienfeld & I. D. Waldman (Eds.), *Psychological science under scrutiny: Recent challenges and proposed solutions* (pp. 323-339). West Sussex, UK: John Wiley & Sons, Inc.

Boccara, N. (2010). *Modelling complex systems* (2nd ed.). New York, NY: Springer Science & Business Media. doi:10.1007/978-1-4419-6562-2

Bogen, J., & Woodward, J. (1988). Saving the phenomena. *Philosophical Review, 97*(3), 303-352. doi:10.2307/2185445

Bolton, D. (2008). *What is mental disorder? An essay in philosophy, science, and values*. Oxford, UK: Oxford University Press.

Bonta, J., & Andrews, D. A. (Eds.) (2017). *The psychology of criminal conduct* (6th ed.). New York, NY: Routledge.

Boon, M., & Van Baalen, S. (2019). Epistemology for interdisciplinary research - shifting philosophical paradigms of science. *European Journal for Philosophy of Science, 9*(1), 16. doi:10.1007/s13194-018-0242-4

Booth, D. A. (2003). Food, nutrition, and appetite. In B. Caballero (Ed.), *Encyclopedia of food sciences and nutrition* (2nd ed., pp. 5098-5102). Oxford, UK: Academic Press.

Borgna-Pignatti, C., & Zanella, S. (2016). Pica as a manifestation of iron deficiency. *Expert Review of Hematology*, *9*(11), 1075-1080. doi:10.1080/17474086.2016.1245136

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry, 16*(1), 5-13. doi:10.1002/wps.20375

Borsboom, D., Cramer, A. O. J., & Kalis, A. (2019). Brain disorders? Not really: Why network structures block reductionism in psychopathology research. *Behavioral and Brain Sciences, 42*, e2. doi:10.1017/S0140525X17002266

Boyd, K. M. (2005). Medical ethics: Principles, persons, and perspectives: From controversy to conversation. *Journal of Medical Ethics, 31*(8), 481-486. doi:10.1136/jme.2003.005710

Brewerton, T. D. (1995). Toward a unified theory of serotonin dysregulation in eating and related disorders. *Psychoneuroendocrinology*, *20*(6), 561-590. doi:10.1016/0306-4530(95)00001-5

Brewin, C. R., Cloitre, M., Hyland, P., Shevlin, M., Maercker, A., Bryant, R. A., Humayun, A., Jones, L. M., Kagee, A., Rousseau, C., Somasundaram, D., Suzuki, Y., Wessely, S., van Ommeren, M., & Reed, G. M. (2017). A review of current evidence regarding the ICD-11 proposals for diagnosing PTSD and complex PTSD. *Clinical Psychology Review, 58*, 1-15. doi:10.1016/j.cpr.2017.09.001.

Broft, A., Berner, L. A., Martinez, D., & Walsh, B. T. (2011). Bulimia nervosa and evidence for striatal dopamine dysregulation: A conceptual review. *Physiology & Behavior, 104*(1), 122-127. doi:10.1016/j.physbeh.2011.04.028

Brooks, D. S. & Eronen, M. I. (2018). The significance of levels of organization for scientific research: A heuristic approach. *Studies in the History and Philosophy of Biological and Biomedical Science, 68-69*, 34-41. doi:10.1016/j.shpsc.2018.04.003

Brown, T. A., Wisniewski, L., & Anderson, L. K. (2020). Dialectical behavior therapy for eating disorders: State of the research and new directions. *Eating Disorders: The Journal of Treatment & Prevention, 28*(2), 97-100. doi:10.1080/10640266.2020.1728204

Browne, M. A. O., Wells, J. E., & Scott, K. M. (2006). *Te Rau Hinengaro: The New Zealand Mental Health Survey*. Wellington, New Zealand: Ministry of Health

Bruch, H. (1973). *Eating disorders: Obesity, anorexia nervosa and the person within.* New York, NY: Basic Books.

Bruch, H. (1982). *The golden cage: The enigma of anorexia nervosa.* Cambridge, MA: Harvard University Press.

Bulik, C. M., Berkman, N. D., Brownley, K. A., Sedway, J. A., & Lohr, K. N. (2007). Anorexia nervosa treatment: A systematic review of randomized controlled trials. *International Journal of Eating Disorders, 40*(4), 310-320. doi:10.1002/eat

Bulik, C. M., Trace, S. E., Kleiman, S. C., & Mazzeo, S. E. (2014). Feeding and eating disorders. In D. C. Beidel, B. C. Frueh, & M. Hersen (Eds.), *Adult psychopathology and diagnosis* (7th ed., pp. 473-522). Hoboken, NJ: John Wiley & Sons, Inc.

Burke, N. L., Kaam, A. M., & Tanofsky-Kraff, M. (2018). Interpersonal therapy for the treatment of eating disorders. In W. S. Agras & A. Robinson (Eds.), *The Oxford handbook of eating disorders* (2nd ed., pp. 287-318). New York, NY: Oxford University Press.

Burkhauser, R. V., & Cawley, J. (2008). Beyond BMI: The value of more accurate measures of fatness and obesity in social science research. *Journal of Health Economics, 27*(2), 519-529. doi:10.1016/j.jhealeco.2007.05.005

Carragher, N., Krueger, R. F., Eaton, N. R., & Slade, T. (2015). Disorders without borders: Current and future directions in the meta-structure of mental disorders. *Social Psychiatry and Psychiatric Epidemiology, 50*(3), 339–350. doi:10.1007/s00127-014-1004-z

Carter, E., Ward, T., & Strauss-Hughes, A. (2020). The classification of crime and its related problems: A pluralistic approach. *Aggression and Violent Behavior, In Press,* 101440. doi:10.1016/j.avb.2020.101440

Cash, T. F., & Deagle, E. A. (1997). The nature and extent of body-image disturbances in anorexia nervosa and bulimia nervosa: A meta-analysis. *International Journal of Eating Disorders, 22*(2), 107-125. doi:10.1002/(SICI)1098-108X(199709)22:2<107::AID-EAT1>3.0.CO;2-J

Casper, R. C. (2006). The 'drive for activity' and "restlessness" in anorexia nervosa: Potential pathways. *Journal of Affective Disorders, 92*(1), 99-107. doi:10.1016/j.jad.2005.12.039

Caspi, A., & Moffitt, T. E. (2018). All for one and one for all: Mental disorders in one dimension. *American Journal of Psychiatry, 175*(9), 831-844. doi:10.1176/appi.ajp.2018.17121383

Cassioli, E., Rossi, E., Squecco, R., Baccari, M. C., Maggi, M., Vignozzi, L., Comeglio, P., Gironi, V., Lelli, L., Rotella, F., Monteleone, A. M., Ricca, V., & Castellini, G. (2020). Reward and psychopathological correlates of eating disorders: The explanatory role of leptin. *Psychiatry Research, 290,* 113071. doi:10.1016/j.psychres.2020.113071

Chami, R., Monteleone, A. M., Treasure, J., & Monteleone, P. (2019). Stress hormones and eating disorders. *Molecular and Cellular Endocrinology, 497,* 110349. doi:10.1016/j.mce.2018.12.009

Chapman, A. L., Gratz, K. L., & Brown, M. Z. (2006). Solving the puzzle of deliberate self-harm: The experiential avoidance model. *Behavior Research and Therapy, 44*(3), 371–394. doi:10.1016/j.brat.2005.03.005

Chater, N. (1999). The search for simplicity: A fundamental cognitive principle? *The Quarterly Journal of Experimental Psychology Section A, 52*(2), 273-302. doi:10.1080/713755819

Chater, N. & Vitányi, P. (2003). Simplicity: A unifying principle in cognitive science? *Trends in Cognitive Sciences, 7*(1), 19-22. doi:10.1016/S1364-6613(02)00005-0

Chatoor, I., & Kurpnick, J. (2001). The role of non-specific factors in treatment outcome of psychotherapy studies. *European Child & Adolescent Psychiatry 10*, S19. doi:10.1007/s007870170004

Chen, E. Y., Yiu, A., & Safer, D. L. (2018). Dialectical behaviour therapy and emotion-focused therapies for eating disorders. In W. S. Agras & A. Robinson (Eds.), *The Oxford handbook of eating disorders* (2nd ed., 334-350). New York, NY: Oxford University Press.

Clack, S., & Ward, T. (2020). Modeling the symptoms of psychopathology: A pluralistic approach. *New Ideas in Psychology*, *59*, 100799. doi:10.1016/j.newideapsych.2020.100799

Cloitre, M., Courtois, C. A., Charuvastra, A., Carapezza, R., Stolbach, B. C., & Green, B. L. (2011). Treatment of complex PTSD: Results of the ISTSS expert clinician survey on best practices. *Journal of Traumatic Stress, 24*(6), 615-627. doi:10.1002/jts.20697

Colles, S. L., Dixon, J. B., & O'Brien, P. E. (2008). Loss of control is central to psychological disturbance associated with binge eating disorder. *Obesity, 16*(3), 608-614. doi:10.1038/oby.2007.99

Colombetti, G. (2013). Psychopathology and the enactive mind. In K.W. M. Fulford, M. Davies, R. G. T. Gipps, G. Graham, J. Z. Sadler, G. Stranghellini, & T. Thornton (Eds.), *The Oxford handbook of philosophy and psychiatry*. Oxford, UK: Oxford University Press. doi:10.1093/oxfordhb/9780199579563.013.0063

Compan, V., Laurent, L., Jean, A., Macary, C., Bockaert, J., & Dumuis, A. (2012). Serotonin signalling in eating disorders. *Wiley Interdisciplinary Reviews: Membrane Transport and Signaling, 1*(6), 715-729. doi:10.1002/wmts.45

Corstorphine, E., Waller, G., Ohanian, V., & Baker, M. (2006). Changes in internal states across the binge-vomit cycle in bulimia nervosa. *The Journal of Nervous and Mental Disease, 194*(6), 446-449. doi:10.1097/01.nmd.0000221303.64098.23

Costello, E. G. (1992). *Symptoms of depression.* Hoboken, NJ: John Wiley & Sons Ltd.

Cowen, P. J., & Browning, M. (2015). What has serotonin to do with depression? *World Psychiatry, 14*(2), 158-160. doi:10.1002/wps.20229

Coyne, J. C., & Kok, R. N. (2014). Salvaging psychotherapy research: A manifesto. *Journal of Evidence-Based Psychotherapies, 14*(2), 105-124. https://psycnet.apa.org/record/2014-42863-001

Craig, A. D. (2002). How do you feel? Interoception: The sense of the physiological condition of the body. *Nature Reviews Neuroscience, 3*(8), 655-666. doi:10.1038/nrn894

Craighead, W. E., Johnson, B. N., Carey, S., & Dunlop, B. W. (2015). Psychosocial treatments for major depressive disorder. In P. E. Nathan & J. M. Gorman (Eds.), *A guide to treatments that work* (4th ed., 381-408). New York, NY: Oxford University Press.

Craske, M. G. (2012). Transdiagnostic treatment for anxiety and depression. *Depression & Anxiety, 29*(9), 749-753. doi:10.1002/da.21992

Craver, C. F. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience.* New York, NY: Oxford University Press.

Craver, C., & Kaplan, D. M. (2020). Are more details better? On the norms of completeness for mechanistic explanations. *The British Journal for the Philosophy of Science, 71*(1), 287-319. doi:10.1093/bjps/axy015

Cresci, G. A. M., & Izzo, K. (2019). The gut microbiome. In M. L. Corrigan, K. Roberts, & E. Steiger (Eds.), *Adult short bowel syndrome: Nutritional, medical, and surgical management* (pp. 45-54). London, UK: Elsevier Academic Press. doi:10.1016/C2017-0-01495-7

Cuijpers, P., Sijbrandij, M., Koole, S. L., Andersson, G., Beekman, A. T., & Reynolds III, C. F. (2013). The efficacy of psychotherapy and pharmacotherapy in treating depressive and anxiety disorders: A meta-analysis of direct comparisons. *World Psychiatry, 12*(2), 137-148. doi:10.1002/wps.20038

Culbert, K. M., Racine, S. E., & Klump, K. L. (2015). Research review: What we have learned about the causes of eating disorders - A synthesis of sociocultural, psychological, and biological research. *The Journal of Child Psychology and Psychiatry, 56*(11), 1141-1164. doi:10.1111/jcpp.12441

Culbert, K. M., Racine, S. E., & Klump, K. L. (2016). Hormonal factors and disturbances in eating disorders. *Current Psychiatry Reports, 18*(7), 65-81. doi:10.1007/s11920-016-0701-6

Curtis, E., Jones, R., Tipene-Leach, D., Walker, C., Loring, B., Paine, S-J., & Reid, P. (2019). Why cultural safety rather than cultural competency is required to achieve health equity: A literature review and recommended definition. *International Journal for Equity in Health, 18*, 174. doi:10.1186/s12939-019-1082-3

Cuthbert, B. N. (2014). The RDoC framework: Facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry, 13*(1), 28-35. doi:10.1002/wps.20087

Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine, 11*(1), 126. doi:10.1186/1741-7015-11-126

Dalal, F. (2018). *CBT: The cognitive behavioral tsunami: Managerialism, politics, and the corruptions of science.* New York, NY: Routledge.

