Unusual Belief

From Eating Disorders to Imposter Syndrome

by

Stephen Gadsby

A thesis presented for the award of

Doctor of Philosophy

School of Philosophical, Historical and International Studies

Monash University

2021

Copyright Notice

Contents

Abstract

People who suffer from eating disorders often believe that they are overweight, despite having normal sized, thin, or, in some cases, dangerously thin bodies. People with imposter syndrome often believe that they lack certain abilities such as intelligence, competence, and talent, despite their overwhelming success and accomplishments. In both cases, these beliefs cause considerable suffering and hardship for those who hold them. In this thesis, I introduce and adopt a resolutely interdisciplinary approach towards the problem of categorising, explaining, and understanding these perplexing and unsettling beliefs.

The approach I adopt emphasises a multi-directional influence between disciplines, wherein philosophers engage with and facilitate the flow of information between disparate bodies of literature, contributing equally to both philosophical debates and scientific explanations. In addition, this approach dictates that philosophers should step outside their methodological comfort zones and employ techniques from the sciences to explore research questions, test hypotheses, and produce valuable data. In line with this interdisciplinary approach, I present a novel account of the false beliefs associated with eating disorders and imposter syndrome.

In the first part of the thesis, I develop, discuss, and defend an account of the false body size beliefs associated with eating disorders. According to this account, these beliefs are driven by two factors. The first factor involves misleading experiences of body size. The second factor involves motivated reasoning, wherein a strong desire to be thin biases these individuals towards negative beliefs about their own body size. Throughout the first five chapters, I develop this account, present evidence in favour of it, and highlight and evaluate its philosophical implications and commitments. This sequence of chapters includes an empirical study of body size perception, demonstrating this project's interdisciplinary reach.

In the final two chapters, I develop a novel account of the false beliefs about abilities that are associated with imposter syndrome. I argue that, similar to eating disorders, these beliefs are underpinned by motivated reasoning, wherein a strong desire to succeed in challenging environments biases these individuals towards negative beliefs about themselves. In accordance with my interdisciplinary approach, I introduce a new experimental paradigm for testing some of the assumptions and predictions of my framework and present novel empirical results aimed at doing just that.

In sum, I demonstrate that the beliefs associated with these conditions are not as perplexing or irrational as they appear and that a resolutely and exhaustively interdisciplinary approach is the most promising for understanding beliefs such as these and, ultimately, improving the lives of those who suffer from them.

Thesis including published works declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes 3 original papers published in peer reviewed journals and 1 submitted publications. The core theme of the thesis is understanding the beliefs associated with eating disorders and imposter syndrome. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the student, working within the School of Philosophical, Historical and International Studies under the supervision of Jakob Hohwy and Jennifer Windt.

The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.

In the case of the chapters that comprise this thesis, my contribution to the work involved the following:

Thesis Chapter	Publication Title	Status *(published, in press, accepted or returned for revision, submitted)*	Nature and % of student contribution	Co-author name(s) Nature and % of Co-author's contribution*	Co-author(s), Monash student Y/N*
5	Action, affordances, and anorexia: body representation and basic cognition	Published	65%	1) Daniel Williams: *Manuscript Writing and Editing* (35%)	No

I have renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Student name: Stephen Gadsby

Student signature: **Date: 30/01/2021**

I hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-authors' contributions to this work. In instances where I am not the responsible author I have consulted with the responsible author to agree on the respective contributions of the authors.

Main Supervisor name: Jakob Hohwy

Main Supervisor signature:  **Date: 30/01/2021**

General Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Signature: 

Print name: Stephen Gadsby

Date: 30/01/2021

Publications During Enrolment

(publications included in this thesis are bolded)

- **Gadsby, S. (2021). Imposter Syndrome and Self-Deception. *Australasian Journal of Philosophy*. 1-15.**

- **Gadsby, S. (2020). Self-deception and the second factor: How desire causes delusion in anorexia nervosa. *Erkenntnis*, 85(3), 609-626.**

- Gadsby, S., & Hohwy, J. (2020). Why use predictive processing to explain psychopathology? The case of anorexia nervosa. in *The Philosophy and Science of Predictive Processing* (Eds.) S. Gouveia, R. Mendonça, & M. Curado. Bloomsbury.

- Gadsby, S. (2019). Body representations and cognitive ontology: Drawing the boundaries of the body image. *Consciousness and Cognition*, 74, 102772.

- Gadsby, S. (2019). Manipulating body representations with virtual reality: clinical implications for anorexia nervosa. *Philosophical Psychology*, 32:6, 898-922.

- Tappin, B., & Gadsby, S. (2019). Biased belief in the bayesian brain: a deeper look at the evidence. *Consciousness and Cognition*, 65, 107-114.

- Gadsby, S. (2018). How are the spatial characteristics of the body represented? A reply to Pitron & de Vignemont. *Consciousness and Cognition*, 62, 183-168.

- **Gadsby, S., & Williams, D. (2018). Action, affordances, and anorexia: body representation and basic cognition. *Synthese*, 195(12), 5297-5317.**

- Gadsby, S. (2017). Explaining body size beliefs in anorexia. *Cognitive Neuropsychiatry*, 22(6), 495-507.

Submitted

- **Engel, M., Gadsby, S., Corcoran, A., Keizer, A., Dijkerman, C., Hohwy, J. Waiting longer, feeling fatter: Effects of response delay on tactile distance estimation and confidence in anorexia nervosa.**

Acknowledgments

This research was supported by an Australian Government Research Training Program (RTP) Scholarship.

In addition to those listed in the acknowledgments at the end of published chapters, I owe my gratitude to a number of different people:

Jakob Hohwy for his invaluable support, encouragement, and advice over the last few years and for working hard to facilitate and promote such an open, engaging, and collegial intellectual environment in the lab.

Jennifer Windt for her prompt, detailed, and insightful comments on numerous drafts, especially during this last year (in the midst of a pregnancy, no less).

Thomas Corbin, for useful feedback on the introduction, conclusion, and linking text.

All of the members of the Cognition and Philosophy Lab—especially Julian, Kelsey, Andrew, Iwan, Andy, Manu, Jonno, Niccolo, Simon—for feedback on various manuscripts and (sometimes half-baked) lab group presentations; help with programming and statistics; many engaging discussions over coffee, beer, and wonton noodle soup; and a willingness to endure the background noise I introduced to the lab.

Anouk Keizer, Chris Dijkerman, and all the members of the Dijkerman Lab for welcoming me into their lab and fostering my (initial and continued) interest in anorexia and body representation research.

Dan Williams, for devoting considerable time and effort to providing feedback on many parts of this thesis, for constant encouragement and comradery, and for providing a consistent source of inspiration through his approach to research.

Manja, as a collaborator, for her generosity and patience; for teaching me the intricacies of experimental design, data collection, and statistics; and for providing invaluable insights into anorexia. As a partner, for her unwavering love, support, and encouragement.

Annette and Andrew, for too many things to mention.

Thesis Introduction

1. Troubling Conditions and Unusual Beliefs

Beliefs are important. They are, to borrow a popular metaphor, "the maps by which we steer". When our beliefs are false, and these maps misrepresent the world, they can steer us in some troubling directions. Consider some examples of false belief that have been well-discussed in the philosophical literature: delusions. People with Fregoli delusion believe that different individuals are in fact the same person, in disguise. People with Capgras delusion believe that a loved one has been replaced by an imposter. These beliefs lead to unusual, and notably detrimental, behaviours. People with Fregoli delusion suffer from significant fear and anxiety about the disguised people that they run into, in many cases leading to confrontations (Ellis & Szulecka, 1996). Tragically, people with Capgras sometimes act violently towards the one's they love (Bourget & Whitehurst, 2004).

This thesis is about two categories of unusual beliefs, associated with two distinct conditions. The first category is the false beliefs about body size associated with eating disorders. People with eating disorders often believe that they are not thin, despite meeting their own standards for thinness. The second category is the false beliefs about abilities associated with imposter syndrome. People with imposter syndrome believe that they lack certain abilities (intelligence, competence, talent), despite their demonstrable talent and success. While certainly unusual—most people do not radically misjudge their body size or intellectual ability—these beliefs are not (usually) classified as delusions. Nevertheless, they are, at least in some cases, just as harmful as Capgras and Fregoli delusions. While people with eating disorders do not intentionally harm others, they cause considerable harm to themselves. And while people with imposter syndrome do not reach the levels of anxiety associated with severe cases of Fregoli, many suffer from considerable distress over the predicament that they believe themselves to be in. In what follows, I provide a more in-depth introduction to both conditions and the unusual beliefs associated with them. I then draw out some similarities between these beliefs, in terms of their philosophical and scientific importance, and I advocate for a resolutely interdisciplinary approach towards understanding them.

2. Eating Disorders, Imposter Syndrome, and Unusual Beliefs

2.1. Eating Disorders

Eating disorders comprise a vast and heterogenous class of disorders. In this thesis, I focus on three categories of eating disorder. The most well-known, and well-researched is *anorexia nervosa*, whose diagnostic criteria, according to the latest diagnostic and statistical manual of mental disorders, DSM-5, are:

> 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 behavior 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. (American Psychiatric Association, 2013, pp. 338-339)

Anorexia nervosa is currently recognised as involving two subtypes: *restrictive* and *binge-purge*. In restrictive anorexia nervosa, the most common subtype, individuals intentionally lose weight using a number of techniques related to managing the amount or types of food eaten, the frequency of meals, or the amount of exercise conducted. Most commonly, these methods are combined. The binge-purge subtype also involves these behaviours, with the addition of recurrent episodes of binge eating, commonly followed by purging, in the form of self-induced vomiting, or the use of laxatives, diuretics, and enemas.

Startlingly, anorexia nervosa has the highest mortality rate of any mental disorder, at around 5.1% (Arcelus et al., 2011). Some studies suggest that less than half of those diagnosed with anorexia nervosa will fully recover (Steinhausen, 2002). This is very likely exacerbated by the lack of success in formulating effective treatments. Indeed, clinicians know it as "one of the most frustrating and recalcitrant forms of psychopathology" (Vitousek et al., 1998, p. 391).

The other major category of eating disorder is *bulimia nervosa*, whose diagnostic criteria are:

> 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 over eating 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 behaviors in order to prevent weight gain, such as self-induced vomiting; misuse of laxatives, diuretics, or other medications; fasting; or excessive exercise.
>
> C. The binge eating and inappropriate compensatory behaviors both occur, on average, at least once a week for 3 months.
>
> D. Self-evaluation is unduly influenced by body shape and weight.
>
> E. The disturbance does not occur exclusively during episodes of anorexia nervosa. (*ibid.* p. 345)

Diagnoses of anorexia nervosa and bulimia nervosa are historically intertwined: bulimia nervosa was originally identified as a "chronic phase of anorexia nervosa" (Russel, 1979). Given that the cycle of binging and purging associated with bulimia nervosa is often a response to a (perceived) failure at restricting, some with a diagnosis of bulimia nervosa think of themselves as "failed anorexics" (Fairburn & Harrison, 2003, p. 407). Many clinicians and researchers do not consider these to be separate disorders, as one prominent researcher states, "[anorexia nervosa] and [bulimia nervosa] are far from being entirely discreet [sic] disorders and can be made to seem so only by dint of a certain sophistry." (Palmer, 2003, p. 2). Consequently, similar theoretical models and treatment methods are often applied to both disorders (see below).

Finally, as with many psychiatric categories, the DSM recognises a category of eating disorders for those whose symptoms fall short of the criteria for either anorexia nervosa or bulimia nervosa. In previous editions of the DSM, this category was known as "eating disorder not otherwise specified" (EDNOS), in the latest DSM it is referred to as "other specified feeding or eating disorder" (OSFED), with two relevant sub-types:

> 1. Atypical anorexia nervosa: All of the criteria for anorexia nervosa are met, except that despite significant weight loss, the individual's weight is within or above the normal range.
>
> 2. Bulimia nervosa (of low frequency and/or limited duration): All of the criteria for bulimia nervosa are met, except that the binge eating and inappropriate compensatory behaviors occur, on average, less than once a week and/or for less than 3 months. (DSM V, p. 353)

Following the literature, I will refer to the two subtypes of EDNOS/OSFED collectively as "atypical eating disorders". Atypical eating disorder diagnoses are as common, if not more common, than anorexia nervosa or bulimia nervosa diagnoses (Fairburn & Cooper, 2007; Fairburn & Harrison, 2003). Many individuals also migrate between these categories, in fact, according to Fairburn and Cooper (2007, p. 107), "migration … is the norm rather than the exception". This migration is predominately one-directional, as

those initially diagnosed with anorexia nervosa or bulimia nervosa are later diagnosed as atypical (Fairburn & Harrison, 2003).

Given the considerable comorbidity between these three diagnostic categories, many researchers argue that they are driven by similar mechanisms (ibid.). Consequently, they adopt a *transdiagnostic* approach, which emphasises identifying and treating these shared mechanisms. In this thesis, I will follow in this transdiagnostic approach, focusing on one particular mechanism, shared by individuals who fall under each of these categories: false beliefs about body size. As I will show, these beliefs are a driving force behind the harmful weight loss behaviours found within each of these categories. While some of the chapters (e.g. chapters 1 and 4), address anorexia nervosa more specifically, the arguments presented in this thesis, overall, are intended to apply to this broader category of individuals.

2.2. Beliefs about Body Size

A large subset of those diagnosed into the foregoing categories (anorexia nervosa, bulimia nervosa, and atypical eating disorders), attempt to lose weight out of a desire to be thin and a belief that they are not. Up until recently, there has been surprisingly little research into the beliefs that people with eating disorders hold about their body size (Phillipou et al., 2017). Instead, the literature has more closely focused on attitudes of dissatisfaction, i.e. how satisfied one is with their current body size and shape (Cash & Deagle, 1997).

A long-held assumption is that body dissatisfaction in eating disorders is driven not by unusual beliefs, but by extreme desires, wherein people who suffer from them desire to be extremely or unnaturally thin. As Wolf (2013) writes, people with eating disorders are "simply trying to maintain an unnatural 'ideal' body shape and weight" (p. 5) or, as Stice and colleagues (1998, p. 131) put it "ideal-body internalization leads to body dissatisfaction because it represents an unrealistic goal". According to these accounts, the relevant body dissatisfaction (and resulting weight loss behaviours) stem from a desire to achieve a body size which is unnatural, unrealistic, or unattainable. Consequently, eating disorders involve unusual desires, rather than unusual beliefs.

Contrary to the foregoing assumption, multiple strands of empirical evidence suggest that eating disorders do not necessarily involve extreme desires or evaluative standards. People with eating disorders do not always exhibit extreme standards for what qualifies as a thin body (Alleva et al., 2013), nor do they necessarily aspire to excessive thinness (Moscone et al., 2017). Additionally, many strands of research suggest that people with eating disorders are unaware of their true body size. For example, in tasks where they are shown line-ups of different depictions of body size and asked to pick the one that most closely matches their current body size, they select depictions that are much larger than themselves (Mölbert et al., 2017). Their willingness to endorse these misjudgements of body size suggests that they hold false beliefs about how large they truly are. A central assumption of this thesis is that many of the problems associated with eating disorders—i.e. dissatisfaction with the body and subsequent attempts to lose weight—stem not from extreme evaluative standards, but false beliefs about body size.

Caution must be taken in assuming how applicable this approach will be to the broader class of individuals diagnosed with eating disorders. These disorders are highly complex, driven by a variety of distinct biological, psychological, developmental and socio-cultural factors (Rikani et al., 2013). Consequently, not all people with eating disorders desire to be thin and believe that they are not. For example, some of those diagnosed with bulimia nervosa engage in their harmful eating patterns as a form of emotion regulation rather than weight loss (Stice et al., 1998). Similarly, some of those diagnosed with anorexia nervosa do not believe that they are overweight, and do not lose weight out of a desire to be thin (Lee & Kwok, 2005). Thus, there is significant heterogeneity in regard to the cause of harmful weight loss behaviours. Nevertheless, the research discussed does suggest that many individuals diagnosed into one of these three categories hold false beliefs about their body size. This specific group of individuals is the focus of this thesis.

2.3. Imposter Syndrome

> It's almost like the better I do, the more my feeling of inadequacy actually increases, because I'm just going, 'Any moment, someone's going to find out I'm a total fraud, and that I don't deserve any of what I've achieved'

The above quote, from the actress Emma Watson, is a characteristic description of imposter syndrome, a phenomenon suffered by many successful, intelligent, and driven individuals. While the first five chapters of the thesis will address eating disorders and the false beliefs associated with them, the last two chapters focus on this distinct condition.

Imposter syndrome was originally identified in the 70s, by the clinical psychologists Clance and Imes (1978). They observed that a significant number of their female clients believed that their success was unwarranted and feared being uncovered as "imposters". While imposter syndrome was originally conceived as uniquely afflicting women, later research discovered that it is prevalent amongst both males and females (Badawy et al., 2018; Bravata et al., 2019).

Unlike eating disorders, imposter syndrome is not currently classified as a mental disorder. While some researchers have called for its inclusion in the DSM (Bravata et al., 2019),[1] most approach it as a form of (maladaptive) personality trait, one that is present (to a greater or lesser degree) in a large proportion of the population (Feenstra et al., 2020, p. 2). This latter assumption is likely due to the influence of Clance and Imes (1978) who were careful not to pathologise the condition, giving it the more neutral label "imposter phenomenon". Today, most psychologists use the term "imposter phenomenon", while (the few) philosophical treatments tend towards the term that is more popular in everyday use: "imposter syndrome" (Hawley, 2019; Paul, 2019).[2]

[1] Bravata and colleagues (2019, p. 21) argue that doing so will "codify" the most appropriate treatment approaches, to the benefit of health care providers.
[2] As chapter 6 addresses a (predominately) philosophical audience, I adopt the term "imposter syndrome". Chapter 7 addresses a (predominately) psychological audience, so I adopt the term "imposter phenomenon". For now, I will continue with the term "imposter syndrome".

Imposter syndrome is typically characterised as comprising three features (Harvey & Katz, 1985). The first, and most commonly discussed, is that people who suffer from imposter syndrome fear being discovered and exposed as imposters, who do not belong in their role and are undeserving of their success. The second feature is the belief that they are inadequate (see below). The final feature is a host of biases in the way in which people with imposter syndrome treat information: they dismiss and ignore forms of information that would validate their competence and talent (Leary et al., 2000, p. 727).

2.4. Beliefs about Ability

People with imposter syndrome commonly believe that they are inadequate because they lack certain qualities, such as intelligence, talent, and competence. As in the case of eating disorders, a core assumption of the thesis is that many of the issues surrounding the imposter syndrome are related to these individuals' (false) beliefs about their own inadequacy. Accordingly, the later parts of this thesis (chapters 6 and 7) will focus on understanding, explaining, and categorising these *inadequacy beliefs*.

Inadequacy beliefs can be present-directed or future-directed—people with imposter syndrome can believe that they lack the requisite abilities for a role that they currently fill (such as PhD candidate), or a role that they desire to fill (such as professor). As noted, the kinds of abilities that people with imposter syndrome believe they lack also differ between case and context. Commonly (especially in the academic context) people with imposter syndrome believe that they lack the requisite intelligence. While there are many different properties that people with imposter syndrome can believe that they lack, they can be generally categorised as beliefs about lacking the ability required to succeed within a particular role (Clark et al., 2014; Hawley, 2019; Leonhardt et al., 2017).

In the case of eating disorders, weight loss behaviours cause harm and draw clinical attention. In the case of imposter syndrome, the feature that causes harm and draws clinical attention is the fear that they will be discovered and exposed, causing considerable anxiety and distress (Leonhardt et al., 2017). Yet, as in the case of eating disorders, this harm can be traced back to the relevant false beliefs. If people with

imposter syndrome were to believe that they were as intelligent, competent, and able as their peers then they would no longer feel as if they were imposters. Consequently, understanding how these individuals come to hold such beliefs is crucial to understanding and addressing the condition.

3. Significance and Methodology

There is good reason to believe that the foregoing beliefs—regarding body size and personal ability—cause considerable suffering for those who hold them. In the case of eating disorders, body dissatisfaction and harmful weight loss behaviours are driven by false beliefs about body size; in the case of imposter syndrome, fear and anxiety (associated with being discovered and exposed) are driven by beliefs about one's own inadequacy. There is thus a need to understand these beliefs, in order to help those who suffer from them.

Beyond this practical benefit, however, there are theoretical advantages to understanding these beliefs, as they represent philosophically and scientifically unique and illuminating case studies. Their philosophical importance stems from the way in which they appear to violate norms of rationality. The first thing to note about both forms of belief is their *epistemic* irrationality. People with imposter syndrome possess overwhelming amounts of evidence in favour of their talents and abilities and must go to great lengths to uncover even a small amount of evidence to the contrary. Further still, people with eating disorders seemingly possess no evidence whatsoever in favour of the belief that they are overweight. Both forms of belief appear then to (severely) violate norms of epistemic rationality—which insist that our beliefs must be (at least somewhat) tethered to the evidence that we encounter.

Perhaps more interesting is the way in which these beliefs appear to violate norms of *pragmatic* rationality, which hold that we ought to act, and reason, in ways that promote our own well-being. In each case, the relevant beliefs and the behaviours that they cause undermine these individuals' own health and happiness. Consequently, these beliefs

represent philosophically important case studies for illuminating the relationship between human psychology and rationality.

Beyond that, both forms of belief are scientifically unique. Pathologically unusual beliefs, like delusions, are commonly linked to some identifiable cause—such as neurological disorder, damage to the nervous system, or domain general reasoning deficit (Gilleen & David, 2005; McLean et al., 2017). In such cases, there are plausible candidates for the cause of the relevant beliefs (though the details often require filling out). In the case of eating disorders and imposter syndrome, this is not the case. Neither condition is reliably accompanied by neurological disorder, nervous system damage, or domain general reasoning deficit.[3] There is thus the question of how to explain these beliefs, and whether the same models that are applied to beliefs like delusions can provide any explanatory traction.

Finally, these beliefs represent a useful example of what can be achieved when philosophers adopt a resolutely interdisciplinary approach. Philosophers have made impressive progress in contributing to our understanding of many unusual beliefs, however, much of this research is limited in both scope and methodology. For example, a significant portion of these efforts have been devoted to just a sub-set of unusual beliefs, most notably, delusions. There are, however, many forms of unusual belief in this world; by focusing on two forms of belief that have mostly been overlooked by philosophers, I will illustrate the benefits of broadening the scope of phenomena that are investigated.

Additionally, I argue for a broadening of our methodological approach. In the past, philosophers were severely restricted in the way in which they engaged with the science of unusual belief, taking from, but not giving back to the relevant scientific literature (Bortolotti, 2010, p. 4). Fortunately, many contemporary philosophers are genuinely concerned with contributing to our scientific understanding of unusual beliefs and bettering the lives of those who suffer from them. These interdisciplinary philosophers

[3] There is mixed evidence of an association between eating disorders and set-shifting deficits, however, these have only been found in participants with anorexia nervosa, not bulimia nervosa (Kanakam & Treasure, 2013), despite the relevant beliefs being associated with both conditions.

provide crucial insights to shape and reform scientific explanations and, in many cases, develop explanatory frameworks of their own. This shift towards two-way interaction between philosophy and science is promising and has delivered many crucial insights. This thesis will follow this shift and extend it, presenting new methodological principles for interdisciplinary research of this kind and illustrating the value of those principles.

First, I advocate for an extensively interdisciplinary research, wherein a wide variety of disciplines are drawn from. This involves philosophers acting as facilitators of the exchange of knowledge, seeking out and collating knowledge from a broad variety of disciplines and drawing undiscovered links between seemingly disparate fields. This allows us to gain new insights, new understanding, and new solutions to the problems that cause suffering for so many individuals. To illustrate this methodological principle, this thesis will draw from a variety of influences, from various disciplines: theories of belief formation from cognitive neuropsychiatry and motivated reasoning from social psychology (chapter 1), accounts of rationality from philosophy (chapter 2), of body representation from cognitive neuroscience (chapter 5), and of motivated reasoning from behavioural economics (chapters 6 and 7). I draw insight from each of these sources, illustrating how they can integrate with and enrich one another.

Another feature of the interdisciplinary methodology I advocate for relates to the methods used to generate knowledge. While philosophers have begun to develop theoretical frameworks of their own to explain mental disorders and their unusual beliefs, they have generally restricted themselves in regard to testing those frameworks. Consequently, they must wait, and hope, for scientists to learn about, adopt, and test the theories that they produce. However, just as interdisciplinary research should be unlimited in the number or kinds of disciplines from which it draws, so too should it be unlimited in the research techniques that it employs. Interdisciplinary researchers who hope to make significant and lasting contributions to our understanding of unusual beliefs should be willing to translate their ideas into experimental paradigms, either by suggesting design principles for experiments, or by building and running experiments, either collaboratively or independently.

This thesis illustrates the benefits of this unbridled interdisciplinary methodology. In the chapters that follow, the thesis employs a number of different methodologies, not only the conceptual analysis and theory building that are commonly employed in contemporary philosophy of mind, but empirical research as well. chapter 3 suggests principles for the design of experimental paradigms, while chapters 4 and 7 adapt prior paradigms (from cognitive neuroscience and behavioural economics, respectively) and employ them to test empirical hypotheses that are crucial to the theoretical proposals developed in other chapters.

4. The Roadmap

Chapter 1 introduces an empiricist account of the false body size beliefs associated with eating disorders. According to this account, these unusual beliefs are caused by unusual experiences, conveying false information about body size. I discuss evidence pointing to three such forms of experience, which people with eating disorders may suffer from, forming the basis of the account. While these experiences provide the content of the beliefs, I suggest that motivational biases (constituting a "second factor") contribute to the maintenance of these beliefs in the face of disconfirmatory evidence.

Chapter 2 describes a novel form of illusory body size experience that people with eating disorders likely suffer from: proprioceptive misperception of their own bodily boundaries. This chapter both adds to the empiricist model outlined in chapter 1 and addresses an issue of philosophical significance related to the addition, regarding the epistemic rationality of eating disorders. I draw out the ramifications of this in addition, in terms of the rationality of eating disorders, and address the philosophical and practical implications of my argument.

Chapter 3 addresses the hypothesis that eating disorders involve visual misperception of body size. If this claim is true, even if only in some cases, this would represent an important addition to the proposed empiricist account. However, there is no previous empirical evidence to convincingly support the hypothesis. I discuss the history of the hypothesis and the attempts to confirm it, through the use of tasks requiring participants

to visually estimate their own body size. I outline the methodological problems that this research program faces and provide some suggestions for future research to overcome those problems.

Chapter 4 presents the results of an experiment assessing whether anorexia nervosa is associated with misperception of touch. Research suggests that, in addition to the forms of misperception outlined in the previous chapters, anorexia nervosa involves misperception of tactile size. In the linking text, I outline the implications of this finding for the proposed empiricist account. As with the hypothesis regarding visual misperception, if true, this would represent a clinically important form of unusual experience: tactile misperception of body size. The results presented suggest that differences in tactile estimation in participants with anorexia nervosa do not stem from misperception of touch, as has been assumed. Consequently, it seems unlikely that people with anorexia nervosa misperceive their body size through touch.

Chapter 5 addresses the philosophical foundations of a key feature of the proposed empiricist model, namely, that eating disorders involve distorted body representations. Many philosophers argue that explanations of psychological phenomena—especially phenomena like movement and affordance processing—should avoid positing mental representations. This chapter addresses that argument, pointing out how the kinds of body representations that feature in this model qualify as legitimately representational. This introduces a general approach for making progress in these debates: by looking to scientific models themselves, the functional nature of the representational states that they posit, and the predictive value derived from positing those states.

Chapter 6 changes focus to the phenomenon of imposter syndrome. People who suffer from imposter syndrome believe that they are incompetent and maintain such beliefs by discounting evidence in favour of their abilities. I suggest that this biased evidence treatment stems from the motivational benefits of downplaying one's own abilities, namely, motivating oneself to exert effort in order to succeed at a challenging task. This presents a new account of imposter syndrome as a form of self-deception.

Chapter 7 introduces an experimental paradigm for understanding the relationship between imposter syndrome and motivated reasoning. In this experiment, post graduate students were required to solve a number of reasoning problems and estimate their own performance. We discovered that those participants who scored highly in a measure of imposter syndrome were more negative in their self-evaluations. We also discovered that, when controlling for depression and low self-esteem, imposter syndrome was only predictive of one particular form of self-evaluation bias: evaluation of one's performance compared to others. This supports a characterisation of imposter syndrome as fundamentally related to one's abilities compared with one's peers.

As a typographical note, for published chapters, Monash University regulations dictate that the printed version (as typeset by the journal) must be inserted. Consequently, chapters 2, 5 and 6 are formatted differently to the rest of the thesis.

Linking Text Between Introduction and Chapter 1

In the following chapter, I introduce a new explanatory account of the false body size beliefs associated with eating disorders. While there are many ways to characterise such beliefs (e.g. "I am fat", "I am not thin", "I need to lose weight"), in what follows I focus on beliefs about not (yet) meeting one's ideal body size. The basic framework is inspired by what are commonly termed *empiricist* accounts, from the literature on delusions. While there is significant diversity amongst the different empiricist accounts, they share a basic assumption: that the unusual beliefs associated with mental disorders are caused by unusual experiences.

To illustrate the approach, consider the case of Capgras delusion—the belief that a loved one has been replaced by an imposter. The standard empiricist account of this delusion refers to a dual-route mechanism for the visual processing of faces: one route underpins the recognition of a familiar face and the other underpins the affective response (Ellis & Young, 1990). Capgras is claimed to involve a breakdown in this system, such that the recognition route is intact while the affective route is dysfunctional. This gives rise to a particular kind of unusual experience: "Conscious recognition of the identity of a face but absence of the affective response that characteristically accompanies the perception of a familiar face" (Coltheart, 2007, p. 1047). This experience of recognising a loved one's face, absent the usual affective response, is thus proposed to cause the belief that a loved one has been replaced by an imposter. The basic empiricist approach then is to identify unusual experiences that might cause unusual beliefs and characterise the forms of dysfunction that underpin those experiences.

In the next chapter, I introduce an empiricist account which does just that. Unusual for most empiricist account, this one identifies three forms of unusual experience: false self-other comparisons, affordance misperception, and spontaneous mental imagery—each of which play a role in grounding the relevant beliefs. I further develop this into a two-factor account, according to which additional forms of bias—in conjunction with the unusual experiences—help to explain the relevant beliefs.

Erkenntnis (2020) 85:609–626
https://doi.org/10.1007/s10670-018-0039-z

ORIGINAL RESEARCH

Self-Deception and the Second Factor: How Desire Causes Delusion in Anorexia Nervosa

Stephen Gadsby[1]

Received: 13 March 2018 / Accepted: 10 July 2018 / Published online: 24 July 2018

Abstract

Empiricist models explain delusional beliefs by identifying the abnormal experiences which ground them. Recently, this strategy has been adopted to explain the false body size beliefs of anorexia nervosa patients. As such, a number of abnormal experiences of body size which patients suffer from have been identified. These *oversized experiences* convey false information regarding the patients' own bodies, indicating that they are larger than reality. However, in addition to these oversized experiences, patients are also exposed to significant evidence suggesting their bodies are in fact thin. This situation poses a conundrum: why do patients appear strongly influenced by the former kinds of evidence while the latter has little effect? To solve this conundrum, I suggest a two-factor account. First, I discuss research on the biases patients exhibit in how they gather, attend to and interpret evidence related to their own body size. Such biases in evidence treatment, I suggest, cause oversized experiences to be sought out, attended to and accepted, while veridical body size experiences are ignored or explained away. These biases constitute the second factor for this empiricist model, accounting for the unwarranted conviction with which these beliefs are held. Finally, in line with recent research into self-deception, I propose that, paradoxically, these biases in evidence treatment arise from patients' own desires.

1 Introduction

Many clinical patients assert the truth of patently implausible statements: some claim that a loved one has been replaced by an imposter (Capgras delusion), others, that a part of their body belongs to someone else (Somatoparaphrenia) and some even insist that (despite appearances) they are, in fact, dead (Cotard delusion).[1]

[1] Herein I adopt the mainstream assumption that such delusions are indeed beliefs (Bortolotti 2010; Clutton 2018; cf. Currie 2000; Gerrans 2014).

✉ Stephen Gadsby
Gadsby.st@gmail.com

[1] Monash, Melbourne, VIC, Australia

These delusions can arise in the context of a diverse number of conditions, such as brain trauma, cerebral accident, degenerative brain conditions and substance use disorders. Overlooking this bewildering diversity, there is an incredibly fertile field of cognitive neuropsychiatry devoted to understanding the nature and cause of the beliefs themselves (Coltheart 2007).

In contrast with these patently bizarre delusions, the false beliefs I will be discussing here pertain to the somewhat mundane topic of *body size* and are generally found in a small subset of related eating disorders. Eating disorders (ED) are characterised by persistent disturbed eating related behaviour of some form—an underspecified characterisation which aggregates a number of disorders with distinct and diverse symptomatology, and likely aetiology.

Perhaps due to the emphasis on eating behaviour, the false body size beliefs of ED patients have avoided the level of scrutiny applied to the aforementioned delusions. Indeed, psychiatrists have traditionally refused adopting the label of delusion to describe these beliefs, instead classifying them as *overvalued ideas*—a form of belief which carries an "air of limited clinical significance", and is regularly associated with abnormal personality, rather than pathology per se (McKenna 1984, 579; see also: Veale 2002). In spite of the orthodox line, a move is now occuring towards classifying at least a sub-group of these beliefs as delusions; with this, comes increased scientific interest into their nature and cause (Konstantakopoulos et al. 2012; McKenna et al. 2014; Mountjoy et al. 2014 Phillipou et al. 2017).