Davis, H., & Attia, E. (2017). Pharmacotherapy of eating disorders. *Current Opinion in Psychiatry, 30*(6), 452-457. doi:10.1097/YCO.0000000000000358

de Haan, S. (2020). *Enactive psychiatry*. Cambridge, UK: Cambridge University Press.

Del Giudice, M. (2018). *Evolutionary psychopathology: A unified approach*. New York, NY: Oxford University Press.

Delaney, C. B., Eddy, K. T., Hartmann, A. S., Becker, A. E., Murray, H. B., & Thomas, J. J. (2015). Pica and rumination behavior among individuals seeking treatment for

eating disorders or obesity. *International Journal of Eating Disorders, 48*(2), 238-248. doi:10.1002/eat.22279

Devlin, M. J., Kissileff, H. R., Zimmerli, E. J., Samuels, F., Chen, B. E., Brown, A. J., Geliebter, A., & Walsh, B. T. (2012). Gastric emptying and symptoms of bulimia nervosa: Effect of a prokinetic agent. *Physiology & Behavior, 106*(2), 238-242. doi:10.1016/j.physbeh.2012.02.009

Dickerson Mayes, S., & Zickgraf, H. (2019). Atypical eating behaviors in children and adolescents with autism, ADHD, other disorders, and typical development. *Research in Autism Spectrum Disorders, 64*, 76-83. doi:10.1016/j.rasd.2019.04.002

DiClemente, C. C., & Prochaska, J. O. (1998). Toward a comprehensive, transtheoretical model of change: Stages of change and addictive behaviors. In W. R. Miller & N. Heather (Eds.), *Treating addictive behaviors* (2nd ed., pp. 3-24). New York, NY, US: Plenum Press.

Dishion, T. J., & Patterson, G. R. (2015). The development and ecology of antisocial behavior in children and adolescents. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology: Volume three: Risk, disorder, and adaptation* (2nd ed., pp. 503-541). Hoboken, NJ: John Wiley & Sons, Inc.

Donnelly, B., Touyz, S., Hay, P., Burton, A., Russell, J., & Caterson, I. (2018). Neuroimaging in bulimia nervosa and binge eating disorder: A systematic review. *Journal of Eating Disorders, 6*, 3. doi:10.1186/s40337-018-0187-1

Dovey, T. M., Kumari, V., Blissett, J., & 'Mealtime Hostage Parent Science Gang'. (2020). Eating behaviour, behavioural problems and sensory profiles of children with avoidant/restrictive food intake disorder (ARFID), autistic spectrum disorders or picky eating: Same or different? *European Psychiatry, 61*, 56-62. doi:10.1016/j.eurpsy.2019.06.008

Dwyer, D. S., Horton, R. Y., & Aamodt, E. J. (2011). Role of the evolutionarily conserved starvation response in anorexia nervosa. *Molecular Psychiatry, 16*(6), 595-603. doi:10.1038/mp.2010.95

Eddy, K. T., Dorer, D. J., Franko, D. L., Tahilani, K., Thompson-Brenner, H., & Herzog, D. B. (2008). Diagnostic crossover in anorexia nervosa and bulimia nervosa: Implications for DSM-V. *American Journal of Psychiatry, 165*(2), 6. doi:10.1176/appi.ajp.2007.07060951

Eddy, K. T., Tabri, N., Thomas, J. J., Murray, H. B., Keshaviah, A., Hastings, E., Edkins, K., Krishna, M., Herzog, D. B., Keel, P. K., & Franko, D. L. (2017). Recovery from anorexia nervosa and bulimia nervosa at 22-year follow-up. *Journal of Clinical Psychiatry, 78*(2), 184–189. doi:10.4088/JCP.15m10393

Ehrenthal, J. C., & Benecke, C. (2019). Tailored treatment planning for individuals with personality disorders: The operationalized psychodynamic diagnosis (OPD) approach. In U. Kramer (Ed.), *Case formulation for personality disorders* (pp. 291-314). Cambridge, MA: Academic Press. doi:10.1016/B978-0-12-813521-1.00015-1.

Eisler, I. (2005). The empirical and theoretical base of family therapy and multiple family day therapy for adolescent anorexia nervosa. *Journal of Family Therapy, 27*(2), 104-131. doi:10.1111/j.1467-6427.2005.00303.x

Evans, D. (2003). Hierarchy of evidence: A framework for ranking evidence evaluating healthcare interventions. *Journal of Clinical Nursing, 12*(1), 77-84. doi:10.1046/j.1365-2702.2003.00662.x

Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). Cognitive behaviour therapy for eating disorders: A "transdiagnostic" theory and treatment. *Behaviour Research and Therapy, 41*(5), 509-528. doi:10.1016/S0005-7967(02)00088-8

Fairburn, C. G., Shafran, R., & Cooper, Z. (1998). A cognitive behavioural theory of anorexia nervosa. *Behaviour Research and Therapy, 37*(1), 1-13. doi:10.1016/S0005-7967(98)00102-8

Faris, P. L., Kim, S. W., Meller, W. H., Goodale, R. L., Oakman, S. A., Hofbauer, R. D., Marshall, A. M., Daughters, R. S., Banerjee-Stevens, D., Eckert, E. D., & Hartman, B. K. (2000). Effect of decreasing afferent vagal activity with ondansetron on symptoms of bulimia nervosa: A randomised, double-blind trial. *The Lancet, 355*(9206), 792-797. doi:10.1016/S0140-6736(99)09062-5

Farstad, S. M., McGeown, L. M., & von Ranson, K. M. (2016). Eating disorders and personality 2004-2016: A systematic review and meta-analysis. *Clinical Psychology Review, 46,* 91-105. doi:10.1016/j.cpr.2016.04.005

Fava, M. (2003). Diagnosis and definition of treatment resistant depression. *Biological Psychiatry, 53*(8), 649-659. doi: 10.1016/S0006-3223(03)00231-2

Fawcett, E. J., Fawcett, J. M., & Mazmanian, D. (2016). A meta-analysis of the worldwide prevalence of pica during pregnancy and the postpartum period. *International Journal of Gynecology & Obstetrics, 133*(3), 277-283. doi:10.1016/j.ijgo.2015.10.012

Faye, J. (2014). *The nature of scientific thinking: On interpretation, explanation, and understanding.* New York, NY: Palgrave MacMillan.

Feinberg, A. P. (2008). Epigenetics at the epicenter of modern medicine. *JAMA, 299*(11), 1345-1350. doi:10.1001/jama.299.11.1345 %J JAMA

Feldman, J. (2016). The simplicity principle in perception and cognition. *WIREs Cognitive Science, 7*(5), 330-340. doi:10.1002/wcs.1406

Fichter, M. M., & Quadflieg, N. (2007). Long-term stability of eating disorder diagnoses. *International Journal of Eating Disorders, 40*(S3), S61-66. doi:10.1002/eat.20443

Fischer, S., Smith, G. T., & Cyders, M. A. (2008). Another look at impulsivity: A meta-analytic review comparing specific dispositions to rash action in their relationship to bulimic symptoms. *Clinical Psychology Review, 28,* 1413-1425. doi:10.1016/j.cpr.2008.09.001

Flament, M. F., Bissada, H., & Spettigue, W. (2012). Evidence-based pharmacotherapy of eating disorders. *International Journal of Neuropsychopharmacology, 15*(2), 189-207. doi:10.1017/S1461145711000381

Fleming, D. (2016). *Lean logic: A dictionary for the future and how to survive it.* S. Chamberlin & J. Porritt (Eds.). White River Junction, VT: Chelsea Green Publishing.

Forbush, K. T., Chen, P., Hagan, K. E., Chapa, D. A. N., Gould, S. R., Eaton, N. R., & Krueger, R. F. (2018). A new approach to eating-disorder classification: Using

empirical methods to delineate diagnostic dimensions and inform care. *International Journal of Eating Disorders, 51*(7), 710-721. doi:10.1002/eat.22891

Ford, B. Q., & Gross, J. J. (2018). Emotion regulation: Why beliefs matter. *Canadian Psychology/Psychologie Canadienne, 59*(1), 1-14. doi:10.1037/cap0000142

Forrest, L. N., Jones, P. J., Ortiz, S. N., Smooth, A. R. (2018). Psychopathology in anorexia nervosa and bulimia nervosa: A network analysis. *International Journal of Eating Disorders, 51*(7), 668-679. doi:10.1002/eat.22871

Forrest, L. N., Sarfan, L. D., Ortiz, S. N., Brown, T. A., & Smith, A. R. (2019). Bridging eating disorder symptoms and trait anxiety in patients with eating disorders: A network approach. *International Journal of Eating Disorders, 52*(6), 701-711. doi:10.1002/eat.23070

Foster, J. A., & McVey Neufeld, K-A. (2013). Gut–brain axis: How the microbiome influences anxiety and depression. *Trends in Neurosciences, 36*(5), 305-312. doi:10.1016/j.tins.2013.01.005

Fortune, C-A., & Heffernan, R. (2019). The psychology of criminal conduct: A consideration of strengths, weaknesses and future directions. *Psychology, Crime & Law, 25*(6), 659–674. doi:10.1080/1068316X.2018.1560445

Frank, G. K. W. (2013). Altered brain reward circuits in eating disorders: Chicken or egg? *Current Psychiatry Reports, 15*(10), 396. doi:10.1007/s11920-013-0396-x

Fredrickson, B. L., & Roberts, T. (1997). Objectification theory: Toward understanding women's lived experiences and mental health risks. *Psychology of Women Quarterly, 21*(2), 173–206. doi:10.1111/j.1471-6402.1997.tb00108.x

Freud, S. (1964). *The standard edition of the complete psychological works of Sigmund Freud*. J. Strachey (Ed.). London, UK: Macmillan.

Fried, E. I., & Nesse, R. M. (2015). Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR* D study. *Journal of Affective Disorders, 172*, 96-102. doi:10.1016/j.jad.2014.10.010

Fried, E. I. & Robinaugh, D. J. (2020). Systems all the way down: Embracing complexity in mental health research. *BMC Medicine, 18*, 205. doi:10.1186/s12916-020-01668-w

Fried, E. I., van Borkulo, C. D., Cramer, A. O. J., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017). Mental disorders as networks of problems: A review of recent insights. *Social Psychiatry and Psychiatric Epidemiology, 52*(1), 1-10. doi:10.1007/s00127-016-1319-z

Friederich, H-C., Wu, M., Simon, J. J., & Herzog, W. (2013). Neurocircuit function in eating disorders. *International Journal of Eating Disorders, 46*(5), 425-432. doi:10.1002/eat.22099

Friedman, M. (1974). Explanation and scientific understanding. *Journal of Philosophy, 71*(1), 5-19. https://www.jstor.org/stable/2024924?seq=1#metadata_info_tab_contents

Frith, U. (2003). *Autism: Explaining the enigma* (2nd ed.). Malden, MA: Blackwell Publishing.

Fuchs, T. (2018). *Ecology of the brain.* New York, NY: Oxford University Press.

Fumerton, R. (2006). *Epistemology.* Oxford, UK: Blackwell.

Fusar-Poli, P., Solmi, M., Brondino, N., Davies, C., Chae, C., Politi, P., Borgwardt, S., Lawrie, S. M., Parnas, J., & McGuire, P. (2019). Transdiagnostic psychiatry: A systematic review. *World Psychiatry, 18*(2), 192-207. doi:10.1002/wps.20631

Gabbard, G. O., & Kay, J. (2001). The fate of integrated treatment: Whatever happened to the biopsychosocial psychiatrist? *American Journal of Psychiatry, 158*(12), 1956-1963. doi:10.1176/appi.ajp.158.12.1956

Gagnon-Girouard, M-P., Chenel-Beaulieu, M-P., Amié, A., Ratté, C., & Bégin, C. (2019). Psychological meanings of eating disorders and their association with symptoms, motivation toward treatment, and clinical evolution among outpatients. *Europe's Journal of Psychology, 15*(2), 367-379. doi:10.5964/ejop.v15i2.1623

Galecki, P., Mossakowska-Wójcik, J., & Talarowska, M. (2018). The anti-inflammatory mechanism of antidepressants – SSRIs, SNRIs. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 80*(part C), 291-294. doi:10.1016/j.pnpbp.2017.03.016

Galmiche, M., Déchelotte, P., Lambert, G., & Tavolacci, M. P. (2019). Prevalence of eating disorders over the 2000–2018 period: A systematic literature review. *The*

American Journal of Clinical Nutrition, 109(5), 1402-1413. doi:10.1093/ajcn/nqy342

Gandini, S., Botteri, E., Iodice, S., Boniol, M., Lowenfels, A. B., Maisonneuve, P., & Boyle, P. (2007). Tobacco smoking and cancer: A meta-analysis. *International Journal of Cancer, 122*(1), 155-164. doi:10.1002/ijc.23033

Garfinkel, P. E., Garner, D. M., Rose, J., Darby, P. L., Brandes, J. S., O'Hanlon, J., & Walsh, N. (2009). A comparison of characteristics in the families of patients with anorexia nervosa and normal controls. *Psychological Medicine, 13*(4), 821-828. doi:10.1017/S0033291700051539

Gasquoine, P. G. (2014). Contributions of the insula to cognition and emotion. *Neuropsychology Review, 24*(2), 77-87. doi:10.1007/s11065-014-9246-9

Gatt, L., Jan, S., Mondraty, N., Horsfield, S., Hart, S., Russell, J., Laba, T. L., & Essue, B. (2014). The household economic burden of eating disorders and adherence to treatment in Australia. *BMC Psychiatry, 14*(1), 338-347. doi:10.1186/s12888-014-0338-0

Gaudiano, B. A. (2011). Evaluating acceptance and commitment therapy: An analysis of a recent critique. *International Journal of Behavioral Consultation and Therapy, 7*(1), 54-65. doi:10.1037/h0100927

Gay, N. G., Wisco, B. E., Jones, E. C., & Murphy, A. D. (2020). Posttraumatic stress disorder symptom network structures: A comparison between men and women. *Journal of Traumatic Stress, 33*(1), 96-105. doi:10.1002/jts.22470

Geddes, J. R., & Miklowitz, D. J. (2013). Treatment of bipolar disorder. *Lancet, 381*(9878), 1672-1682. doi:10.1016/S0140-6736(13)60857-0

Gillett, C. (2021). Using compositional explanations to understand compositional levels: An integrative account. In D. S. Brooks, J. DiFrisco, & W. C. Wimsatt (Eds.), *Levels of organization in the biological sciences.* Massachusetts, MA: MIT Press.

Gillon, R. (1994). Medical ethics: Four principles plus attention to scope. *BMJ, 309*(6948), 184-188. doi:10.1136/bmj.309.6948.184

Giordano, S. (2005). *Understanding eating disorders: Conceptual and ethical issues in the treatment of anorexia and bulimia nervosa.* New York, NY; Oxford University Press.

Granic, I., & Patterson, G. R. (2006). Toward a comprehensive model of antisocial development: A dynamic systems approach. *Psychological Review, 113*(1), 101-131. doi:10.1037/0033-295X.113.1.101

Gratz, K.L. (2003). Risk factors for and functions of deliberate self-harm: An empirical and conceptual review. *Clinical Psychology: Science and Practice, 10*(2), 192-205. doi:10.1093/clipsy.bpg022

Grave, R. D., Conti, M., Sartirana, M., Sermattei, S., & Calugi, S. (2021). Enhanced cognitive behaviour therapy for adolescents with eating disorders: A systematic review of current status and future perspectives. *Italian Journal of Eating Disorders and Obesity, 3*, 1-11. doi:10.32044/ijedo.2021.01

Greeno, C. G., Wing, R. R., & Shiffman, S. (2000). Binge antecedents in obese women with and without binge eating disorder. *Journal of Consulting and Clinical Psychology, 68*(1), 95-102. doi:10.1037/0022-006X.68.1.95

Gregertsen, E. C., Mandy, W., & Serpell, L. (2017). The egosyntonic nature of anorexia: An impediment to recovery in anorexia nervosa treatment. *Frontiers in Psychology, 8*, 2273. doi:10.3389/fpsyg.2017.02273

Guillaume, S., Gorwood, P., Jollant, F., Van den Eynde, F., Courtet, P., & Richard-Devantoy, S. (2015). Impaired decision-making in symptomatic anorexia nervosa and bulimia nervosa patients: A meta-analysis. *Psychological Medicine, 45*(16), 3377-3391. doi:10.1017/S003329171500152X

Hacker, P. (2009) Agential reasons and the explanation of human behaviour. In C. Sandis (Ed.), *New essays on the explanation of action*. London, UK: Palgrave Macmillan.

Haig, B. D. (2014). *Investigating the psychological world: Scientific method in the behavioural sciences*. Cambridge, MA: The MIT Press.

Hadassah Cheng, Z., Perko, V. L., Fuller-Marashi, L., Gau, J. M., & Stice, E. (2018). Ethnic differences in eating disorder prevalence, risk factors, and predictive effects of

risk factors among young women. *Eating Behaviors, 32,* 23-30. doi:10.1016/j.eatbeh.2018.11.004

Harrison, A., O'Brien, N., Lopez, C., & Treasure, J. (2010). Sensitivity to reward and punishment in eating disorders. *Psychiatry Research, 177*(1-2), 1-11. doi:10.1016/j.psychres.2009.06.010

Hay, P. (2020). Current approaches to eating disorders: A clinical update. *International Medicine Journal, 50*(1), 24-29. doi:10.1111/imj.14691

Hayes, S. C., Luoma, J. B., Bond, F. W, Masuda, A., & Lillis, J. (2006). Acceptance and commitment therapy: Model, processes and outcomes. *Behavior Research and Therapy, 44*(1), 1-25. doi:10.1016/j.brat.2005.06.006

Hayes S. C., Strosahl K. D., & Wilson K. G. (1999). *Acceptance and commitment therapy: An experiential approach to behavior change.* New York, NY: The Guilford Press.

Heatherton, T. F., & Baumeister, R. F. (1991). Binge eating as escape from self-awareness. *Psychological Bulletin, 110*(1), 86-108. doi:10.1037/0033-2909.110.1.86

Hedges, D., Farrer, T. J., Bigler, E. D., & Hopkins, R. O. (2019). Cognition in anorexia nervosa and bulimia nervosa. In D. Hedges, T. J. Farrer, E. D. Bigler, & R. O. Hopkins (Eds.), *The brain at risk: Associations between disease and cognition* (pp. 67-83). Cham, Switzerland: Springer Nature Switzerland.

Heffernan, R., & Ward, T. (2017). A comprehensive theory of dynamic risk and protective factors. *Aggression and Violent Behavior, 37,* 129–141. doi:10.1016/j.avb.2017.10.003

Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science, 15*(2), 135-175. doi:10.1086/286983

Henrich, J., Heine, S. J., Norenzayan, A. (2010). Beyond WEIRD: Towards a broad-based behavioral science. *Behavioral and Brain Sciences, 33*(2-3), 111-135. doi:10.1017/S0140525X10000725

Herbert, B. M., & Pollatos, O. (2018). The relevance of interoception for eating behavior and eating disorders. In M. Tsakiris & H. De Preester (Eds.), *The interoceptive mind:*

From homeostasis to awareness (pp. 165-186). Oxford, UK: Oxford University Press.

Hiroi, R., & Neumaier, J. F. (2011). Complex roles of estrogen in emotion: Sex matters. *Biological Psychiatry, 70*(10), 908-909. doi:10.1016/j.biopsych.2011.09.017

Hochstein, E. (2016). One mechanism, many models: A distributed theory of mechanistic explanation. *Synthese, 193*(5), 1387-1407. doi:10.1007/s11229-015-0844-8

Hochstein, E. (2017). Why one model is never enough: A defense of explanatory holism. *Biological Philosophy, 32*, 1105-1125. doi:10.1007/s10539-017-9595-x

Hoek, H. W. (2016). Review of the worldwide epidemiology of eating disorders. *Current Opinion in Psychiatry, 29*(6), 336-339. doi:10.1097/yco.0000000000000282

Hofmann, S. G., & Asmundson, G. J. G. (2008). Acceptance and mindfulness-based therapy: New wave or old hat? *Clinical Psychology Review, 28*(1), 1-16. doi:10.1016/j.cpr.2007.09.003

Hofmann, S. G., Asnaani, A., Vonk, I. J. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research, 36*(5), 427-440. doi:10.1007/s10608-012-9476-1

Hollon, S. D., DeRubeis, R. J., Fawcett, J., Amsterdam, J. D., Shelton, R. C., Zajecka, J., Young, P. R., & Gallop, R. (2014). Effect of cognitive therapy with antidepressant medications vs antidepressants alone on the rate of recovery in major depressive disorder: A randomized clinical trial. *JAMA Psychiatry, 71*(10), 1157-1164. doi:10.1001/jamapsychiatry.2014.1054

Holmes, E. A., Ghaderi, A., Harmer, C. J., Ramchandani, P. G., Cuijpers, P., Morrison, A. P., Roiser, J. P., Bockting, C. L. H., O'Connor, R. C., Shafran, R., Moulds, M. L., & Craske, M. G. (2018). Psychological treatments research in tomorrow's science. *The Lancet Psychiatry, 5*(3), 237-286. doi: 10.1016/S2215-0366(17)30513-8

Hooker, C. A. (1987). *A realistic theory of science*. New York, NY: State University of New York Press.

Howard, M. S. & Medway, F. J. (2004). Adolescent's attachment and coping with stress. *Psychology in the Schools, 41*(3), 391-402. doi:10.1002/pits.10167

Howe-Martin, L. S., Murrell, A. R., & Guarnaccia, C. A. (2012). Repetitive nonsuicidal self-injury as experiential avoidance among a community sample of adolescents. *Journal of Clinical Psychology, 68*(7), 809-829. doi:10.1002/jclp.21868

Howick, J. (2011). *The philosophy of evidence-based medicine*. UK: John Wiley & Sons.

Hudson, J. L., & Rapee, R. M. (2004). From anxious temperament to disorder: An etiological model. In R. G. Heimberg, C. L. Turk, & D. S. Mennin (Eds.), *Generalized anxiety disorder: Advances in research and practice* (pp. 51–74). New York, NY: The Guilford Press.

Hyman, S. E. (2010). The diagnosis of mental disorders: The problem of reification. *Annual Review of Clinical Psychology, 6*, 155-179. doi:10.1146/annurev-clinpsy.3.022806.091532

Hyman S. E. (2011). Diagnosing the DSM: Diagnostic classification needs fundamental reform. *Cerebrum: The Dana Forum on Brain Science, 2011*, 6. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3574782/

Illari, P. M., & Russo, F. (2014). *Causality: Philosophical theory meets scientific practice*. New York, NY: Oxford University Press.

Illari, P. M., & Williamson, J. (2012). What is a mechanism? Thinking about mechanisms across the sciences. *European Journal for Philosophy of Science, 2*, 119-135. doi:10.1007/s13194-011-0038-2

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Sanislow, C., & Wang, P. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry, 167*(7), 748-751. doi:10.1176/appi.ajp.2010.09091379

Jablensky, A. (2012). The nosological entity in psychiatry: A historical illusion or a moving target? In K. S. Kendler & J. Parnas (Eds.), *Philosophical issues in psychiatry II: Nosology* (pp. 77-94). Oxford, UK: Oxford University Press.

Jacobi, C., Hutter, K., & Fittig, E. (2018). Psychosocial risk factors for eating disorders. In W. S. Agras & A. Robinson (Eds.), *The Oxford handbook of eating disorders* (2nd ed., pp. 106-125). New York, NY: Oxford University Press.