While some of the evidence I will discuss pertains to ED patients more generally, the lion's share pertains to anorexia nervosa (AN) patients specifically. For the time being then, I restrict my claims to these specific patients. Nevertheless, I remain open to the possibility that the following model might be adapted to account for the abnormal body size beliefs of others, such as bulimia nervosa (BN) patients (see Sect. 4). Despite undeniable differences in symptomatology, the various eating disorders may share important commonalities in the aetiology of the patients' harmful body size beliefs.

In order to pin down the relevant target of scientific interest, a first task is to describe the content of such beliefs. At first glance, this appears problematic, as *specific* belief content undoubtedly varies from patient to patient. Despite common misconception, not *all* AN patients believe they are "fat" so much as that they aren't "thin", or even that they aren't "excessively/dangerously/too thin" (Konstantakopoulos et al. 2012; O'Connell et al. 2017). The inherent issue is that *specific* belief content is largely determined by idiosyncratic standards regarding concepts such as "thinness" and "fatness".

That said, this diversity in belief content need not concern us. In order to pin down a specific framing of belief content shared by AN patients, I focus here on beliefs regarding not yet meeting one's *ideal* body size. This highlights the fundamental irrationality of patients' beliefs, as the central problem is not that they have unrealistic desires or extreme evaluative standards, but rather that they do not realise they have already met and surpassed their own ideal standards for thinness (Moscone et al. 2017). Further, while specific belief content—e.g. "I am not thin",

"I am fat", "I look awful" etc.—might vary, most patients can be said to believe they aren't yet their ideal size (Gardner and Brown 2010).[2]

2 The Empiricist Approach to Anorexia Nervosa

One approach which has seen considerable success in explaining delusions is termed *empiricism* (Bayne and Pacherie 2004). Emerging from the work of Maher (1974), empiricist models aim to explain delusional beliefs by identifying abnormal experiences with related content which might ground them. To take an example from the aforementioned delusions, in the standard empiricist model of Capgras it's proposed that a breakdown in the system responsible for autonomic responses to familiar faces causes these patients to lack the usual sense of familiarity when seeing a loved one (Ellis and Young 1990). This abnormal experience (recognising a loved one but lacking the expected affective response) is assumed to ground the belief that they are an imposter.

Adapting this explanatory strategy to the case of AN has also proven fruitful as, on inspection, there are a number of abnormal experiences of body size which patients suffer from that may ground their abnormal body size beliefs—what have been termed *oversized experiences* (Gadsby 2017a; 2017b). It should be noted that these oversized experiences are largely visual in nature. While there is a significant body of evidence tying eating disorders to dysfunctional interoceptive processing (Kaye et al. 2009; Pollatos et al. 2008), it seems unlikely that one could experience their bodies as larger than reality through interoception alone—as interoceptive signals do not deliver content related to body size or shape.[3] In what follows, I briefly review this empiricist model and the oversized experiences it posits.

2.1 Oversized Experiences

One form of oversized experience AN patients likely suffer from is *spontaneous recurrent mental imagery*. Most commonly seen in anxiety disorders, this involves constant, intrusive mental imagery representing an individual's own fears (Hackmann et al. 1998). There is some evidence to suggest that ED patients suffer from recurrent mental imagery, whereby the images are of their own bodies as overweight (Cooper et al. 1998; Somerville and Cooper 2007; Somerville et al. 2007).[4] It's possible that such recurrent imagery could ground false beliefs about body size, as long as patients took this imagery to be veridical. While those who

[2] Evidence of this comes from figural drawing scale experiments: participants are asked to select body size silhouettes that match both their current and ideal body sizes, with the difference in size taken to be an indication of "body dissatisfaction" (Moussally et al. 2017).

[3] This isn't to say that the oversized experiences discussed are in no way related to the deficits in interoceptive processing that have been uncovered. A link between these two forms of dysfunction may emerge, though it's not yet clear what this link might be.

[4] Unlike with the other forms of oversized experiences, most of the research into spontaneous mental imagery in ED has been conducted on BN, rather than AN, patients (cf. Cooper et al. 1998). This lends further support to the possibility that the model discussed here might apply to a number of eating disorders (see Sect. 4).

suffer from anxiety disorders appear to interpret their imagery as veridical (Hackmann et al. 2000, 602), it's not yet clear whether this also holds for ED patients.

Another form of oversized experience stems from *perceptual body image* distortion. Following Schilder (1935, 11), the perceptual body image can be defined as "the picture of our own body which we form in our mind". In contrast to spontaneous episodes of mental imagery, the content of the perceptual body image derives from a *stored* representation we have of the shape, size and form of our own bodies (de Vignemont 2010; Gadsby 2017c). In the case of AN, this mental image is said to be distorted, as patients' bodies are represented as *larger* than reality. Evidence of this comes from tasks requiring patients to estimate the size of their bodies, whereby overestimation is considered reflective of an oversized perceptual body image (Smeets 1997).

How could distortion of the perceptual body image cause oversized experiences? One possibility is through a mode of experience similar to the previously discussed episodes of spontaneous mental imagery. AN patients may *consciously* mentally picture this distorted body representation and, similar to spontaneous mental imagery, take it to be veridical. Although this story seems plausible, there is, as of yet, no evidence of such behaviour.

Nevertheless, there is a relevant kind of activity ED patients regularly engage in: mentally comparing themselves with their "thin" peers (Alleva et al. 2013, 99; Corning et al. 2006; Espeset et al. 2012, 524; Hamel et al. 2012). This process of size comparison relies on the perceptual body image, as judging whether a visually presented body is larger or smaller (without looking at our own bodies) requires mental comparison with one's own perceptual body image (Longo 2015; Longo and Haggard 2012). In fact, this process is markedly similar to that of many body size estimate tasks, which require patients to compare their own body size to a number of visually presented bodies (Gardner and Brown 2010). Patients' misjudgement in these tasks suggests similar misjudgement may occur throughout regular life: when comparing their body size with others, patients would falsely judge themselves to be larger. These regularly occurring (mis)judgments might plausibly ground beliefs about not yet reaching one's ideal size.

Finally, it has been proposed that patients falsely misjudge the *affordances* of their environments. Affordances can broadly be defined as the actions provided by an agent's environment (Gibson 1979). In the case of AN, false affordance judgment arises from distortion of *the body schema*—a sub-personal representation of the body relied on for motor control and affordance processing (de Vignemont 2010; Gadsby 2017c; Keizer et al. 2013).

Certain affordances, such as fitting through apertures, or into clothes, are determined by the size of one's body. However, the processing of these affordances is determined not by body size directly but body size as *represented* by the body schema (Gadsby and Williams 2018). Given that, in AN, the body schema is oversized, it's suggested that patients falsely judge such affordances: they judge that they are *too large* to accomplish certain actions, when in fact they are sufficiently small. Again, this is supported by evidence from experiments whereby patients are asked to judge whether they could accomplish certain size determined actions

(Engel and Keizer 2017; Guardia et al. 2010, 2012; Metral et al. 2014); misjudgement of these affordances suggests a similar phenomenon could occur throughout regular life. These misjudgements convey content such as "I couldn't fit into her top" or "I can't fit in between those chairs" (Gadsby 2017a, 12), providing patients with *false* information about their own body size (i.e. that it is *too large* for particular actions). These affordance-based oversized experiences are proposed to further ground false beliefs about not yet meeting one's ideal size.

A further step within empiricist models is to identify the neuropsychological abnormalities which give rise to the relevant experiences. In this regard, there is still much work to be done. False self-other and affordance judgments are suggested to arise from body image and body schema distortion, respectively—though it's not yet clear the cause of such distortion (Gadsby 2017c). In terms of tracing spontaneous mental imagery back to a single cognitive or neurological factor, this task has yet to begin. There is still much work to be done before the empiricist model of false body size belief in AN is complete; nevertheless, I put this aside for now, in order to discuss a potential challenge to the account.

2.2 Searching for a Second Factor

While it seems likely that AN patients suffer from oversized experiences—providing evidence towards the belief "I am not my ideal size"—there is more to the story, in that such patients are also exposed to *contradictory* evidence, suggesting that their bodies are in fact *thinner* than their ideal size. This evidence comes in two varieties: patients have *experiential* evidence of their own body size through weight scale and clothes size readings; further, they are exposed to *testimonial* evidence from family, friends and clinicians attempting to convince them of their true body size.[5]

This evidential situation poses a particular kind of conundrum for the basic empiricist account: why, in the face of this disconfirmatory evidence, do patients maintain their delusional beliefs? Exposure to veridical evidence of body size *should* serve to dislodge the beliefs, at very least putting patients into a state of significant uncertainty regarding their own body size, yet in many cases this evidence appears to have little effect. Attempting to solve this maintenance problem is ground well-trodden within empiricist research, whereby the most common strategy is to propose an *additional* cognitive deficit that delusional patients suffer from, beyond those that gives rise to the relevant abnormal experiences; these are generally termed "two-factor" accounts (Coltheart 2007; Davies et al. 2001).[6]

[5] It might be that patients also experience their body size accurately through *direct* visual perception, though this is a source of some contention. While some patients claim they (directly) see themselves as thin, others claim the opposite (Espeset et al. 2011). Based on evidence from mirror exposure research, it has recently been argued that AN patients' direct perception of their bodies must be veridical (Gadsby 2017c, 27). Nevertheless, this is still an open question (cf. Mohr et al. 2016).

[6] Within delusion literature, the maintenance problem just posed is sometimes distinguished from the *adoption* problem, which requires an explanation for why the abnormal content was adopted as belief in the first place (Davies and Egan 2013). That said, I won't delve into the specifics of this distinction here as, although interesting, it's orthogonal to my central thesis (cf. Gadsby 2017b, 501–503).

Bolstering the demand for an additional cognitive factor are the findings of *non-delusional analogues*. These individuals suffer from the same abnormal experiences as delusional patients albeit lacking the associated beliefs (Coltheart 2007). For example, there are cases of patients with ventromedial frontal lesions who show the same lack of autonomic response as Capgras patients when presented with familiar faces (Tranel et al. 1995). Given this evidence, it's assumed these patients have the same abnormal experience of familiar faces that Capgras patients do (recognition without affective response), despite lacking the associated beliefs (Davies et al. 2001, 144).[7] Such cases are taken as further evidence that a second cognitive factor, beyond abnormal experience, is required. This factor would be present in delusional patients but absent in their non-delusional counterparts, dissociating the two groups.

There is some tentative evidence in favour of the existence of non-delusional individuals who undergo oversized experiences. For example, many healthy controls (especially those with high body concern) exhibit oversized perceptual body images (Baker et al. 1995; Plies and Florin 1992; Taylor and Cooper 1992). Similarly, in one study, experimenters found that roughly 50% of their healthy controls reported spontaneous mental imagery of an overweight body (Somerville et al. 2007, 439). These individuals might constitute non-delusional analogues, suffering from oversized experiences without the associated false beliefs. Although it's too early for any definitive claims regarding the existence of non-delusional analogues, if indeed they do exist then this would constitute further evidence in favour of a second factor in the case of AN.

What could constitute this second factor, accounting for why AN patients maintain their beliefs in the face of contrary evidence and also, perhaps, dissociating them from non-delusional analogues? Earlier formulations of the two-factor account dictated a few essential properties of the second factor. First, it was suggested to be a *domain general* belief system dysfunction: a failure to inhibit pre-potent doxastic response, causing patients to unreflectively endorse their experiences as belief (Davies et al. 2001, 153). Second, its neurological basis is right frontal hemisphere damage (Coltheart 2007).

This early proposal is irreconcilable with the case of AN. Most obviously, there is no evidence to suggest AN patients exhibit right frontal hemisphere damage. Further, the domain general condition entails an unlikely consequence: that delusional patients should develop delusions in response to any strange experience e.g. perceptual illusions; yet there is no evidence that this is the case in AN—or any other delusion, for that matter (Davies et al. 2001, 153; Hohwy and Rajan 2012, 8; cf. Coltheart 2007, 1056).

Nevertheless, two-factor adherents often advocate a *more general* two-factor approach, which simply specifies two desiderata for any potential theory of delusion: answering *how* the content of the delusion arises and *why* the delusion is maintained (Langdon et al. 2008). Within this less stringent framework, the suggestion was made that, at least for *some* delusions, motivational biases might

[7] In more recent two-factor literature, there is some debate over whether indeed there is a conscious experiential element to this autonomic dysfunction; I won't delve into those debates here though (see: Coltheart et al. 2010; Young 2011).

perform the second factor role (Davies 2009; Langdon et al. 2008; McKay et al. 2005; 2007). Under this proposal, the existence of motivational biases explain why the belief is adopted and maintained in the face of contrary evidence. Such a suggestion not only unshackles us from restrictive claims regarding neurological damage, it also avoids the aforementioned criticism against the domain general condition, as motivational biases only apply to *certain kinds* of incoming sensory input. Further, as I will demonstrate, this suggestion seems particularly compatible with the case of AN.

3 Self-Deception and Biased Evidence Treatment

Insight into the role that motivational biases might play as a second factor comes from the literature on *self-deception*—largely considered a paradigmatic example of individuals holding false beliefs as a result of motivational influences (Deweese-Boyd 2017). Self-deception occurs in two varieties: *straight* and *twisted*. In straight self-deception, one deceives oneself into believing something they desire to be the case, such as the parent who deceives themselves into believing their child isn't experimenting with illicit drugs. This clearly doesn't fit the case of AN—patients don't desire to *not* be their ideal size, quite the opposite! Instead, AN has been suggested as an instance of twisted self-deception, whereby one deceives oneself into believing something one desires to be *false* (Sullivan-Bissett et al. 2016). As a non-pathological example, consider a jealous husband who, eagerly desiring his wife to be faithful, becomes convinced that she isn't, despite all evidence to the contrary (Mele 2001).

In recent years, the relationship between self-deception and delusions has gained considerable attention (Bayne and Fernández 2009; Bortolotti and Mameli 2012; McKay et al. 2005). Of particular relevance is Mele's (2001) *minimal* notion of self-deception. Mele claims that motivations can contribute to false belief through biasing the way in which individuals gather, attend to and interpret evidence (26–27). Such beliefs are, for Mele, instances of self-deception. That said, I won't concern myself here with the conceptual issue of whether AN patients truly count as self-deceived. Such a proposal rests on much deeper debates regarding the philosophical underpinnings of the notion. Instead—following similar attempts aimed at other delusional beliefs (Davies 2009; Mele 2006)—I merely commandeer explanatory power from Mele's analysis to reveal how motivational bias could constitute the second factor in the case of AN.

First, I review evidence of AN patients' biases in evidence treatment. These biases, I suggest, cause evidence from oversized experiences to be amplified, while veridical body experience is neglected. Then, to proffer an explanation for how these biases arise, I adopt Mele's (2001) proposal regarding the relationship between desires and hypothesis testing in self-deception.

3.1 Evidence Treatment Biases

3.1.1 Selective Evidence Gathering

Biased evidence gathering is considered to be a fundamental feature of AN. To start, the standard cognitive behavioural model of AN posits a repeated process of *body checking* (Fairburn et al. 1998). Aspen et al. (2013, 821) describe this process and its epistemic consequences:

> Frequent checking of body parts (e.g., checking weight and/or the way clothes fit, intense scrutiny of particular [disliked] body parts; pinching skin to assess fatness) leads to strengthened AN-related behaviours. This repeated scrutiny of body parts serves as a confirmation bias in which individuals with AN seek out supporting evidence of their AN-related beliefs (e.g., ''I am enormous'')…

This body checking constitutes a form of biased evidence gathering as it is targeted to *reinforce* false body size beliefs. Indeed, some patients even admit to engaging in such behaviour in order to "induce distress and hence increase their motivation to maintain dietary restraint" (Shafran et al. 2004, 100).

The standard cognitive behavioural model also claims that AN patients sometimes adopt *body avoidance* strategies:

> Over time, the hypervigilant monitoring of shape and weight ('body checking') may become highly aversive. Some individuals are no longer able to tolerate the repeated self-examination. As a consequence they come to actively avoid monitoring altogether. *This avoidance maintains their shape and weight concerns in part because they no longer have a potential means of disconfirmation*. (Fairburn et al. 1998, 7, my emphasis)

Both these strategies, despite appearing contradictory, subserve the same evidential purpose (Tuschen-Caffier et al. 2015, 12). Body-checking behaviour is targeted at evidence which supports false body size beliefs (e.g. visually attending to "fatter" parts of the body, see Sect. 3.1.2), while avoidance is targeted at evidence which disconfirms these beliefs.

Another form of evidence gathering bias relates to the *self-comparison* behaviour of ED patients. As discussed, ED patients are known to engage in self-comparison significantly more than healthy controls. In addition to this increase in frequency, there is a further bias to this self-comparison behaviour. In terms of self-comparisons, there are two possible kinds: upwards comparison, with those one believes are *better* than themselves (in this case, those with thinner bodies) or downwards comparison, with those one believes are *worse*. Not only do ED patients engage in significantly more self-comparison than neurotypical individuals, they engage in significantly more *upwards* self-comparison (Blechert et al. 2009; Corning et al. 2006; Green et al. 2009; Thompson et al. 1999). In other words, they actively seek out comparisons that will result in "they are thinner than me" judgments, reinforcing "I am not my ideal size" beliefs.

3.1.2 Attentional Biases

The belief confirming nature of patients' evidence gathering practises is likely buttressed by biases at the attentional level. While much of the research into attentional biases in ED pertains to the study of *semantic* processing—e.g. through the use of Stroop colour-naming and dichotic listening tasks (for review, see Williamson et al. 2000, 562–563)—most relevant to our current purposes is research from eye-gaze studies. This shows that when viewing themselves, those with eating disorders (or high eating disorder symptomatology) allocate their attention more towards self-identified "ugly" body parts (Freeman et al. 1991; Jansen et al. 2005; Tuschen-Caffier et al. 2015). Such biases likely work in tandem with body checking behaviour, specifically allocating attention to evidence which reinforces false body size beliefs.

In a similar vein, attentional biases are posited as relevant to the role of affordance based oversized experiences. Specifically, it's claimed that an individual's *preoccupations* contribute to the salience of certain affordances (Gadsby 2017a, 608). Consider that for any given environment-agent combination there are an infinite number of affordances that can potentially be processed. As such, the individual must *filter* these affordances, only attending to those which are personally relevant. Short-term desires (e.g. to engage in certain actions), of course, help to filter which affordances will be attended to but more general themes of preoccupation also play a role.

Consider the example of someone who obsessively thinks about skateboarding. The skateboarding related affordances of an environment (jumping those stairs, sliding down that rail) will have increased salience for such an individual: these affordances will *jump out*, commandeering greater attention. This is regardless of whether they indeed *desire* to skateboard at that particular time. Similarly, for someone with *less* preoccupation with this general theme—thinking about it less often—salience of these affordances will be decreased, leaving them less regularly attended to and perceived. In this way, mental preoccupation partly determines affordance salience.

In the case of AN, patients exhibit intense preoccupation with ideas about their own body size (Mountjoy et al. 2014). As such, it has been suggested that this preoccupation leads to a heightened salience of body size related affordances (Gadsby 2017a, 609). In this way, size-related affordances which would otherwise remain unnoticed come to play a stronger role in grounding the relevant body size beliefs.

3.1.3 Interpretational Biases

There is also evidence of bias in how evidence is *interpreted* by AN patients. A common clinical observation is that they tend to infer negative meaning when presented with even slightly ambiguous information about their own body size. For example, "if someone comments to AN patients that they are 'looking healthy,' the patient may interpret that statement as 'I am getting fat'" (Jackman et al. 1995,

342); similarly, feelings of fullness are often interpreted as "feeling fat" (Espeset et al. 2012, 523; Williamson et al. 2004, 714).

Biases in interpretation of ambiguous evidence have also been shown within experimental settings, through tasks which require participants to imagine themselves in situations described by ambiguous body size related sentences. Upon recall of the imagined situations, ED patients and those high in body dysphoria show a bias towards negative (i.e. "fat") interpretations (Cooper 1997; Jackman et al. 1995; Williamson et al. 2000). Beyond the negative interpretation of ambiguous evidence, AN patients have also been known to *explain away* contradictory evidence. For example, Espeset et al. (2011, 183) discuss a number of AN patients who interpret their low weight scale readings as evidence that their scales are broken, or that their bones are "lighter than usual", rather than that their beliefs about not yet meeting their ideal body size are misguided.[8]

Finally, interpretational bias is posited to play a role in the aforementioned empiricist model. Indeed, for spontaneous mental imagery to play a role in grounding body size beliefs, an additional interpretational bias *must* be present: the mental imagery must be *interpreted* as veridical. That said, it's not yet clear how this bias might be characterised or even if it is present. It does, however, represent a potentially fruitful target for future empirical investigation.[9]

3.1.4 Evidence Treatment Biases as Second Factor

In the preceding sections, I discussed a number of biases in the way AN patients gather, attend to and interpret body size related evidence (see Table 1). Evidence of these biases is nothing new and many have incorporated them into maintenance models of eating disorders before (Fairburn et al. 1998; Williamson et al. 2004). However, such biases are uniquely relevant to the proposed empiricist model: they solve the maintenance problem, accounting for why patients maintain their false body size beliefs despite being exposed to significant disconfirmatory evidence.

Specifically, under the proposed model, evidence treatment biases serve to emphasize evidence from oversized experiences and deemphasize evidence from veridical body size experiences. AN patients would actively seek out evidence from oversized experiences, while avoiding veridical body size experiences—a process which is further driven by attentional biases, whereby oversized experiences exhibit increased salience. Furthermore, interpretational biases would cause the evidence provided by oversized experiences to be endorsed, while veridical body size experiences are disregarded or explained away. Such biases, in conjunction with the oversized experiences enumerated in Sect. 2, present a persuasive two-factor account of how AN patients adopt and maintain their false body size beliefs.

[8] This example coheres with research into the confabulatory practices of delusional patients, who often arrive at patently implausible explanations for evidence which conflicts with their delusional beliefs (Langdon and Bayne 2010, 323).

[9] See Holmes and Mathews (2010, 354–355) for a discussion of some different hypotheses for why clinical patients might come to interpret spontaneous mental imagery as veridical.

Table 1 Biases in evidence treatment

Evidence gathering	Attentional	Interpretational
Active body checking/ avoidance	Visual biases towards disliked body parts	Negative interpretation of ambiguous evidence
Upwards self-comparison	Heightened affordance salience	Rationalising contradictory evidence

3.2 The Role of Desire

3.2.1 The FTL Model of Hypothesis Testing

While there is significant evidence of biases in evidence treatment in AN, a remaining question pertains to the *cause* of these various biases. Following other two-factor theorists, I proposed that *motivational* biases constitute the second factor. However, it's not yet clear where motivations fit into the picture. If we take a fairly standard definition of motivational bias as: "[the influence on judgment by] the desirability or undesirability of events, consequences, outcomes, or choices" (Montibeller and Winterfeldt 2015, 1235), the task at hand is to show how *desirability* relates. This is where Mele's (2001) theory of self-deception comes into play, the central insight of which is that biases in evidence treatment can arise from the undesirability of holding certain false beliefs.

Mele's proposal sits within the context of what he terms the "Friedrich-Trope-Liberman" (FTL) model of hypothesis testing (Friedrich 1993; Trope and Liberman 1996). While we *generally* all desire to hold only true beliefs, the central insight of the FTL model is that holding some false beliefs is represented as less desirable than holding others. On this view, agents aren't neutral truth trackers, steadfastly seeking true belief while avoiding falsity; rather, when it comes to belief formation, they seek to minimise "costly errors" (Mele 2001, 31).

In this model, *errors* are the false beliefs themselves and the *cost* of a false belief is "the cost, including missed opportunities for gains, that it would be reasonable for the person to expect the belief—if false—to have, given his desires and beliefs" (Mele 2001, 58). For example, falsely believing my child is experimenting with illicit drugs has a great cost associated with it, in terms of unnecessary psychological discomfort. Falsely believing the girl next door *isn't* interested in me has a great cost also, in terms of missing out on a date. As such, falsely believing that my child is experimenting with drugs and falsely believing that the girl next door isn't interested are both costly errors.

Now that the essential ingredients are laid on the counter, the best way to sample the flavour of this model is by contemplating how it might work in some typical instances of self-deception. Consider the examples of the pharmaceutically curious child and the romantically interested girl next door. As stated, falsely believing these statements come with costs. However, falsely believing the opposite of these—that my child *isn't* experimenting with drugs and that the girl *is* interested—comes with costs also: missing the opportunity to counsel the child and landing oneself in an embarrassing situation with the girl.

What determines which costs are greater? This comes down to the psychological profile of the individual. For people with certain psychological profiles—such as a significant adversity to psychological discomfort or an extreme desire to find a mate, coupled with extreme attraction to the girl next door—the costs associated with the former errors may outweigh the costs associated with the latter. When cost disparity occurs, hypothesis testing (i.e. evidence treatment) is influenced: one tests hypotheses in ways that minimise costly errors (Mele 2001, 41). This biased hypothesis testing in turn contributes to the adoption and maintenance of false beliefs.

3.2.2 Application to the Case of Anorexia Nervosa

While the FTL model appears adequate in accounting for standard cases of self-deception—whereby biased evidence treatment results in individuals believing something they desire to be true—to be applicable to AN, it must also account for twisted cases, whereby the belief is something the subjects desire to be *false*. Fortunately, Mele (Mele 2001, chapter 5) offers a description for how the FTL model can also apply to the twisted variation.

In many putative cases of twisted self-deception, holding the relevant false beliefs—e.g. "my wife is unfaithful", "I am not my ideal size"—undoubtedly causes psychological distress. As such, we might suppose that such beliefs would be classified as costly, causing the belief to be avoided. Yet in cases of twisted self-deception, Mele argues, cost disparity still occurs; for example, the costs associated with falsely believing "my wife is unfaithful" can be outweighed by the costs associated with falsely believing the opposite. Mele describes how particular psychological profiles might result in such an imbalance:

> It certainly is conceivable that, given a certain psychological profile, a strong desire to maintain one's relationship with one's spouse plays a role in rendering the potential error of falsely believing one's spouse to be innocent of infidelity a "costly" error, in the FTL sense, and more costly than the error of falsely believing one's spouse to be guilty. After all, the former error may reduce the probability that one takes steps to protect the relationship against an intruder. (2006, 114)

A similar situation may hold in the case of AN. Undoubtedly, falsely believing that one *isn't* their ideal size causes psychological discomfort but falsely believing one *is* their ideal size may be regarded as even more costly. Indeed, given the strong social element involved in AN patients' drive for thinness (i.e. the desire to be *seen* as thin) a situation in which one *wasn't* thin but believed they were would be significantly undesirable and therefore classified as significantly costly. Such a situation would render one unwilling to amend their current state (dieting to decrease size) due to ignorance of their own body size—unbeknownst to them, all their peers would see them as overweight.

In addition to costs derived from the desire to not be *seen* as overweight, the pathological desire to not *be* overweight might also play a role, through *lowering* the costs associated with falsely believing that one *hasn't* reached their ideal size.

Indeed, falsely holding such a belief is in many ways advantageous for someone with a pathological desire for thinness: it helps to motivate weight loss efforts, ensuring that one never slips accidently into being overweight. This final point coheres with the aforementioned testimony from AN patients who claim they often engage in biased evidence gathering *in order to* increase motivation for dietary restriction (Shafran et al. 2004, 100).[10]

Given these associated costs and the relative imbalance between them, the FTL model predicts that such agents will engage in biased forms of hypothesis testing, a natural result of the tendency to avoid costly errors. Further, we might suppose that the *greater* the cost imbalance (as might be the case in AN), the *more* biased the hypothesis testing practises become. Given the pathological nature of patients' desire to neither *be* overweight nor be *seen as* overweight, we might thus expect significant evidence treatment biasing to occur. To conclude, the FTL model offers an explanation for how AN patients' own fears about being overweight bias their evidence treatment practices, via a high cost associated with falsely believing "I am my ideal size" and a low cost associated with falsely believing the opposite.

It seems plausible that individuals who pathologically value thinness might consider falsely believing they are their ideal size to be vastly more costly than falsely believing the opposite. However, we needn't rely on speculation alone here: the issue is also open to empirical investigation. Specifically, questionnaires could be designed which aim at establishing the costs patients associate with holding certain false beliefs. While offering a specific proposal for what form such questionnaires would take is beyond the purview of this essay, one possibility is that participants are read short vignettes—whereby an individual holds the relevant false belief—and asked to rate how undesirable such scenarios would be. It might even be found that differences in undesirability ratings go some way towards accounting for the differences in belief conviction between patients (Phillipou et al. 2017).

Of course, confirming that AN patients exhibited the relevant cost imbalances would not settle the issue: there is still the matter of empirically validating the FTL model itself and confirming that costs associated with false belief outcomes do indeed cause biased evidence treatment, of the kind AN patients exhibit. Nevertheless, the suggested experimental approach certainly represents a step in the right direction.

Within the context of the FTL model, I have suggested that AN patients' strong desire to neither be overweight nor be seen as overweight may result in a high cost associated with falsely accepting the hypothesis "I am my ideal size" and a low cost associated with the opposite. According to this model of hypothesis testing, evidence treatment is biased towards avoiding costly false beliefs. This explains

[10] An interesting point arises here regarding whether the relationship between desires and biased hypothesis testing must be *consciously* mediated by a belief that this form of hypothesis testing will avoid the relevant costly error (Mele 2001, 31–32, 42–46). In some cases, such as the mentioned excerpt, patients clearly are *aware* that certain evidence treatment practices (i.e. body checking) will aid in avoiding undesirable situations. Yet this needn't be the case with all instances of biased hypothesis testing. For example, it seems less likely that attentional and interpretational biases are consciously mediated and indeed the FTL model allows that much of this biasing is "automatic and inflexible … reflecting the operation of evolved cognitive adaptations to a range of biologically significant problems" (Friedrich 1993, p. 317).

how the biases in evidence gathering, interpretation and attention reviewed in Sect. 3.1 might arise.[11] This (pathologically) strong desire to neither be overweight nor be seen as overweight might then be said to constitute the second-factor, causing biased evidence treatment practises, contributing to the maintenance of patients' false body size beliefs. Similarly, this strong desire may dissociate AN patients from non-delusional analogues who undergo similar oversized experiences, without the associated body size beliefs. While this proposal is in need of appropriate empirical validation, it appears credible as an explanation for how patients' treatment of body size related evidence comes to be biased.

4 Conclusion

While the empiricist approach demonstrates promise as an explanation for the false body size beliefs of AN patients, as with other delusions, a second factor is needed. Following a common proposal within two-factor research, I suggested that motivational biases might fill this role. Specifically, I proposed that biases in evidence treatment cause oversized experiences to be amplified, while veridical size experiences are discounted and, further, that these biases result from motivational influences—via cost imbalances related to beliefs about meeting one's ideal body size.

It remains to be seen how this model can be applied to other eating disorders. The most likely candidate is, of course, BN. Not only do the two disorders frequently overlap (Vitousek et al. 1998, 396) but BN is also associated with perceptual body image distortion (Norris 1984; Whitehouse et al. 1986), recurrent spontaneous mental imagery (Somerville and Cooper 2007; Somerville et al. 2007) and systematic biases in body size evidence treatment (Williamson et al. 2004). Despite a relative lack of body representation and delusional belief research targeting BN, it seems likely that the same model will apply to many of these patients, even if a few modifications are needed. Ultimately, I suspect that the kinds of behavioural (and weight) differences that distinguish AN from BN patients won't prove relevant to whether the proposed model applies; rather, the important factor will be the existence and prevalence of the relevant beliefs. As such—and in following with the cognitive neuropsychiatric tradition—it may prove beneficial to overlook the apparent diversity in the conditions themselves and instead narrow our focus onto the beliefs.

Even restricting our scope to AN, there is still much work to be done. While some of the discussed biases in evidence treatment in AN are supported by robust bodies of evidence, others clearly need further empirical validation and some are little more than tentative suppositions. Further, both the FTL model of hypothesis testing and the costs patients associate with holding certain false beliefs call for

[11] It's worth highlighting that this story would markedly differ from the discussed hypothesis regarding affordance salience in AN. While that hypothesis claims *all* size-determined affordances have increased salience (due to patients' mental preoccupation with body size related themes), the proposed self-deception hypothesis suggests a particular subset of size-determined affordances (those likely to reinforce beliefs about being overweight) would exhibit increased salience. Such a bias would manifest in patients attending to affordances they believed their bodies were too large for.

meticulous empirical probing. Finally, there might be other factors at play here—perhaps some yet to be identified cognitive deficit(s)—which contribute to the maintenance of these false beliefs. It's still early days in terms of this empiricist approach, with disagreement and debate undoubtedly on the horizon. Nevertheless, once the dust settles and a finalised model is left standing, it seems likely that motivational biases will play a decisive role.

Acknowledgements Thanks to Peter Clutton and Jakob Hohwy for feedback on earlier drafts. This research was supported by an Australian Government Research Training Program (RTP) Scholarship.

References

Alleva, J., Jansen, A., Martijn, C., Schepers, J., & Nederkoorn, C. (2013). Get your own mirror. Investigating how strict eating disordered women are in judging the bodies of other eating disordered women. *Appetite, 68,* 98–104.

Aspen, V., Darcy, A. M., & Lock, J. (2013). A review of attention biases in women with eating disorders. *Cognition and Emotion, 27*(5), 820–838.

Baker, J. D., Williamson, D. A., & Sylve, C. (1995). Body image disturbance, memory bias, and body dysphoria: Effects of negative mood induction. *Behavior Therapy, 26*(4), 747–759.