Jenkinson, P. M., Taylor, L., & Laws, K. R. (2018). Self-reported interoceptive deficits in eating disorders: A meta-analysis of studies using the eating disorder inventory. *Journal of Psychosomatic Research, 110*, 38-45. doi:10.1016/j.jpsychores.2018.04.005

Johnstone, L., & Dallos, R. (2014). Introduction to formulation. In L. Johnstone & R. Dallos (Eds.), *Formulation in psychology and psychotherapy: Making sense of people's problems*. New York, NY: Routledge.

Jones, T. (2008). Unification. In M. Curd & S. Psillos (Eds.), *The Routledge companion to the philosophy of science.* Malden, MA: Routledge

Kafetsios, K. (2004). Attachment and emotional intelligence abilities across the life course. *Personality and Individual Differences, 37*(1), 129-145. doi:10.1016/j.paid.2003.08.006

Kaiser, M. L. (2015). *Reductive explanations in the biological sciences.* New York, NY: Springer.

Kaiser, M. I., & Krickel, B. (2016). The metaphysics of constitutive mechanistic phenomena. *The British Journal for the Philosophy of Science, 68*(3), 745–779. doi:10.1093/bjps/axv058

Kanakam, N., & Treasure, J. (2013). A review of cognitive neuropsychiatry in the taxonomy of eating disorders: State, trait, or genetic? *Cognitive Neuropsychiatry, 18*(1-2), 83-114. doi:10.1080/13546805.2012.682362

Kaye, W. H., Frank, G. K., Bailer, U. F., Henry, S. E., Meltzer, C. C., Price, J. C., Mathis, C. A., & Wagner, A. (2005). Serotonin alterations in anorexia and bulimia nervosa: New insights from imaging studies. *Physiology & Behavior, 85*(1), 73-81. doi:10.1016/j.physbeh.2005.04.013

Kazdin, A. E., & Blase, S. L. (2011). Rebooting psychotherapy research and practice to reduce the burden of mental illness. *Perspectives on Psychological Science, 6*(1), 21-37. doi:10.1177/1745691610393527

Keel, P. K. (2018). Epidemiology and course of eating disorders. In W. S. Agras & A. Robinson (Eds.), *The Oxford handbook of eating disorders* (2nd ed., pp. 34-36). New York, NY: Oxford University Press.

Keel, P. K., & Brown, T. A. (2010). Update on course and outcome in eating disorders. *International Journal of Eating Disorders*, *43*, 195-204. doi:10.1002/eat.20810

Keel, P. K., Brown, T. A., Holm-Denoma, J. M., & Bodell, L. P. (2011). Comparison of DSM-IV versus proposed DSM-5 diagnostic criteria for eating disorders: Reduction of eating disorder not otherwise specified and validity. *International Journal of Eating Disorders, 44*(6), 553-560. doi:10.1002/eat.20892

Keel, P. K., & Klump, K. L. (2003). Are eating disorders culture-bound syndromes? Implications for conceptualizing their etiology. *Psychological Bulletin, 129*(5), 747-769. doi:10.1037/0033-2909.129.5.747

Keijsers, G. P. J., Schaap, C. P. D. R., & Hoogduin, C. A. L. (2000). The impact of interpersonal patient and therapist behavior on outcome in cognitive-behavior therapy: A review of empirical studies. *Behavior Modification, 24*(2), 264-297. doi:10.1177/0145445500242006

Kendler, K. S. (2005). Toward a philosophical structure for psychiatry. *The American Journal of Psychiatry, 162*(3), 433-440. doi:10.1176/appi.ajp.162.3.433

Kendler, K. S., Parnas, J., & Zachar, P. (2020) (Eds.). *Levels of analysis in psychopathology: Cross-disciplinary perspectives.* Cambridge, UK: Cambridge University Press.

Kendler, K. S., Zachar, P., & Craver, C. (2011). What kinds of things are psychiatric disorders? *Psychological Medicine, 41*(6), 1143-1150. doi:10.1017/S0033291710001844

Keys, A., Brožek, J., Henschel, A., Mickelsen, O., & Taylor, H. L. (1950). *The biology of human starvation. (2 vols)*. Minneapolis, Min: University of Minnesota Press.

Khalsa, S. S., Portnoff, L. C., McCurdy-McKinnon, D., & Feusner, J. D. (2017). What happens after treatment? A systematic review of relapse, remission, and recovery in anorexia nervosa. *Journal of Eating Disorders*, *5*(1), 20. doi:10.1186/s40337-017-0145-3

Khan, A., Faucett, J., Lichtenberg, P., Kirsch, I., & Brown, W. A. (2012). A systematic review of comparative efficacy of treatments and controls for depression. *PLoS ONE, 7*(7), 11. doi:10.1371/journal.pone.0041778

Kim, K. R., Ku, J., Lee, J.-H., Lee, H., & Jung, Y.-C. (2012). Functional and effective connectivity of anterior insula in anorexia nervosa and bulimia nervosa. *Neuroscience Letters, 521*(2), 152-157. doi:10.1016/j.neulet.2012.05.075

Kirsch, I. (2014). Antidepressants and the placebo effect. *Journal of Psychology, 222*(3), 128-134. doi:10.1027/2151-2604/a000176

Kitcher, P. (1989). Explanatory unification and the causal structure of the world. In P. Kitcher, W. C. Salmon (Eds.), *Scientific explanation* (pp. 410-505). Minneapolis, Min: University of Minnesota Press.

Klabunde, M., Collado, D., & Bohon, C. (2017). An interoceptive model of bulimia nervosa: A neurobiological systematic review. *Journal of Psychiatric Research, 94*, 36-46. doi:10.1016/j.jpsychires.2017.06.009

Klein, D. A., & Walsh, B. T. (2004). Eating disorders: Clinical features and pathophysiology. *Physiology & Behavior, 81*(2), 359-374. doi:10.1016/j.physbeh.2004.02.009

Konjarski, M., Murray, G. Lee, V. V., & Jackson, M. L. (2018). Reciprocal relationships between daily sleep and mood: A systematic review of naturalistic prospective studies. *Sleep Medicine Reviews, 42*, 47-58. doi:10.1016/j.smrv.2018.05.005

Koplan, B. A., & Stevenson, W. G. (2007). Sudden arrhythmic death syndrome. *Heart, 93*, 547-548. doi:10.1136/2Fhrt.2006.108134

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., Brown, T. A., Carpenter, W. T., Caspi, A., Clark, L. A., Eaton, N. R., Forbes, M. K., Forbush, K. T., Goldberg, D., Hasin, D., Hyman, S. E., Ivanova, M. Y., Lynam, D. R., Markon, K., . . . Zimmerman, M. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology, 126*(4), 454–477. doi:10.1037/abn0000258

Kotov, R., Krueger, R. F., & Watson, D. (2018). A paradigm shift in psychiatric classification: The Hierarchical Taxonomy of Psychopathology (HiTOP). *World Psychiatry, 17*(1), 24-25. doi:10.1002/wps.20478

Krueger, R. F., & Eaton, N. R. (2015). Transdiagnostic factors of mental disorders. *World Psychiatry, 14*(1), 27–29. doi:10.1002/wps.20175

Kuikka, J. T., Tammela, L., Karhunen, L., Rissanen, A., Bergström, K. A., Naukkarinen, H., . . . Uusitupa, M. (2001). Reduced serotonin transporter binding in binge eating women. *Psychopharmacology, 155*(3), 310-314. doi:10.1007/s002130100716

Ladyman, J., Lambert, J., & Wiesner, K. (2013). What is a complex system? *European Journal of Philosophical Science, 3*, 33-67. doi:10.1007/s13194-012-0056-8

Laible, D. J., Carlo, G., & Roesch, S. C. (2004). Pathways to self-esteem in late adolescence: The role of parent and peer attachment, empathy, and social behaviours. *Journal of Adolescence, 27*(6), 703-716. doi:10.1016/j.adolescence.2004.05.005

Lam, D. D., Garfield, A. S., Marston, O. J., Shaw, J., & Heisler, L. K. (2010). Brain serotonin system in the coordination of food intake and body weight. *Pharmacology Biochemistry and Behavior, 97*(1), 84-91. doi:10.1016/j.pbb.2010.09.003

Larsen, P. S., Strandberg-Larsen, K., Micali, N., & Andersen, A. M. (2015). Parental and child characteristics related to early-onset disordered eating: A systematic review. *Harvard Review Psychiatry, 23*(6), 395-412. doi:10.1097/hrp.0000000000000073

Laughlin, R. B., & Pines, D. (2000). The theory of everything. *Proceedings of the National Academy of Sciences of the United States of America, 97*(1), 28-31. doi:10.1073/pnas.97.1.28

Lavagnino, L., Arnone, D., Cao, B., Soares, J. C., & Selvaraj, S. (2016). Inhibitory control in obesity and binge eating disorder: A systematic review and meta-analysis of neurocognitive and neuroimaging studies. *Neuroscience & Biobehavioral Reviews, 68*, 714-726. doi:10.1016/j.neubiorev.2016.06.041

Lavis, A. (2018). Not eating or tasting other ways to live: A qualitative analysis of 'living through' and desiring to maintain anorexia. *Transcultural Psychiatry, 55*(4), 454-474. doi:10.1177/1363461518785796

Leahy, R. L. (2008). A closer look at ACT. *The Behavior Therapist, 31*(8), 148–150. https://psycnet.apa.org/record/2009-02716-004

Leehr, E. J., Krohmer, K., Schag, K., Dresler, T., Zipfel, S., & Giel, K. E. (2015). Emotion regulation model in binge eating disorder and obesity: A systematic review. *Neuroscience & Biobehavioral Reviews, 49*, 125-134. doi:10.1016/j.neubiorev.2014.12.008

Le Grange, D., Lock, J., Loeb, K., & Nicholls, D. (2010). Academy for eating disorders position paper: The role of the family in eating disorders. *International Journal of Eating Disorders, 43*(1), 1-5. doi:10.1002/eat.20751

Le Grange, D., & Rienecke, R. (2018). Family therapy for eating disorders. In W. S. Agras & A. Robinson (Eds.), *The Oxford handbook of eating disorders* (2nd ed., pp. 319-333). New York, NY: Oxford University Press.

Lehman, B. J., David, D. M., & Gruber, J. A. (2017). Rethinking the biopsychosocial model of health: Understanding health as a dynamic system. *Social Personality Psychology Compass, 11*(8), e12328. doi:10.1111/spc3.12328

Leibowitz, S. F. (1990). The role of serotonin in eating disorders. *Drugs, 39*(3), 33-48. doi:10.2165/00003495-199000393-00005

Levinson, C. A., Brosof, L. C., Vanzhula, I., Christian, C., Jones, P., Rodebaugh, T. L., Langer, J. K., White, E. K., Warren, C., Weeks, J. W., Menatti, A., Lim, M. H. & Fernandez, K. C. (2018). Social anxiety and eating disorder comorbidity and underlying vulnerabilities: Using network analysis to conceptualize comorbidity. *International Journal of Eating Disorders, 51*(7), 693-709. doi:10.1002/eat.22890

Lilienfeld, S. O. (2007). Psychological treatments that cause harm. *Perspectives on Psychological Science, 2*(1), 53-70. doi:10.1111/j.1745-6916.2007.00029.x

Lilienfeld, S. O. (2014). The Research Domain Criteria (RDoC): An analysis of methodological and conceptual challenges. *Behaviour Research and Therapy, 62*, 129-139. doi:10.1016/j.brat.2014.07.019

Lilienfeld, S. O., & Marino, L. (1995). Mental disorder as a Roschian concept: A critique of Wakefield's "harmful dysfunction" analysis. *Journal of Abnormal Psychology, 104*(3), 411-420. doi:10.1037/0021-843X.104.3.411

Lilienfeld, S. O., Ritschel, L. A., Lynn, S. J., Cautin, R. L., & Latzmane, R. D. (2013). Why many clinical psychologists are resistant to evidence-based practice: Root causes and constructive remedies. *Clinical Psychology Review, 33*(7), 883-900. doi:10.1016/j.cpr.2012.09.008

Lilienfeld, S. O., Ritschel, L. A., Lynn, S. J., Cautin, R. L., & Latzmane, R. D. (2014). Why ineffective psychotherapies appear to work: A taxonomy of spurious therapeutic

effectiveness. *Perspectives on Psychological Science, 9*(4), 355-387. doi:10.1177/1745691614535216

Lilienfeld, S. O., & Treadway, M. T. (2016). Clashing diagnostic approaches: DSM-ICD versus RDoC. *Annual Review of Clinical Psychology, 12,* 435-463. doi:10.1146/annurev-clinpsy-021815-093122

Linardon, J., & Wade, T. D. (2018). How many individuals achieve symptom abstinence following psychological treatments for bulimia nervosa? A meta-analytic review. *International Journal of Eating Disorders*, *51*(4), 287-294. doi:10.1002/eat.22838

Linehan, M. M. (1987). Dialectical behaviour therapy for borderline personality disorder. *Bulletin of the Menninger Clinic*, *51*(3), 261-276. https://psycnet.apa.org/record/1987-32134-001

Linehan, M. M., & Chen, E. Y. (2005). Dialectical behavior therapy for eating disorders. In A. Freeman, S. H. Felgoise, C. M. Nezu, & M. A. Reinecke (Eds.), *Encyclopedia of cognitive behavior therapy* (pp. 168-171). New York, NY: Springer.

Lock, J., & Le Grange, D. (2005). Family-based treatment of eating disorders. *International Journal of Eating Disorders*, *37,* S64-S67. doi:10.1002/eat.20122

Lombrozo, T., & Carey, S. (2006). Functional explanation and the function of explanation. *Cognition, 99*(2), 167-204. doi:10.1016/j.cognition.2004.12.009

Long, C. G., Fitzgerald, K-A., & Hollin, C. R. (2012). Treatment of chronic anorexia nervosa: A 4-year follow-up of adult patients treated in an acute inpatient setting. *Clinical Psychology & Psychotherapy, 19*(1), 1-13. doi:10.1002/cpp.738

Lopiccolo, C. J., Eldon Taylor, C., Clemence, C., & Eisdorfer, C. (2005). Split treatment: A measurement of coordination between psychiatrists. *Psychiatry (Edgmont), 2*(1), 43–46. https://pubmed.ncbi.nlm.nih.gov/21179642/

Lo Sauro, C., Ravaldi, C., Cabras, P. L., Faravelli, C., & Ricca, V. (2008). Stress, hypothalamic-pituitary-adrenal axis and eating disorders. *Neuropsychobiology, 57*(3), 95-115. doi:10.1159/000138912

Lowe, M. R., & Butryn, M. L. (2007). Hedonic hunger: A new dimension of appetite? *Physiology & Behavior*, *91*(4), 432-439. doi:10.1016/j.physbeh.2007.04.006

Lydecker, J. A., & Grilo, C. M. (2019). I didn't want them to see: Secretive eating among adults with binge-eating disorder. *International Journal of Eating Disorders, 52*(2), 153-158. doi:10.1002/eat.23002

Machado, P. P., Gonçalves, S., & Hoek, H. W. (2013). DSM-5 reduces the proportion of EDNOS cases: Evidence from community samples. *International Journal of Eating Disorders,* 46, 60-65. doi:10.1002/eat.22040

Maiese, M. (2016). *Embodied selves and divided minds.* New York, NY: Oxford University Press.

Majuri, J., Joutsa, J., Johansson, J., Voon, V., Alakurtti, K., Parkkola, R., Lahti, T., Alho, H., Hirvonen, J., Arponen, E., Forsback, S., & Kaasinen, V. (2017). Dopamine and opioid neurotransmission in behavioral addictions: A comparative PET study in pathological gambling and binge eating. *Neuropsychopharmacology, 42*(5), 1169-1177. doi:10.1038/npp.2016.265

Mantzavinos, C. (2016). *Explanatory pluralism.* Cambridge, UK: Cambridge University Press.

Maon, I., Horesh, D., & Gvion, Y. (2020). Siblings of individuals with eating disorders: A review of the literature. *Frontiers in Psychiatry, 11*, 604. doi:10.3389/fpsyt.2020.00604

Marks, A. (2019). The evolution of our understanding and treatment of eating disorders over the past 50 years. *Journal of Clinical Psychology, 75*(8), 1380-1391. doi:10.1002/jclp.22782

Marques, L., Alegria, M., Becker, A. E., Chen, C., Fang, A., Chosak, A., & Diniz, J. B. (2011). Comparative prevalence, correlates of impairment, and service utilization for eating disorders across US ethnic groups: Implications for reducing ethnic disparities in health care access for eating disorders. *International Journal of Eating Disorders, 44*(5), 412-420. doi:10.1002/eat.20787

Martin, E., Dourish, C. T., Rotshtein, P., Spetter, M. S., & Higgs, S. (2019). Interoception and disordered eating: A systematic review. *Neuroscience & Biobehavioral Reviews, 107*, 166-191. doi:10.1016/j.neubiorev.2019.08.020

Mattu, J., & Sullivan, J. A. (2020). Classification, kinds, taxonomic stability and conceptual change. *Aggression and Violent Behavior*, 101477. doi:10.1016/j.avb.2020.101477

Mayer, E. A., Tillisch, K., & Gupta, A. (2015). Gut/brain axis and the microbiota. *The Journal of Clinical Investigation, 125*(3), 926-938. doi:10.1172/JCI76304

Mayes, R., & Horwitz, A. V. (2005). DSM-III and the revolution in the classification of mental illness. *Journal of the History of the Behavioral Sciences, 41*(3), 249-267. doi:10.1002/jhbs.20103

McElroy, S. L., Guerdjikova, A. I., Mori, N., & Keck Jr., P. E. (2018). Pharmacotherapy for eating disorders. In W. S. Agras & A. Robinson (Eds.), *The Oxford handbook of eating disorders* (2nd ed., pp. 359-394). New York, NY: Oxford University Press.

McElroy, S. L., Guerdjikova, A. I., Mori, N., & O'Melia, A. M. (2012). Current pharmacotherapy options for bulimia nervosa and binge eating disorder. *Expert Opinion on Pharmacotherapy, 13*(14), 2015-2026. doi:10.1517/14656566.2012.721781

McLeod, B. D., Wood, J. J., & Avny, S. B. (2011). Parenting and child anxiety disorders. In D. McKay & E. A. Storch (Eds.), *Handbook of child and adolescent anxiety disorders* (pp. 213–228). Berlin, Germany: Springer Science & Business Media.

McMeekin, N., Wu, O., Germeni, E., & Briggs, A. (2020). How methodological frameworks are being developed: Evidence from a scoping review. *BMC Medical Research Methodology, 20*, 173. doi:10.1186/s12874-020-01061-4

Mechanic, D. (2007). Mental health services then and now. *Health Affairs, 26*(6), 1548-1550. doi:10.1377/hlthaff.26.6.1548

Meehan, K. G., Loeb, K. L., Roberto, C. A., & Attia, E. (2006). Mood change during weight restoration in patients with anorexia nervosa. *International Journal of Eating Disorders, 39*(7), 587-589. doi:10.1002/eat.20337

Mehler, P. S. (2018). Medical complications of anorexia nervosa and bulimia nervosa. In W. S. Agras & A. Robinson (Eds.), *The Oxford handbook of eating disorders* (2nd ed., pp. 222-228). New York, NY: Oxford University Press.

Mehler, P. S., & Brown, C. (2015). Anorexia nervosa - medical complications. *Journal of Eating Disorders, 3*(11), 1-8. doi:10.1186/s40337-015-0040-8

Meier, M., Kossakowski, J. J., Jones, P. J., Kay, B., Riemann, B. C., & McNally, R. J. (2020). Obsessive–compulsive symptoms in eating disorders: A network investigation. *International Journal of Eating Disorders, 53*(3), 362-371. doi:10.1002/eat.23196

Merikangas, K. R., & Kalaydjian, A. (2007). Magnitude and impact of comorbidity of mental disorders from epidemiologic surveys. *Current Opinion in Psychiatry, 20*(4), 353-358. doi:10.1097/YCO.0b013e3281c61dc5

Merwin, R. M., Moskovich, A. A., Wagner, R., Ritschel, L. A., Craighead, L. W., Zucker, N. L. (2013). Emotion regulation difficulties in anorexia nervosa: Relationship to self-perceived sensory sensitivity. *Cognition and Emotion, 27*(3), 441-452. doi:10.1080/02699931.2012.719003

Meyer, C., Leung, N., Barry, L., & De Feo, D. (2009). Emotion and eating psychopathology: Links with attitudes towards emotional expression among young women. *International Journal of Eating Disorders, 43*(2), 187-189. doi:10.1002/eat.20659

Meyer, D. J., & Simon, R. I. (1999a). Split treatment: Clarity between psychiatrists and psychotherapists, part I. *Psychiatry Annals, 29*(5), 241-245. doi:10.3928/0048-5713-19990501-04

Meyer, D. J., & Simon, R. I. (1999b). Split treatment: Clarity between psychiatrists and psychotherapists, part II. *Psychiatry Annals, 29*(6), 241-245. doi:10.3928/0048-5713-19990601-03

Ministry of Health. (2016). *Mental Health and Addiction: Service use 2012/13*. Wellington, New Zealand: Ministry of Health

Ministry of Health. (2018). *Mental Health and Addiction: Service use 2015/16*. Wellington, New Zealand: Ministry of Health

Minuchin, S., Baker, L., Rosman, B. L., Liebman, R., Milman, L., & Todd, T. C. (1975). A conceptual model of psychosomatic illness in children: Family organization and family therapy. *Archives of General Psychiatry*, *32*(8), 1031-1038. doi:10.1001/archpsyc.1975.01760260095008

Minuchin, S., Rosman, B. L., & Baker, L. (1978). *Psychosomatic families: Anorexia nervosa in context*. Boston, MA: Harvard University Press.

Mitchell, S. D. (2002). Integrative pluralism. *Biology and Philosophy, 17*, 55-70. doi:10.1023/A:1012990030867

Mitchell, S. D., & Dietrich, M. R. (2006). Integration without unification: An argument for pluralism in the biological sciences. *The American Naturalist, 168*(Suppl 6), S73-S79. doi:10.1086/509050

Mojtabai, R., & Rieder, R. O. (1998). Limitations of the symptom-oriented approach to psychiatric research. *British Journal of Psychiatry, 173*, 198-202. doi:10.1192/bjp.173.3.198

Mölbert, S. C., Klein, L., Thaler, A., Mohler, B. J., Brozzo, C., Martus, P., Karnath, H., Zipfel, S, & Giel, K. E. (2017). Depictive and metric body size estimation in anorexia nervosa and bulimia nervosa: A systematic review and meta-analysis. *Clinical Psychology Review, 57*, 21-31. doi:10.1016/j.cpr.2017.08.005

Monroe, S. M., & Cummins, L. F. (2015). Diathesis-stress models. In R.L. Cautin and S.O. Lilienfeld (Eds.), *The Encyclopedia of clinical psychology* (pp. 1-6). doi:10.1002/9781118625392.wbecp466

Moradi, B., & Huang, Y. (2008). Objectification theory and psychology of women: A decade of advances and future directions. *Psychology of Women Quarterly*, *32*(4), 377-398. doi:10.1111/j.1471-6402.2008.00452.x

Morrow, J. (2010). *Hollow: An unpolished tale.* Chicago, IL: Moody Publishers.

Muang, H. H. (2016). Diagnosis and causal explanation in psychiatry. *Studies in History and Philosophy of Biological and Biomedical Science, 60*, 15-24. doi:10.1016/j.shpsc.2016.09.003

Mulder, R., Murray, G., & Rucklidge, J. (2017). Common versus specific factors in psychotherapy: Opening the black box. *Lancet Psychiatry, 4*(12), 953-962. doi:10.1016/S2215-0366(17)30100-1

Murialdo, G., Casu, M., Falchero, M., Brugnolo, A., Patrone, V., Cerro, P. F., . . . Ferro, A. M. (2007). Alterations in the autonomic control of heart rate variability in patients with anorexia or bulimia nervosa: Correlations between sympathovagal activity,

clinical features, and leptin levels. *Journal of Endocrinological Investigation, 30*(5), 7. doi:10.1007/BF03346310

Murphy, D. (2006). *Psychiatry in the scientific image*. Cambridge, MA: MIT Press.