Bayne, T., & Fernández, J. (2009). *Delusion and self-deception. Mapping the terrain in delusion and self-deception*. New York: Psychology Press.

Bayne, T., & Pacherie, E. (2004). Experience, belief, and the interpretive fold. *Philosophy Psychiatry & Psychology, 11*(1), 81–86.

Blechert, J., Nickert, T., Caffier, D., & Tuschen-Caffier, B. (2009). Social comparison and its relation to body dissatisfaction in bulimia nervosa: Evidence from eye movements. *Psychosomatic Medicine, 71*(8), 907–912.

Bortolotti, L. (2010). *Delusions and other irrational beliefs*. Oxford: Oxford University Press.

Bortolotti, L., & Mameli, M. (2012). Self-deception, delusion and the boundaries of folk psychology. *Humana. mente, 20,* 203.

Clutton, P. (2018). A new defence of doxasticism about delusions: The cognitive phenomenological defence. *Mind and Language, 33*(2), 198–217.

Coltheart, M. (2007). The 33rd Sir Frederick Bartlett Lecture: Cognitive neuropsychiatry and delusional belief. *The Quarterly Journal of Experimental Psychology, 60*(8), 1041–1062. https://doi.org/10.1080/17470210701338071.

Coltheart, M., Menzies, P., & Sutton, J. (2010). Abductive inference and delusional belief. *Cognitive Neuropsychiatry, 15*(1–3), 261–287. https://doi.org/10.1080/13546800903439120.

Cooper, M. (1997). Bias in interpretation of ambiguous scenarios in eating disorders. *Behaviour Research and Therapy, 35*(7), 619–626.

Cooper, M., Todd, G., & Wells, A. (1998). Content, origins, and consequences of dysfunctional beliefs in anorexia nervosa and bulimia nervosa. *Journal of Cognitive Psychotherapy, 12*(3), 213–230.

Corning, A. F., Krumm, A. J., & Smitham, L. A. (2006). Differential social comparison processes in women with and without eating disorder symptoms. *Journal of Counseling Psychology, 53*(3), 338.

Currie, G. (2000). Imagination, delusion and hallucinations. *Mind and Language, 15,* 168–183.

Davies, M. (2009). *Delusion and motivationally biased belief: Self deception in the two factor framework* (pp. 71–86). Delusion and Self deception: Affective and Motivational Influences on Belief Formation.

Davies, M., Coltheart, M., Langdon, R., & Breen, N. (2001). Monothematic delusions: Towards a two-factor account. *Philosophy, Psychiatry & Psychology, 8*(2), 133–158.

Davies, M., & Egan, A. (2013). Delusion: Cognitive approaches–Bayesian inference and compartmentalization. In *The oxford handbook of philosophy and psychiatry* (pp. 88–727).

de Vignemont, F. (2010). Body schema and body image—Pros and cons. *Neuropsychologia, 48*(3), 669–680.

Deweese-Boyd, I. (2017). Self-deception. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy (fall 2017 edition)*.

Ellis, H. D., & Young, A. W. (1990). Accounting for delusional misidentifications. *British Journal of Psychiatry*. https://doi.org/10.1192/bjp.157.2.239.
Engel, M. M., & Keizer, A. (2017). Body representation disturbances in visual perception and affordance perception persist in eating disorder patients after completing treatment. *Scientific Reports, 7*(1), 16184.
Espeset, E., Gulliksen, K. S., Nordbø, R. H., Skårderud, F., & Holte, A. (2012). Fluctuations of body images in anorexia nervosa: Patients' perception of contextual triggers. *Clinical Psychology & Psychotherapy, 19*(6), 518–530.
Espeset, E., Nordbø, R. H., Gulliksen, K. S., Skårderud, F., Geller, J., & Holte, A. (2011). The concept of body image disturbance in anorexia nervosa: An empirical inquiry utilizing patients' subjective experiences. *Eating Disorders, 19*(2), 175–193.
Fairburn, C. G., Shafran, R., & Cooper, Z. (1998). A cognitive behavioural theory of anorexia nervosa. *Behaviour Research and Therapy, 37*(1), 1–13.
Freeman, R., Touyz, S., Sara, G., Rennie, C., Gordon, E., & Beumont, P. (1991). In the eye of the beholder: Processing body shape information in patients and bulimic patients. *International Journal of Eating Disorders, 10*(6), 709–714.
Friedrich, J. (1993). Primary error detection and minimization (PEDMIN) strategies in social cognition: A reinterpretation of confirmation bias phenomena. *Psychological Review, 100*(2), 298.
Gadsby, S. (2017a). Anorexia nervosa and oversized experiences. *Philosophical Psychology, 30*(5), 594–615.
Gadsby, S. (2017b). Explaining body size beliefs in anorexia. *Cognitive Neuropsychiatry, 22*(6), 495–507.
Gadsby, S. (2017c). Distorted body representations in anorexia nervosa. *Consciousness and Cognition, 51,* 17–33.
Gadsby, S., & Williams, D. (2018). Action, affordances, and anorexia: Body representation and basic cognition. *Synthese*. https://doi.org/10.1007/s11229-018-1843-3.
Gardner, R. M., & Brown, D. L. (2010). Body image assessment: A review of figural drawing scales. *Personality and Individual Differences, 48*(2), 107–111.
Gerrans, P. (2014). *The measure of madness: Philosophy of mind, cognitive neuroscience, and delusional thought*. Cambridge, MA: MIT Press.
Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
Green, M. A., Scott, N. A., Cross, S. E., Liao, K. Y. H., Hallengren, J. J., Davids, C. M., et al. (2009). Eating disorder behaviors and depression: A minimal relationship beyond social comparison, self-esteem, and body dissatisfaction. *Journal of Clinical Psychology, 65*(9), 989–999.
Guardia, D., Conversy, L., Jardri, R., Lafargue, G., Thomas, P., Dodin, V., et al. (2012). Imagining one's own and someone else's body actions: Dissociation in anorexia nervosa. *PLoS ONE, 7*(8), e43241.
Guardia, D., Lafargue, G., Thomas, P., Dodin, V., Cottencin, O., & Luyat, M. (2010). Anticipation of body-scaled action is modified in anorexia nervosa. *Neuropsychologia, 48*(13), 3961–3966.
Hackmann, A., Clark, D. M., & McManus, F. (2000). Recurrent images and early memories in social phobia. *Behaviour Research and Therapy, 38*(6), 601–610.
Hackmann, A., Surawy, C., & Clark, D. M. (1998). Seeing yourself through others' eyes: A study of spontaneously occuring images in social phobia. *Behavioural and Cognitive Psychotherapy, 26*(1), 3–12.
Hamel, A. E., Zaitsoff, S. L., Taylor, A., Menna, R., & Grange, D. L. (2012). Body-related social comparison and disordered eating among adolescent females with an eating disorder, depressive disorder, and healthy controls. *Nutrients, 4*(9), 1260–1272.
Hohwy, J., & Rajan, V. (2012). Delusions as forensically disturbing perceptual inferences. *Neuroethics, 5*(1), 5–11. https://doi.org/10.1007/s12152-011-9124-6.
Holmes, E. A., & Mathews, A. (2010). Mental imagery in emotion and emotional disorders. *Clinical Psychology Review, 30*(3), 349–362.
Jackman, L. P., Williamson, D. A., Netemeyer, R. G., & Anderson, D. A. (1995). Do weight-preoccupied women misinterpret ambiguous stimuli related to body size? *Cognitive Therapy and Research, 19*(3), 341–355.
Jansen, A., Nederkoorn, C., & Mulkens, S. (2005). Selective visual attention for ugly and beautiful body parts in eating disorders. *Behaviour Research and Therapy, 43*(2), 183–196.
Kaye, W. H., Fudge, J. L., & Paulus, M. (2009). New insights into symptoms and neurocircuit function of anorexia nervosa. *Nature Reviews Neuroscience, 10*(8), 573.

Keizer, A., Smeets, M. A., Dijkerman, H. C., Uzunbajakau, S. A., van Elburg, A., & Postma, A. (2013). Too fat to fit through the door: First evidence for disturbed body-scaled action in anorexia nervosa during locomotion. *PLoS ONE, 8*(5), e64602.

Konstantakopoulos, G., Varsou, E., Dikeos, D., Ioannidi, N., Gonidakis, F., Papadimitriou, G., et al. (2012). Delusionality of body image beliefs in eating disorders. *Psychiatry Research, 200*(2), 482–488.

Langdon, R., & Bayne, T. (2010). Delusion and confabulation: Mistakes of perceiving, remembering and believing. *Cognitive Neuropsychiatry, 15*(1–3), 319–345. https://doi.org/10.1080/13546800903000229.

Langdon, R., McKay, R., & Coltheart, M. (2008). The cognitive neuropsychological understanding of persecutory delusions. In D. Freeman, R. Bentall & P. Garety (Eds.), *Persecutory delusions: Assessment, theory, and treatment* (pp. 221–236).

Longo, M. R. (2015). Three-dimensional coherence of the conscious body image. *The Quarterly Journal of Experimental Psychology, 68*(6), 1116–1123.

Longo, M. R., & Haggard, P. (2012). Implicit body representations and the conscious body image. *Acta Psychologica, 141*(2), 164–168.

Maher, B. A. (1974). Delusional thinking and perceptual disorder. *Journal of Individual Psychology, 30*(1), 98–113.

McKay, R., Langdon, R., & Coltheart, M. (2005). Sleights of mind: Delusions, defences, and self-deception. *Cognitive Neuropsychiatry, 10*(4), 305–326. https://doi.org/10.1080/13546800444000074.

McKay, R., Langdon, R., & Coltheart, M. (2007). Models of misbelief: Integrating motivational and deficit theories of delusions. *Consciousness and Cognition, 16*(4), 932–941.

McKenna, P. (1984). Disorders with overvalued ideas. *The British Journal of Psychiatry: The Journal of Mental Science, 145,* 579.

McKenna, G., Fox, J. R., & Haddock, G. (2014). Investigating the 'jumping to conclusions' bias in people with anorexia nervosa. *European Eating Disorders Review, 22*(5), 352–359.

Mele, A. (2001). *Self-deception unmasked.* Princeton: Princeton University Press.

Mele, A. (2006). Self-deception and delusions. *European Journal of Analytic Philosophy, 2*(1), 109–124.

Metral, M., Guardia, D., Bauwens, I., Guerraz, M., Lafargue, G., Cottencin, O., et al. (2014). Painfully thin but locked inside a fatter body: Abnormalities in both anticipation and execution of action in anorexia nervosa. *BMC research Notes, 7*(1), 707.

Mohr, H. M., Rickmeyer, C., Hummel, D., Ernst, M., & Grabhorn, R. (2016). Altered visual adaptation to body shape in eating disorders: Implications for body image distortion. *Perception, 45*(7), 725–738.

Montibeller, G., & Winterfeldt, D. (2015). Cognitive and motivational biases in decision and risk analysis. *Risk Analysis, 35*(7), 1230–1251.

Moscone, A.-L., Amorim, M.-A., Le Scanff, C., & Leconte, P. (2017). A model-driven approach to studying dissociations between body size mental representations in anorexia nervosa. (Report). *Body Image, 20,* 40. https://doi.org/10.1016/j.bodyim.2016.11.003.

Mountjoy, R. L., Farhall, J. F., & Rossell, S. L. (2014). A phenomenological investigation of overvalued ideas and delusions in clinical and subclinical anorexia nervosa. *Psychiatry Research, 220*(1), 507–512.

Moussally, J. M., Grynberg, D., Goffinet, S., Simon, Y., & Van der Linden, M. (2017). Novel assessment of own and ideal body perception among women: Validation of the computer-generated figure rating scale. *Cognitive Therapy and Research, 41*(4), 632–644.

Norris, D. (1984). The effects of mirror confrontation on self-estimation of body dimensions in anorexia nervosa, bulimia and two control groups. *Psychological Medicine, 14*(4), 835–842.

O'Connell, J. E., Bendall, S., Morley, E., Huang, C., & Krug, I. (2017). Delusion-like beliefs in anorexia nervosa: An interpretative phenomenological analysis. *Clinical Psychologist*. https://doi.org/10.1111/cp.12137.

Phillipou, A., Mountjoy, R. L., & Rossell, S. L. (2017). *Overvalued ideas or delusions in anorexia nervosa?* (Vol. 51, pp. 563–564).

Plies, K., & Florin, I. (1992). Effects of negative mood induction on the body image of restrained eaters. *Psychology and Health, 7*(3), 235–242.

Pollatos, O., Kurz, A. L., Albrecht, J., Schreder, T., Kleemann, A. M., Schöpf, V., et al. (2008). Reduced perception of bodily signals in anorexia nervosa. *Eating Behaviors, 9*(4), 381–388.

Schilder, P. (1935). *The image and appearance of the human body: Studies in the constructive energies of the psyche*. London: K. Paul, Trench, Trubner.

Shafran, R., Fairburn, C. G., Robinson, P., & Lask, B. (2004). Body checking and its avoidance in eating disorders. *International Journal of Eating Disorders, 35*(1), 93–101.
Smeets, M. A. (1997). The rise and fall of body size estimation research in anorexia nervosa: A review and reconceptualization. *European Eating Disorders Review, 5*(2), 75–95.
Somerville, K., & Cooper, M. (2007). Using imagery to identify and characterise core beliefs in women with bulimia nervosa, dieting and non-dieting women. *Eating Behaviors, 8*(4), 450–456.
Somerville, K., Cooper, M., & Hackmann, A. (2007). Spontaneous imagery in women with bulimia nervosa: An investigation into content, characteristics and links to childhood memories. *Journal of Behavior Therapy and Experimental Psychiatry, 38*(4), 435–446.
Sullivan-Bissett, E., Bortolotti, L., Broome, M., & Mameli, M. (2016). Moral and legal implications of the continuity between delusional and non-delusional beliefs. In G. Keil, L. Keuck & R. Hauswald (Eds.), *Vagueness in psychiatry* (chapt. 10). Oxford University Press.
Taylor, M. J., & Cooper, P. J. (1992). An experimental study of the effect of mood on body size perception. *Behaviour Research and Therapy, 30*(1), 53–58.
Thompson, J. K., Coovert, M. D., & Stormer, S. M. (1999). Body image, social comparison, and eating disturbance: A covariance structure modeling investigation. *International Journal of Eating Disorders, 26*(1), 43–51.
Tranel, D., Damasio, H., & Damasio, A. R. (1995). Double dissociation between overt and covert face recognition. *Journal of Cognitive Neuroscience, 7*(4), 425–432. https://doi.org/10.1162/jocn.1995.7.4.425.
Trope, Y., & Liberman, A. (1996). Social hypothesis-testing: Cognitive and motivational mechanisms. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles*. New York: Guilford Press.
Tuschen-Caffier, B., Bender, C., Caffier, D., Klenner, K., Braks, K., & Svaldi, J. (2015). Selective visual attention during mirror exposure in anorexia and bulimia nervosa. *PLoS ONE, 10*(12), e0145886.
Veale, D. (2002). Over-valued ideas: A conceptual analysis. *Behaviour Research and Therapy, 40*(4), 383–400.
Vitousek, K., Watson, S., & Wilson, G. (1998). Enhancing motivation for change in treatment- resistant eating disorders. *Clinical Psychology Review, 18*(4), 391–420.
Whitehouse, A., Freeman, C., & Annandale, A. (1986). Body size estimation in bulimia. *The British Journal of Psychiatry: The Journal of Mental Science, 149,* 98–103.
Williamson, D. A., Perrin, L., Blouin, D., & Barbin, J. (2000). Cognitive bias in eating disorders: Interpretation of ambiguous body-related information. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity, 5*(3), 143–151.
Williamson, D. A., White, M. A., York-Crowe, E., & Stewart, T. M. (2004). Cognitive-behavioral theories of eating disorders. *Behavior Modification, 28*(6), 711–738.
Young, G. (2011). On abductive inference and delusional belief: Why there is still a role for patient experience within explanations of Capgras delusion. *Cognitive Neuropsychiatry, 16*(4), 303–325.

Linking Text Between Chapters 1 and 2

In chapter 1, I reviewed core debates and introduced a novel explanatory account whose assumptions are the focus of, and inspiration for, many of the following chapters. The assumptions are as follows. First, contrary to popular portrayals, many individuals with eating disorders believe that they are "fat" not because they exhibit extreme evaluative standards (regarding what constitutes a "fat" or "thin" body), but because they are unaware that they have already met and surpassed their ideal standards for body size. Second, these individuals' false beliefs about meeting their ideal size are grounded in and reinforced by misleading experiences of body size—*oversized experiences*. Third, distorted body representations are responsible for many of these oversized experiences.

In chapter 1, I developed this empiricist model into a two-factor account, reviewing evidence that suggests eating disorders involve biases in the treatment of evidence related to body size. I suggested that these biases constitute a "second factor" for the proposed model, contributing to the maintenance of body size beliefs in the face of disconfirmatory evidence. This two-factor model can be graphically represented:

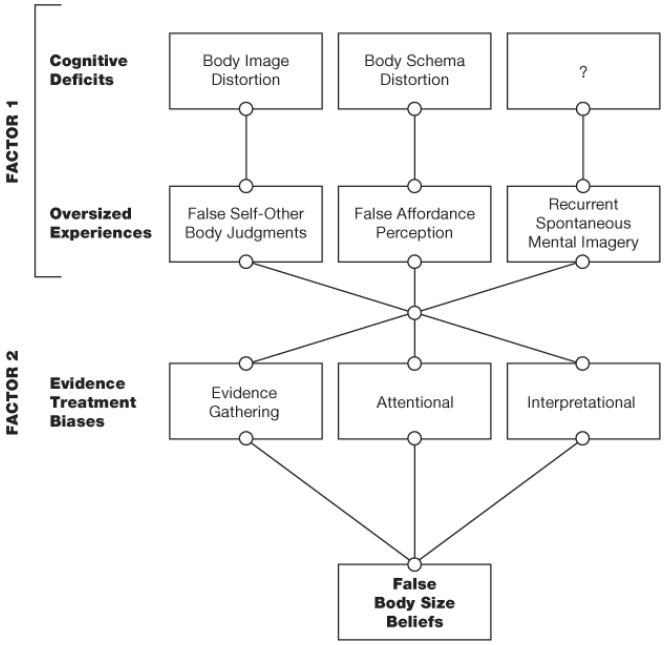


Figure 1: A graphic representation of the two-factor model. Three forms of cognitive deficit (two forms of body representation distortion and one currently unknown factor, responsible for the relevant mental imagery) underpin three distinct forms of oversized experience. These deficits and experiences fall under the first factor category, providing the content of body size beliefs. Evidence from these oversized experiences is given additional doxastic weight due to biases in evidence treatment, comprising a second factor in the model. Each factor jointly contributes to the false body size beliefs associated with eating disorders.

I went on to propose a unifying explanation of these biases, arguing that they can be understood within the "Friedrich-Trope-Liberman" (FTL) model of hypothesis testing, according to which we treat evidence in ways that minimises costly errors. I suggested that, in the case of eating disorders, it is costlier to falsely believe that one has reached one's ideal size, than it is to falsely believe the opposite (as those with eating disorders do). According to the FTL model, this cost imbalance should cause people with eating disorders to seek out, attend to, and positively interpret evidence that supports the hypothesis "I am not my ideal size" and avoid, ignore, and dismiss evidence suggesting the opposite—precisely as the evidence suggests. This provides a novel hypothesis for the underlying cause of evidence treatment biases in eating disorders and fills out a crucial element of the proposed two-factor model.

One promising aspect of the proposed account is its promise in illuminating the core principles of, and debates over, empiricist frameworks. For example, McKay, Langdon, Coltheart (2007) note that, traditionally, models of delusions that reference deficits and dysfunctions are seen as competing with those that reference motivational factors. Many researchers beholden to deficit-based explanations of delusions associate motivational factors with psychodynamics accounts of delusion which have, for good reason, fallen into disrepute. In contrast to this, McKay and colleagues claim that researchers can, and should, draw from both kinds of explanation. One way in which to do this, they suggest, is to introduce motivational bias "at the level of the second factor" (p. 938). The presented model conforms to this suggestion, validating its usefulness and reinforcing an important lesson that belief researchers should heed, regarding the "potent doxastic force" of motivations (p. 939). This very lesson guides some of the subsequent research presented in this thesis (see: chapter 6).

Another methodological lesson to draw from this chapter is that, in developing empiricist accounts, we should avoid limiting ourselves in the number of experiences identified as grounding the relevant beliefs. Empiricist accounts of delusions standardly focus on one particular kind of unusual experience which grounds the relevant belief. However, on close inspection of the available evidence from eating disorders, there are a number of important experiences that might ground false beliefs about body size. While the dysfunctions that cause such experiences may be aetiologically linked (Gadsby, 2017b), each of these experiences should be individually investigated and considered in terms of the doxastic effect that they bestow. This principle guides the next few chapters (2, 3, and 4), where I address the possibility of adding additional forms of oversized experience to the proposed model.

Finally, eating disorders represent a pivotal avenue through which many of the core claims of empiricist models can be tested. As noted, up until this point, empiricist accounts have predominately been proposed for different forms of monothematic delusions (the most common example of which is Capgras) (Coltheart, 2007). Such accounts are often motivated by how individuals would react in response to unusual experiences (for example, in response to seeing a loved one without the usual affective

response) (Coltheart et al., 2010; McKay, 2012). However, how individuals do in fact react to particularly unusual experience is an empirical question, one in need of testing. In the case of monothematic delusions, this presents a practical problem, as these delusions are not only considerably rare, but usually abate without need for clinical intervention (Coltheart, 2007). Most of what we know about these conditions stems from (rapidly conducted) single case studies. This renders it difficult to gather data on how individuals react to unusual and unlikely experiences (data which can be useful for confirming the assumptions of different accounts).

Unlike monothematic delusions, eating disorders are exceptionally widespread, indeed, there are a number of journals devoted exclusively to experiments conducted on this population.[4] So while the literature on empiricist models can help to guide and constrain an incredibly active (but often theoretically unsophisticated) research program (i.e. on eating disorders), these conditions can also help to collect data on issues of importance to empiricist models, such as people's reactions to evidence from unusual experiences.

The next step for this model is to explore the nature and cause of the relevant deficits, confirm the presence and assess the regularity of the relevant experiences (for example, through qualitative research, interviewing people with eating disorders regarding their day-to-day experiences of body size), and try to build on the model, by exploring the possibility of additional oversized experiences. This is, quite clearly, an immense undertaking, requiring the concerted efforts of a considerable number of researchers from different fields. Nevertheless, I will make some progress on this task in subsequent chapters.

In the next chapter, I extend the model by discussing another form of oversized experience that stems from distortion of the body schema (or "body model", as it is referred to in the next chapter): proprioceptive misperception of bodily boundaries. In chapters 3 and 4, I address the possibility that two additional forms of oversized

[4] For example: Eating Disorders, International Journal of Eating Disorders, Eating Behaviours, Eating and Weight Disorders.

experience might be included in the model, stemming from visual and tactile misperception, respectively.

In the previous chapter, I argued that people with eating disorders possess evidence in favour of their body size beliefs. I did not, however, address the question of whether it is rational to form and maintain those beliefs, based off of this evidence. The next chapter is more closely focused on that question. I argue that the new form of oversized experience I introduce vindicates the epistemic rationality of eating disorders, and I draw out some philosophical and practical implications of that vindication.

2. The Rationality of Eating Disorders

Abstract

Those who suffer from eating disorders often hold false beliefs about their own body size. Such beliefs appear to violate norms of epistemic rationality, being neither grounded by nor responsive to appropriate forms of evidence. Contrary to appearances, I defend the rationality of these beliefs. I argue that they are grounded in and reinforced by appropriate forms of evidence, emanating from proprioceptive misperception of bodily boundaries. This argument has far reaching implications for the treatment of eating disorders, empathy for those who suffer from them, and philosophical debates regarding the relationship between rationality and human psychology.

1. Introduction

Contemporary philosophers have shown considerable interest in real world cases of irrationality, such as delusions (Bortolotti, 2010), superstitions (Ichino, 2020), conspiracy theories (Levy, 2019), and self-deception (Funkhouser, 2019), examining them in the hope that they will furnish insights into the relationship between rationality and human psychology. Within this steadily growing body of literature, many philosophers invoke the irrationality of eating disorders (specifically, anorexia nervosa and bulimia nervosa) to support their arguments (Adler, 2002; Chislenko, 2016; Draper, 2000; Funkhouser, 2019; Radden, 2010). Such philosophers generally assume, rather than investigate, the irrationality of these disorders (cf. Tan et al., 2006). To list some examples, Radden states that eating disorders are irrational "because of their behavioral effects and their often palpably inaccurate or implausible content (such as the emaciated anorexic woman's conviction that she is fat)" (2010, p. 34). Similarly, Draper writes "The anorexic's determination to starve in the face of abundance is essentially seen as irrational - what ever psychological theory is used to explain this behaviour" (2000, p. 129).

This paper presents a novel philosophical approach towards the rationality of these eating disorders, by investigating and defending the view that those who suffer from them are rational. Specifically, I focus on norms of epistemic rationality, which specify that our

beliefs must be grounded in and responsive to appropriate forms of evidence (Bortolotti, 2014). While sufferers of eating disorders certainly appear to violate such norms—by holding seemingly ungrounded beliefs about their body size and maintaining such beliefs in the face of contradictory evidence—I will argue that such appearances are misleading.

I begin by introducing the false beliefs about body size associated with eating disorders and the norms of rationality such beliefs appear to violate (Section 2). I outline the implications of this irrationality, which include philosophical implications regarding the relationship between rationality and human psychology and practical implications related to treating eating disorders and empathising with those who suffer from them (Section 3). I propose an account that vindicates the epistemic rationality of these beliefs, by illustrating how they are grounded in and reinforced by appropriate forms of evidence, emanating from proprioception (Sections 4 and 5). I finish by addressing how this account bears relevance to the implications outlined (Section 6).

2. Eating Disorders and Epistemic Rationality

Epistemic rationality pertains to the relationship between beliefs and evidence—to be epistemically rational is to proportion one's beliefs to the available evidence (Hume, 2000). Contemporary philosophical accounts distinguish two norms of epistemic rationality, related to belief formation and belief maintenance, respectively (Bortolotti, 2010). The first norm is to "form new beliefs that are firmly grounded on the available evidence" (*ibid.*, p. 17). If one sees that there is a carton of milk in the fridge then it is rational to form the belief that there is a carton of milk in the fridge; it is irrational to form the belief that there is no milk in the fridge, ten cartons of milk in the fridge, or a pair of shoes in the fridge. I will refer to this as the norm of *grounding*.

The second norm, which I will refer to as the norm of *responsivity*, is to "update existing beliefs when relevant evidence becomes available" (*ibid.*). If I return to the fridge and see that the carton of milk is gone, the norm of responsivity demands that I update my belief about the presence of milk in the fridge. To continue believing that there is a carton of milk in the fridge is to violate this norm.

People with eating disorders appear to violate both of these norms, in virtue of the beliefs they hold about their own body size. These disorders are associated with certain evaluative beliefs about body size—for example, "I am too fat", "I am not thin enough"—and it is these evaluative beliefs that philosophers generally refer to when they refer to the irrationality of eating disorders.[5] In contrast, I focus here on beliefs about bodily dimensions. Decades of research suggests that people who suffer from eating disorders hold false beliefs about their own bodily dimensions. The clearest example of this comes from experiments that present participants with line-ups of different sized bodies, asking them to identify the image that best represents their current body size (for a recent review, see: Mölbert et al., 2017). When faced with this task, participants with eating disorders consistently indicate body sizes much larger than their own. If they are willing to endorse such judgments ("that silhouette matches my body size"), this suggests that they believe that their bodies are larger; in other words, they hold false beliefs about their bodily dimensions.

At first glance, these beliefs appear to violate both the aforementioned norms of epistemic rationality. First, they appear *ill-grounded*: given that people with eating disorders' bodies are not as large as they believe them to be, there seems to be no evidence that could ground their false beliefs. Consequently, those who hold such beliefs violate the norm of grounding. Second, these beliefs appear *incorrigible*: resistant to counter evidence. A well-known and well-discussed feature of eating disorders is that those who suffer from them are "extraordinarily resistant to efforts to persuade them to think anything else" (Vitousek, 1996, p. 388) and "strongly resistant to social feedback about their physical appearance" (Vandereycken & Van Humbeeck, 2008, p. 113). Despite the attempts of family, friends, and clinicians to convince them of their true body size, people with eating disorders maintain their false beliefs. Consequently, they violate the norm of responsivity.

Note that what is up for debate is not whether eating disorders are *ideally* rational. Psychologists have persuasively demonstrated that human beings, by and large, do not

[5] While not all individuals diagnosed with anorexia nervosa or bulimia nervosa hold such beliefs, my focus here is on those who do.

live up to ideal norms of rationality (Kahneman, 2011). As Bayne & Pacherie (2005, p. 180) put it, "Given our finitary predicament—the computational, memory, and time limitations we are subject to—it is actually irrational for us to aspire to ideal rationality". What stands out about people with eating disorders, then, is not that they violate norms of rationality, but that they do so severely. They appear to have no evidence whatsoever in favour of their beliefs about body size, but considerable evidence to the contrary, thus these beliefs appear entirely ungrounded by and irresponsive to appropriate forms of evidence.[6] What is interesting and important about these conditions, then, is how severely they seem to violate norms of rationality. In virtue of this severity, a number of important ramifications follow, which I address next.

3. The Implications of Epistemic Irrationality

3.1. The Functional Characteristics of Belief

One reason that philosophers have taken such interest in cases of irrationality is for their promise in illuminating the necessary functional characteristics of belief. Beliefs are regularly characterised as exhibiting four functional properties (Ganapini, 2019; Levy, 2018). First, beliefs are appropriately sensitive to evidence—in other words, they are epistemically rational (Funkhouser, 2019, p. 37). Second, beliefs produce appropriate behavioural effects. Believing that there is a carton of milk in the fridge gives rise to behaviours such as asserting that there is a carton of milk in the fridge and walking to the fridge when it is time to add milk to one's cereal. Third, beliefs are, as Stich (1978) phrased it, inferentially promiscuous: serving as premises to further inferences. Believing that there is no milk in the fridge leads one to infer that it is time to buy more. Finally, beliefs are practical-setting independent, producing the relevant inferential, affective, and behavioural effects regardless of context (Van Leeuwen, 2014).

Much philosophical debate surrounds the possibility that beliefs, or belief-like states, might exhibit some of these characteristics but not others (Ganapini, 2019; Levy, 2018;

[6] There is a line-drawing issue here regarding what constitutes a sufficiently severe violation of the norms of epistemic rationality. Nevertheless, we need not concern ourselves with that issue. It is intuitively obvious that the described characteristics of these beliefs represent severe violations of the two outlined norms. By the end of this paper, I will show that it is intuitively obvious that they do not.

Schwitzgebel, 2001). For example, Levy (2020) argues that some beliefs are not setting independent, while Frankish (2009) argues that some beliefs do not produce appropriate behavioural effects. What stands out as philosophically important, then, about the body size beliefs associated with eating disorders, is that they exhibit a highly unconventional functional profile. These beliefs generate the appropriate behavioural and inferential consequences—weight loss behaviour and inferences such as "those clothes will not fit" (Gadsby, 2017a, p. 607)—and do so regardless of context. However, they lack belief's (arguably) most important feature: being appropriately guided and constrained by evidence. If these beliefs are as epistemically irrational as they appear, then this demands a radical reshuffling of current mental-state taxonomies, along with an explanation for what kinds of beliefs (or belief-like states) they comprise.

In fact, some philosophers already assume that eating disorders involve philosophically important violations of the usual functional characteristics of belief. This is seen in debates over the possibility of epistemic akrasia. In epistemic akrasia an agent holds a belief while simultaneously believing that the available evidence does not support that belief (Owens, 2002). In these debates, eating disorders are taken as a potential example of epistemic akrasia, wherein people with eating disorders believe that they need to lose weight but also that the available evidence does not support that belief (Adler, 2002, 2006; Chislenko, 2016). Such debates assume that eating disorders involve severe violations of epistemic rationality, wherein these individuals possess evidence against their belief, recognise the significance of this evidence, but nevertheless fail to update their belief—violating the norm of responsivity.

3.2. Empathy

Beyond the foregoing philosophical implications, the epistemic rationality of eating disorders bears many practical implications, one of which relates to empathy for those who suffer from them. Family, friends, and clinicians often find themselves unable to empathise with those who suffer from eating disorders—"few outsiders can empathize with the plight of an emaciated adolescent distraught over the prospect of eating ice cream or stepping on a scale" (Vitousek et al., 1998, p. 398). As one former

neuropsychologist notes, in recalling their first encounter with patients with eating disorders:

> … I am now sorry to say, I found them tiresome: I'd just finished a placement at a neurorehab centre working with people who were struggling to adjust to the cruel effects of stroke, traumatic brain injury or some other neurological catastrophe. They had what I considered to be 'real' problems. Now here were these (I thought) precious young women just refusing to eat. (Broks, 2020)

One barrier to empathy for those with irrational beliefs is that we cannot imagine ourselves believing such things (Currie & Jureidini, 2001; Frankish, 2009; Gerrans, 2014).[7] By dismissing individuals as irrational, we disengage from imagining ourselves in their situation and empathising with them. In this way, judgments of irrationality undermine the possibility of empathy. This may lead to further problems, as some clinicians insist that empathy is necessary for successful treatment (Vitousek et al., 1998).