Murray, S. B. (2020). Updates in the treatment of eating disorders in 2019: A year in review in Eating Disorders: The Journal of Treatment & Prevention. *Eating Disorders: The Journal of Treatment & Prevention, 28*(1), 21-31. doi:10.1080/10640266.2020.1723373

Murray, H. B., Juarascio, A. S., Di Lorenzo, C., Drossman, D. A., & Thomas, J. J. (2019). Diagnosis and treatment of rumination syndrome: A critical review. *The American Journal of Gastroenterology, 114*(4), 562-578. doi:10.14309/ajg.0000000000000060

Murray, S. B., Quintana, D. S., Loeb, K. L., Griffiths, S., & Le Grange, D. (2018). Treatment outcomes for anorexia nervosa: A systematic review and meta-analysis of randomized controlled trials. *Psychological Medicine, 49*(4), 535-544. doi:10.1017/S0033291718002088

Mustafa, Z., & Ghaffari, M. (2020). Diagnostic methods, clinical guidelines, and antibiotic treatment for group A streptococcal pharyngitis: A narrative review. *Frontiers in Cellular and Infection Microbiology, 10,* 563627. doi:10.3389/fcimb.2020.563627

Nemani, K., Hosseini Ghomi, R., McCormick, B., & Fan, X. (2015). Schizophrenia and the gut–brain axis. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 56,* 155-160. doi:10.1016/j.pnpbp.2014.08.018

Newton-Smith, W. (2002). *A companion to the philosophy of science*. Malden, MA: Blackwell.

Ng, L., & Davis, C. (2013). Cravings and food consumption in binge eating disorder. *Eating Behaviors, 14*(4), 472-475. doi:10.1016/j.eatbeh.2013.08.011

Nielsen, K., & Ward, T. (2018). Towards a new conceptual framework for psychopathology: Embodiment, enactivism and embedment. *Theory & Psychology, 28*(6), 800-822. doi:10.1177/0959354318808394

Nielsen, K., & Ward, T. (2020a). Phenomena complexes as targets of explanation in psychopathology: The relational analysis of phenomena (RAP) approach. *Theory & Psychology, 30*(2), 164-185. doi:10.1177/0959354320906462

Nielsen, K., & Ward, T. (2020b). Mental disorder as both natural and normative: Developing the normative dimension of the 3e conceptual framework for psychopathology. *Journal of Theoretical and Philosophical Psychology, 40*(2), 107-123. doi:10.1037/teo0000118

Nordbø, R. H. S., Espeset, E. M. S., Gulliksen, K. S., Skårderud, F., & Holte, A. (2006). The meaning of self-starvation: Qualitative study of patients' perception of anorexia nervosa. *International Journal of Eating Disorders, 39*(7), 556-564. doi:10.1002/eat.20276

Norton, P. J., & Paulus, D. J. (2017). Transdiagnostic models of anxiety disorder: Theoretical and empirical underpinnings. *Clinical Psychology Review, 56*, 122-137. doi:10.1016/j.cpr.2017.03.004

Nowakowski, M. E., McFarlane, T., & Cassin, S. E. (2013). Alexithymia and eating disorders: A critical review of the literature. *Journal of Eating Disorders, 1*(1), 21. doi:10.1186/2050-2974-1-21

Nutt, D. J., Forshall, S., Bell, C., Rich, A., Sandford, J., Nash, J., & Argyropoulos, S. (1999). Mechanisms of action of selective serotonin reuptake inhibitors in the treatment of psychiatric disorders. *European Neuropsychopharmacology, 9*(S3), S81-S86. doi:10.1016/S0924-977X(99)00030-9

Osman, J. L., & Sobal, J. (2006). Chocolate cravings in American and Spanish individuals: Biological and cultural influences. *Appetite, 47*(3), 290-301. doi:10.1016/j.appet.2006.04.008

Panayiotou, G., Pantelli, M., & Vlemincx, E. (2021). Adaptive and maladaptive emotion processing and regulation, and the case of alexithymia. *Cognition and Emotion, 35*(3), 488-499. doi:10.1080/02699931.2019.1671322

Pearson, C. M., Wonderlich, S. A., & Smith, G. T. (2015). A risk and maintenance model for bulimia nervosa: From impulsive action to compulsive behavior. *Psychology Review, 122*(3), 516-535. doi:10.1037/a0039268

Peat, C. M., Berkman, N. D., Lohr, K. N., Brownley, K. A., Bann, C. M., Cullen, K., Quattlebaum, M. J., & Bulik, C. M. (2017). Comparative effectiveness of treatments for binge-eating disorder: Systematic review and network meta-analysis. *European Eating Disorders Review, 25*(5), 317-328. doi:10.1002/erv.2517

Peebles, R., Wilson, J. L., & Lock, J. D. (2011). Self-injury in adolescents with eating disorders: Correlates and provider bias. *Journal of Adolescent Health, 48*(3), 310-313. doi:10.1016/j.jadohealth.2010.06.017

Pender, S., Gilbert, S. J., & Serpell, L. (2014). The neuropsychology of starvation: Set-shifting and central coherence in a fasted nonclinical sample. *PLoS ONE, 9*(10), e110743. doi:10.1371/journal.pone.0110743

Perpiñá, C., Segura, M., & Sánchez-Reales, S. (2017). Cognitive flexibility and decision-making in eating disorders and obesity. *Eating and Weight Disorders, 22*, 435-444. doi:10.1007/s40519-016-0331-3

Peschel, S. K. V., Feeling, N. R., Vögele, C., Kaess, M., Thayer, J. F., & Koenig, J. (2016). A meta-analysis on resting state high-frequency heart rate variability in bulimia nervosa. *European Eating Disorders Review, 24*(5), 355-365. doi:10.1002/erv.2454

Pike, K. M., Hoek, H. W., & Dunne, P. E. (2014). Cultural trends and eating disorders. *Current Opinion in Psychiatry, 27*(6), 436-442. doi:10.1097/yco.0000000000000100

Pinaquy, S., Chabrol, H., Simon, C., Louvet, J. P., & Barbe, P. (2003). Emotional eating, alexithymia, and binge-eating disorder in obese women. *Obesity research, 11*(2), 195–201. doi:10.1038/oby.2003.31

Pollice, C., Kaye, W. H., Greeno, C. G., & Weltzin, T. E. (1998). Relationship of depression, anxiety, and obsessionality to state of illness in anorexia nervosa. *International Journal of Eating Disorders, 21*(4), 367-376. doi:10.1002/(SICI)1098-108X(1997)21:4<367::AID-EAT10>3.0.CO;2-W

Polivy, J., & Herman, C. P. (1985). Dieting and binging: A causal analysis. *American Psychologist, 40*(2), 193-201. doi:10.1037/0003-066X.40.2.193

Potochnik, A. (2010). Levels of explanation reconceived. *Philosophy of Science, 77*(1), 59-72. doi:10.1086/650208

Potochnik, A. & McGill, B. (2012). The limitations of hierarchical organization. *Philosophy of Science, 79*(1), 120-140. doi:10.1086/663237

Povich, M. (forthcoming). Mechanistic explanation in psychology. In H. Stam & H. L. De Jong (Eds.), *The SAGE Handbook of Theoretical Psychology*. Sage. https://philarchive.org/archive/POVMEI

Prentice, A. M., & Jebb, S. A. (2001). Beyond body mass index. *Obesity Reviews, 2*(3), 141-147. doi:10.1046/j.1467-789x.2001.00031.x

Prince, A. C., Brooks, S. J., Stahl, D., & Treasure, J. (2009). Systematic review and meta-analysis of the baseline concentrations and physiologic responses of gut hormones to food in eating disorders. *American Journal of Clinical Nutrition, 89*(3), 755-765. doi:10.3945/ajcn.2008.27056

Pringle, A., Browning, M., Cowen, P. J., & Harmer, C. J. (2011). A cognitive neuropsychological model of antidepressant drug action. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 35*(7), 1586-1592. doi:10.1016/j.pnpbp.2010.07.022

Reas, D. L., & Grilo, C. M. (2015). Pharmacological treatment of binge eating disorder: Update review and synthesis. *Expert Opinion of Pharmacotherapy, 16*(10), 1463-1478. doi:10.1517/14656566.2015.1053465

Resick, P. A., Bovin, M. J., Calloway, A. L., Dick, A. M., King, M. W., Mitchell, K. S., Suvak, M. K., Wells, S. Y., Stirman, S. W., & Wolf, E. J. (2012). A critical evaluation of the complex PTSD literature: Implications for DSM-5. *Journal of Traumatic Stress, 25*(3), 241-251. doi:10.1002/jts.21699

Rieger, E., Van Buren, D. J., Bishop, M., Tanofsky-Kraff, M., Welch, R., & Wilfley, D. E. (2010). An eating disorder-specific model of interpersonal psychotherapy (IPT-ED): Causal pathways and treatment implications. *Clinical Psychology Review, 30*, 400-410. doi:10.1016/j.cpr.2010.02.001

Ricca, V., Castellini, G., Fioravanti, G., Lo Sauro, C., Rotella, F., Ravaldi, C., Lazzeretti, L., & Faravelli, C. (2012). Emotional eating in anorexia nervosa and bulimia

nervosa. *Comprehensive Psychiatry, 53*(3), 245-251. doi:10.1016/j.comppsych.2011.04.062

Roberts, M. E., Tchanturia, K., Stahl, D., Southgate, L., & Treasure, J. (2007). A systematic review and meta-analysis of set-shifting ability in eating disorders. *Psychological Medicine, 37*, 1075-1084. doi:10.1017/S0033291707009877

Robinson, E., Marty, L., Higgs, S., & Jones, A. (2021). Interoception, eating behaviour and body weight. *Physiology & Behavior, 237*, 113434. doi:10.1016/j.physbeh.2021.113434

Rogers, K. H., Luton, R., Biggs, H., Biggs, R., Blignaut, S., Choles, A. G., Palmer, C. G., & Tangwe, P. (2013). Fostering complexity thinking in action research for change in social-ecological systems. *Ecology and Society, 18*(2), 31. doi:10.5751/ES-05330-180231

Roiser, J. P., Elliott, R., & Sahakian, B. J. (2012). Cognitive mechanisms of treatment in depression. *Neuropsychopharmacology, 37*(1), 117-136. doi:10.1038/npp.2011.183

Rowa, K., Kerig, P. K., & Geller, J. (2001). The family and anorexia nervosa: Examining parent-child boundary problems. *European Eating Disorders Review, 9*(2), 97-114. doi:10.1002/erv.383

Ruphy, S. (2016). *Scientific pluralism reconsidered*. Pittsburgh, PA: University of Pittsburgh Press.

Russell, G. (1979). Bulimia nervosa: An ominous variant of anorexia nervosa. *Psychological Medicine*, *9*(3), 429-448. doi:10.1017/s0033291700031974.

Russell, G., Szmukler, G. I., Dare, C., & Eisler, I. (1987). An evaluation of family therapy in anorexia nervosa and bulimia nervosa. *Archives of General Psychiatry*, *44*(12), 1047-1056. doi:10.1001/archpsyc.1987.01800240021004

Rzevski, G., & Skobelev, P. (2014). *Managing complexity*. Southhampton, Boston: WIT Press.

Samnaliev, M., Noh, H. L., Sonneville, K. R., & Austin, S. B. (2015). The economic burden of eating disorders and related mental health comorbidities: An exploratory

analysis using the U.S. Medical Expenditures Panel Survey. *Preventive Medicine Reports, 2*, 32-34. doi:10.1016/j.pmedr.2014.12.002

Savulescu, J., Roache, R., & Davies, W. (Eds.). (2020) *Psychiatry reborn: Biopsychosocial psychiatry in modern medicine.* Oxford, UK: Oxford University Press.

Schaefer, L. M., Burke, N. L., & Thompson, J. K. (2018). Thin-ideal internalization: How much is too much? *Eating and Weight Disorders, 24*, 933–937. doi:10.1007/s40519-018-0498-x

Schag, K., Teufel, M., Junne, F., Preissl, H., Hautzinger, M., Zipfel, S., & Giel, K. E. (2013). Impulsivity in binge eating disorder: Food cues elicit increased reward responses and disinhibition. *PLoS ONE, 8*(10), e76542-e76542. doi:10.1371/journal.pone.0076542

Schaumberg, K., Welch, E., Breithaupt, L., Hübel, C., Baker, J. H., Munn-Chernoff, M. A., Yilmaz, Z., Ehrlich, S., Mustelin, L., Ghaderi, A., Hardaway, A. J., Bulik-Sullivan, E. C., Hedman, A. M., Jangmo, A., Nilsson, I. A. K., Wiklund, C., Yao, S., Seidel, M., & Bulik, C. M. (2017). The science behind the academy for eating disorders' nine truths about eating disorders. *European Eating Disorders Review, 25*(6), 432-450. doi:10.1002/erv.2553

Schienle, A., Schäfer, A., Hermann, A., & Vaitl, D. (2009). Binge-eating disorder: Reward sensitivity and brain activation to images of food. *Biological Psychiatry, 65*(8), 654-661. doi:10.1016/j.biopsych.2008.09.028

Schmidt, U., Adan, R., Böhm, I., Campbell, I. C., Dingemans, A., Ehrlich, S., Elzakkers, I., Favaro, A., Giel, K., Harrison, A., Himmerich, H., Hoek, H. W., Herpertz-Dahlmann, B., Kas, M. J., Seitz, J., Smeets, P., Sternheim, L., Tenconi, E., van Elburg, A., van Furth, E., & Zipfel, S. (2016). Eating disorders: The big issue. *The Lancet Psychiatry*, *3*(4), 313-315. doi:10.1016/S2215-0366(16)00081-X

Seitz, J., Trinh, S., & Herpertz-Dahlmann, B. (2019). The microbiome and eating disorders. *Psychiatric Clinics of North America, 42*(1), 93-103. doi:10.1016/j.psc.2018.10.004

Sfärlea, A., Dehning, S., Keller, S. K., & Schulte-Körne, G. (2019). Alexithymia predicts maladaptive but not adaptive emotion regulation strategies in adolescent girls

with anorexia nervosa or depression. *Journal of Eating Disorders, 7*, 41. doi:10.1186/s40337-019-0271-1

Shapiro, D. L., Berkman, N. D., Brownley, K. A., Sedway, J. A., Lohr, K. N., & Bulik, C. M. (2007). Bulimia nervosa treatment: A systematic review of randomized controlled trials. *International Journal of Eating Disorders, 40*(4), 321-336. doi:10.1002/eat.20372

Shelton, R. C., Osuntokun, O., Heinloth, A. N., & Corya, S. A. (2010). Therapeutic options for treatment-resistant depression. *CNS Drugs, 24*(2), 131-161. doi:10.2165/11530280-000000000-00000

Simmons, W. K., & DeVille, D. C. (2017). Interoceptive contributions to healthy eating and obesity. *Current Opinion in Psychology, 17*, 106-112. doi:10.1016/j.copsyc.2017.07.001

Skinner, B. F. (1976). *About behaviorism*. New York, NY: Random house, Inc.

Skodol, A. E. (2018). Impact of personality pathology on psychosocial functioning. Current *Opinion in Psychology, 21*, 33-38. doi:10.1016/j.copsyc.2017.09.006

Slater, M. H., & Borghini, A. (2013). Introduction: Lessons from the scientific butchery. In J. K. Campbell, M. O'Rourke, & M. H. Slater (Eds.), *Carving nature at its joints: Natural kinds in metaphysics and science*. Cambridge, MA: MIT Press.

Smith, B. (2003). Ontology. In Luciano Floridi (Ed.), *Blackwell Guide to the Philosophy of Computing and Information* (pp. 155-166). Oxford, UK: Blackwell.

Smith, K. E., Crosby, R. D., Wonderlich, S. A., Forbush, K. T., Mason, T. B., & Moessner, M. (2018). Network analysis: An innovative framework for understanding eating disorder psychopathology. *International Journal of Eating Disorders, 51*(3), 214-222. doi: 10.1002/eat.22836

Smith, T. E., Lee, C. A., Martel, M. M., & Axelrad, M. E. (2017). ODD symptom network during preschool. *Journal of Abnormal Child Psychology, 45*, 743-748. doi:10.1007/s10802-016-0196-y

Solmi, M., Collantoni, E., Meneguzzo, P., Degortes, D., Tenconi, E., & Favaro, A. (2018). Network analysis of specific psychopathology and psychiatric symptoms in

patients with eating disorders. *International Journal of Eating Disorders, 51*(7), 680-692. doi:10.1002/eat.22884

Solmi, M., Collantoni, E., Meneguzzo, P., Tenconi, E., & Favaro, A. (2019). Network analysis of specific psychopathology and psychiatric symptoms in patients with anorexia nervosa. *European Eating Disorders Review, 27*(1), 24-33. doi:10.1002/erv.2633

Souery, D., Papakostas, G. I., & Trivedi, M. H. (2006). Treatment-resistant depression. *The Journal of Clinical Psychiatry, 67*(Suppl 6), 16–22. https://pubmed.ncbi.nlm.nih.gov/16848672/

Spring, B., & Neville, K. (2011). Evidence-based practice in clinical psychology. In D. H. Barlow (Ed.), *The Oxford handbook of clinical psychology* (pp. 128-149). New York, NY: Oxford University Press.

Springer, K. S., Levy, H. C., & Tolin, D. F. (2018). Remission in CBT for adult anxiety disorders: A meta-analysis. *Clinical Psychology Review, 61*, 1-8. doi:10.1016/j.cpr.2018.03.002

Stack-Sullivan, H. (1955). *The interpersonal theory of psychiatry*. London, UK: Tavistock Publications Limited.

Staddon, J. E. R., & Cerutti, D. T. (2003). Operant conditioning. *Annual Review of Psychology, 54*(1), 115-144. doi:10.1146/annurev.psych.54.101601.145124

Steiger, H. (2004). Eating disorders and the serotonin connection: State, trait and developmental effects. *Journal of Psychiatry & Neuroscience*, *29*(1), 20-29. https://www.ncbi.nlm.nih.gov/pubmed/14719047

Stein, D. J., Phillips, K. A., Bolton, D., Fulford, K. W. M., Sadler, J. Z., & Kendler, K. S. (2010). What is a mental/psychiatric disorder? From DSM-IV to DSM-V. *Psychological Medicine, 40*(11), 1759-1765. doi:10.1017/S0033291709992261

Steward, T., Mestre-Bach, G., Vintro-Alcaraz, C., Aguera, Z., Jimenez-Murcia, S., Granero, R., & Fernandez-Aranda, F. (2017). Delay discounting of reward and impulsivity in eating disorders: from anorexia nervosa to binge eating disorder. *European Eating Disorders Review, 25*(6), 601-606. doi:10.1002/erv.2543

Stewart, R. E., Stirman, S. W., & Chambless, D. L. (2012). A qualitative investigation of practicing psychologists' attitudes toward research-informed practice: Implications for dissemination strategies. *Professional Psychology Research and Practice, 43*(2), 100-109. doi:10.1037/a0025694

Stiegler, L. N. (2005). Understanding pica behavior: A review for clinical and education professionals. *Focus on Autism and Other Developmental Disabilities, 20*(1), 27-38. doi:10.1177/10883576050200010301

Stinson, C. (2019). The absent body in psychiatric diagnosis, treatment, and research. *Synthese, 196*, 2153-2176. doi:10.1007/s11229-017-1507-8

Stockford, C., Stenfert Kroese, B., Beesley, A., & Leung, N. (2018). Severe and enduring anorexia nervosa: The personal meaning of symptoms and treatments. *Women's Studies International Forum, 68*, 129-138. doi:10.1016/j.wsif.2018.03.003

Strauss-Hughes, A., Heffernan, R., & Ward, T. (2019). A cultural-ecological perspective on agency and offending behaviour. *Psychiatry, Psychology and Law, 26*(6), 938-958. doi:10.1080/13218719.2019.1644250

Striegel Weissman, R., & Rosselli, F. (2017). Reducing the burden of suffering from eating disorders: Unmet treatment needs, cost of illness, and the quest for cost-effectiveness. *Behaviour Research and Therapy, 88*, 49-64. doi:10.1016/j.brat.2016.09.006

Strober, M. (2004), Managing the chronic, treatment-resistant patient with anorexia nervosa. *International Journal of Eating Disorders, 36*(3), 245-255. doi:10.1002/eat.20054

Strober, M., Freeman, R., Lampert, C., Diamond, J., & Kaye, W. H. (2000). Controlled family study of anorexia nervosa and bulimia nervosa: evidence of shared liability and transmission of partial syndromes. *The American Journal of Psychiatry, 157*(3), 393-401. doi:10.1176/appi.ajp.157.3.393

Stuhldreher, N., Konnopka, A., Wild, B., Herzog, W., Zipfel, S., Löwe, B., & König, H.-H. (2012). Cost-of-illness studies and cost-effectiveness analyses in eating disorders: a systematic review. *International Journal of Eating Disorders, 45*(4), 476-491. doi:10.1002/eat.20977

Sullivan, J. A. (2017). Coordinated pluralism as a means to facilitate integrative taxonomies of cognition. *Philosophical Explorations, 20*(2), 129-145. doi:10.1080/13869795.2017.1312497

Svaldi, J., Schmitz, F., Baur, J., Hartmann, A., Legenbauer, T., Thaler, C., von Wietersheim, J., de Zwann, M., & Tuschen-Caffier, B. (2018). Efficacy of psychotherapies and pharmacotherapies for bulimia nervosa. *Psychological Medicine, 49*(6), 898-910. doi:10.1017/S0033291718003525

Swanson, S. A., Crow, S. J., Le Grange, D., Swendsen, J., & Merikangas, K. R. (2011). Prevalence and correlates of eating disorders in adolescents. *Archives of General Psychiatry, 68*(7), 714-723. doi:10.1001/archgenpsychiatry.2011.22

Sysko, R., Devlin, M. J., Walsh, B. T., Zimmerli, E., & Kissileff, H. R. (2007). Satiety and test meal intake among women with binge eating disorder. *International Journal of Eating Disorders, 40*(6), 554-561. doi:10.1002/eat.20384

Szasz, T. S. (1960). The myth of mental illness. *American Psychologist, 15*(2), 113-118. doi:10.1037/h0046535

Tafet, G. E., & Nemeroff, C. B. (2020). Pharmacological treatment of anxiety disorders: The role of the HPA axis. *Frontiers in Psychiatry, 11*, 443. doi:10.3389/fpsyt.2020.00443

Tanner-Smith, E. E., Wilson, S. J., & Lipsey, M. W. (2013). The comparative effectiveness of outpatient treatment for adolescent substance abuse: A meta-analysis. *Journal of Substance Abuse Treatment, 44*(2), 145-158. doi:10.1016/j.jsat.2012.05.006

Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences, 12*(3), 435-467. doi:10.1017/S0140525X00057046

Thagard, P. (1992). *Conceptual revolutions*. Princeton, NJ: Princeton University Press.

Thomas, J. G., & Sharp, P. B. (2019). Mechanistic science: A new approach to comprehensive psychopathology research that relates psychological and biological phenomena. *Clinical Psychological Science, 7*(2), 196–215. doi:10.1177/2167702618810223

Thompson, J. K. (2004). The (mis)measurement of body image: Ten strategies to improve assessment for applied and research purposes. *Body Image, 1*(1), 7-14. doi:10.1016/S1740-1445(03)00004-4

Tiihonen, J., Keski-Rahkonen, A., Löppönen, M., Muhonen, M., Kajander, J., Allonen, T., . . . Rissanen, A. (2004). Brain serotonin 1A receptor binding in bulimia nervosa. *Biological Psychiatry, 55*(8), 871-873. doi:10.1016/j.biopsych.2003.12.016

Tolin, D. F. (2010). Is cognitive-behavioral therapy more effective than other therapies? A meta-analytic review. *Clinical Psychology Review, 30*(6), 710-720. doi:10.1016/j.cpr.2010.05.003

Tozzi, F., Thornton, L. M., Klump, K. L., Fichter, M. M., Halmi, K. A., Kaplan, A. S., Strober, M., Woodside, D. B., Crow, S., Mitchell, J., Rotondo, A., Mauri, M., Cassano, G.., Keel, P., Plotnicov, K. H., Pollice, C., Lilenfeld, L. R., Berrettini, W. H., Bulik, C. M., & Kaye, W. H. (2005). Symptom fluctuation in eating disorders: Correlates of diagnostic crossover. *American Journal of Psychiatry, 162*(4), 732-740. doi:10.1176/appi.ajp.162.4.732

Trout, J. D. (2016). *Wondrous truths: The improbably triumph of science*. New York, NY: Oxford University Press.