Judgments of irrationality may also reduce empathy towards people with eating disorders by inducing epistemic blame. As a number of philosophers have highlighted, when people violate norms of epistemic rationality—for example, "dogmatically continuing to believe a claim even after receiving evidence which undermines it" (Brown, 2020, p. 3596)—we often blame them for doing so (Rettler, 2018).[8] Perhaps, then, some find it difficult to empathise with those who suffer from eating disorders because they consider them epistemically blameworthy, given the way in which they flout norms of rationality. Given these considerations, the assumption that people with eating disorders are irrational may be the key to understanding why it is so difficult to empathise with them.

[7] While this claim is more commonly seen in philosophical discussions of irrational belief, it is consistent with recent empirical evidence outlining the importance of imagination to empathy (Vollberg et al., 2021).
[8] While there is significant philosophical debate over whether we should blame people for their beliefs, philosophers generally accept that we do.

3.3. Treatment

Another issue relates to how we attempt to treat eating disorders and whether some forms of treatment are justified. Many treatment methods are premised on beliefs about body size being amenable to change. For example, some methods expose clients with eating disorders to perceptual evidence regarding their body size, in an attempt to dislodge their false beliefs (Delinsky & Wilson, 2006; Keizer et al., 2019). However, if the body size beliefs of people with eating disorders are, as they appear, entirely irresponsive to counter-evidence, then such techniques are doomed to fail.[9]

Another issue related to treatment—directly tied to the question of epistemic rationality—pertains to the ethics of coercive treatment. Coercive treatment for eating disorders involves either confining someone to an inpatient treatment facility or, in some cases, force feeding them (i.e. through nasogastric tube). This issue holds deep practical importance: not only do eating disorders, most notably anorexia nervosa, exhibit a high mortality rate (Arcelus et al., 2011), they are strongly associated with treatment refusal (Goldner, 1989). While clinicians often resort to coercive treatment for eating disorders, its legal and ethical justification is a contentious issue (Draper, 2000; Tan et al., 2006).

There is one scenario in which, most agree, coercive treatment is justified. This is in cases where individuals are incompetent to make decisions regarding their own treatment. One necessary condition for decisional competence (in most European and North American legal contexts) is "the ability to reason in certain basic ways" (Grisso & Appelbaum, 1998). While there is some ambiguity over what forms of basic reasoning abilities are required (Hawkins & Charland, 2020), many specify them as those that facilitate understanding and appreciation of the relevant medical facts (Matthews, 2000, p. 63). Being unable to appropriately respond to evidence regarding one's own physical condition (body size), and thus form accurate beliefs regarding this domain, is

[9] One might argue that it is an empirical question whether these treatment methods work. If they do, then, in itself, this validates the rationality of eating disorders. However, assessing whether these techniques work in virtue of changing the relevant beliefs is practically difficult. Eating disorder treatment generally involves combining a variety of methods (Shafran et al., 2009) so when success does occur, it is difficult to pin down which methods contributed. Additionally, many studies show little to no difference between eating disorder treatments, even when compared with other, non-eating disorder specific, treatments (Murray et al., 2019; van den Berg et al., 2019)

undoubtedly crucial to understanding and evaluating what is at stake.[10] In this way, the question of epistemic rationality is directly relevant to that of decisional competence.

If people with eating disorders are entirely irresponsive to evidence regarding their true body size, then they cannot reason appropriately regarding medical advice related to their physical health. Consequently, they qualify as decisionally incompetent, and coercive treatment is justified (at least according to highly influential views regarding coercive treatment). This bears significant importance, as the irrationality of beliefs about own-body size is often cited by clinicians in justifying coercive treatment for these individuals (Draper, 2000, p. 129).

4. False Belief and Proprioceptive Misperception

4.1. Feeling Fat: The Misdescription View

A common complaint from people with eating disorders is that they "feel fat" (Calugi et al., 2018; Linardon et al., 2018; Mehak & Racine, 2020). This experience is a well-recognised aspect of such disorders: it features in many theoretical models (of both anorexia nervosa and bulimia nervosa) (Fairburn, 2008) and is measured in the eating disorders examination-questionnaire, one of the most commonly used measures for assessing eating disorder symptomatology (Fairburn & Beglin, 1994).[11] Feeling fat occurs with striking regularity. For example, Linardon and colleagues (2018) studied the prevalence of the phenomenon in a sample of 123 participants with anorexia nervosa and 51 participants with bulimia nervosa. 54% of participants reported that they "felt fat" every day of the last 28 days (Linardon, personal communication). Many people with eating disorders describe feeling fat as occurring at multiple times throughout the day. Some describe feeling this way "whenever they eat food, or meet friends (Espeset et al., 2012, pp. 526-527), others, "all day long" (Keizer, 2014, p. 10). Feeling fat is considered to be clinically important, as it drives the relevant bodily attitudes. As one clinical handbook notes "Feeling fat is a target for treatment because it tends to be equated with being fat (irrespective of the patient's actual shape and weight) and hence maintains body

[10] Indeed, people who suffer from eating disorders sometimes reject medical advice because it only applies to those who are "thin" (Tan et al., 2006, p. 7).
[11] The questionnaire asks: "On how many of the past 28 days have you felt fat?"

dissatisfaction" (Murphy et al., 2010, p. 622). Simply put, people with eating disorders believe that they are fat because they feel that way.

Despite this feeling's noted role in driving attitudes towards body size, there has been surprisingly little research into it. One likely reason is the clinical consensus that reports of feeling fat involve a kind of misdescription: "… feeling fat is a result of mislabeling certain emotions and bodily experiences. … These typically are negative mood states (e.g., feeling bored or depressed) or physical sensations that heighten body awareness (e.g., feeling full, bloated, or sweaty)" (*ibid.*). According to clinicians, when their clients report "feeling fat", they are misdescribing entirely distinct bodily experiences or emotions (McFarlane (McFarlane et al., 2011; Mehak & Racine, 2020; Zhang et al., 2014).

Note that, on this misdescription account, the false body size beliefs associated with eating disorders would violate norms of epistemic rationality, as being bored, depressed, or sweaty are simply the wrong forms of evidence to justify beliefs about bodily dimensions. Just as seeing milk and forming the belief that there is a pair of shoes in the fridge is irrational, so too is feeling sweaty and forming the belief that one is fat.

Contrary to this misdescription account, I will argue that the forms of evidence referred to in reports of "feeling fat" are not always distinct and irrelevant emotional or bodily experiences. Rather, "feeling fat" often refers to a form of proprioceptive misperception of bodily boundaries. In other words, when people with eating disorders report feeling fat, these are accurate descriptions of their experience.

4.2. Feeling Fat: The Misperception View

My argument that "feeling fat" involves misperception draws on three pieces of evidence. First, when people with eating disorders describe feeling fat, they often provide concrete, physical descriptions, of a form that seem unlikely to refer to unrelated emotional or bodily experiences. Second, the literature on illusions and mental disorders suggests that illusory experiences of body size are not only possible, but remarkably common, lending plausibility to the claim that such experiences could be implicated in the case of eating

disorders. Third, empirical evidence suggests that eating disorders involves aberrant proprioceptive processing, of a form that would give rise to misperception of body size.

In discussion with someone with a prior diagnosis of anorexia nervosa, I asked her to describe her experience of "feeling fat". She replied: "I feel as if my stomach extends to *this* point", indicating, with her hands, a point 5 cm beyond the boundary of her abdomen.[12] What stands out about this description is its concrete nature: she could indicate, quite exactly, a difference in the felt dimensions of her body. In fact, concrete, physical descriptions of this kind are a common feature of first-person reports:

> I feel fat all day long. I feel fat and fat rolls all over my body, and especially after I eat something *it feels as if my face, stomach and legs are blown up.* … When I'm around others, for example when we're sitting on a couch, or when we're eating, *it feels as if I take up too much space.* In these situations *I feel big and plump* … (Keizer, 2014, p. 10, my emphasis)

> "I feel huge. I feel so goddamn fat … I feel like a big blob … It feels like I'm overflowing." (Wooldridge, 2018, pp. 196-197)

> Heidi: I don't want to live like this for the rest of my life. But something happens when I eat. *It feels as my thighs immediately expand.* I know it isn't possible but. . .
> Interviewer: But that is how you feel.
> Heidi: Yeah, *physically* … (Nordbø et al., 2012, p. 64, my emphasis)

According to the misdescription account, these reports refer to emotional or bodily experiences that do not involve differences in body size. In contrast, I will argue that we should take these reports at face value and accept that they properly refer to matching experiences.

[12] Thanks to Manja Engel for discussing this experience with me and allowing me to relay her description.

4.2.1. Body Size Misperception is Common

Proprioception provides us with an awareness of our own bodily boundaries—where our bodies end, and the world begins. This informs us about bodily location (my boundaries are located at this point) and also body size (my body takes up this much space). While proprioception is, in typical cases, a reliable source of information about body size, there are a remarkable variety of ways in which it malfunctions, providing misleading information about the body.

Consider a few common examples. If you have ever hit your thumb with a hammer—or caused yourself immediate pain some other way—you might have felt it increase in size, despite looking the same (Valenzuela-Moguillansky, 2013). Similarly, those who undergo local anaesthesia report changes in the experienced size of the anesthetised body part (Gandevia & Phegan, 1999). Indeed, proprioceptive illusions of body size are remarkably simple to induce. For example, consider the phantom nose illusion (Ramachandran & Hirstein, 1998). In this illusion, one participant sits, blindfolded, behind another. An experimenter takes the blindfolded participants' finger and taps it on the nose of the participant in front, while at the same time tapping the blindfolded participants' nose (see figure 1). If the illusion is successful, the blindfolded participant will report that they feel their nose extended out to where their finger is tapping.

Figure 1: The phantom nose illusion (Kilteni et al., 2015, p. 5)

Misperception of body size is also a common feature of numerous syndromes. For example, those who suffer from Alice in wonderland syndrome—a condition which is commonly associated with migraines—experience their bodies as growing larger or smaller:

> … I have a very peculiar feeling of being very close to the ground as I walk along. It is as though I were short and wide, as the reflection in one of those broadening mirrors one sees in carnivals, etc. Of course I know it isn't true. (Lippman, 1952, p. 349; cited in Pitron & de Vignemont, 2017, p. 118)

> A feeling that I was very tall. When walking down the street I would think I would be able to look down on the tops of others' heads, and it was very frightening and annoying not to see as I was feeling. The sensation was so real that when I would see myself in a window or full-length mirror, it was quite a shock to realize that I was still my normal height of under five feet. (*ibid.*)

These reports refer to proprioceptive misperception of body size, inconsistent with these individuals' beliefs or visual experience. The phantom limb phenomenon also involves proprioceptive misperception of body size. Here, amputees describe feeling the spatial presence of their former limb (Ramachandran & Hirstein, 1998). As in the cases of eating disorders, many can describe this feeling in "fairly precise spatial terms" (Ratcliffe, 2019, p. 82).

Researchers take the foregoing reports at face value, accepting that they refer to genuine (albeit illusory) experiences of body size. I suggest that we take the same approach towards the reports of feeling fat associated with eating disorders. Doing so leads to the hypothesis that eating disorders involve proprioceptive misperception of body size. As I will argue, this hypothesis is eminently plausible, in light of the available evidence.

4.2.2. Eating Disorders and Distorted Body Models

Proprioception is underpinned by a consistent stream of afferent signals, emanating from a wide variety of receptors tracking the properties of our skin, tendons, muscles, and

joints. In order to calculate our bodily boundaries, these signals must be combined with information regarding the spatial properties of the body itself (Proske & Gandevia, 2012). This information is stored in a neural representation, referred to as the *body model* (Longo & Haggard, 2010). Because our perception of bodily boundaries is determined by the content of the body model, if the body model misrepresents body size, we misperceive these boundaries.

There is strong evidence to suggest that, in eating disorders, the body model misrepresents body size. First, note that this model not only underpins our perception of our bodily boundaries, but also our ability to control our bodies and assess potential actions (Gadsby, 2019a; Peviani & Bottini, 2018). This is because in order to process body size appropriate motor commands information about the size and shape of the body (derived from the body model) is required.[13] Considerable evidence suggests that those who suffer from eating disorders both act and assess their ability to act as if they had larger bodies. For example, when passing through doorways, they turn their shoulders as if their bodies were wider than reality (Beckmann et al., 2020; Keizer et al., 2013; Metral et al., 2014). Similarly, they (passively) judge their ability to pass through doorways, and other apertures, as if their bodies were wider (Engel & Keizer, 2017; Guardia et al., 2012; Guardia et al., 2010; Metral et al., 2014). In the psychological literature on eating disorders, this is taken as strong evidence in favour of the claim that the body models of people with eating disorders represent them as larger (Gadsby, 2017b).[14]

If, as researchers assume, the body model underpins perception of bodily boundaries, then people with eating disorders would misperceive these boundaries—consistent with the concrete, physical descriptions of feeling fat that many of these individuals provide.

[13] In the literature on motor control, this representation is more commonly referred to as the 'body schema' (or 'long-term body schema') (Gadsby & Williams, 2018). Nevertheless, there is good reason to assume the labels 'body model' and 'body schema' refer to the same representation (Gadsby, 2019a, p. 7).
[14] It is worth noting that, thus far, these experiments have only been conducted with participants diagnosed with anorexia nervosa or "eating disorders not otherwise specified". Nevertheless, given the close relations between the diagnostic categories (Fairburn, 2008, p. 17), it seems likely that, if tested, participants with bulimia nervosa would exhibit similar behaviours.

5. Returning to Rationality

5.1. Grounding

The norm of grounding claims that our beliefs must be formed in response to appropriate forms of evidence. At first glance, the body size beliefs associated with eating disorders appeared to violate this norm; because these individuals are much thinner than they believe themselves to be, it seemed as if there was no evidence appropriate for grounding their beliefs. Contrary to this view, I argued that these individuals do possess appropriate evidence, provided by proprioception.

The issue at hand, however, is normative rather than descriptive: while people with eating disorders may possess evidence suggesting that their bodies are larger, the question is whether they should form beliefs based off such evidence? As I will show, there are a number of reasons why they should.

First, note that proprioception is a generally reliable source of information about our bodies. For most of us, being severely misled about the boundaries of our bodies is rare. Because proprioception is a consistently reliable source of information about body size, there is good reason to invest trust in what we feel, endorsing evidence provided by this sense. However, this is not to say that proprioception should always be trusted. Consider the aforementioned examples of misperception. Those who suffer from Alice in wonderland syndrome do not believe that they are suddenly 8 feet tall, and people undergoing the Pinocchio illusion do not believe that their nose is suddenly 3 feet long. Nor should they. It would be epistemically irrational to endorse these illusory experiences, or any experience where the likelihood of the relevant state of affairs is particularly low (Bermúdez, 2001). This is because, in such cases, the kinds of evidence provided by these experiences is implausible: people do not suddenly grow 3 feet, nor do their noses. Given the implausibility of the evidence provided by such experiences, the more rational response is to believe that the experience is illusory—which is precisely what people who suffer from these illusions do.

There is, however, an important difference between the aforementioned illusions and the kinds of proprioceptive misperception associated with eating disorders: in many cases, the experience of feeling fat is entirely plausible. Unlike with phantom noses, phantom limbs, and Alice in wonderland syndrome, many of these experiences simply convey the content that particular body parts are larger than reality (though still within a reasonable range for humans).[15] In cases where proprioception provides us with plausible evidence regarding our body size, it is epistemically rational—or at least not severely epistemically irrational—to form beliefs based on that evidence.

5.2. Revisability

Contrary to appearances, the false body size beliefs associated with eating disorders are grounded in appropriate forms of evidence and therefore, in holding them, these individuals do not violate the norm of grounding. This leaves the norm of revisability. The issue here is that although these beliefs may be formed in response to appropriate evidence, they nevertheless appear inappropriately resistant to counter evidence. To assess whether this feature violates the norm of revisability, we must separately consider two different forms of counter evidence—emanating from testimony and sensory experience, respectively—and assess the effect they ought to have, given the co-occurrence with proprioceptive misperception.

5.2.1. Testimonial Counter Evidence

As noted, the incorrigibility of false body size beliefs in eating disorders is a well-known source of frustration for clinicians—despite their attempts to provide clients with accurate information regarding body size, those clients persist with their false beliefs. The important question is whether, given the proprioceptive evidence that they possess, this lack of response to testimonial evidence constitutes a (severe) violation of the norm of revisability. As I will show, there are a few features of the relevant evidential context that

[15] This is not to say that all instances of feeling fat are plausible, for example, experiencing one's body suddenly expanding after eating is not (Nordbø et al., 2012, p. 64). Rather different instances of misperception would fall on different ends of a spectrum of plausibility, with some (one's thighs suddenly expanded) falling closer to the implausible end and others (one's abdomen extending to a certain point in space) falling closer to the plausible end.

explain why testimonial evidence fails to produce the expected effect, with important ramifications for the norm of revisability.

The first thing to note is that proprioception provides a form of *first-person* evidence, which cannot be made available to others (Bayne & Pacherie, 2005, p. 183). One can only proprioceptively experience the size of their own body, and this experience cannot be shared. This does not entail that it is epistemically inferior—many of our beliefs are based on first-person evidence, for example, memories of past events. However, it does entail that this evidence cannot be shared with others. Because it cannot be shared, it cannot be held up for scrutiny and dispute, and therefore it cannot be easily contradicted by third parties (Hohwy & Rosenberg, 2005, p. 146). Beliefs grounded by proprioceptive evidence are thus difficult to refute via testimony.

While such evidence cannot be shared, it can still be referred to. And—on inspection of the relevant first-person reports—it is clear that many of these individuals do refer to this evidence when called on to justify their beliefs, for example, responding "…I can feel my body … it just feels big" (O'Connell et al., 2018, p. 5) . While these individuals may appear to be simply dismissing testimonial evidence, my account suggests a different story: they are responding to challenges with evidence of their own, emanating from first-person experience. The issue is simply that this evidence cannot be shared and appropriately scrutinised.

What appears like an outright dismissal of testimonial evidence may instead be an instance of a much more common phenomenon: people trusting their own first-person experience over the testimony of others. Indeed, this interpretation is consistent with observations from those who have interviewed people with eating disorders. As O'Connell and colleagues note, "As experts of their own bodies, they did not trust input from others that suggested their perceptual experience may be incorrect" (*ibid.*).

5.2.2 Perceptual Counter Evidence

A more significant challenge to the epistemic rationality of eating disorders refers to the perceptual counter evidence that these individuals possess. The most appropriate form of

counter evidence stems from vision, for example, during mirror exposure.[16] There is some ambiguity regarding what people with eating disorders see when they look in the mirror. Some insist that they visually perceive themselves as larger than reality, though strong behavioural evidence in support of this has not yet emerged (Gadsby, in prep.). Others claim that they see themselves as thin, despite this contradicting how they feel (Espeset et al., 2012). As one clinician describes, "although usually they may perceive their wasted body visually, they do not 'feel' the emaciation" (Vandereycken, 2006, p. 344). For present purposes, I'll assume that some who suffer from eating disorders accurately visually perceive their body size, and I'll discuss what is expected of such individuals, according to the norm of revisability.

First, note that, in many cases, visual perception of body size is much less regular than proprioceptive misperception. Apart from the fact that mirror viewing does not, generally, take place multiple times a day, many people with eating disorders avoid viewing themselves in the mirror (Fairburn et al., 1999), or if they do engage in mirror viewing, they focus on "trouble areas", rather than the body as a whole, in order to assess body size in a limited way (Tuschen-Caffier et al., 2015) So, if accurate visual perception does occur, it does so less commonly than the proprioceptive misperception described, consequently, it is less influential. When considered simply in terms of the amount of evidence, this suggests that people with eating disorders may not dismiss visual evidence, but rather that such evidence is unable to outweigh proprioception.

More importantly, it is difficult to specify the rationally mandated response when someone with an eating disorder is exposed to sensory evidence (visual or otherwise) regarding their true body size. Psychologists often approach the question of whether mental disorders involve irrationality by proposing rationally expected responses to new evidence and assessing whether certain groups respond in this way (Coltheart et al., 2010;

[16] There is some debate over whether, and in which ways, people with eating disorders accurately perceive their own body size. For example, evidence suggests that eating disorders may involve sensory disturbance in a number of modalities, which may further impede these individuals' abilities to accurately perceive their own body size (Gadsby, 2017c). Some of these other forms of sensory disturbance may even reinforce false beliefs about bodily dimensions, in combination with proprioception. However, to address that possibility is beyond the scope of this paper, so I have narrowed my focus here to proprioception alone.

McKay, 2012). However, while there is a way to answer the theoretical question of expected rational responses, it is overwhelmingly difficult to apply this to real world cases. This is because we do not possess all the relevant psychological facts, such as the individuals' background beliefs, the kinds of hypotheses that they are likely to generate in the face of the evidence, how trustworthy they rate different sensory modalities, and so on. In real-world cases, we do not possess this information, and it is difficult to practically obtain. Thus, we cannot make appropriately informed predictions about how individuals should react to highly unusual evidential circumstances.

Nevertheless, there is a prediction we can make about what epistemic rationality demands here. In assessing the epistemic rationality of beliefs, the appropriate relationship to consider is between a belief and one's total evidence, gathered over time (Worsnip, 2018). We must consider how beliefs about body size in eating disorders shift, once a substantial body of contradictory evidence has been gathered. Consider someone who is far along in the course of their eating disorder. They have suffered from reoccurring proprioceptive misperception of their body, providing evidence suggesting that their bodies are large. At the same time, they have been exposed to considerable evidence to the contrary: testimonial evidence from family, friends and clinicians, as well as perceptual evidence, such as visual perception of their body. What do the norms of epistemic rationality demand when someone is faced with significant, contradictory evidence of this kind?

One epistemically rational response, highlighted by philosophers, is to suspend one's judgment entirely (Friedman, 2013; Raleigh, 2019; Worsnip, 2018). This should not be conflated with believing nothing, rather, suspending one's judgment is a genuine doxastic attitude, and one that is rational in the face of equivocal evidence. Suspension of judgment is precisely what we see in many cases of eating disorders. Those who are further along in the course of their disorders—having amassed significant contradictory bodies of evidence—often suspend their judgment, claiming 'I don't know how I really look', or 'I've lost my sense of reality', (Espeset et al., 2012, p. 522).

While eating disorders are well known for the strength with which body size beliefs are initially held (at the point of diagnosis and early treatment) (Vitousek et al., 1998),

research conducted on individuals in the later stages of their disorders suggests that a significant number exhibit low confidence regarding the true size of their own bodies (Phillipou et al., 2017). This suggests that, contrary to appearances, many who suffer from eating disorders respond rationally to the evidential circumstances they find themselves in; at first, trusting their proprioceptive evidence and, after time, suspending judgment about their own body size.

6. Implications

According to the foregoing account, the false beliefs about body size associated with eating disorders are not ill-grounded, rather, they are grounded by re-occurring proprioceptive misperception of body size. Further, these beliefs are not incorrigible: counter evidence is, at first, simply outweighed by proprioceptive evidence. However, after time—once sufficient contradictory evidence has been gathered—many who suffer from eating disorders suspend judgment about their own body size—just as norms of rationality dictate.

This account suggests that eating disorders may not represent the case study in irrationality that many philosophers have assumed. Consider philosophical debates over the possibility of epistemic akrasia. In those debates, eating disorders are taken as a potential example of epistemic akrasia, wherein people with eating disorders both believe that they need to lose weight, and also that the evidence contradicts that belief. Philosophers engaged in these debates propose different explanations for this apparent akrasia, in terms of how people with eating disorders attend to the relevant thoughts (Adler, 2006), or attribute their own beliefs (Chislenko, 2016). In both cases, these philosophers take their ability to explain eating disorders as an explanatory benefit of their accounts. My argument suggests a different position: that the assumption that eating disorders involve epistemic irrationality (and consequently epistemic akrasia) is unwarranted. According to my account, people with eating disorders have sufficient evidence in favour of their "I need to lose weight" beliefs and do not consider their evidence as favouring the opposite. Consequently, eating disorders never arise as a potential instance of epistemic akrasia.

The position I have argued for also bears a number of practical implications. First, in terms of the ethics and legality of coercive treatment: it suggests that—contrary to the assumptions of many clinicians—the false body size beliefs held by people with eating disorders do not render them decisionally incompetent. My account suggests that the problem lies not in these individuals' ability to respond to evidence but in the unfortunate evidential circumstances that they find themselves in. This is not to say that people with eating disorders are ideally rational. The important point for the issue of decisional competence, however, is whether they suffer from severe breakdowns in their rational capacity (Grisso & Appelbaum, 1998). My account suggests that they do not. Consequently, the assumption that these individuals are decisionally incompetent, in virtue of being irrational, is unwarranted, and coerceive treatment cannot be justified in this way. This is not to say that people with eating disorders can never qualify as decisionally incompetent, there are many avenues through which they might (Tan et al., 2006).[17] However, my account suggests that clinicians and judges must justify their verdict with reference to these alternative reasons—they cannot assume that holding false beliefs about body size is sufficient evidence of irrationality.

My account also vindicates treatment methods aimed at changing those beliefs, including those that focus on exposing people with eating disorders to accurate information regarding their own body size. My account suggests that these beliefs can be changed, though doing so may be difficult, so long as the relevant body model distortion—and the false proprioceptive evidence it creates—remains (Gadsby, 2019b). In order to make progress on treatment, researchers might then focus on how to address body model distortion, as some are already doing (Keizer et al., 2019; Keizer et al., 2016). By adjusting the body models of people with eating disorders, we can correct these individuals' perception of their bodies.

[17] For example, severe starvation often reduces cognitive capacity, such that concentrating becomes difficult. In such cases, clinicians often consider coercive treatment warranted (Draper, 2000). This is partly justified in terms of saving the patient's life, but also for restoring mental capacity, so that the individual is able to reason appropriately about their own treatment.

Another promising route forward is through training people with eating disorders to understand, interpret, and reject their illusory experiences of body size. A first step in doing so—one which has not yet been explored—would be not to dismiss references to them as misdescriptions but to provide these individuals with comprehendible psychological explanations of their personal experience (Frankish 2009, p. 271). This would involve describing the phenomenon of body model distortion, and its perceptual consequences, and teaching these individuals to be wary of their proprioceptive awareness.

Finally, the proposed account characterises those who suffer from eating disorders not as precious, unreasonable, or irrational, but as victims of highly troubling evidential circumstances, outside of their will or control. Consequently, they are neither blameworthy for their beliefs, nor unsuitable as recipients of our empathy. By educating not just eating disorders sufferers, but clinicians and the general public about body model distortion and its perceptual consequences, we will hopefully engender some much-needed compassion for those who suffer from these debilitating conditions.

7. Conclusion

In this paper, I outlined a novel approach towards the rationality of eating disorders. Specifically, I defended the view that people with eating disorders are not (severely) irrational, in the way they form and maintain beliefs about their bodies. According to this account, the body size beliefs associated with eating disorders are grounded and reinforced by certain forms of proprioceptive evidence, and the nature of this evidence helps explain why these beliefs appear irresponsive to counter evidence. This position bears a number of crucial ramifications. First, the beliefs that people with eating disorders hold regarding their body size are not different in kind from typical beliefs, thus, contrary to the assumptions of philosophers, they cannot be used to support arguments regarding epistemic akrasia. Second, treatment methods premised on these beliefs being responsive to evidence are not misguided, so long as these techniques take into account the relevant forms of proprioceptive evidence. Finally, while people with eating disorders may be

decisionally incompetent (and thus appropriate recipients of coercive treatment), they are not so in virtue of irrationally forming and maintaining beliefs about their bodies.

Note a caveat regarding the preceding argument, worth repeating. I have not argued that people with eating disorders are ideally rational. They likely exhibit many forms of more subtle, non-optimal reasoning processes.[18] I have, however, shown that eating disorders do not involve severe violations of the norms of epistemic rationality, as those who suffer from them do not hold entirely ill-grounded and incorrigible beliefs. For the implications outlined, this is the key point.

There is still much work to be done in developing, extending, and verifying the proposed account. It may turn out that this account only vindicates the rationality of some of those who hold these beliefs, or for some periods of their illness. This is ultimately an empirical question—more information about the beliefs and the experiences that ground them is required. Nevertheless, given the foregoing arguments, I am hopeful that this framework will vindicate the epistemic rationality of eating disorders and usher in a new way of understanding those who suffer from these debilitating conditions—not as epistemically irrational, but epistemically unfortunate.

[18] For example, I have argued elsewhere that eating disorders involve motivated reasoning, which deviates from ideal norms of rationality (though not severely so) (Gadsby, 2020).

Linking Text Between Chapters 2 and 3

In the previous chapter, I introduced a new form of oversized experience, proprioceptive misperception of bodily boundaries, and discussed the ramifications of this form of experience, in terms of the epistemic rationality of eating disorders. This proposal augments and extends the model outlined in chapter 1, illustrating the significant amount of evidence that people with eating disorders possess in support of their false body size beliefs. It also vindicates the epistemic rationality of eating disorders: while people with eating disorders are not ideally rational (as illustrated in chapter 1), neither are they severely irrational—contrary to the assumptions of many philosophers and clinicians.

There is still much work to be done in developing this model and drawing out its philosophical implications. A crucial issue, not yet addressed, is whether and how the other forms of oversized experiences contribute to the argument that eating disorders are epistemically rational. In some cases, they clearly would. Consider the experience of (mis)perceiving an affordance: for example, not being able to fit into a chair one desired to sit in. Affordance perception is a generally reliable process, it is rare that we are wrong about where our bodies can and cannot fit, thus there is good reason to trust affordance perception to provide information about our (comparative) body size (Gadsby, 2017a, p. 610). Consequently, this form of evidence contributes to rationally justifying the relevant beliefs (in conjunction with direct proprioceptive misperception). On the other hand, it is more difficult to defend the rationality of forming beliefs based on mental imagery, given that (unlike proprioception) mental imagery is not a generally reliable source of information about the body. There are thus some important differences in how the different forms of oversized experiences relate to the rationality of eating disorders. I will not address this issue further here. However, it represents an important task for future research into the philosophical implications of the proposed empiricist model.

It is worth briefly addressing a potential contradiction between the arguments in the previous two chapters. In the first chapter, I claimed that the body size beliefs associated with eating disorders are best explained with reference to a two-factor account. In the second chapter, I claimed that, in virtue of their proprioceptive awareness, people with

eating disorders are rational to adopt and maintain their beliefs about body size. In the literature on two-factor theories, it is often assumed that a second factor is required in cases where the beliefs are not rationally justified by the relevant experiences (Coltheart et al., 2010; McKay, 2012). This second factor explains why the individuals adopt and/or maintain their delusions, despite the irrationality of doing so. So, are the relevant experiences sufficient to justify false body size beliefs (as suggested in chapter 2), or is a second factor required (as suggested in chapter 1)? This seeming contradiction can be avoided by noting how the two-factor framework from chapter 1 is motivated.

In standard two-factor accounts (as applied to monothematic delusions), the proposal for a second factor is speculative: theorists argue that the relevant experiences are insufficient to explain why the relevant beliefs are adopted or maintained, they thus assume the presence of a second factor to fill this explanatory gap. This process is akin to an inference to the best explanation: some form of (second factor) deficit is the best explanation for why delusional beliefs are held, despite them being irrational. In the case of eating disorders, however, an inference to the best explanation is not required. There is considerable independent evidence in favour of the relevant second factor (biases in evidence treatment), and it is this evidence which justifies their inclusion in the model. Thus, irrespective of whether those biases are necessary for the false body size beliefs to be adopted and maintained, they exist. If the biases exist, then they play a role in maintaining the relevant beliefs, and this role warrants their inclusion in the model. Because we need not assume that adopting and maintaining the relevant beliefs is irrational in order to posit a second factor, the epistemic rationality and one-factor vs two-factor issues come apart.

This leads to the following question: would people with eating disorders hold the relevant beliefs if they did not exhibit the second factor (motivational biases)? Or would they belong to a category of non-delusional analogues, people who suffer from the relevant oversized experiences but do not hold the relevant beliefs (I suggested the possibility of this in chapter 1). I suspect that, even if the relevant motivational biases were not present, many people with eating disorders would still hold false beliefs about their own body size. While they would not believe as strongly as if they exhibited the relevant biases, they

would believe, nonetheless. That said, some may also fall into (something akin to) a non-delusional analogue group, suffering the unusual experiences without the associated beliefs. The difference maker here would likely be the regularity of the oversized experiences (i.e. how much evidence an individual is exposed to in support of the relevant beliefs).

The next two chapters explore two additional avenues through which people with eating disorders might misperceive their own body size: through vision (chapter 4) and touch (chapter 5). People with eating disorders often claim that they cannot "see" how thin they have become. If these reports are accurate, they refer to an important form of oversized experience, one which must be incorporated into the proposed model and acknowledged as grounding and reinforcing the relevant body size beliefs. If they are not, then visual self-perception remains an important stream of disconfirmatory evidence—one that should be exploited in treatment. While first-person reports from those who suffer from eating disorders suggest that they visually misperceive their own bodies, behavioural research has yet to confirm them. I outline the issues that researchers face in accomplishing this task and provide some suggestions for overcoming these problems.

A few systematic principles will emerge from the next two chapters, related to how researchers should approach confirming that a certain population suffer from unusual experiences. While I will illustrate this point with reference to eating disorders, many of the issues I highlight, as well as the solutions, apply more generally to confirming that certain groups misperceive the world. These chapters thus bear crucial insights for researchers hoping to empirically confirm the central tenet of empiricist models: that people with unusual beliefs perceive the world in unusual ways.

3. Visual Misperception and the Science of Eating Disorders

Abstract

Many who suffer from eating disorders report that they see themselves as "fat". Despite decades of research into this phenomenon, scientists have failed to empirically confirm that people with eating disorders visually misperceive their own body size. I illustrate the importance of this issue, not only for the explanation and treatment of eating disorders, but also for recent debates regarding the cognitive penetrability of vision. I review attempts to experimentally confirm the presence of visual misperception in eating disorders and outline the problems that this research program faces. Finally, I propose potential solutions to each of the outlined problems and sketch a way forward for the science of eating disorders to make progress on this issue.

1. Introduction

A common complaint from people with eating disorders—both anorexia nervosa and bulimia nervosa—is that they see themselves as "fat". This complaint has been a feature of clinical reports since the 70s (see: section 1), leading at least some researchers to assume that visual misperception of body size is a key feature of those disorders. For example, Brooks and colleagues (2016, p. 1) write that people with eating disorders "often view themselves as much fatter than they really are". Similarly, one clinicians' handbook states, "In [anorexia nervosa], a sufferer does not see her skeletal body as it really is" (Mondraty & Sachdev, 2011, p. 3257). The claim can also be found in research pitched at general audiences. For example, in a recent article for "the conversation"—an online outlet for academics to share their knowledge with the general public—Cazzato & Sachetti (2020) write:

> [For] many people with an eating disorder … what they see in the mirror is different to what other people see when they look at them. Very often, a person with an eating disorder will perceive their body as too fat or too imperfect, despite them seeming emaciated to others.

Despite some scientists assuming that people with eating disorders visually misperceive themselves, this claim has yet to be empirically confirmed. In this paper, I discuss the importance of the claim and the problems faced by scientists attempting to confirm it. In section 1, I outline what the claim entails, with reference to the strongest evidence in its favour: first-person reports from those who suffer from eating disorders. In section 2, I describe the claim's implications, for both our understanding of eating disorders and visual perception in general. In section 3, I review attempts to experimentally verify the claim, involving tasks that require participants to visually estimate the size of their own bodies. I describe the problems that scientists face in using such tasks to confirm the presence of visual misperception. Finally, in section 4, I propose a number of solutions to these problems, in the way of experimental design principles.

2. The Visual Misperception Hypothesis

In this section, I introduce the visual misperception hypothesis and distinguish it from two related—albeit less controversial and important—hypotheses, involving cognitive and attentional differences associated with eating disorders. I also draw out an underappreciated but important feature of the phenomenon: its heterogeneity.

2.1. What Does the Visual Misperception Hypothesis Entail?

The idea that eating disorders involve visual misperception of the body can be traced back to the well-known clinician Hilde Bruch (1974, p. 89), who claimed that anorexia nervosa involves a "disturbed size awareness". Bruch based this claim on her observations of her own clients and their complaints about their inability to accurately perceive their own body size.

> …even after a good therapeutic relationship has developed, when they appear to be actively interested in understanding the background of their condition, they will complain, with a certain bewilderment, that they cannot "see" how thin they are (*ibid.*, p. 89-90)[19]

[19] This excerpt is followed by an anecdote:

Beyond Bruch's work, complaints about being unable to accurately see one's own body size have featured in a number of first-person reports (see below).

I will refer to the claim that people with eating disorders visually misperceive their own body size as the *visual misperception hypothesis*. Understanding this hypothesis—and how scientists can experimentally confirm it—is the goal of this paper.

The first thing to note is that visual misperception in eating disorders is assumed to be self-specific, only affecting perception of one's own body size. People with eating disorders do not report misperception of other people's bodies (Smeets, 1997, p. 78), indeed, they are often surprised at just how thin their peers appear (Espeset et al., 2011, p. 185). While some researchers have proposed that eating disorders involve domain general misperception of body size, I will not address this hypothesis here (Brooks et al., 2019).

It is important to distinguish the visual misperception hypothesis from some related hypotheses, with which it is sometimes conflated. The first related hypothesis pertains to cognitive differences associated with eating disorders. Consider the experience of seeing a banana: when looking at them, bananas appear to us as a certain size, shape, and colour. We can contrast this perceptual experience with the kinds of cognitive states—thoughts, judgments, and beliefs—that we form about bananas, such as the judgment: "that looks delicious", or the belief "there is a banana in front of me". The visual misperception hypothesis is not about the kinds of cognitive states (thoughts, judgments, and beliefs) that people with eating disorders form when they look at themselves.

Research shows that eating disorders are associated with negative judgments and beliefs about body size (Cash & Deagle, 1997). Why such judgments occur is what researchers are attempting to understand. They might occur due to visual misperception—wherein

---

> A woman of 20, who seemed to be making good progress, admitted "I really cannot see how thin I am. I look into the mirror and I cannot see it; I know I am thin because when I feel myself I notice that there is nothing but bones"

those with eating disorders see their bodies as larger—or they might occur despite accurate perception of body size. For example, people with eating disorders might simply apply particularly harsh standards, judging themselves as "fat", "wide", or "overweight" where others would not. The important question is what these individuals see when they look in the mirror, not what they think, judge, or believe.

The visual misperception hypothesis must also be distinguished from a hypothesis about attentional differences. Evidence suggests that when looking in the mirror, people with eating disorders selectively attend to the (self-rated) "unattractive" parts of their own bodies (Tuschen-Caffier et al., 2015). While this difference in attention does change their perception of themselves—they perceive fatter as opposed to thinner parts, in virtue of attending to those parts—it is also not what is at stake in the current debate.

Rather, the visual misperception hypothesis claims when people with eating disorders view themselves (for example, in a mirror) they misperceive their own bodily dimensions, such that the size of the body they perceive is larger. First-person reports support this strong interpretation, as many describe seeing their bodies as larger, in ways that cannot be reduced to judgments or thoughts about the body. For example, when asked what they see when they look at themselves in the mirror, many of these individuals draw a distinction between what they see and what they think, or believe:

> With my eyes I actually saw myself as big, but my intellect told me that it couldn't be true (Espeset et al., 2011, p. 184)

> In the mirror I see that I'm fat, but I do actually understand that it can't be true (Espeset et al., 2012, p. 523)

Consistent with the visual misperception hypothesis, these reports distinguish perception of the body from judgments about the body.[20] While such reports are promising evidence

[20] One might respond that these reports do not distinguish differences in visual perception from differences in attention. However, it seems unlikely that one would report seeing themselves as larger when in fact they were simply attending to those parts of their bodies that they judged to be larger.

in favour of the visual misperception hypothesis, they are only anecdotal and thus should be treated with caution. Nevertheless, for the time being, I will take these reports at face value. The overarching goal of this paper is to assess whether we can move past such reports and empirically confirm the phenomenon, as many scientists have attempted.

2.2. The Heterogeneity of Visual Misperception

Before discussing the different methods through which scientists have attempted to verify the visual misperception hypothesis, I will introduce an important feature of the phenomenon: its heterogeneity. This heterogeneity comes in two forms, *between-person* and *within-person*, both of which are evident in the first-person reports from people with eating disorders.

While some individuals with eating disorders insist that they see themselves as larger, others deny ever having misperceived their own body size:

> … I've always seen that I'm thin. And that's quite unusual. (Espeset et al., 2011, p. 181)

> I've always seen that; gosh, I'm getting too thin now, I'm sure people will notice it. (*ibid.*, p. 182)

This suggests that if visual self-misperception occurs, it does not do so uniformly—some people with eating disorders experience it, while others do not. In addition to this between-person heterogeneity, the phenomenon exhibits within-person heterogeneity: for those who do experience it, they do not do so consistently:

> It depends on your age and how long you've been sick because when I was younger I think I saw myself bigger in the mirror than I actually was in reality. But now I feel that I'm quite realistic, that I actually see my body as a childish body. I'm not big and fat… (*ibid.*)

> One day I see myself as too thin, and the next day too fat, and it can also vary from hour to hour. (*ibid.*, p. 181)

> Sometimes I actually see that I'm underweight, but then other times, I can't see it at all. (*ibid.*)

These reports suggest that visual misperception exhibits within-person heterogeneity, shifting over the course of an illness and possibly over shorter periods, such as the course of a day.

3. Implications

Some who suffer from eating disorders complain that they see themselves as larger and take care to recognize and distinguish this experience from associated thoughts and attitudes. However, not all report this, and for those who do, they often report it occurring inconsistently. Taking these reports at face value presents the following hypothesis: eating disorders involve misperception of own body size, albeit not in all cases and not consistently. In what follows, I outline a number of ramifications that this hypothesis holds. The first set of ramifications are clinical, related to how we understand and treat eating disorders. The second set are scientific, related to how we understand visual processing.

3.1. Clinical Implications

The clearest implication that the visual misperception hypothesis holds relates to one of the most important symptoms of eating disorders: false beliefs about body size. Perhaps the most well-known feature of eating disorders is that those who suffer from them believe, contrary to reality, that they are overweight, and that they must lose weight in order to achieve an acceptable body size (American Psychiatric Association, 2013). Understanding why people with eating disorders hold such beliefs is crucial to effective treatment. The visual misperception hypothesis provides a simple explanation: people with eating disorders believe that they are overweight because they see themselves as such. This explains the origin of these beliefs, as well as the remarkably tenacity that they

exhibit. Seeing oneself as larger than reality would cause people with eating disorders to believe that they were overweight, and such a belief would be consistently reinforced by regular mirror viewing. On this account, visual misperception would be a driving factor behind eating disorders. Such experiences would also contribute to these individuals' rejection of their diagnoses, as evident in the following quote:

> I've had big problems accepting that I've been diagnosed with anorexia. 'Cause people with anorexia are very thin, and I'm not. So then it doesn't fit me. And when I look at myself in the mirror I really can't understand where I have anorexia. It's nowhere! (Espeset et al., 2011, p. 185)

This would have important implications for how we attempt to treat eating disorders. One common treatment method involves mirror exposure, wherein clients are encouraged to view themselves in the mirror and objectively evaluate their own body size (Delinsky & Wilson, 2006). If the visual misperception hypothesis is true, then these treatment methods would be counterproductive, solidifying these individuals' conviction that they are overweight. If, on the other hand, people with eating disorders accurately visually perceive their own body size, then this represents a crucial form of counter evidence against their false body size beliefs, vindicating its use in treatment.

Another important issue within the science of eating disorders relates to the phenomenon of distorted body representations (Gadsby, 2017b). In order to appropriately process proprioceptive signals, our brains rely on representations of our own body size. Mounting behavioural evidence suggests that, in the case of eating disorders, these representations are distorted, representing these individuals' bodies as larger than reality. A key task for this research program is to explain the cause of this distortion (*ibid.*). The visual misperception hypothesis provides a solution here: if people with eating disorders visually misperceive themselves, then this misperception may provide the inaccurate content exhibited by their body representations.

3.2. Scientific Implications

Beyond its importance for understanding eating disorders and the false beliefs associated with them, the visual misperception hypothesis holds important ramifications for vision research more generally. Our understanding of vision has been greatly advanced through studying cases where it malfunctions. For example, cerebral achromatopsia (loss of colour perception) and motion blindness (loss of motor perception) were both instrumental in providing scientific insights regarding the visual processing of colour and motion (McCloskey & Chaisilprungraung, 2017). If eating disorders involve visual misperception of body size, then they can inform us about how visual processing of body size functions.

Note that the visual misperception hypothesis does not suggest that people with eating disorders exhibit a domain general form of perceptual dysfunction. Unlike cerebral achromatopsia and motion blindness, eating disorders do not involve an inability to perceive the relevant properties (bodily dimensions). Rather, they are assumed to involve a qualitative difference in the way in which those properties are perceived: as larger. Further still, this difference does not occur in all cases of body size perception, only in perception of one's own body size. It thus qualifies as an incredibly unique form of misperception, one that has ramifications for perhaps the most controversial issue in vision research: the cognitive penetrability of vision (Firestone & Scholl, 2016; Stokes, 2013).

Think back to the distinction between perception and cognition. It is uncontroversial to state that perception causally influences cognition. Seeing a banana causes all kinds of cognitions, for example, judgments about its taste ("how disgusting") and beliefs about its presence ("there's a banana in front of me"). What is controversial, however, is whether cognition can influence perception; for example, whether believing "there's a banana in front of me" could cause one to actually see a banana. This is referred to as cognitive penetrability, and it comes in many forms, each of which is the topic of their own debates. Two particularly controversial forms of cognitive penetrability involve beliefs influencing perception (Stokes, 2014) and emotions influencing perception (Niedenthal & Wood, 2019).

As noted, visual misperception exhibits within-person heterogeneity, fluctuating over time. Some first-person reports suggest the factors that might be responsible for such fluctuations. Specifically, they suggest that emotional states and the expectation of self-viewing can modulate whether visual misperception occurs:

> Yesterday I was really sad and today I'm in a better mood. Today when I look in the mirror, I see myself differently, more positive. Yesterday I only saw a big hippo. Actually, I've never thought about it before, that I see my feelings in the mirror. I think that's exactly how it is; when I have a bad day, I always see a big hippo. (Espeset et al., 2012, p. 524)

> I remember one occasion, I was passing an open door and saw myself in the mirror, but actually, I didn't know that I saw myself. I just saw the image of a person in the mirror and thought; "Oh gosh, she is thin!" But then, when I understood that it was actually me, I didn't see me as thin anymore. But then I actually saw a glimpse of it. (Espeset et al., 2011, p. 183)

These reports suggest that, in the case of eating disorders, vision can be modulated by features of the context of self-viewing, such as the current emotional states of the individual or the expectation of seeing themselves—contradicting the claim that vision is cognitively impenetrable. As before, we ought to exercise caution in how much trust we put in these anecdotal reports as providing evidence of visual misperception. Nevertheless, taken at face value, they do suggest that visual misperception might qualify as an instance of cognitive penetration, with significant ramifications for our understanding of visual processing.

Even if we do not accept that visual misperception can be modulated in this way, the form of visual misperception under discussion (specific to one's own body size), undoubtedly represents a compelling form of aberrant visual processing, one which is sure to provide many important scientific insights regarding the principles by which vision operates.

4. Measuring Visual Misperception

In this section, I describe the history of scientific research into the visual misperception hypothesis. Starting in the 70s, this research program attempted to measure visual misperception of the body using tasks that required participants to visually estimate their own body size. I discuss the principles behind these tasks and review the problems they face in uncovering evidence of visual misperception in eating disorders.

4.1. Body Size Estimate Tasks

Directly inspired by Bruch's clinical observations, psychologists began attempting to verify the visual misperception hypothesis, by conducting experiments requiring participants to estimate the size of their own bodies (Slade, 1985). The first of these experiments, by Slade and Russell (1973), instructed participants to manipulate the distance between two lights (mounted onto a track) until it corresponded to the width of different body parts (chest, face, waist, and hips). While the control group estimated body size accurately, the eating disorder group overestimated the width (but not height) of a number of body parts. The participants with eating disorders did not, however, overestimate the size of inanimate objects (see also: Garner et al., 1976), and only slightly overestimated other women's body parts (see also: Smeets, 1997, p. 79).

Inspired by this study, psychologists developed and employed many variations of these body size estimate (BSE) tasks, comparing participants with eating disorders (both anorexia nervosa and bulimia nervosa) against controls. These experiments' results suggest that, on average, participants with eating disorders overestimate their bodies compared to controls (for review and theoretical discussion, see: Farrell et al., 2005; Gardner & Brown, 2014; Mölbert et al., 2017; Smeets, 1997; Smeets et al., 1997).

There are a wide variety of BSE tasks. A useful distinction, for our purposes, refers to the kinds of stimuli used for estimating body size. For most BSE tasks, participants indicate their body size using stimuli that do not realistically represent themselves, such as silhouettes of bodies, computer generated bodies, or abstract distances indicated with lines, callipers, or points on a wall (for example, see figure 1). However, a newer form of

BSE task presents participants with photographs of their own body, digitally manipulated to be different sizes, and asks them to indicate the image that most closely matches their own (see figure 2). This distinction, between BSE tasks that use non-realistic stimuli and those that use realistic stimuli (such as manipulated photographs), will prove important (section 3.2.)

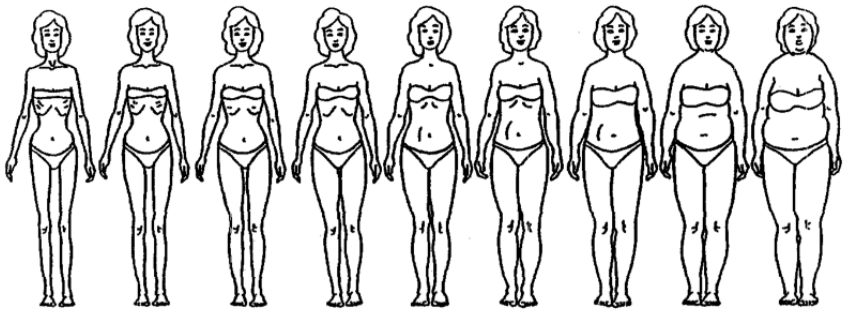


Figure 1. The contour drawing rating scale (Thompson & Gray, 1995). Experiments that use this scale instruct participants to select the image that most accurately depicts their current body size.

Figure 2. Stimuli representing a participants' own body, manipulated to appear at different sizes (Brooks et al., 2016). Though removed here (for anonymity), the participant's own face is visible in the experiment.

Early on, scientists who conducted BSE experiments interpreted their results—participants with eating disorders overestimating their own bodies—as confirmation of the visual misperception hypothesis, assuming that participants estimated themselves as larger because they saw themselves as larger (Garner & Garfinkel, 1982; Slade & Russell, 1973). This assumption requires some unpacking. Because BSE tasks restrict participants from viewing their own bodies, it is not immediately clear how visual self-perception could be implicated (Smeets, 1999, p. 46) (cf. Section 3.3). These tasks were not assumed

to directly measure participants' visual perception of their bodies. Rather, they did so indirectly, via the medium of visual memory. Scientists assumed that in order to estimate one's own body size, participants would rely on a visual memory of their own body, sometimes referred to as "the body image" (Schilder, 1935). Thus, BSE tasks were taken to be a measure of the body image's dimensions—"if overestimation takes place, the body image has to be fatter" (Smeets, 1997, p. 88).

However, even if overestimation stems from a distorted visual memory of the body, this does not entail that participants visually misperceive their bodies. For that conclusion, an additional assumption is required, namely, that the content of the body image is derived from visual perception. Smeets and Panhuysen (1995) describe this assumption:

> The body image was believed to be a snapshot of the body percept (the body as it is directly perceived), in which size and shape characteristics have been accurately preserved. Thus, overestimation of body size was interpreted as indicating how a patient perceives herself when looking in the mirror, the body image functioning as an intermediary station between this percept and the size estimate (p. 113, endnote 1)

By the 90s, most scientists had abandoned the assumption that BSE tasks could provide evidence in favour of the visual misperception hypothesis (for review, see: Smeets & Panhuysen, 1995; Smeets, 1999). I will focus here on two specific concerns that arose. The first concern regards the role of additional factors in influencing body size estimates. The second concern regards the role of non-perceptual factors in influencing visual memory of the body (the body image).

Consider the first concern. While participants may rely on their visual memory in order to estimate their body size, other factors could also influence such estimates. Thinking back to the distinction between perception and cognition, participants with eating disorders might allow their thoughts, judgments, and beliefs about their bodies to influence their estimates. Demand characteristics might also play a role, wherein participants overestimate in an attempt to "help" or "sabotage" the experiment (Smeets et al., 2009, p.

158). Another possibility is that participants with eating disorders estimate their own body size based off proprioceptive awareness. As noted, eating disorders are associated with distorted proprioceptive representations of body size. Such distortion may cause proprioceptive misperception of body size (Gadsby, in prep.). If this were the case, and participants with eating disorders relied on proprioception to estimate their bodies, it would explain their tendency to overestimate. Taking a step back, the broader issue is that responses in these tasks may be biased by factors distinct from visual memory. I will refer to this as the *response bias challenge*.

Consider the second concern. Even if body size estimates are directly reflective of visual memory—and those with eating disorders do remember themselves as larger than reality—we need not also assume that memory accurately reflects experience. Doing so would assume (as the above quote reflects) a notion of memory as operating much like a camera, storing faithful "snapshots" of experience for later recall. However, researchers no longer hold such a view: memories are not so much recalled as reconstructed—a process which is itself prone to bias and error (Kourken & Sutton, 2017, section 4.2). Scientists now assume that this process of reconstruction is a more likely culprit for body image distortion. As one of the early pioneers of BSE research, Gardner, concedes, "abnormalities in body image perception likely occur during the brain's reconstruction of the visual body image" (2012, p. 529). According to this view, while overestimation may stem from distorted visual memory, these memories become distorted during reconstruction. I will refer to this as the *reconstruction challenge*.

4.2. The el Greco Fallacy

In the previous section, I identified two problems plaguing the use of BSE tasks to test the visual misperception hypothesis: the problem of response bias and the problem of reconstruction. In this section, I introduce yet another problem, related to the so-called el Greco fallacy (Rock, 1966), and apply it to the interpretation of a certain form of BSE task, involving direct perception of body size.

El Greco was a Spanish renaissance artist who painted subjects with elongated fingers. Some art historians suggested that this painting style was due to El Greco suffering from astigmatism, an ocular defect resulting in a stretched-out perception of the world (Firestone, 2013). As Firestone & Scholl illustrate, such an interpretation is fallacious:

> If El Greco truly experienced a stretched-out world, then he would also have experienced a stretched-out canvas. In that case, the distortions should have canceled each other out: Just as El Greco would have seen real-word figures as elongated, so too would he have seen his paintings as elongated, and so the real-world distortions he experienced would never have transferred to his reproductions. The distortions in El Greco's paintings, then, must have some alternative explanation beyond a literal perceptual distortion (2014, p. 39)

The El Greco fallacy can be reduced to a simple principle, "When a constant-error distortion should affect equally the means of reproduction and the item reproduced, the effects should cancel each other out" (*ibid.*, p. 45). One way of thinking clearly about this fallacy is in terms of a difference between one's perception of what is reproduced and one's perception of the means of reproduction, which I will label: the input and output. If El Greco suffered from astigmatism, both his perception of the world (the input) and his perception of the lines on his canvas (output) would be susceptible to the same distortion. Because of this match in distortion, the strokes of his paintbrush would remain faithful to the lines of the reproduced objects—cancelling out any distortion. If El Greco's perception was distorted, his painting would not reflect it.

Note a few things about this fallacy. It is not important what kinds of actions are used to reproduce the input. El Greco could be painting, drawing, manipulating a photograph, or answering a set questions in response to different stimuli, so long as the same distortion affects both input and output, cancellation should occur. Thus, the relevance of the El Greco fallacy to BSE results is not determined by the kinds of actions participants do to recreate their body size. What does matter, however, is the appearance of the reproduced images. El Greco's assumed deficit (astigmatism) is domain general, causing him to

perceive everything as elongated. Because of this, he would misperceive both the world and the canvas, leading to a cancellation effect.

This is importantly different to the visual misperception hypothesis, wherein the relevant distortion only affects visual perception of one's own body size. This places an important constraint on how the El Greco fallacy applies to BSE results. For cancellation to take place, we must expect the same distortion to apply equally to the input (the visually perceived body) and the output (the stimuli used to estimate the body).

Consider again the distinction between different kinds of stimuli used for BSE tasks. As discussed, many of these tasks require participants to estimate their body size using non-realistic stimuli (lines, callipers, silhouettes). There is no reason to assume that participants with eating disorders would misperceive such stimuli, so no cancellation would occur. However, as noted, many contemporary variants of BSE tasks involve realistic stimuli, wherein the images represent the participants' own bodies. Such tasks lead to a match between input (the participant's visual perception of their own body, via the intermediary of visual memory) and output (their perception of the stimuli used to estimate their size). The visual misperception hypothesis therefore predicts that distortion would apply to both the perception of the body and the perception of the photograph, leading to cancellation.

4.3. Direct Perception Tasks and the el Greco Fallacy

Consider how this issue plays out in the context of a particular form of BSE task, specifically designed to avoid the reconstruction problem: *direct perception* BSE tasks. Recapping, the reconstruction problem arises because BSE tasks do not test visual perception of the body but visual memory of the body, and this memory might become distorted through non-perceptual influences. To remedy this, Shafran & Fairburn (2002) designed a task to measure visual self-perception in a more direct manner. In their paradigm, participants estimate their body size using their mirror reflection as a guide. A crucial feature of this paradigm is that it employs realistic stimuli: participants estimate size by manipulating a realistic image of themselves, one that appears identical to the

mirror reflection. This is achieved by photographing the participant's reflection, cropping out the background, adjusting it to the appropriate size, and projecting it directly next to the mirror (with the head and feet at the same height) (*ibid.*, p. 461). As with other BSE tasks, when participants with eating disorders complete this task, they systematically overestimate their own body size, manipulating the image to a size much larger than their own (Shafran & Fairburn, 2002; Øverås et al., 2014). In fact, one of these experiments found that participants with eating disorders overestimated their body size more in this direct task compared to a standard, memory-based BSE task (Øverås et al., 2014).

In these tasks, the stimuli used to estimate size are highly realistic. Given the similarity between input (mirror reflection) and output (estimating stimuli), this suggests that misperception would apply to both the mirror reflection and stimuli, causing a cancellation effect. When the participants attempt to recreate the dimensions of their reflection, they would do so faithfully, creating an accurate representation of their own body size. Thus, the finding that participants overestimate their own body size in these tasks is not evidence in favour of the visual misperception hypothesis: this misestimation cannot be explained with reference to visual misperception, as such misperception would cancel out (Gadsby, 2017b, p. 27).

Note that, even if these results were different, and these experiments found no differences in estimation between groups, such a finding would neither confirm nor disconfirm the visual misperception hypothesis. A finding of no differences could be due to a cancellation effect (suggesting misperception occurs) or due to no visual misperception occurring. The lesson to take from this is that BSE tasks must involve a mismatch between input and output, otherwise a cancellation effect will occur and results cannot be used to determine the visual misperception hypothesis.

5. Salvaging Body Size Estimate Tasks

In the previous section, I outlined a number of problems related to the use of BSE tasks for testing the visual misperception hypothesis. Most eating disorder researchers took these problems as insurmountable and abandoned their attempts to use BSE tasks to

measure visual misperception (Mölbert et al., 2017; Smeets, 1997). In this section, I show that there is still hope for BSE tasks to illuminate the visual misperception hypothesis, by providing suggestions on how to overcome these issues.

In section 1, I identified an important feature of visual misperception in eating disorders: its heterogeneity. If visual misperception occurs, it does not occur for everyone, and for those for whom it does, it may not occur consistently. This is problematic because any sample of participants with eating disorders may include many who do not suffer from visual misperception, leading to noisy data. If researchers hope to validate the visual misperception hypothesis, then they must modify their inclusion criteria to ensure that their sample only includes participants who report suffering from misperception. Further still, given the way in which visual misperception appears to fluctuate over time, researchers must ensure that their participants suffer from misperception at the time of testing. As far as I am aware, none of the research conducted thus far has taken this into account.

In section 3, I identified a number of problems related to the use of BSE tasks to test the visual misperception hypothesis (see table 1). Different problems plague different task designs. Designs that rely on participants' memory of their own bodies fall victim to the problem of reconstruction: different factors influence reconstruction of our memories, and therefore misremembering is not necessarily indicative of misperception. Scientists can avoid this problem by requiring participants to directly estimate their own body size, as is the case with direct perception BSE tasks.

For BSE tasks where there is a match in the input and output (the stimuli being estimated and the stimuli being used to estimate), the el Greco fallacy predicts a cancellation effect. This will occur for BSE tasks that use realistic stimuli for participants to estimate their size with, such as the direct perception tasks thus far conducted. In cases where the el Greco fallacy applies, results cannot be used to confirm or disconfirm the visual misperception hypothesis. Problems related to the el Greco fallacy can be avoided by creating a mismatch between input and output, such that the stimuli used to estimate are not realistic depictions of the participant's own body.

Finally, the problem of response bias plagues any task where participants are explicitly instructed to estimate their own body size. Once this instruction is in place, beliefs, thoughts, and judgments about body size, as well as expectations about what the experimenter expects, can influence the participants response. One way to approach the problem of response bias is to remove the explicit instruction to estimate body size. The issue then is how to accomplish this while still measuring visual perception of the body. In order to solve this problem, I introduce a variation on BSE tasks that use realistic stimuli: the transposed BSE task. This variation involves switching the input and output of a BSE task in an unexpected way, requiring participants to estimate the size of an object (a rectangle), by manipulating a realistic depiction of themselves (a photograph) (figure 3).

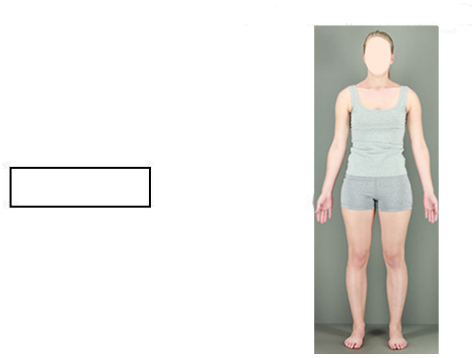


Figure 3. Example stimuli: perception of the rectangle is the input, perception of the photograph is the output. In the experiment, the participants own face would be shown. Potential task instructions: "manipulate the size of your body until the width of your abdomen matches the width of the rectangle". Photograph taken from Brooks et al. (2016).

First, note that this task design avoids the reconstruction problem. The experiment is designed to measure how participants (directly) perceive the stimuli used to estimate the size of the rectangle (the picture of their body), therefore visual memory is not involved. The experiment also avoids the el Greco fallacy, as there is a mismatch between input (rectangle) and output (photograph), so no cancellation is predicted to occur. What is notable about this design is that it obscures the goal of the experiment, in a way that avoids response bias. It does so by predicting an unintuitive result: that participants with

eating disorders will underestimate the size of the rectangle. Consider two possible outcomes:

> Outcome 1: No between-group differences in size estimation
>
> Explanation: This suggests that, just like the control group, people with eating disorders accurately perceive both the input and output, and accurately estimate the size of the rectangle, without some additional form of response bias. This finding is consistent with the hypothesis that there are no perceptual differences between these groups, in contradiction to the visual misperception hypothesis.
>
> Outcome 2: The eating disorder group underestimate the rectangle's size, compared to the control group.
>
> Explanation: This finding is consistent with the visual misperception hypothesis. Imagine that both the stimuli (rectangle and body) are 20cm wide. If participants accurately perceived the rectangle but misperceived their own body—say, as 25 cm wide, rather than 20—then, in attempting to match the perceived size of the body to the perceived size of the box, they would need to manipulate the body to be (objectively) smaller than reality.

Note how the expected outcome avoids the response bias problem. Given that the task does not require participants to estimate their own body size, it is unlikely that their response would be influenced by beliefs, judgments, or thoughts about body size. Similarly, it is unlikely that demand characteristics could cause such an outcome: it would require extremely careful thinking on the behalf of the participant to conclude that the experimenter desired them to underestimate the size of the rectangle. Because this desired outcome is much less obvious, the design is less susceptible to these influences. Thus, unlike in the previously discussed BSE tasks, the findings of this task do illuminate the visual misperception hypothesis.

This leaves a third potential outcome:

Outcome 3: The eating disorder group overestimate the rectangle's size, compared to the control group.

Explanation: While it is not clear why such an outcome would occur, it is consistent with some form of response bias wherein participants incorrectly assume the goal of the experiment (i.e. to see whether the appearance of their own body causes them to misestimate).

This experimental design is only schematic. There are many details to be worked out and improvements to be made (for example, some form of debrief, where participants are asked whether they suspected the aim of the experiment, would be useful). In conducting such an experiment, scientists must also consider the issues surrounding inclusion criteria, ensuring that their eating disorders sample reports suffering from visual misperception. By designing BSE tasks to include all of the solutions discussed in this section (table 1), scientists can salvage these tasks as a way in which to test the visual misperception hypothesis.

Problem	**Solution**
Heterogeneity	Inclusion Criteria
Reconstruction	Direct Body Size Estimation
El Greco Fallacy	Mismatched Input and Output
Response Bias	Transposed Body Size Estimation

Table 1. List of identified problems and proposed solutions.

6. Conclusion

I discussed the hypothesis that people with eating disorders visually misperceive their own body size. While first-person reports suggest that visual misperception does occur, experimental research has yet to confirm the phenomenon. This is because of the considerable number of methodological issues this research faces, which, thus far, have not been overcome.

Two of these issues—related to memory reconstruction and the el Greco fallacy—can be overcome by requiring participants to directly estimate their body size and ensuring that there is a mismatch between input (the stimuli being estimated) and output (the stimuli used to make the estimate). The issue of response bias is more difficult to surpass, plaguing every kind of BSE task thus far conducted. To address this issue, I introduced a new form of BSE task (the transposed BSE task) wherein the input and output of a BSE task are switched, and participants are required to estimate the size of an object by manipulating a photograph of themselves. When combined with solutions to the other issue, there is a hope for making progress on the question of visual self-misperception in eating disorders.

The outcome of this research has important ramifications for our understanding of eating disorders and the false beliefs associated with them. Whether or not those who suffer from eating disorders visually misperceive their own bodies is crucial to explaining their false beliefs about their body size and knowing how to treat such beliefs. The outcome is also important for our understanding of visual processing in general. If visual misperception occurs in eating disorders, then this may constitute evidence for the cognitive penetrability of vision, greatly impacting our understanding of the architecture and functional characteristics of visual processing.

Linking Text Between Chapters 3 and 4

In the previous chapter, I addressed the possibility that people with eating disorders visually misperceive their own body size. While many who suffer from eating disorders claim to—and their reports appear to refer to genuine perception (rather than attitudes or judgments)—we have yet to empirically confirm the phenomenon. Given the lack of empirical evidence, it is too early to know for sure what implications this form of experience holds for the proposed empiricist account. Nevertheless, I highlighted an important task for building on the proposed account, in terms of identifying and evaluating evidence of misperception. I outlined a number of problems related to interpreting behavioural results as confirmation of perceptual differences.