Trull, T. J., & Widiger, T. A. (2013). Dimensional models of personality: The five-factor model and the DSM-5. *Dialogues in Clinical Neuroscience, 15*(2), 135–146. doi:10.31887/DCNS.2013.15.2/ttrull

Tsou, J. Y. (2016). Natural kinds, psychiatric classification and the history of the DSM. *History of Psychiatry, 27*(4), 406–424. doi:10.1177/0957154X16656580

US Department of Health and Human Services. (2010). *How tobacco causes disease: The biology and behavioral basis for smoking-attributable disease: A report of the surgeon general.* Rockville, MD: Dept. of Health and Human Services, Public Health Service, Office of Surgeon General. https://www.ncbi.nlm.nih.gov/books/NBK53017/pdf/Bookshelf_NBK53017.pdf

Van den Eynde, F., Koskina, A., Syrad, H., Guillaume, S., Broadbent, H., Campbell, I. C., & Schmidt, U. (2012). State and trait food craving in people with bulimic eating disorders. *Eating Behaviors, 13*(4), 414-417. doi:10.1016/j.eatbeh.2012.07.007

Van den Eynde, F., & Schmidt, U. (2008). Treatment of bulimia nervosa and binge eating disorder. *Psychiatry*, *7*(4), 161-166. doi:10.1016/j.mppsy.2008.02.001

van Hoeken, D., & Hoek, H. W. (2020). Review of the burden of eating disorders: Mortality, disability, costs, quality of life, and family burden. *Current Opinion in Psychiatry*, *33*(6), 521-527. doi:10.1097/YCO.0000000000000641

Vanzhula, I. A., Calebs, B., Fewell, L., & Levinson, C. A. (2019). Illness pathways between eating disorder and post-traumatic stress disorder symptoms: Understanding comorbidity with network analysis. *European Eating Disorders Review*, *27*(2), 147-160. doi:10.1002/erv.2634

Verlinde, E. (2011). On the origin of gravity and the laws of Newton. *Journal of High Energy Physics, 29*, 1-26. doi:10.1007/JHEP04(2011)029

Vervaet, M., Audenaert, K., & van Heeringen, C. (2003). Cognitive and behavioural characteristics are associated with personality dimensions in patients with eating disorders. *European Eating Disorders Review, 11*, 363-378. doi:10.1002/erv.532

Voneida, T. J. (1998). Sperry's concept of mind as an emergent property of brain function and its implications for the future of humankind. *Neuropsychologia, 36*(10), 1077-1082. doi:10.1016/s0028-3932(98)00061-x

Wade, D. T., & Halligan, P. W. (2017). The biopsychosocial model of illness: A model whose time has come. *Clinical Rehabilitation, 31*(8), 995–1004. doi:10.1177/0269215517709890

Wai Sze Chan, T., Ahn, W-Y., Bates, J. E., Busemeyer, J. R., Guillaume, S., Redgrave, G. W., Danner, U. N., & Courtet, P. (2014). Differential impairments underlying decision making in anorexia nervosa and bulimia nervosa: A cognitive modeling analysis. *International Journal of Eating Disorders, 47*(2), 157-167. doi:10.1002/eat.22223

Wakefield, J. C. (1992). The concept of mental disorder: On the boundary between biological facts and social values. *American Psychologist, 47*(3), 373-388. doi:10.1037/0003-066X.47.3.373

Wang, G.-J., Geliebter, A., Volkow, N. D., Telang, F. W., Logan, J., Jayne, M. C., Galanti, K., Selig, P. A., Han, H., Zhu, W., Wong, C. T., & Fowler, J. S. (2011). Enhanced

striatal dopamine release during food stimulation in binge eating disorder. *Obesity, 19*(8), 1601-1608. doi:10.1038/oby.2011.27

Ward, T. (2013). Addressing the dual relationship problem in forensic and correctional practice. *Aggression and Violent Behavior, 18*(1), 92-100. doi:10.1016/j.avb.2012.10.006

Ward, T. (2019). Why theory matters in correctional psychology. *Aggression and Violent Behavior, 48*, 36-45. doi:10.1016/j.avb.2019.08.015

Ward, T., & Brown, M. (2006). The good lives model and conceptual issues in offender rehabilitation. *Psychology, Crime & Law, 10*(3), 243-257. doi:10.1080/10683160410001662744

Ward, T., & Clack, S. (2019). From symptoms of psychopathology to the explanation of clinical phenomena. *New Ideas in Psychology, 54*, 40-49. doi:10.1016/j.newideapsych.2019.01.004

Ward, T., Clack, S., & Haig, B. D. (2016). The abductive theory of method: Scientific inquiry and clinical practice. *Behaviour Change, 33*(4), 212-231. doi:10.1017/bec.2017.1

Ward, T., & Durrant, R. (2021). Practice frameworks in correctional psychology: Translating causal theories and normative assumptions into practice. *Aggression and Violent Behavior, 58*, 101612. doi:10.1016/j.avb.2021.101612

Ward, T., Hudson, S. M., & Bulik, C. M. (1993). The abstinence violation effect in bulimia nervosa. *Addictive Behaviors, 18*(671-680). doi:10.1016/0306-4603(93)90020-A

Ward, T., & Fischer, R. (2019). Families of network structures – we need both phenomenal and explanatory models. *Behavioral and Brain Sciences, 42*, e31. doi:10.1017/S0140525X1800122X

Ward, T., Polaschek, D., & Beech, A. R. (2006). *Theories of sexual offending*. West Sussex, England: John Wiley & Sons Ltd.

Ward, T., Wilshire, C. E. (2020). Science informed practice in clinical psychology [Manuscript in preparation]. School of Psychology, Victoria University of Wellington.

Waters, A., Hill, A., & Waller, G. (2001). Bulimics' responses to food cravings: Is binge-eating a product of hunger or emotional state? *Behaviour Research and Therapy, 39*(8), 877-886. doi:10.1016/S0005-7967(00)00059-0

Watson, H. J., & Bulik, C. M. (2013). Update on the treatment of anorexia nervosa: Review of clinical trials, practice guidelines and emerging interventions. *Psychological Medicine, 43*(12), 2477-2500. doi:10.1017/S0033291712002620

Waxman, S. E. (2009). A systematic review of impulsivity in eating disorders. *European Eating Disorders Review, 17*, 408-425. doi:10.1002/erv.952

Weinbach, N., Sher, H., & Bohon, C. (2018). Differences in emotion regulation difficulties across types of eating disorders during adolescence. *Journal of Abnormal Child Psychology, 46*, 1351-1358. doi:10.1007/s10802-017-0365-7

Weintraub, M. J., Schneck. C. D., & Miklowitz, D. J. (2020). Network analysis of mood symptoms in adolescents with or at high risk for bipolar disorder. *Bipolar Disorders, 22*(2), 128-138. doi:10.1111/bdi.12870

Westwood, H., Eisler, I., Mandy, W., Leppanen, J., Treasure, J., & Tchanturia, K. (2016). Using the autism-spectrum quotient to measure autistic traits in anorexia nervosa: A systematic review and meta-Analysis. *Journal of Autism and Developmental Disorders, 46*(3), 964-977. doi:10.1007/s10803-015-2641-0

Westwood, H., Kerr-Gaffney, J., Stahl, D., & Tchanturia, K. (2017). Alexithymia in eating disorders: Systematic review and meta-analyses of studies using the Toronto Alexithymia Scale. *Journal of Psychosomatic Research, 99*, 66-81. doi:10.1016/j.jpsychores.2017.06.007

Wewetzer, C., Deimel, W., Herpertz-Dahlmann, B., Mattejat, F., & Remschmidt, H. (1996). Follow-up investigation of family relations in patients with anorexia nervosa. *European Child & Adolescent Psychiatry, 5*, 18–24. doi:10.1007/BF00708210

Whooley, O. (2014). Nosological reflections: The failure of DSM-5, the emergence of RDoC, and the decontextualization of mental distress. *Society and Mental Health, 4*(2), 92-110. doi:10.1177/2156869313519114

Widiger, T. A., & Costa, P. T., Jr. (Eds.). (2013). *Personality disorders and the five-factor model of personality* (3rd ed.). American Psychological Association. doi:10.1037/13939-000

Wierenga, C. E., Ely, A., Bischoff-Grethe, A., Bailer, U. F., Simmons, A. N., & Kaye, W. H. (2014). Are extremes of consumption in eating disorders related to an altered balance between reward and inhibition? *Frontiers in Behavioral Neuroscience, 8,* 410. doi:10.3389/fnbeh.2014.00410

Williams, D. E., & McAdam, D. (2012). Assessment, behavioral treatment, and prevention of pica: Clinical guidelines and recommendations for practitioners. *Research in Developmental Disabilities, 33*(6), 2050-2057. doi:10.1016/j.ridd.2012.04.001

Wilsch, T. (2016). The deductive-nomological account of metaphysical explanation. *Australasian Journal of Philosophy, 94*(1), 1-23. doi:10.1080/00048402.2015.1033637

Wilshire, C. E., Ward, T., & Clack, S. (2021). Symptom descriptions in psychopathology: How well are they working for us? *Clinical Psychological Science, 9*(3), 323-339. doi:10.1177/2167702620969215

Wilson, G. T. (2018). Cognitive-behavioural therapy for eating disorders. In W. S. Agras & A. Robinson (Eds.), *The Oxford Handbook of Eating Disorders* (2nd ed., pp. 271-286). New York, NY: Oxford University Press.

Witt, A. A., & Lowe, M. R. (2014). Hedonic hunger and binge eating among women with eating disorders. *International Journal of Eating Disorders, 47*(3), 273-280. doi:10.1002/eat.22171

Wolkowitz, O. M., Wolf, J., Shelly, W., Rosser, R., Burke, H. M., Lerner, G. K., Reus, V. I., Nelson, J. C., Epel, E. S., & Mellon, S. H. (2011). Serum BDNF levels before treatment predict SSRI response in depression. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 35*(7), 1623-1630. doi:10.1016/j.pnpbp.2011.06.013

World Health Organization (WHO). (2018). *The International Classification of Diseases and Related Health Problems* (11th ed.). World Health Organization. https://icd.who.int/browse11/l-m/en

Wright, A., & Simms, L. (2015). A metastructural model of mental disorders and pathological personality traits. *Psychological Medicine, 45*(11), 2309-2319. doi:10.1017/S0033291715000252

Wu, M., Brockmeyer, T., Hartmann, M., Skunde, M., Herzog, W., & Friederich, H. (2016). Reward-related decision making in eating and weight disorders: a systematic review and meta-analysis of the evidence from neuropsychological studies. *Neuroscience and Biobehavioral Reviews, 61*, 177-196. doi:10.1016/j.neubiorev.2015.11.017

Zachar, P. (2008). Real kinds but no true taxonomy: An essay in psychiatric systematics. In K. S. Kendler & J. Parnas (Eds.), *Philosophical issues in psychiatry: Explanation, phenomenology, and nosology* (pp. 327–367). Baltimore, MD: Johns Hopkins University Press.

Zachar, P. (2014). *A metaphysics of psychopathology.* Cambridge, MA, US: MIT Press.

Zachar, P., & Kendler, K. S. (2017). The philosophy of nosology. *Annual Review of Clinical Psychology, 13*, 49-71. doi:10.1146/annurev-clinpsy-032816-045020

Zanchi, D., Depoorter, A., Egloff, L., Haller, S., Mählmann, L., Lang, U. E., Drewe, J., Beglinger, C., Schmidt, A., & Borgwardt, S. (2017). The impact of gut hormones on the neural circuitry of appetite and satiety: Systematic review. *Neuroscience and Biobehavioral Reviews, 80*, 457-475. doi:10.1016/j.neubiorev.2017.06.013

Zeeck, A., Stelzer, N., Linster, H. W., Joos, A., & Hartmann, A. (2011). Emotion and eating in binge eating disorder and obesity. *European Eating Disorders Review, 19*(5), 426-437. doi:10.1002/erv.1066

Zimmerman, J., & Fisher, M. (2017). Avoidant/restrictive food intake disorder (ARFID). *Current Problems in Pediatric and Adolescent Health Care, 47*(4), 95-103. doi:10.1016/j.cppeds.2017.02.005