Note that while I introduced these problems in reference to a claim about visual misperception, many of them are applicable to the verification of other putative forms of misperception. Given that empiricist models are premised on the claim that certain individuals misperceive the world (in particular ways), these methodological problems are of considerable importance to the general empiricist approach. By expanding their scope and developing principles for the design of experimental paradigms, philosophers can make progress on verifying the foundational assumption of empiricist models.

In the next chapter, I apply this methodology to a novel form of misperception, putatively associated with eating disorders. Specifically, I address the possibility that people with eating disorders misperceive their own body size through touch. If they do, as the available evidence suggests, then this represents an integral form of oversized experience, which should be incorporated into the proposed model. In order to first develop this possibility and situate it within the theme of the thesis, I will first discuss the relevant research on this matter and outline its relevance to the idea of oversized experiences.

In the previous chapters, I discussed evidence that people with eating disorders exhibit distortion of the perceptual body image and body schema. In addition to this, a growing

body of evidence suggests that eating disorders may be associated with another kind of distortion, which affects a representation of the body used to process tactile size. I will refer to this representation as the *tactile form* (Gadsby, 2017b).

When an object comes in contact with our skin, our brain processes the size of that object by mapping the location of pressure on to a mental representation of the body. For example, when an object presses against the skin from point a to point b, the brain calculates the size of this object with reference to information about the distance between points a and b (Longo et al., 2010). This information is stored in the tactile form; consequently, the tactile form underpins our experience of tactile size.

As in the case of the perceptual body image and body schema, evidence suggests that anorexia nervosa is associated with distortion of the tactile form. This evidence stems from experiments requiring participants to estimate the distance between different points of tactile stimulation, without visual input (Keizer et al., 2011; Keizer et al., 2012; Risso et al., 2020; Spitoni et al., 2015). These experiments show that participants with anorexia nervosa overestimate distances compared to controls, suggesting that their tactile signals are mapped onto a distorted tactile form (Gadsby, 2017b).[21] Additionally, participants with anorexia nervosa overestimate distances horizontally, on the thigh and abdomen, but not vertically, and not on the sternum (Risso et al., 2020; Spitoni et al., 2015). This suggests that rather than being generally enlarged, tactile form distortion reflects the commonsense dimensions of a fatter body (wider along the hips and abdomen). This form of misperception would, in some situations, provide these individuals with misleading information about their own body size

Tactile perception of one's own body size is a salient aspect of many common experiences. Most will remember a time that, when putting on a particularly tight t-shirt, they became acutely aware of the size of their torso, through the feeling of tactile pressure. Indeed, many common tactile experiences convey information about our body size: not only wearing tight clothing, but also resting our body against surfaces, such as

[21] Or, at least, one that is distorted compared to neurotypical controls.

when we sit on chairs or lie on beds. In such instances, touching other objects provides us with feedback about our own body size. Distortion of the tactile form would cause one to misperceive their own body size during such experiences. If, as the evidence suggests, eating disorders involve tactile form distortion, then this would cause them to misperceive their own body size, when sitting on chairs, lying on beds, wearing tight clothing. This would qualify as an important form of oversized experience, with relevance to the proposed empiricist model.

With this in mind, in chapter 4 we test the hypothesis that people with eating disorders—specifically, anorexia nervosa—misperceive tactile size. In line with some of the conceptual arguments regarding visual misperception of body size, canvassed in the previous chapter, we assessed whether tactile size misjudgments associated with anorexia nervosa stem from some form of non-perceptual influence (for example, beliefs, emotions, or demand characteristics). To do so, we amended an often-used experimental paradigm for assessing tactile misperception in anorexia nervosa (Engel & Keizer, 2017; Keizer et al., 2011; Keizer et al., 2012). This paradigm requires blindfolded participants to estimate the distance between points of tactile contact on their skin.

In the previous experiments using this paradigm, participants were allowed as much time as desired to judge tactile distances. We reasoned that this feature of the experiments may facilitate influence by non-perceptual factors, as participants allow their thoughts and feelings about their bodies, as well as judgments about the goal of the experiment to influence their estimates of size. Consequently, we modified the paradigm by manipulating the time participants were allowed to provide their estimates. In a *direct* condition, participants were required to estimate the distance immediately after presentation; in a *delayed* condition, participants were required to estimate the distance after a 5 second delay.

When comparing these conditions, we found that while participants with anorexia nervosa overestimated distances (compared to controls) in the delayed condition, they did not do so in the direct condition. This suggests that, contrary to previous assumptions, differences in tactile size judgments in anorexia nervosa do not stem from

differences in perception. Given these findings, it seems unlikely that people with anorexia nervosa misperceive their bodies through touch.

In the previous chapter, I argued that, while it may be difficult, researchers can overcome methodological obstacles and make progress on the question of whether certain populations misperceive the world, as opposed to simply misjudging it. By developing and employing a novel form of tactile distance estimation task, we put this claim to the test, uncovering evidence that tactile distance overestimation in eating disorders does not stem from misperception of tactile size. This approach is in line with the resolutely interdisciplinary approach that I advocate for, and it illustrates the benefits of collaborative research between philosophers and scientists.

4. Waiting longer, feeling fatter: Effects of response delay on tactile distance estimation and confidence in females with anorexia nervosa

Manja M. Engel*, Stephen Gadsby*, Andrew W. Corcoran, Anouk Keizer, H. Chris Dijkerman[1], Jakob Hohwy

*Joint first author

Abstract

Background: Research evidence suggests that anorexia nervosa (AN) patients exhibit differences in the perceptual processing of their own bodies. However, some researchers suggest that these differences are better explained with reference to non-perceptual factors. In this study, we investigated whether overestimation of tactile distances in AN patients results from differences in tactile processing, or non-perceptual factors, by measuring the role of allowed response time in an adapted version of the Tactile Distance Estimation task (TDE-D). We further investigated the relationship between allowed response time and participants' confidence in their tactile judgments.

Method: Our sample consisted of females: AN patients (N = 30), recovered AN (REC) patients (N = 29) and healthy controls (N=31). Participants were asked to estimate tactile distances presented on the skin of either a salient (abdomen) or non-salient (arm) body part, either directly after stimulus presentation (direct condition) or after a 5 second delay (delayed condition). Confidence of estimation accuracy was measured after each response.

Results: Results showed that allowing AN and REC more time to respond caused them to estimate tactile distances as larger than in the direct condition. Additionally, AN patients became less confident when given more time to respond.

Conclusions: These results suggest that attitudinal influences cause AN patients to increase their estimates of tactile distances and become less certain of these estimates. Based on these results, we suggest that previous findings—where AN patients estimate

tactile distances as larger than HC—may be due to attitudinal, rather than perceptual, differences.

1. Introduction

Evidence suggests that anorexia nervosa (AN) patients exhibit differences in perceptual processing, in visual, proprioceptive, and tactile domains. Specifically, the evidence suggests that patients perceptually represent themselves as larger than reality. For example, patients visually estimate their bodies as larger, as well as judging action affordances and moving as if their bodies were larger (Engel & Keizer, 2017; Gadsby, 2017; Guardia et al., 2012; Guardia et al., 2010; Keizer et al., 2013; Metral et al., 2014; Mölbert et al., 2017). These biases are generally considered to be specific to patients' own bodies, as they are not present when patients are required to estimate the size of other bodies (Guardia et al., 2012), or inanimate objects (Engel et al., 2020). Many of these differences have also been found, albeit to lesser degrees, in recovered (REC) patients (Engel & Keizer, 2017).

Apart from differences in visual size estimation, affordance perception, and movement, AN patients appear to perceive touch as if their bodies were larger than reality (Keizer et al., 2011; Keizer et al., 2012; Risso et al., 2020; Spitoni et al., 2015 cf. Mergen et al., 2018). Specifically, research shows that AN patients estimate tactile distances as larger than healthy controls (HC). This difference is more pronounced on body parts that are salient in AN (e.g. abdomen, waist) compared to those that are not (e.g. sternum), and on the horizontal rather than vertical axis (Risso et al., 2020; Spitoni et al., 2015). The standard interpretation of these findings attributes them to perceptual differences between AN patients and HC. For example, one hypothesis claims that, in the case of AN, tactile signals are mapped onto a representation of skin surface which is distorted in ways stereotypical of an overweight body (wider along the thighs and waist). This distortion causes patients to experience distances as wider on salient body parts, along the horizontal axis (Spitoni et al., 2015; Gadsby, 2017; cf. Keizer, et al., 2012). In turn, this research has been used to inform theoretical models of body representation (Gadsby, 2017; Riva, 2014). Correct understanding of these models is important for the

development of effective treatments for AN, and related disorders of body representation.

One relevant issue pertains to the interpretation of behavioural differences between AN and healthy controls on these perceptual judgment tasks. While many assume that overestimation stems from perceptual misrepresentation of the body, an alternative interpretation is that overestimation stems from non-perceptual factors, such as differences in cognitive and evaluative attitudes regarding body size (Smeets, 1997). The possibility problematises findings of tactile overestimation in AN. Rather than perceiving tactile distances to be wider, non-perceptual differences may cause AN patients to *estimate* tactile distances as wider than HC. These factors could take many forms, for example: patients' belief that certain body parts are overweight, their emotions associated with certain body parts (Øverås et al., 2014), or demand characteristics involved in the setup (i.e. patients' beliefs about how the experimenter desires them to perform) (Proctor & Morley, 1986). This undermines the ability of these findings to inform models of body representation and tactile processing.

One fruitful way in which perceptual and attitudinal influences on body size estimation can be experimentally teased apart is through the manipulation of allowed response time. In all previous studies which found differences in tactile distance estimation between AN and HC, responses were self-paced: participants were given as much time as they desired to respond (Spitoni, personal communication). This design allows participants time to reflect on beliefs and emotional attitudes regarding their own body size, which may bias the response in distance judgment (D'Amour & Harris, 2017). Therefore, these findings may confound perceptual with attitudinal influences.

In addition to evidence suggesting that patients misestimate their own body size, multiple strands of evidence suggest that AN patients exhibit low confidence in their body attitudes and perceptions. For example, a number of studies employ semi-structured interviews to investigate the confidence patients exhibit in their beliefs about being overweight (Konstantakopoulos et al., 2012; Mountjoy et al., 2014). These results suggest that many patients exhibit low confidence in relation to such beliefs. Other research

focuses on the confidence patients exhibit in their perception of their own bodily states. These results suggest that patients exhibit low confidence in their perception of interoceptive sensations such as hunger and heartbeat (Fassino et al., 2004; Jenkinson et al., 2018; Kinnaird et al., 2020). However, the confidence patients hold in their tactile experiences of their own bodies has yet to be explored.

This study investigates the extent to which estimates of tactile distances in AN and REC patients are influenced by non-perceptual factors, by investigating the role of allowed response time in tactile size estimation tasks. In this design, participants were asked to estimate tactile distances presented on the skin of either a salient (abdomen) or non-salient (arm) body part, either directly after stimulus presentation (direct condition) or after a 5 second delay (delayed condition). By asking for a direct response, we minimize the opportunity for the cognitive evaluation that can occur during a longer response window (delayed response) (Rubinstein, 2007). We thus test whether tactile size estimation is, in part, influenced by attitudinal factors. We also include a confidence measure to investigate whether between-group differences in estimation confidence could be found—in line with recent proposals regarding eating disorders and confidence in bodily perception and belief.

We expected an influence of delayed response on tactile size estimates in AN patients. Specifically, we expected delayed estimates to be larger in our AN group, and for this effect to be amplified in salient body parts. However, we expected no differences between direct and delayed conditions for the HC and REC groups, as neither of these groups exhibit negative body attitudes of the same severity as AN patients (Engel & Keizer, 2017). Over both delayed and direct conditions, we expected REC patients to estimate tactile distances as larger than HC (albeit smaller than AN), consistent with evidence of persistent perceptual dysfunction within this group.

We further expected that AN patients would report lower confidence ratings on average compared to HC and REC. We also expected group differences in confidence ratings to interact with response delay, such that confidence ratings increase in HC, but decline in AN over time.

2. Method

2.1. Ethics statement

This study was approved by Monash University Human Research Ethics Committee (MUHREC ProjectID: 19265&19131). All participants provided signed informed consent before taking part.

2.2. Preregistration

Experimental hypotheses, methods, and planned analyses were preregistered prior to data collection: https://aspredicted.org/blind.php?x=sf2w3e. We aimed to recruit 30 participants into each group.

2.3. Participants

Participants were females recruited through the Eating Disorders Victoria Facebook page, Twitter, and posters distributed around various higher education institutions in the Melbourne metropolitan area. All participants were compensated for their time.

AN and REC patients were included if they had a present or past diagnosis of AN and REC, respectively, obtained from a psychiatrist or general physician. REC patients were considered recovered if they had successfully completed treatment for their eating disorder and reported that they were no longer in need of treatment. HC were included if they had no history of ED. Self-reported diagnosis was checked with the EDE-Q.

We recruited 98 individuals who fit our inclusion criteria. Eight were excluded as they failed to discriminate two simultaneously presented points on the skin of the arm or abdomen at a distance of 40mm (Weinstein, 1968). For a thorough description of the two-point discrimination task, see Supplementary Material. Our final sample consisted of 30 AN patients, 29 REC patients, and 31 HCs (see Table 1 for demographics).

2.4. Materials and procedure

After providing signed, informed consent, participants completed a demographic questionnaire and the following questionnaires and tasks in the order they are listed below.

2.4.1. Eating Disorder Examination Questionnaire (EDE-Q)

We used the EDE-Q (Fairburn & Beglin, 1994) to asses severity of ED pathology between groups (Aardoom et al., 2012). Global scores in our sample were compared against normative data of a large community sample of Australian adult women (Mond et al., 2004). Higher scores reflect higher levels of eating disorder pathology. The EDE-Q is recommended as a replacement of the EDE interview (Berg et al., 2011; Black & Wilson, 1995; Fairburn & Beglin, Anderson & Williamson, 2002).

2.4.2. Body Attitude Test (BAT)

The BAT (Probst et al., 1995) was used to compare the extent of negative body attitudes between groups. The BAT was developed for female patients suffering from eating disorders, and consists of four subscales. The total score across these scales is indicative of body attitudes, with higher scores reflecting more negative attitudes.

2.4.3. Tactile Distances Estimation Questionnaire (TDE-Q)

The TDE-Q was especially designed for this study to measure individuals' attitudes towards their own body parts. The TDE-Q consists of four questions; two questions tapping the evaluation of the salient (abdomen) and non-salient (arm) body part, another two questions of how this feeling varies over time. For example, regarding the salient body part, participants were asked '*How do you feel about the size of your abdomen?'*, responding on a visual analogue scale (VAS) ranging from '*too thin'* to '*too fat'*. To assess variation over time, they were asked *'Does your evaluation of the size of your abdomen change during the day?'* Again, participants answered on a VAS with '*it is stable'* and '*it varies all the time*' as anchors. Responses were measured on a 0-100 scale. Note, the TDE-Q has not been formally validated.

2.4.4. Tactile Distance Estimation–Delayed (TDE-D)

An adapted version of the TDE task (Keizer et al., 2011) was developed to manipulate the time delay between tactile stimulation and estimation. In this *TDE-D* task, participants were presented with two tactile points applied with a calliper on the skin of the left forearm (Non-salient), on the proximo-distal axis, or the left side of the abdomen (Salient), on the medial-lateral axis. Participants were asked to estimate the distance between these two points by placing the index finger and thumb of their right hand on a tablet. Participants made their estimate when they heard a sound cue (10,000Hz tone). This cue was played directly after stimulus presentation (direct) or 5s post-stimulus (delayed).

Participants were instructed to close their index finger and thumb before each trial, and to respond as soon as they heard the sound. Practice trials were performed until participants responded directly on the audio cue. The right arm was placed on an elevated surface with the wrist above the tablet so that estimations could be made quickly (see figure 1). Distances of 50, 60, and 70mm were presented in a randomized order (total 5 trial repetitions per distance) for each Body-part (salient, non-salient) and Response-delay (direct, delayed) condition. The order of the Response-delay and Body-part condition blocks was counterbalanced across participants. All tactile stimuli were presented for 300ms. The experimenter maintained presentation time consistency by applying stimulation concurrent with a 300ms tone played to them through an earpiece.

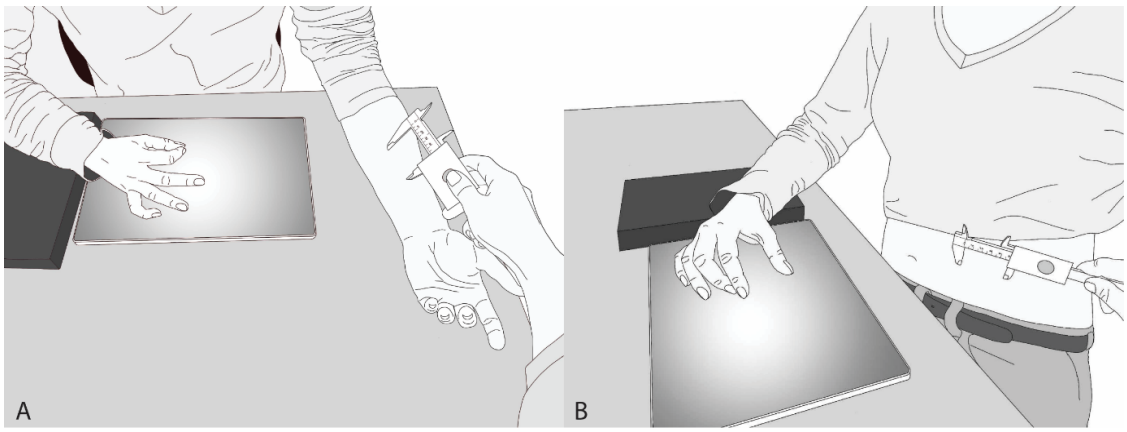


Figure 1. Set up of estimates on arm (A) and abdomen (B). A calliper was used to present distances. The experimenter pressed a button on the calliper that was connected to an earpiece where an audio sound was played for 300ms (duration of stimulus presentation).

2.4.5. Confidence Rating

Our confidence questionnaire was hosted on Gorilla Experiment Builder (www.gorilla.sc). After each distance estimate, participants were asked *'How confident are you that your estimate is correct?'* They responded on a VAS, with 'total guess' anchoring a rating of 0, and 'complete confidence' anchoring a rating of 100.

2.5 Data preparation and analysis

2.5.1. Planned analyses

EDE-Q subscale scores were derived by averaging item scores; EDE-Q global score was calculated from the average of subscale scores. BAT items were summed to derive subscale and total scores. Mixed ANOVAs were conducted in IBM SPSS Statistics for Windows, Version 25.0, to test between-subjects differences in EDE-Q and BAT scores. A mixed ANOVA was used to test for between-group differences in mean VAS scores on each item of the TDE-Q. P-values for planned comparisons were Bonferroni corrected.

TDE-D and Confidence tasks were modelled using mixed ANOVAs, with Group (AN; REC; HC) included as a between-subjects independent variable, and Body-part (arm; abdomen), Distance (50, 60, and 70mm), and Response-delay (direct; delayed) as within-subjects independent variables. In order to facilitate comparison of estimation accuracy across Distance levels, we normalised estimates using the following formula: *Percentage misestimation=(Estimated Distance–Actual Distance)/Actual Distance*100.*

One-tailed planned comparisons were used to test prespecified hypotheses. Missing data were handled by listwise deletion.

Shapiro-Wilk tests and data plots were used to assess normality. Levene's test was used to assess homogeneity of variance. Where this test indicated heteroskedasticity, *Welch's F* was computed (Field, 2009). Mauchly's test was conducted for all mixed ANOVAs. Where this test indicated violation of the sphericity assumption, the Greenhouse-Geisser

correction was used for epsilons ranging .50 –.75, otherwise the Huynh-Feldt correction was used (Field, 2009).

2.5.2. Additional analysis

In addition to our planned analyses, we fit linear mixed-effects models (LMMs) to model distance estimation and confidence at the trial level. These analyses were performed on account of the large degree of individual variation observed within the dataset, and to mitigate loss of information due to listwise deletion of missing data. In order to replicate the structure of our planned ANOVAs, random effects were limited to by-participant random intercepts (more complex models are included in the supplementary materials). The key advantages of this approach over traditional ANOVA are (1) increased power on account of trial-level estimation and (2) 'partial-pooling' of information across individual and group terms. Together, these innovations improve the accuracy and reliability of parameter estimates (Gelman & Hill, 2006).

LMM analyses were conducted in *R* (v3.6.2; R Core Team, 2019) with *RStudio* (v1.2.5033; RStudio Team, 2015). LMMs were fit using the package *lme4* (Bates et al., 2015). Diagnostic plots revealed no evidence of violated assumptions. Although distance can be construed as a random effect, the low number of sampled levels led us to include it as an ordered factor (polynomial contrasts); all remaining factors were unordered and sum-to-zero contrast-coded. Main effect and interaction terms were assessed using Kenward-Roger F tests (Satterthwaite degrees of freedom) from Type-II ANOVA tables obtained from the *car* package (Fox & Weisberg, 2018a). Planned (Bonferroni-corrected) and post-hoc (Tukey-corrected) comparisons were evaluated using the *emmeans* package (Lenth, 2020). *Effects* (Fox & Weisberg, 2018a, 2018b) and *ggplot2* (Wickham, 2016) were used to visualise model predictions (visualisations of response distributions are included in the supplementary materials).

3. Results

3.1. Planned analysis

3.1.1. EDE-Q

A mixed ANOVA showed significant between-group effects for global EDE-Q score, $F(2,87)=26.20, p<.001, \eta^2=.37$. Significant differences were also found for all subscales: Restraint, *Welch's* $F(2,55.42)=25.09, p<.001, \omega=.30$; Eating Concern, *Welch's* $F(2,55.83)=21.81, p<.001, \omega=.30$; Shape Concern, $F(2,87)=18.49, p<.001, \eta^2=.30$; Weight Concern, $F(2,87)=17.04, p<.001, \eta^2=.28$. Post-hoc comparisons indicate that our AN sample demonstrated a higher level of ED psychopathology than the REC and HC groups, which is consistent with the self-reported diagnosis. For means, standard deviations and post-hoc comparisons, see Table 1.

3.1.2. BAT

A mixed ANOVA showed significant between-group differences for total BAT score, $F(2,87)=24.54, p<.001, \eta^2=.36$. Significant differences were also found for all subscales: Negative appreciation with one's own body size, $F(2,87)=18.06, p<.001, \eta^2=.29$; Lack of familiarity with one's own body, $F(2,87)=26.40, p<.001, \eta^2=.38$; General body dissatisfaction, $F(2,87)=12.19, p<.001, \eta^2=.22$. Post-hoc comparisons indicate that, overall, AN and REC patients have more negative body attitudes than HC, and that negative body attitudes (apart from negative appreciation) are higher in AN patients than REC patients. For means, standard deviations and post-hoc comparisons see Table 1.

Table 1.

Demographics, clinical assessment, EDE-Q and BAT scores.

	HC	**REC**	**AN**
	N = 31	*N* = 29	*N* = 30
Demographics			

Age	22.55 ± 3.79	22.76 ± 3.99	22.17 ± 4.14
Age range	19 - 34	19 - 35	18 - 32
BMI	21.06 ± 3.42	21.60 ± 2.74	19.55 ± 3.21
BMI range	16.23 – 32.37	17.26 – 30.11	14.15 – 26.27
Right - Handedness	27	26	28
Reported diagnosis			
AN – Restrictive	--	18	28
AN – Binge/Purge	--	10	1
OSFED	--	1	1
Other diagnoses	--	27	28
OSFED	--	1	1
Other diagnoses	--	27	28
Age: symptom onset	--	13.90 ± 2.58	13.97 ± 3.10
Age: AN diagnosis	--	16.14 ± 2.52	17.23 ± 2.98
Age: start of treatment	--	14.72 ± 5.74	14.47 ± 8.18
Duration treatment	--	2.48 ± 2.79	3.77 ± 3.78
EDE-Q			
Global Score	1.17 ± 0.99	1.86 ± 1.22	**3.33 ± 1.35**
Restraint	0.80 ± 1.03	1.26 ± 1.19	**3.16 ± 1.79**
Eating Concern	0.62 ± 1.02	1.38 ± 1.22	**2.78 ± 1.59**
Shape Concern	1.83 ± 1.26	2.55 ± 1.54	**3.94 ± 1.33**
Weight Concern	1.41 ± 1.32	2.25 ± 1.38	**3.45 ± 1.42**
BAT			
Total score	23.00 ± 16.83	**35.93 ± 13.90**	**47.40 ± 13.33**
Negative appreciation	7.32 ± 6.19	13.14 ± 6.09	**16.80 ± 6.36**
Lack of familiarity	7.68 ± 4.52	**12.28 ± 5.12**	**17.53 ± 6.15**

General dissatisfaction	8.00 ± 4.16	10.52 ± 4.19	**13.07 ± 3.65**

Note. Other diagnoses are other lifetime diagnosis; for a full overview see Table S.1. Duration of treatment is in years (± *SD*). Boldface indicates significant difference from corresponding estimate in the preceding column.

3.1.3. TDE-Q

A mixed ANOVA showed significant between-group differences for evaluation of arm size, $F(2,87)=3.88, p<.05, \eta^2=.08$. Post-hoc comparisons revealed that AN patients evaluate their arm as more fat ($M=60.30$, $SD=22.09$) compared to HC ($M=46.90$, $SD=5.77$) (*p<.05*). No differences were found between AN and REC patients ($M=56.07$, $SD=19.83$) (*p=1.0*), or REC patients and HC (*p=.21*).

Significant differences were also found for change in evaluation of arm size during the day, *Welch's* $F(2,51.93)=4.72, p<.05, \omega=.07$. AN patients ($M=33.47$, $SD=38.15$) report more fluctuation during the day compared to HC ($M=11.74$, $SD=16.73$) (*p<.05*). No significant differences were found between AN and REC patients ($M=18.72$, $SD=25.47$ (*p=.14*), or REC patients and HC (*p=1.0*).

Significant differences were also found for change in evaluation of abdomen size during the day, $F(2,87)=13.65, p<.001, \eta^2=.24$. AN patients ($M= 79.10$, $SD= 28.47$) and REC patients *(*$M=65.59$, $SD=26.98$) report more fluctuation compared to HC ($M=42.74$, $SD=27.03$; *p<.001* and *<.05*, respectively). No differences were found between AN and REC patients (*p=.19*). No significant differences were found for evaluation of abdomen size, $F(2,87)=2.82, p=.065$.

Taken together, these results imply that AN patients rate their arm as fatter compared to REC patients and HC. AN patients also report more change in their evaluation of arm and abdomen size compared to HC. The latter was also apparent in REC patients compared to HC.

3.1.4. TDE-D

Two participants were excluded from analysis due to missing data (technical error). The full ANOVA table is presented in Table 2 (see Table S.2 for descriptive statistics).

AN patients' performance on the TDE-D during the Salient condition did not significantly differ as a function of Response-delay compared to REC patients and HC, $F(1,85)=1.104, p=.296$. REC patients and HC also showed no significant difference in estimation accuracy on this contrast, $F(1,85)=1.950, p=.166$. TDE-D did not significantly differ between groups, $F(1,85)=1.968, p=164$.

3.1.5. Confidence Ratings

Eight participants were excluded from analysis due to missing data (technical error). The full ANOVA table is presented in Table 2 (see Table S.3 for descriptive statistics).

TDE-D confidence ratings did not significantly differ between AN patients and REC patients or HC ($F(1,79)=2.758, p=.101$), nor REC patients and HC ($F(1,79)=0.055, p=.815$). As predicted, AN participants' confidence ratings declined significantly from the Direct to the Delay condition, $F(1,79)=4.194, p<.05$. However, Response-delay did not significantly modulate confidence ratings in HC, $F(1,79)=1.574, p=.213$.

Table 2.

Planned analysis: ANOVAs for TDE-D and Confidence Ratings

Source	*df*	*F*	*p*	*partial* η^2
TDE-D				
Body-part	**1, 85**	**10.68**	**.002**	**.11**
Response-delay	**1, 85**	**14.63**	**< .001**	**.15**
Distance	**1.46, 170[1]**	**26.83**	**< .001**	**.24**
Group	1, 85	2.09	.13	
Group*Distance	**4, 85**	**4.30**	**.002**	**.09**
Group*Body-part	2, 85	1.77	.18	

Group*Response-delay	2, 85	1.64	.20	
Body-part*Response-delay	1, 85	.84	.36	
Body-part*Distance	2, 170	1.81	.17	
Response-delay*Distance	1.90, 170[2]	2.76	.07	
Body-part*Response-delay*Group	2, 85	.53	.59	
Body-part*Distance*Group	2, 85	.55	.70	
Response-delay*Distance*Group	4, 85	.52	.72	
Body-part*Response-delay*Distance	1.91, 170[2]	.42	.65	
Body-part*Response-delay*Distance*Group	8, 85	.42	.80	
Confidence Ratings				
Distance	**1,79**	**13.66**	**< .001**	**.15**
Body-part	1, 79	.024	.876	
Response-delay	1, 79	3.00	.087	
Group	1, 79	1.40	.254	
Distance*Group	**4, 158**	**3.16**	**.016**	**.07**
Distance*Body-part	**1.93, 152.48[2]**	**4.94**	**.009**	**.06**
Group*Body-part	2, 79	.012	.989	
Group*Response-delay	2, 79	2.12	.127	
Body-part*Response-delay	1, 79	.481	.490	
Response-delay*Distance	1.84,145.15[2]	.69	.49	
Body-part*Response-delay*Distance	1.95, 153.9[2]	1.72	.183	
Body-part*Response-delay*Group	2, 79	.25	.777	
Body-part*Distance*Group	4, 158	1.86	.120	
Response-delay*Distance*Group	4, 158	1.97	.101	
Body-part*Response-delay*Distance*Group	4, 158	.87	.486	

Note. [1]Greenhouse-Geisser correction, [2]Huynh-Feldt correction

3.2. Additional analysis

3.2.1. TDE-D

The ANOVA table from the random-intercepts LMM of TDE-D estimates is presented in Table 3 (see Table S.4 for model summary).

TDE-D in AN did not significantly interact with Response-delay and Body-part, $t(5210)=1.04, p=.650$. In the Salient condition, delay-induced changes in estimation performance did not significantly differ between AN and REC patients, $t(5210)=0.40, p=1$, but did significantly differ between REC patients and HC, $t(5210)=6.45, p=.008$. No differences were found in the Non-Salient condition (ps>.09). Modelling Group as a linear trend supported the hypothesis that TDE-D estimates were largest for AN and smallest for HC, $t(93.54)=2.01, p=.048$.

The LMM further revealed significant two-way interactions between Group*Body-part, Group*Response-delay, and Group*Distance. Post-hoc contrasts showed that AN patients and HC underestimated distances less on the abdomen compared to the arm (both ps<.001; see figure 2A). AN and REC patients both estimated distances as larger in the delayed condition compared to the direct condition (ps<.001; see figure 2B), while their distance estimates in the 50mm condition differed significantly from those in the 60mm and 70mm condition (all ps<.001; see figure 2C).

3.1.2. Confidence Ratings

The ANOVA table from the random-intercepts LMM of confidence ratings is presented in Table 3 (see Table S.5 for model summary).

TDE-D confidence ratings did not significantly differ between AN and REC patients, $t(85.9)=1.12, p=.265$, nor REC patients and HC, $t(86)=0.60, p=.553$. As predicted, AN patients reported lower confidence after a delayed response, $t(5038)=5.26, p<.001$. By contrast, HC showed no significant difference in confidence ratings between Direct and Delayed trials, $t(5044)=0.16, p=.44$ (see figure 2E).

The LMM revealed additional significant interactions between Body-part*Distance and Group*Distance. Post-hoc analysis of the Distance*Body-part interaction revealed that participants were more confident about their estimation accuracy on the arm for 70mm compared to 50mm (p=.005). On the abdomen, participants were more confident at 70mm compared to 50mm (p<.001) and at 70 mm compared to 60mm (p<.001; see figure 2D). Contrasts for the Group*Distance interaction revealed that HC were more confident in distance estimates at 70mm compared to 60 mm (p=.005), 70mm compared to 50mm (p<.001), and 60mm compared to 50mm (p=.016; See figure 2F).

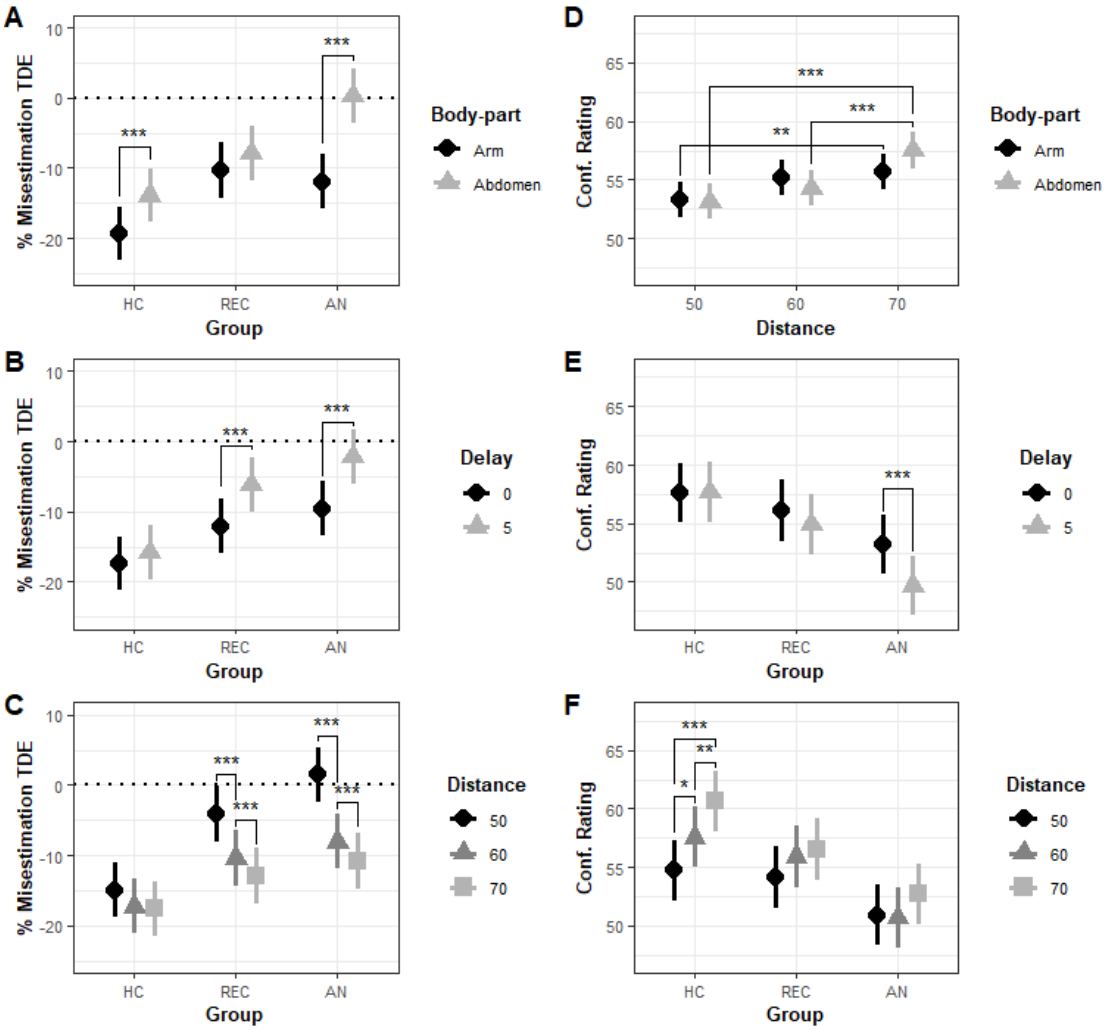


Figure 2. Significant interactions from TDE-D (left column) and Confidence Ratings (right column) LMMs. (A) Group*Body-Part interaction. (B) Group*Delay interaction. (C) Group*Distance interaction. (D) Distance*Body-part interaction. (E) Group*Delay interaction. (F) Group*Distance interaction. *** = p <.001, ** = p <.01, * = p <.05, error bars depict S.E.

Table 3.

Additional analysis: LMMs for TDE-D and Confidence Ratings

Source	*df*	*F*	*p*
TDE-D			
Body-part	**1, 5210.6**	**103.32**	**< .001**
Response-delay	**1, 5210.6**	**54.98**	**< .001**
Distance	**2, 5210.0**	**50.79**	**< .001**
Group	2, 87	2.12	0.126
Group*Body-part	**2, 5210.6**	**18.40**	**< .001**
Group*Response-delay	**2, 5210.6**	**7.03**	**< .001**
Group*Distance	**4, 5210.0**	**6.59**	**< .001**
Body-part*Response-delay	1, 5210.7	1.61	0.205
Body-part*Distance	2, 5210.0	1.65	0.192
Response-delay*Distance	2, 5210.0	2.74	0.064
Group*Body-part*Response-delay	2, 5210.0	0.94	0.39
Group*Body-part*Distance	4, 5210.0	0.57	0.68
Group*Response-delay*Distance	4, 5210.0	0.60	0.66
Body-part*Response-delay*Distance	2, 5210.0	0.38	0.68
Group*Body-part*Response-delay*Distance	4, 5210.0	0.19	0.94
Confidence Ratings			
Distance	**2, 5037**	**26.44**	**< .001**
Response-delay	**1, 5039.7**	**15.96**	**< .001**
Body-part	1, 5041.4	0.46	.497
Group	2, 86.0	1.57	.215
Body-part*Distance	**2, 5037**	**4.55**	**.011**
Group*Response-delay	**2, 5039.8**	**7.52**	**< .001**
Group*Distance	**4, 5037**	**4.33**	**< .001**
Group*Body-part	2, 5041.4	0.74	.477
Body-part*Response-delay	1, 5039.8	0.23	.633
Response-delay*Distance	2, 5037	1.87	.155

Group*Body-part*Response-delay	2, 5039.8	0.05	.949
Group*Body-part*Distance	4, 5037	0.99	.410
Group*Response-delay*Distance	4, 5037	1.28	.277
Body-part*Response-delay*Distance	2, 5037	0.82	.441
Group*Body-part*Response-delay*Distance	4, 5037	0.60	.665

Finally, distance estimates and confidence ratings were correlated with clinical variables (duration of illness, treatment, BMI, etc.). No significant results survived Bonferroni correction (alpha = .003).

4. Discussion

A much-debated issue in AN research is whether overestimation of body size reflects a difference in perceptual processing, or stems from non-perceptual factors, such as the cognitive and affective attitudes patients hold towards their own bodies (Smeets, 1997). This study aimed to investigate the comparative influence of perceptual and non-perceptual factors on tactile size estimation in AN and REC patients, by manipulating response delay. Here, a delayed response allowed participants to deliberate, increasing the influence of non-perceptual factors (e.g. attitudes and emotions) on their final estimation, while a direct response did not include this deliberation period, more directly reflecting their perception of the distance. Additionally, we investigated if this response delay affects confidence in the accuracy of these judgments.

We expected a longer response time would increase tactile distance estimates in AN patients, especially in salient body parts. We also expected REC patients to estimate distances as larger than HC, but smaller than AN patients. While our planned analyses indicated that tactile distance estimation varied as a function of both body-part and response-delay, group-level interactions were only revealed when we modelled participant responses on the trial-level, using linear mixed-effects modelling (LMM). The results of this additional analysis revealed that AN and REC patients, but not HC, judge distances as larger when responding after a 5s delay period. Further, tactile distances were judged as

larger on the forearm compared to the abdomen in AN patients and HC (but not REC). Contrary to our expectations, response-delay did not interact with body part salience.

In terms of estimation confidence, we expected AN patients to report lower confidence than HC and REC. Additionally, we expected AN patients to be less confident in the delayed, compared to the direct condition, while we expected the opposite for HC. Although our ANOVA failed to provide statistical evidence of these effects, the additional sensitivity afforded by the LMM furnished partial support: AN patients reported lower confidence in their estimation accuracy in the delayed condition, compared to the direct condition. Contrary to our expectations, we did not see an influence of response-delay on confidence ratings in REC or HC.

Our findings are consistent with the claim that increased tactile distance estimations in AN patients (relative to HC) are due to non-perceptual factors. In other words, AN patients estimate tactile distances as larger not because they perceive them as larger, but due to processes that occur after the perceptual experience. Previous research suggested that either attitudes about body size (e.g. beliefs about being overweight) or demand characteristics (beliefs about what the experimenter expects) may cause AN patients to overestimate their bodies. Such explanations predict that AN patients would only increase their estimates (when given more time to respond) in the salient (abdomen, medio-lateral) as opposed to non-salient (arm, proximo-distal) condition. This is because patients do not hold false beliefs about their arm length and would not assume the experimenter expected them to misestimate those dimensions. Our results contradict this prediction, as when looking at the AN group, there is no delay*body-part interaction (see: fig. S.7). Given this finding, whichever non-perceptual factors drive misestimation, they must affect both body-part conditions. Explanations of this finding may refer to changes in the memory of the touch experience (Williamson et al., 2004); or judgment biases driven by time to reflect on anxiety (Øverås et al., 2014) or task difficulty (Waller & Hodgson, 1996). While our data cannot distinguish between the various possibilities, this represents a promising avenue for future research.

Another novel finding of the study stems from our use of the TDE-Q, designed to measure fluctuations in bodily attitudes. The results of this questionnaire suggest AN patients show a higher fluctuation in their evaluative attitudes towards their arms and abdomens, compared to HC. Interestingly, this measure also discovered that REC patients' evaluation of their own abdomen size fluctuates significantly more compared to HC. These daily fluctuations in evaluation of body size are additional indicators of uncertainty in the body size attitudes of AN and REC patients (Espeset et al., 2011).

In further support of a difference in confidence between AN, REC, and HC—we found that while HC were more confident for larger compared to smaller distances, this was not the case for AN and REC, who exhibited no between-distance differences in confidence ratings. The results from our HC group can be explained with reference to perceptual ambiguity: smaller tactile distances are closer to the 2-point discrimination threshold and are thus more perceptually ambiguous, causing participants to be less confident in such judgments. While it is not clear how to explain the homogeneity of confidence judgments in AN and REC patients, future research might clarify this issue by using more precise confidence measures (Matthews et al., 2020).

These findings bear significant theoretical relevance, especially related to ongoing debates regarding models of body representation. Such debates focus on the number of body representations, their functional relationships with one another, and the kinds of behavioural capacities they underpin (de Vignemont, 2017; Gadsby, 2018; 2019; Longo, 2015; Pitron & de Vignemont, 2018; Pitron et al., 2018). One proposal within these debates is that multiple body representations (including those that underpin tactile size perception) all derive their spatial content from a shared representation (Gadsby, 2017). This finding is largely motivated by evidence that, in the case of AN, each representation is distorted (as evident in visual size estimate, affordance processing, and tactile distance estimation tasks). However, the finding that overestimation of tactile distances by AN patients (compared to HC) may reflect non-perceptual disturbances undermines this justification for the model.

In order to test and replicate previous findings, we adopted a TDE paradigm employed in (Keizer et al., 2011; Keizer et al., 2012; Engel & Keizer, 2017). This introduces an important caveat to our interpretation. While our results suggest that AN and REC overestimate due to non-perceptual factors, this interpretation may be specific to the paradigm used here. For example, in visual body size estimate research, the type of measure employed has been found to modulate body size estimates (Cornelissen et al., 2017). Some studies which found differences in TDE between AN and HC, used alternative methods (Risso et al., 2020; Spitoni et al., 2015). These methods may be more suited to capturing perceptual differences between these groups (see also: Tosi & Romano, 2020). Accordingly, future research should further investigate our claims, employing various methods of assessing tactile size perception.

We acknowledge that the critical interactions between group and response-time were uncovered in the unplanned component of our analysis. While such findings are usually caveated with the need for cautious interpretation, it should be stressed that our LMMs aimed to emulate the ANOVAs specified in our preregistered analysis with greater power and precision. Although LMMs are commonly regarded in psychology as more complex and difficult to interpret than traditional methods (Meteyard & Davies, 2020), our findings showcase the advantages of such techniques in the context of substantial individual variation and unbalanced data—common occurrences in clinical research settings.

In summary, we showed that allowing AN and REC patients more time to respond during tactile distance estimation causes them to make larger distance estimates. This finding lends support for a non-perceptual explanation of tactile distance overestimation in AN, suggesting that previous findings that AN patients estimate tactile distances as larger than HC may be due to non- perceptual, differences. We also discovered that, in contrast to HC and REC, AN patients became less confident when given more time to respond, contributing of differences in confidence associated with the disorder.

Supplementary Material

Description: Two Point Discrimination task

The TPD task was used to assess the minimum spatial distance needed between two simultaneously presented tactile points on the skin for participants to report feeling two distinct tactile stimuli (Lundborg & Rosén, 2004; Weinstein, 1968). Participants were asked to close their eyes and were presented with one or two tactile points by a calliper. These tactile points presented on the skin of the inside of the left forearm and left side of the abdomen. Participants were asked if they felt one or two points on the skin. In accordance with Weinstein (1968), the starting point was 37mm on the forearm and 33mm on the abdomen. For responses, the forced choice one up, two down staircase method was used with 5 reversals. In each trial either one (33% of the trials) or two (66%) of the trials) tactile stimuli were presented with a calliper. Only responses to the two tactile stimuli trials were used. The TPD threshold was calculated as the average of the last five correct responses. The order of the body parts was counterbalanced. It is worth noting that the reliability of this standard TPD task has been criticized due to a number of factors (Craig & Johnson, 2000), therefore for assessing basic tactile perception researchers should employ alternative methods (Tong et al., 2013).

Table S.1.
Demographic and clinical information AN and REC patients.

Participant	AN type	Other diagnoses
1-AN	Restrictive	
2-AN	Restrictive	
3-AN	Restrictive	
4-AN	Restrictive	Depression, anxiety
5-AN	Restrictive	Depression, anxiety
6-AN	Restrictive	Depression, anxiety
7-AN	Restrictive	Depression, anxiety
8-AN	Restrictive	Depression, anxiety
9-AN	Restrictive	Depression, anxiety
10-AN	Restrictive	Depression, anxiety
11-AN	Restrictive	Depression, anxiety
12-AN	Binge/Purge	Depression, anxiety
13-AN	Restrictive	Depression, anxiety
14-AN	Restrictive	Depression, anxiety

15-AN	Restrictive	Depression, anxiety
16-AN	Restrictive	Depression, anxiety
17-AN	Restrictive	Depression, anxiety
18-AN	Restrictive	Depression, anxiety
19-AN	Restrictive	Depression, anxiety
20-AN	Restrictive	Depression, anxiety
21-AN	OSFED	Depression, anxiety
22-AN	Restrictive	Depression, anxiety
23-AN	Restrictive	Depression, anxiety
24-AN	Restrictive	Depression, anxiety
25-AN	Restrictive	Depression, anxiety
26-AN	Restrictive	Depression, anxiety
27-AN	Restrictive	Depression, anxiety
28-AN	Restrictive	Depression, anxiety
29-AN	Restrictive	Depression, anxiety
30-AN	Restrictive	Depression, anxiety
1-REC	OSFED	Bulimia
2-REC	Binge/Purge	
3-REC	Binge/Purge	
4-REC	Restrictive	
5-REC	Restrictive	
6-REC	Binge/Purge	Depression, anxiety
7-REC	Binge/Purge	Depression, anxiety
8-REC	Binge/Purge	Anxiety
9-REC	Restrictive	OCD traits, anxiety, ADHD
10-REC	Binge/Purge	
11-REC	Restrictive	Borderline, depression, anxiety, aspergers
12-REC	Binge/Purge	Bulimia, depression
13-REC	Restrictive	
14-REC	Binge/Purge	Depression, anxiety
15-REC	Restrictive	OCD
16-REC	Restrictive	
17-REC	Binge/Purge	Binge eating, depression, anxiety
18-REC	Restrictive	Borderline, anxiety, depression, PTSD
19-REC	Restrictive	Depression, anxiety
20-REC	Restrictive	Anxiety
21-REC	Restrictive	
22-REC	Restrictive	Anxiety
23-REC	Restrictive	Depression, anxiety, chronic fatigue, fibromyalgia
24-REC	Restrictive	Depression, anxiety, insomnia
25-REC	Restrictive	
26-REC	Restrictive	OCD

27-REC	Binge/Purge	
28-REC	Restrictive	Depression
29-REC	Restrictive	Depression, anxiety

Table S.2.
Descriptives of the TDE

TDE	**HC** *N* = 29	**REC** *N* = 29	**AN** *N* = 30
Arm			
Direct 50mm	-21.69 ± 20.14	-9.13 ± 30.10	-7.24 ± 26.67
Direct 60mm	-21.74 ± 18.54	-14.48 ± 28.81	-19.68 ± 23.17
Direct 70mm	-23.57 ± 18.46	-16.11 ± 27.61	-21.12 ± 21.83
Delayed 50mm	-14.42 ± 20.64	-1.10 ± 30.73	-0.21 ± 24.90
Delayed 60mm	-19.22 ± 19.00	-8.17 ± 25.20	-10.75 ± 19.62
Delayed 70mm	-19.87 ± 20.12	-13.10 ± 24.70	-13.49 ± 20.44
Abdomen			
Direct 50mm	-13.54 ± 27.10	-8.08 ± 32.22	1.40 ± 41.10
Direct 60mm	-14.15 ± 23.39	-12.75 ± 28.02	-4.65 ± 34.12
Direct 70mm	-12.35 ± 21.96	-11.73 ± 28.66	-6.65 ± 31.71
Delayed 50mm	-13.55 ± 26.97	2.05 ± 32.81	10.41 ± 34.82
Delayed 60mm	-14.86 ± 22.61	-6.34 ± 23.98	2.54 ± 31.11
Delayed 70mm	-14.36 ± 22.14	-10.48 ± 22.62	-2.11 ± 29.10

Table S.3.
Descriptives of the Confidence Ratings

Confidence	**HC** *N* = 26	**REC** *N* = 29	**AN** *N* = 27
Arm			
Direct 50mm	53.87 ± 16.67	54.02 ± 9.34	52.64 ± 17.87
Direct 60mm	57.29 ± 15.01	56.53 ± 13.27	53.26 ± 17.08
Direct 70mm	59.16 ± 14.27	56.31 ± 12.41	51.88 ± 18.05
Delayed 50mm	53.73 ± 15.36	52.84 ± 12.53	50.10 ± 20.81
Delayed 60mm	57.75 ± 15.48	60.86 ± 22.03	46.50 ± 19.11
Delayed 70mm	58.82 ± 15.15	54.63 ± 11.04	50.46 ± 17.92
Abdomen			
Direct 50mm	52.77 ± 17.00	55.72 ± 14.57	50.51 ± 18.69
Direct 60mm	56.81 ± 17.71	56.66 ± 14.02	54.14 ± 18.48
Direct 70mm	60.60 ± 17.97	57.79 ± 14.47	55.28 ± 16.17
Delayed 50mm	54.96 ± 19.12	53.97 ± 15.60	48.77 ± 17.42
Delayed 60mm	55.19 ± 17.16	54.19 ± 16.26	46.80 ± 17.55
Delayed 70mm	60.36 ± 16.00	57.36 ± 12.10	51.51 ± 19.17

Table S.4.
Linear mixed-model summary: TDE-D Percentage misestimation (REC dom intercepts)

Predictors	*Estimates*	*S.E.*	*95% CI*	*Statistic*	*p-value*
(Intercept)	-10.52	2.22	-14.93 – -6.12	-4.75	**<0.001**
Group [AN]	4.70	3.13	-1.53 – 10.92	1.50	0.137
Group [HC]	-6.10	3.11	-12.28 – 0.07	-1.96	0.053
BodyPart [Abdomen]	3.36	0.33	2.70 – 4.01	10.09	**<0.001**
Delay [0]	-2.49	0.33	-3.14 – -1.84	-7.48	**<0.001**
Distance.L	-5.59	0.58	-6.72 – -4.46	-9.70	**<0.001**
Distance.Q	1.71	0.58	0.58 – 2.84	2.96	**0.003**
Group [AN] * BodyPart [Abdomen]	2.72	0.47	1.80 – 3.65	5.80	**<0.001**
Group [HC] * BodyPart [Abdomen]	-0.60	0.47	-1.52 – 0.32	-1.29	0.198
Group [AN] * Delay [0]	-1.23	0.47	-2.15 – -0.31	-2.62	**0.009**
Group [HC] * Delay [0]	1.68	0.47	0.77 – 2.60	3.60	**<0.001**
BodyPart [Abdomen] * Delay [0]	0.41	0.33	-0.24 – 1.07	1.24	0.215
Group [AN] : Distance.L	-3.08	0.81	-4.67 – -1.48	-3.79	**<0.001**
Group [HC] : Distance.L	3.74	0.81	2.15 – 5.33	4.62	**<0.001**
Group [AN] : Distance.Q	1.06	0.81	-0.53 – 2.66	1.31	0.192
Group [HC] : Distance.Q	-0.90	0.81	-2.48 – 0.69	-1.11	0.268
BodyPart [Abdomen] : Distance.L	1.02	0.58	-0.11 – 2.15	1.77	0.077
BodyPart [Abdomen] : Distance.Q	-0.21	0.58	-1.34 – 0.92	-0.37	0.712
Delay [0] : Distance.L	1.36	0.58	0.23 – 2.49	2.37	**0.018**
Delay [0] : Distance.Q	-0.01	0.58	-1.14 – 1.12	-0.01	0.988

(Group [AN] * BodyPart [Abdomen]) * Delay [0]	-0.15	0.47	-1.07 – 0.77	-0.32	0.746
(Group [HC] * BodyPart [Abdomen]) * Delay [0]	0.62	0.47	-0.30 – 1.54	1.32	0.187
Group [AN] : BodyPart [Abdomen] : Distance.L	0.24	0.81	-1.35 – 1.84	0.30	0.764
Group [HC] : BodyPart [Abdomen] : Distance.L	0.34	0.81	-1.25 – 1.93	0.42	0.675
Group [AN] : BodyPart [Abdomen] : Distance.Q	-1.08	0.81	-2.68 – 0.51	-1.33	0.184
Group [HC] : BodyPart [Abdomen] : Distance.Q	0.48	0.81	-1.11 – 2.07	0.59	0.554
Group [AN] : Delay [0] : Distance.L	-0.65	0.81	-2.24 – 0.95	-0.80	0.426
Group [HC] : Delay [0] : Distance.L	-0.44	0.81	-2.03 – 1.15	-0.55	0.586
Group [AN] : Delay [0] : Distance.Q	0.39	0.81	-1.21 – 1.98	0.47	0.636
Group [HC] : Delay [0] : Distance.Q	-0.64	0.81	-2.23 – 0.95	-0.79	0.428
BodyPart [Abdomen] : Delay [0] : Distance.L	0.50	0.58	-0.63 – 1.63	0.87	0.386
BodyPart [Abdomen] : Delay [0] : Distance.Q	0.09	0.58	-1.04 – 1.22	0.16	0.876
Group [AN] : BodyPart [Abdomen] : Delay [0] : Distance.L	0.38	0.81	-1.21 – 1.98	0.47	0.640
Group [HC] : BodyPart [Abdomen] : Delay [0] : Distance.L	-0.53	0.81	-2.12 – 1.06	-0.65	0.513

Group [AN] : BodyPart [Abdomen] : Delay [0] : Distance.Q	-0.32	0.81	-1.91 – 1.28	-0.39	0.697
Group [HC] : BodyPart [Abdomen] : Delay [0] : Distance.Q	0.44	0.81	-1.15 – 2.02	0.54	0.589

Random Effects

σ^2	589.22
$\tau_{00\ Ppn}$	431.53
ICC	0.42
$N_{\ Ppn}$	90
Observations	5333
Marginal R^2 / Conditional R^2	0.055 / 0.455

Table S.5.
Linear mixed-model summary: TDE-D Confidence rating (random intercepts)

Predictors	*Estimates*	*S.E.*	*95% CI*	*Statistic*	*p-value*
(Intercept)	54.86	1.47	51.94 – 57.79	37.28	**<0.001**
Group [AN]	-3.42	2.08	-7.55 – 0.70	-1.65	0.103
Group [HC]	2.79	2.08	-1.34 – 6.91	1.34	0.183
BodyPart [Abdomen]	0.13	0.19	-0.25 – 0.50	0.65	0.516
Delay [0]	0.76	0.19	0.38 – 1.14	3.96	**<0.001**
Distance.L	2.40	0.33	1.75 – 3.05	7.24	**<0.001**
Distance.Q	0.17	0.33	-0.48 – 0.82	0.50	0.615
Group [AN] * BodyPart [Abdomen]	-0.05	0.27	-0.59 – 0.48	-0.20	0.842
Group [HC] * BodyPart [Abdomen]	-0.26	0.27	-0.79 – 0.28	-0.93	0.350
Group [AN] * Delay [0]	0.99	0.27	0.46 – 1.52	3.64	**<0.001**
Group [HC] * Delay [0]	-0.81	0.27	-1.35 – -0.28	-2.99	**0.003**
BodyPart [Abdomen] * Delay [0]	0.09	0.19	-0.29 – 0.47	0.47	0.637
Group [AN] : Distance.L	-1.09	0.47	-2.01 – -0.17	-2.33	**0.020**
Group [HC] : Distance.L	1.78	0.47	0.86 – 2.70	3.79	**<0.001**
Group [AN] : Distance.Q	0.73	0.47	-0.19 – 1.64	1.56	0.120
Group [HC] : Distance.Q	-0.06	0.47	-0.98 – 0.86	-0.13	0.894
BodyPart [Abdomen] : Distance.L	0.72	0.33	0.07 – 1.37	2.17	**0.030**

BodyPart [Abdomen] : Distance.Q	0.70	0.33	0.05 – 1.35	2.10	**0.035**
Delay [0] : Distance.L	0.10	0.33	-0.55 – 0.75	0.29	0.771
Delay [0] : Distance.Q	-0.63	0.33	-1.28 – 0.02	-1.90	0.057
(Group [AN] * BodyPart [Abdomen]) * Delay [0]	-0.03	0.27	-0.56 – 0.50	-0.10	0.921
(Group [HC] * BodyPart [Abdomen]) * Delay [0]	-0.06	0.27	-0.59 – 0.48	-0.22	0.828
Group [AN] : BodyPart [Abdomen] : Distance.L	0.42	0.47	-0.50 – 1.34	0.90	0.370
Group [HC] : BodyPart [Abdomen] : Distance.L	0.08	0.47	-0.84 – 1.00	0.16	0.871
Group [AN] : BodyPart [Abdomen] : Distance.Q	-0.76	0.47	-1.67 – 0.16	-1.61	0.106
Group [HC] : BodyPart [Abdomen] : Distance.Q	0.27	0.47	-0.65 – 1.19	0.58	0.563
Group [AN] : Delay [0] : Distance.L	-0.08	0.47	-1.00 – 0.84	-0.17	0.864
Group [HC] : Delay [0] : Distance.L	0.30	0.47	-0.62 – 1.22	0.63	0.528
Group [AN] : Delay [0] : Distance.Q	-1.01	0.47	-1.92 – -0.09	-2.15	**0.031**

Group [HC] : Delay [0] : Distance.Q	0.36	0.47	-0.56 – 1.28	0.76	0.447
BodyPart [Abdomen] : Delay [0] : Distance.L	0.17	0.33	-0.48 – 0.82	0.51	0.608
BodyPart [Abdomen] : Delay [0] : Distance.Q	-0.39	0.33	-1.04 – 0.26	-1.18	0.237
Group [AN] : BodyPart [Abdomen] : Delay [0] : Distance.L	0.37	0.47	-0.55 – 1.29	0.79	0.430
Group [HC] : BodyPart [Abdomen] : Delay [0] : Distance.L	0.14	0.47	-0.78 – 1.06	0.31	0.760
Group [AN] : BodyPart [Abdomen] : Delay [0] : Distance.Q	0.48	0.47	-0.43 – 1.40	1.03	0.302
Group [HC] : BodyPart [Abdomen] : Delay [0] : Distance.Q	-0.32	0.47	-1.24 – 0.60	-0.68	0.495

Random Effects

σ^2	188.61
$\tau_{00\ Ppn}$	189.34
ICC	0.50
N_{Ppn}	89
Observations	5159
Marginal R^2 / Conditional R^2	0.029 / 0.516

Table S.6. Pearson correlations between TDE-Q and the TDE-D.

		Feeling about fore arm	Change evaluation forearm	Feeling about abdomen	Change evaluation abdomen
AN patients					
Mean arm direct	r.	0.152	0.106	0.167	-0.041
	p	0.424	0.578	0.378	0.829
Mean arm delayed	r.	0.271	0.101	0.247	0.093
	p	0.148	0.596	0.189	0.625
Mean abdomen direct	r.	0.143	0.243	0.144	0.214
	p	0.450	0.195	0.449	0.256
Mean abdomen delayed	r.	0.244	0.317	0.215	0.211
	p	0.195	0.088	0.255	0.263
REC patients					
Mean arm direct	r.	0.218	-0.139	0.162	0.031
	p	0.257	0.473	0.402	0.875
Mean arm delayed	r.	0.119	-0.117	0.142	0.025
	p	0.538	0.547	0.461	0.896
Mean abdomen direct	r.	**406***	-0.025	0.298	0.249
	p	**0.029**	0.898	0.116	0.192
Mean abdomen delayed	r.	0.209	0.101	0.124	0.209
	p	0.277	0.603	0.522	0.277
HC					
Mean arm direct	r.	0.273	-0.002	0.009	0.092
	p	0.144	0.991	0.961	0.628
Mean arm delayed	r.	0.285	0.121	0.016	-0.059
	p	0.127	0.523	0.934	0.758
Mean abdomen direct	r.	0.131	0.029	-0.227	0.049
	p	0.482	0.877	0.220	0.795
Mean abdomen delayed	r.	0.120	0.144	-0.251	-0.034
	p	0.519	0.439	0.172	0.854

Note. Critical p value is .006

Visualisations of Raw Data Distributions

Figure S.1.

*Distribution of TDE-D percentage misestimation: Group * BodyPart*

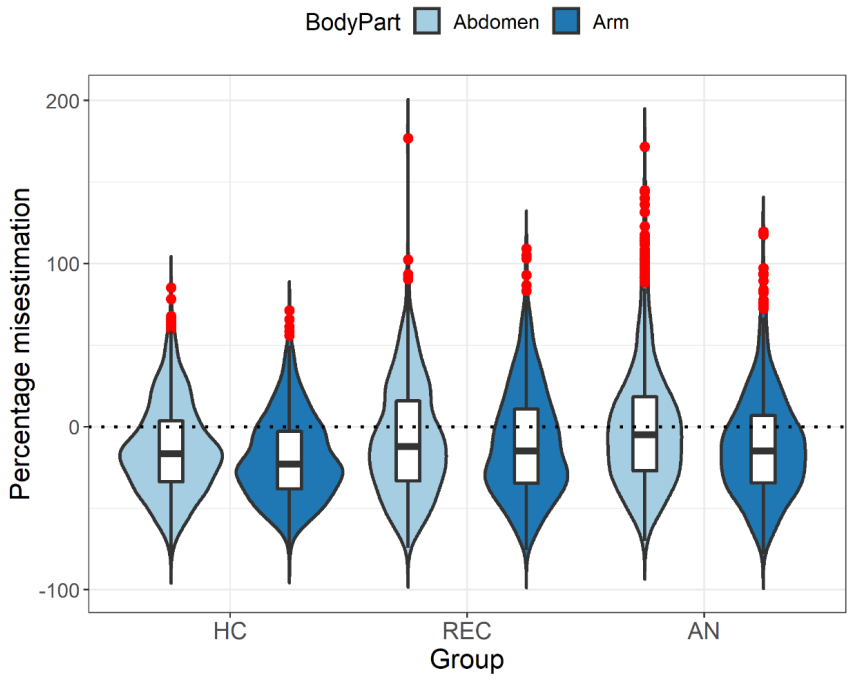


Figure S.2.

*Distribution of TDE-D percentage misestimation: Group * Delay*

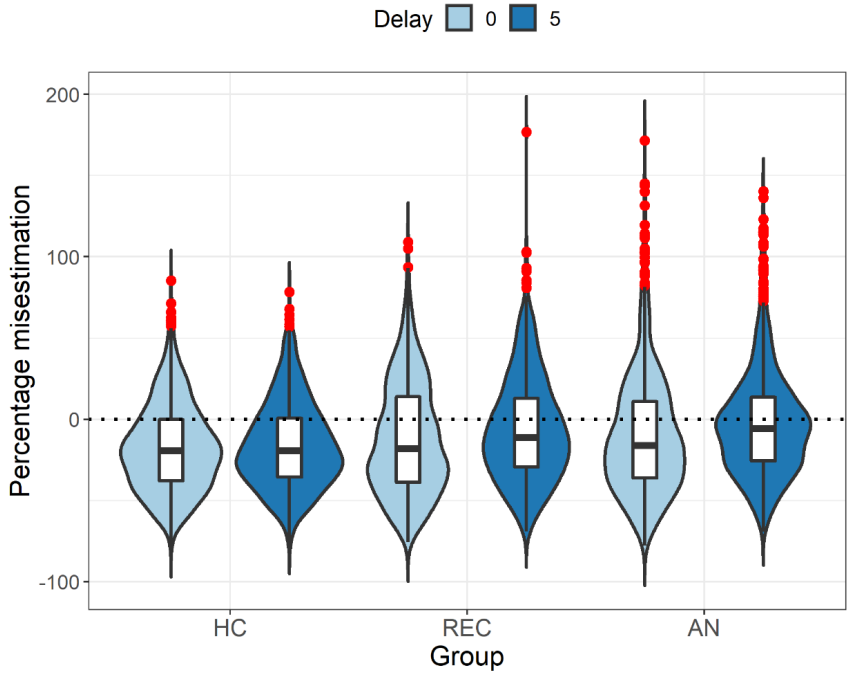


Figure S.3.
*Distribution of TDE-D percentage misestimation: Group * Distance*

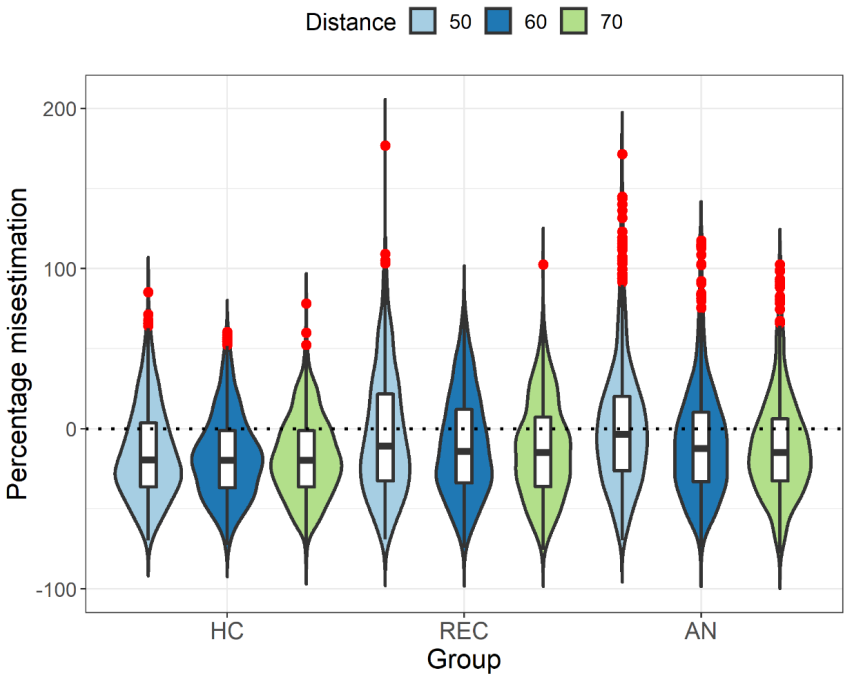


Figure S.4.
*Distribution of TDE-D confidence rating: Group * BodyPart*

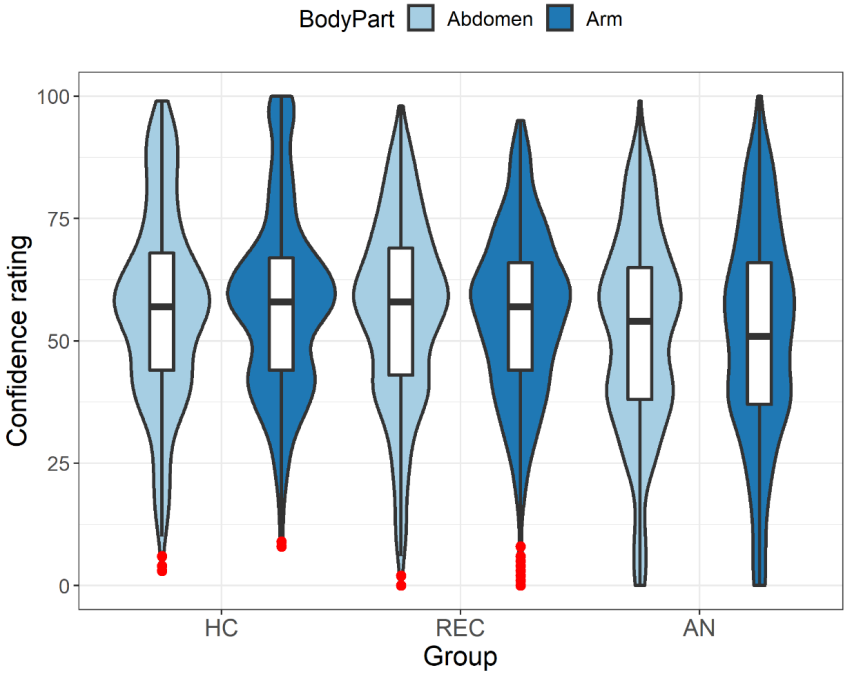


Figure S.5.
*Distribution of TDE-D confidence rating: Group * Delay*

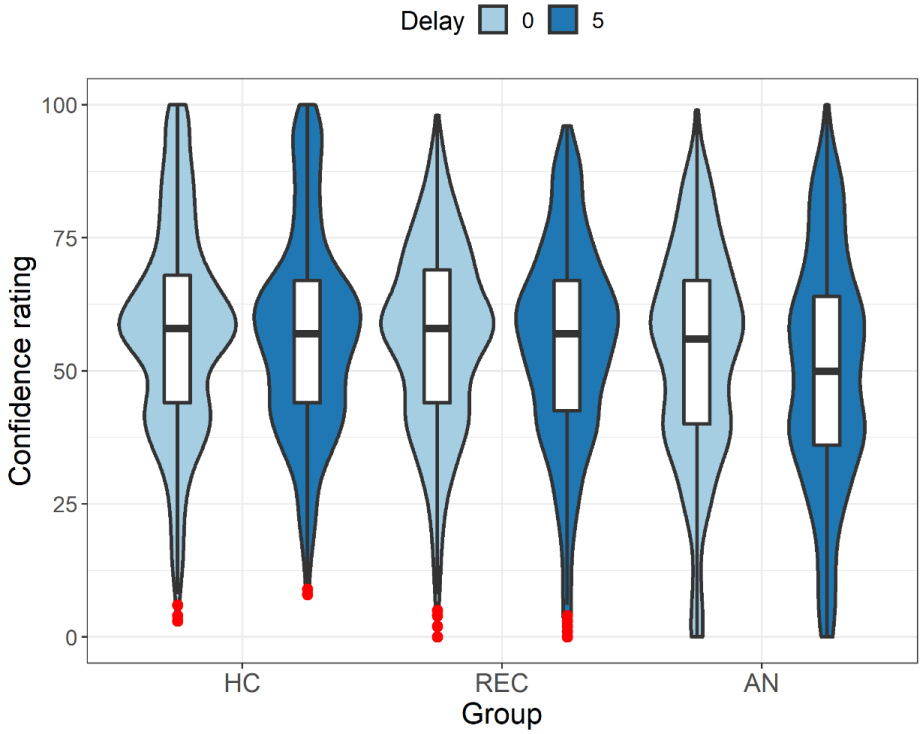


Figure S.6.
*Distribution of TDE-D confidence rating: Group * Distance*

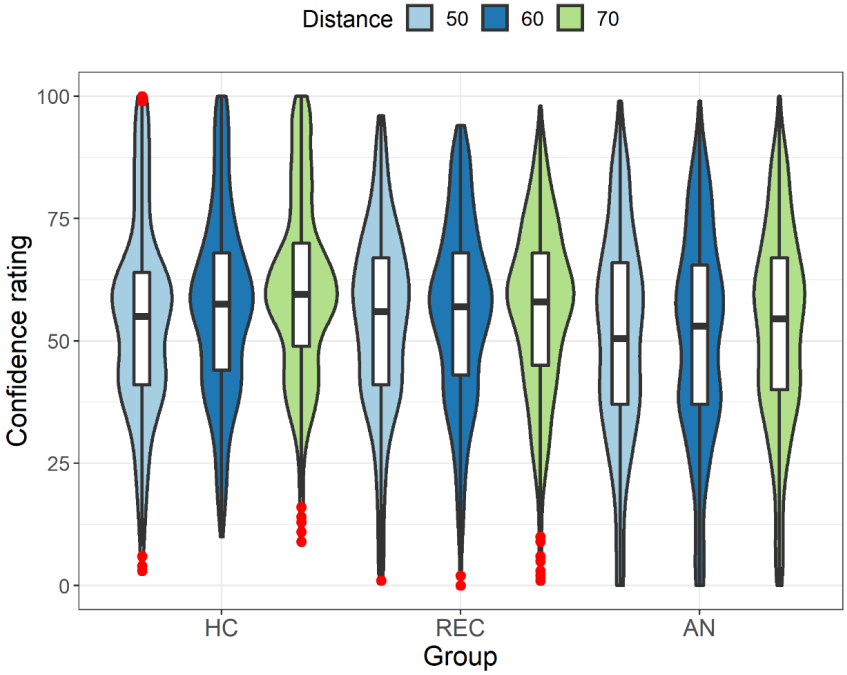


Table S7.
Linear mixed-model summary: TDE-D Percentage misestimation (random slopes)

	Percentage misestimation				
Predictors	*Estimates*	*S.E.*	*95% CI*	*Statistic*	*p-value*
(Intercept)	-10.49	2.22	-14.90 – -6.08	-4.73	**<0.001**
Group [AN]	4.64	3.14	-1.60 – 10.88	1.48	0.143
Group [HC]	-6.06	3.11	-12.25 – 0.12	-1.95	0.055
BodyPart [Abdomen]	3.32	1.06	1.22 – 5.42	3.14	**0.002**
Delay [0]	-2.49	0.65	-3.78 – -1.20	-3.83	**<0.001**
Distance.L	-5.56	0.50	-6.54 – -4.57	-11.09	**<0.001**
Distance.Q	1.69	0.50	0.71 – 2.67	3.38	**0.001**
Group [AN] * BodyPart [Abdomen]	2.75	1.49	-0.22 – 5.72	1.84	0.069
Group [HC] * BodyPart [Abdomen]	-0.67	1.48	-3.62 – 2.28	-0.45	0.652
Group [AN] * Delay [0]	-1.22	0.92	-3.04 – 0.60	-1.33	0.186
Group [HC] * Delay [0]	1.62	0.91	-0.20 – 3.43	1.77	0.080
BodyPart [Abdomen] * Delay [0]	0.46	0.57	-0.67 – 1.60	0.81	0.420
Group [AN] : Distance.L	-3.07	0.71	-4.45 – -1.68	-4.33	**<0.001**
Group [HC] : Distance.L	3.74	0.70	2.36 – 5.12	5.31	**<0.001**
Group [AN] : Distance.Q	1.07	0.71	-0.31 – 2.46	1.52	0.129
Group [HC] : Distance.Q	-0.89	0.70	-2.27 – 0.49	-1.27	0.205
BodyPart [Abdomen] : Distance.L	1.04	0.50	0.06 – 2.02	2.08	**0.038**
BodyPart [Abdomen] : Distance.Q	-0.21	0.50	-1.19 – 0.77	-0.42	0.678
Delay [0] : Distance.L	1.37	0.50	0.39 – 2.35	2.73	**0.006**
Delay [0] : Distance.Q	-0.05	0.50	-1.04 – 0.93	-0.11	0.914

(Group [AN] * BodyPart [Abdomen]) * Delay [0]	-0.18	0.81	-1.78 – 1.43	-0.22	0.828
(Group [HC] * BodyPart [Abdomen]) * Delay [0]	0.66	0.81	-0.94 – 2.26	0.82	0.413
Group [AN] : BodyPart [Abdomen] : Distance.L	0.23	0.71	-1.16 – 1.62	0.33	0.744
Group [HC] : BodyPart [Abdomen] : Distance.L	0.34	0.70	-1.04 – 1.72	0.48	0.632
Group [AN] : BodyPart [Abdomen] : Distance.Q	-1.09	0.71	-2.47 – 0.30	-1.54	0.124
Group [HC] : BodyPart [Abdomen] : Distance.Q	0.47	0.70	-0.91 – 1.85	0.67	0.506
Group [AN] : Delay [0] : Distance.L	-0.66	0.71	-2.05 – 0.73	-0.94	0.349
Group [HC] : Delay [0] : Distance.L	-0.45	0.70	-1.83 – 0.93	-0.64	0.522
Group [AN] : Delay [0] : Distance.Q	0.41	0.71	-0.98 – 1.80	0.58	0.563
Group [HC] : Delay [0] : Distance.Q	-0.62	0.70	-2.00 – 0.76	-0.88	0.380
BodyPart [Abdomen] : Delay [0] : Distance.L	0.50	0.50	-0.48 – 1.49	1.01	0.314
BodyPart [Abdomen] : Delay [0] : Distance.Q	0.07	0.50	-0.92 – 1.05	0.13	0.896
Group [AN] : BodyPart [Abdomen] : Delay [0] : Distance.L	0.35	0.71	-1.03 – 1.74	0.50	0.616
Group [HC] : BodyPart [Abdomen] : Delay [0] : Distance.L	-0.55	0.70	-1.93 – 0.83	-0.78	0.436
Group [AN] : BodyPart [Abdomen] : Delay [0] : Distance.Q	-0.31	0.71	-1.70 – 1.08	-0.44	0.662

Group [HC] : BodyPart [Abdomen] : Delay [0] : Distance.Q	0.45	0.70	-0.94 – 1.83	0.63	0.527

Random Effects

σ^2	445.75
$\tau_{00\ Ppn}$	435.21
$\tau_{11\ Ppn.Delay0}$	30.19
$\tau_{11\ Ppn.BodyPartAbdomen}$	92.75
$\tau_{11\ Ppn.Delay0:BodyPartAbdomen}$	21.70
ρ_{01}	0.09
	0.27
	-0.01
ICC	0.56
N_{Ppn}	90
Observations	5333
Marginal R^2 / Conditional R^2	0.056 / 0.581

Table S.8.
Linear mixed-model summary: TDE-D Confidence rating (Random slopes)

	Confidence rating				
Predictors	*Estimates*	*S.E.*	*95% CI*	*Statistic*	*p-value*
(Intercept)	54.85	1.47	51.92 – 57.78	37.24	**<0.001**
Group [AN]	-3.44	2.08	-7.57 – 0.69	-1.66	0.101
Group [HC]	2.77	2.08	-1.36 – 6.90	1.33	0.186
BodyPart [Abdomen]	0.12	0.43	-0.74 – 0.99	0.28	0.781
Delay [0]	0.76	0.40	-0.03 – 1.55	1.91	0.060
Distance.L	2.41	0.31	1.81 – 3.01	7.88	**<0.001**
Distance.Q	0.15	0.31	-0.45 – 0.75	0.50	0.619
Group [AN] * BodyPart [Abdomen]	-0.02	0.61	-1.24 – 1.20	-0.03	0.973
Group [HC] * BodyPart [Abdomen]	-0.27	0.62	-1.49 – 0.95	-0.44	0.663
Group [AN] * Delay [0]	0.96	0.56	-0.15 – 2.08	1.72	0.089
Group [HC] * Delay [0]	-0.81	0.56	-1.93 – 0.31	-1.45	0.152
BodyPart [Abdomen] * Delay [0]	0.11	0.29	-0.46 – 0.68	0.38	0.708
Group [AN] : Distance.L	-1.06	0.43	-1.91 – -0.21	-2.45	**0.014**
Group [HC] : Distance.L	1.77	0.43	0.92 – 2.62	4.09	**<0.001**
Group [AN] : Distance.Q	0.72	0.43	-0.13 – 1.56	1.67	0.095
Group [HC] : Distance.Q	-0.05	0.43	-0.90 – 0.80	-0.12	0.906
BodyPart [Abdomen] : Distance.L	0.72	0.31	0.12 – 1.32	2.34	**0.019**
BodyPart [Abdomen] : Distance.Q	0.70	0.31	0.10 – 1.30	2.30	**0.022**
Delay [0] : Distance.L	0.09	0.31	-0.51 – 0.69	0.30	0.768
Delay [0] : Distance.Q	-0.63	0.31	-1.23 – -0.03	-2.06	**0.040**

(Group [AN] * BodyPart [Abdomen]) * Delay [0]	0.01	0.41	-0.80 – 0.81	0.01	0.989
(Group [HC] * BodyPart [Abdomen]) * Delay [0]	-0.05	0.41	-0.86 – 0.75	-0.13	0.893
Group [AN] : BodyPart [Abdomen] : Distance.L	0.43	0.43	-0.42 – 1.27	0.98	0.325
Group [HC] : BodyPart [Abdomen] : Distance.L	0.06	0.43	-0.79 – 0.90	0.13	0.898
Group [AN] : BodyPart [Abdomen] : Distance.Q	-0.76	0.43	-1.60 – 0.09	-1.75	0.079
Group [HC] : BodyPart [Abdomen] : Distance.Q	0.27	0.43	-0.58 – 1.12	0.63	0.531
Group [AN] : Delay [0] : Distance.L	-0.08	0.43	-0.93 – 0.77	-0.19	0.851
Group [HC] : Delay [0] : Distance.L	0.30	0.43	-0.55 – 1.15	0.70	0.483
Group [AN] : Delay [0] : Distance.Q	-1.00	0.43	-1.85 – -0.16	-2.32	**0.020**
Group [HC] : Delay [0] : Distance.Q	0.36	0.43	-0.49 – 1.21	0.82	0.410
BodyPart [Abdomen] : Delay [0] : Distance.L	0.17	0.31	-0.43 – 0.77	0.55	0.582
BodyPart [Abdomen] : Delay [0] : Distance.Q	-0.39	0.31	-0.98 – 0.21	-1.26	0.206
Group [AN] : BodyPart [Abdomen] : Delay [0] : Distance.L	0.37	0.43	-0.47 – 1.22	0.87	0.386
Group [HC] : BodyPart [Abdomen] : Delay [0] : Distance.L	0.12	0.43	-0.72 – 0.97	0.29	0.774
Group [AN] : BodyPart [Abdomen] : Delay [0] : Distance.Q	0.47	0.43	-0.37 – 1.32	1.10	0.273

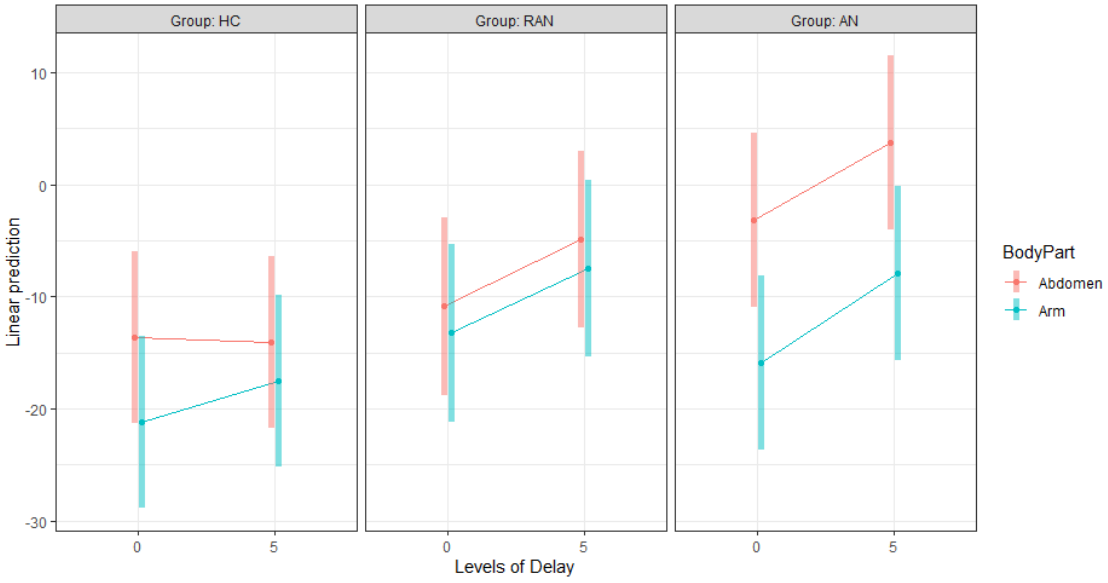
Group [HC] : BodyPart [Abdomen] : Delay [0] : Distance.Q	-0.30	0.43	-1.15 – 0.55	-0.70	0.484

Table S9.

Random Effects	
σ^2	160.30
$\tau_{00\ Ppn}$	189.76
$\tau_{11\ Ppn.Delay0}$	10.91
$\tau_{11\ Ppn.BodyPartAbdomen}$	13.58
$\tau_{11\ Ppn.Delay0:BodyPartAbdomen}$	4.25
ρ_{01}	-0.08
	0.15
	-0.01
ICC	0.57
N $_{Ppn}$	89
Observations	5159
Marginal R^2 / Conditional R^2	0.030 / 0.585

Figure S.7.

*TDE Group*Bodypart*Delay*



Linking Text Between Chapters 4 and 5

In chapter 3, I discussed the interpretational issues that face research into bodily misperception in eating disorders. It also offered suggestions for how to overcome these issues. In chapter 4, we empirically addressed one of these issues: the influence of non-perceptual factors in biasing participant responses—what I referred to as "the response bias challenge". We explored this challenge as applied to the hypothesis that people with anorexia nervosa misperceive touch, uncovering evidence suggesting that the differences in tactile size judgments associated with anorexia nervosa do not stem from differences in perception.

The next chapter zooms out and changes focus, addressing a long-standing debate in philosophy of mind over the role of mental representation in explanations of psychological phenomena. Mental representations—specifically, representations of the body—have played a crucial role in the foregoing chapters. However, many philosophers argue that mental representations should not feature in psychological explanations, especially those targeted at phenomena such as bodily movement and affordance processing. The next chapter argues against this position, claiming that the evidence of body schema distortion associated with anorexia nervosa vindicates the role of mental representation in psychological explanation.

This chapter represents an important instance of the form of interdisciplinary research advocated for in this thesis. In it, I argue that philosophers engaged in debates over the role of mental representation in psychological explanation must pay closer attention to the discipline that they hope to influence. If philosophers desire to shape psychology and, in this case, remove it of certain types of theoretical entities, then they will need to provide replacement explanations, which improve on what is currently on offer. In the following chapter, we introduce this approach as the most appropriate for making progress on these questions and illustrate how adopting it vindicates the use representation in explaining the mind.

Synthese (2018) 195:5297–5317
https://doi.org/10.1007/s11229-018-1843-3

Action, affordances, and anorexia: body representation and basic cognition

Stephen Gadsby[1] · **Daniel Williams**[2]

Received: 4 October 2017 / Accepted: 4 June 2018 / Published online: 6 June 2018

Abstract We evaluate a growing trend towards anti-representationalism in cognitive science in the context of recent research into the development and maintenance of anorexia nervosa in cognitive neuropsychiatry. We argue two things: first, that this research relies on an explanatorily robust concept of representation—the concept of a long-term body schema; second, that this body representation underlies our most basic environmental interactions and affordance perception—the psychological phenomena supposed to be most hospitable to a non-representationalist treatment.

Keywords Representation · Anorexia · Affordances · Content · Basic cognition · Enactivism · Action · Action-oriented · Embodied · Hard problem of content · Radical

1 Introduction

The concept of *internal representation* is both foundational and ubiquitous in the contemporary cognitive sciences, and it has been for a long time (Bermúdez 2010; Von Eckardt 2012). Despite this, recent decades have seen a growing chorus of voices calling for either the marginalisation or outright elimination of this concept in cognitive theorizing (Anderson 2014; Brooks 1991; Barrett 2011; Chemero 2009; Garzón 2008; Hutto and Myin 2013; Noë 2004; Ramsey 2007; Van Gelder 1995). Drawing on insights from phenomenology, embodied cognition, robotics, theoretical biology and

✉ Stephen Gadsby
gadsby.st@gmail.com

Daniel Williams
dw473@cam.ac.uk

1 Monash University, Melbourne, Australia

2 University of Cambridge, Cambridge, UK

dynamical systems theory, this eclectic group of anti-representationalists argue that internal representations are metaphysically problematic and explanatorily unenlightening, masking the embodied and dynamic character of our basic cognitive transactions with the world (Chemero 2009; Hutto and Myin 2013). Given these alleged deficiencies, they call for a full-blooded *revolution* (Hutto and Myin 2013, p. 1) in the mind sciences—a genuine *paradigm shift* (Engel et al. 2015, p. 1) in scientific thought that many argue is already under way.

In this paper, we evaluate these claims about representation in the context of recent research in cognitive neuropsychiatry on the nature and cause of anorexia nervosa. We argue two things: first, that a robustly representational concept—the *body schema*—is explanatorily central within this research; and second, that this body schema underlies even our most basic environmental interactions and affordance perception. Given that such psychological phenomena are supposed to be most hospitable to a non-representationalist treatment (Anderson 2014; Chemero 2009; Hutto and Myin 2013), we conclude that the prospects for a radical non-representationalist shift in cognitive theorizing are dim.

We structure the paper as follows. In Sect. 2, we outline two conditions for the attribution of genuine representational status: first, that the relevant structure passes what Ramsey (2007) calls the "job description challenge"; and second, that it has satisfaction conditions of some kind. For the sake of our argument, we designate these conditions as jointly necessary and sufficient. In Sect. 3, we introduce the concept of the *body schema*, a sensori-motor representation of the body exploited in action, and we explain how distortion of this representation plays an integral role within contemporary research into anorexia nervosa. In Sect. 4, we validate the representational credentials of the body schema: not only does it pass the job description challenge (Sect. 4.1), but its content is essential to its explanatory role (Sect. 4.2). Finally, in Sect. 5 we argue that this research is of central relevance to the "representation wars" of recent cognitive science: it strongly suggests that this body representation underpins basic environmental interactions and affordance perception in neurotypical subjects, and it undercuts some of the most influential arguments for wholly non-representational treatments of these domains. We conclude in Sect. 6 by briefly summarising our argument.

2 Adjudicating the representation wars

At least since the 1960s, one of the defining characteristics of mainstream cognitive science has been an emphasis on the importance of internal representations in psychological theorizing (Bermúdez 2010). Since then, there have been dramatic shifts in scientific thought about the nature of such representations—their vehicles, formats, and systemic roles, for example. There is—to put it mildly—much less enthusiasm for the original vision of cognition as rule-governed operations defined over arbitrary symbol structures (Williams and Colling 2017). Nevertheless, a broad commitment to the idea that cognitive processes implicate content-bearing states remains firm, helping to define orthodoxy in an otherwise extremely heterogenous multidisciplinary effort

to understand the mind (Bechtel 2008; Bermúdez 2010). As Ramsey (2007, p. xi) puts it,

> It has become almost a cliché to say that the most important explanatory posit today in cognitive research is the concept of representation. Like most clichés, it also happens to be true.

When a growing chorus of anti-representationalists in philosophy and cognitive science characterize their view as "radical," then, they're not kidding (Chemero 2009; Hutto and Myin 2013; Ramsey 2007). Whilst the motivations, views, and prescriptions of such theorists are extremely eclectic, they are united in calling for either the marginalisation or outright elimination of the concept of internal representation in cognitive theorizing. The heated debates instigated by this anti-representationalist movement have given rise to what might reasonably be called the "representation wars" in cognitive science (Clark 2015; Williams 2017).

In this section we briefly outline a framework for adjudicating the representation wars. Specifically, we assume that a structure qualifies as representational if and only if it satisfies two conditions: first, the structure must pass what Ramsey (2007) calls the "job description challenge"; second, it must possess satisfaction conditions of some kind. In both cases, we have chosen conditions that are both widely accepted in the broader literature and—more importantly—accepted by anti-representationalists themselves (Hutto and Myin 2013; Ramsey 2007). Specifically, given that we are arguing in favour of the representational credentials of a previously neglected aspect of cognition, we have deliberately chosen conditions that are deemed acceptable by our opponents.

2.1 The job description challenge

The *job description challenge* (henceforth *JDC*) is "the challenge of explaining how a physical state actually fulfils the role of representing in a physical or computational process" (Ramsey 2007, p. xv). Specifically, Ramsey (2007, p. 28) argues that the concept of representation is a "functional notion," and that the representation wars thus concern to what extent the structures implicated in the production of psychological capacities perform this function (see Haugeland 1991, p. 69). To answer the JDC, then, one must show that "the states characterized as representations in explanatory framework X actually serve as representations, given the processes and mechanisms put forth" (Ramsey 2007, p. xiii). That is, one must show that the representational status of the relevant states or structures contributes to their systemic role within the cognitive mechanism or mechanisms of which they are a part.

What is it to function as a representation? Ramsey (2007) argues that our *commonsense pre-theoretic* understanding of representation constrains any acceptable answer to this question, such that "exploring how a representational posit is thought to operate in a system" is a matter of "assessing this role in terms of our ordinary, intuitive understanding of what a representation is and does" (Ramsey 2007, p. 10). According to Ramsey (2007, p. 14), this intuitive understanding clusters around "two different families of representational notions": first, "various notions of mental representation";

and second, "different types of non-mental representation." Because it will be most relevant in what follows, and it is the family of representational notions that Ramsey (2007, p. 22) believes has been most central in cognitive-scientific research, we will focus mostly on the second family here.

One way of passing the JDC, then, is to show that a given physical structure or mechanism functions analogously to public representations with which we are familiar. As Ramsey points out, however, this is no easy feat. Public representations are "employed by cognitive agents as a type of tool"—namely, a tool that enables them to coordinate their behaviour with a domain to which they have restricted access (Ramsey 2007, p. 23; see also Williams and Colling 2017). As such, Ramsey argues, it is not obvious how an internal cognitive structure (a state of one's brain, say) *could* function as a public representation without an inner homunculus to use and interpret it (2007, pp. 26–27).

Unlike philosophers such as Ryle (1949) and Wittgenstein and Anscombe (1953), Ramsey does not argue that this challenge cannot be met. For example, he argues that a legitimately representational approach in cognitive science is the appeal to cognitive systems that exploit internal models of specific domains to solve various cognitive problems (Ramsey 2007, Chap. 3). In this case the relevant cognitive structure genuinely gets used by the system of which it is a part as a proxy or surrogate for something else (Ramsey 2007, p. 23). Nevertheless, he does argue that this explanatory strategy is an increasingly rare example in the contemporary cognitive sciences, and that many appeals to the term "representation" in cognitive theorizing today—especially in cognitive neuroscience—fail the JDC.

2.2 Content

Ramsey's outline of the JDC deliberately focuses on functional questions of the sort that are directly relevant to the explanatory concerns of cognitive science. At least since Quine's (1960) attack on the scientific credentials of meaning, however, a second major focus of the representation wars has been on representational *content*, where "content" is typically understood to minimally include satisfaction conditions of some kind (see Hutto and Myin 2013; Fodor 1987; Rosenberg 2015). For example, Hutto and Myin (2013, p. x) write:

> Just what is content? At its simplest, there is content wherever there are specified conditions of satisfaction. And there is true or accurate content wherever the conditions specified are, in fact, instantiated.

In the case of world-directed representations, at least, the appeal to satisfaction conditions codifies the widespread intuition that an essential property of representations is the ability to *mis*represent—to be in error relative to how things are (Hutto and Myin 2013; Shapiro 2011, p. 143). As Ramsey (2007, p. 12) puts it, "the possibility of misrepresentation is built into our ordinary way of understanding what it is to represent."

An enormous amount of philosophical energy has been expended in trying to answer what Von Eckardt (2012) calls the "foundational problem of cognitive science," namely

the "content determination problem": what in the natural world determines the contents of cognitive representations (see Fodor 1987)? A growing number of philosophers have argued that this "hard problem of content" is unsolvable for most (Hutto and Myin 2013) if not all (Rosenberg 2015) cognitive structures—that such structures do not have satisfaction conditions of any kind. As Ramsey (2007, p. 30) notes, this problem of "understanding how a state's content is grounded in some set of conditions is not the same thing as understanding how the state actually serves as a representation." Nevertheless, the stipulation that representations must be *contentful* has served as a foundational assumption in the representation wars on both sides (Hutto and Myin 2013; Fodor 1987; Rescorla 2013).[1] As such, we will follow others in taking this criterion for representational status as given in what follows.

2.3 Summary

Given the foregoing overview, we assume that a cognitive structure qualifies as a representation if and only if: (i) it performs a genuinely representational function; and (ii) it has satisfaction conditions of some kind. Of course, there is something slightly artificial about distinguishing these conditions. After all, if we have good reason to believe that a structure genuinely functions as a representation, presumably this provides us with good reason to believe that it is contentful as well (see Sect. 4). Nevertheless, the focus on function and content have generated distinguishable threads in the representation wars, so it is useful to separate them in the foregoing manner.

Importantly, we have characterised these conditions as collectively necessary and sufficient for the attribution of genuine representational status. Some philosophers and cognitive scientists might argue that this analysis is too stringent—for example, because a representational structure need not pass the JDC, or because it need not possess satisfaction conditions (Williams 2018). Strictly, all that our argument below *requires* is that the conditions are collectively sufficient—that if a structure passes the JDC and possesses satisfaction conditions, it is a representation. Nevertheless, we will also assume throughout that if a structure does not satisfy such conditions, it is *not* a representation—that these conditions are necessary. This ensures that the case for a representational interpretation of the body schema is maximally difficult: if it could be shown that this cognitive structure fails just one of the conditions outlined in this section, we concede that it should not be treated as representational.

In Sects. 3 and 4 we outline the nature of a cognitive structure—the body schema—and show that it satisfies *both* conditions, thus authenticating its representational status. In Sect. 5, we finish by arguing that the representational nature of the body schema is of central relevance to the representation wars.

[1] An unstated corollary of this assumption that we will take for granted is that the contents of cognitive representations must be an objective property of the representations, and not derived from the interpretation of the theorist.

3 The body schema and anorexia nervosa

3.1 The body schema

Within cognitive neuroscience, the notion of a body schema has had a long and fascinating history. Through the influence of Head and Holmes (1911), the body schema gained traction as a fundamentally contrastive notion to the well-known *body image*—a broad set of intentional states (perceptions, beliefs, attitudes) in which the object is one's body (Gallagher and Cole 1995, p. 371). The body schema was considered distinct from the body image in two important ways: first, its functional role—the body image subserves perception, whilst the body schema subserves action; second, its relationship to consciousness—the body image is conscious, whilst the body schema is unconscious (de Vignemont 2010). What emerged was the notion of the body schema as an "unconscious, sensorimotor, representation of the body that is invoked in action" (Keizer et al. 2013, p. 1). As this quote illustrates, the body schema is simply assumed to be representational in the cognitive neuroscientific literature (Alsmith and de Vignemont 2012, p. 3), and descriptions of its role are saturated with appeals to phenomena such as content and information. We will initially follow orthodoxy in taking this representational terminology at face value. In Sect. 4, we provide a systematic defence of why this terminology is warranted.

Within cognitive neuroscience, the most well-known and fine-grained synthesis of the body schema is provided by de Vignemont (2010). Her description of the way in which body schema content is utilised to enable holistic bodily movement draws on the standard, forward model-based characterisation of motor control. Within this explanation, there are two kinds of "models" to consider (Wolpert and Ghahramani 2000). The inverse model computes the necessary motor commands to achieve a desired state, given the body's current state. Further along the causal chain, the forward (aka emulator) model predicts the sensory feedback of those same motor commands. Whilst there is a fascinating body of literature regarding the representational credentials of the forward models' simulated states (Grush 1997), we need not concern ourselves with that here. Our focus is instead on the *information* needed for these models to control movement.

As de Vignemont describes, inverse models are fed by the *initial body schema*, a collection of body-related information containing "long-term information like the size of the limbs, and short-term information like the joint angles and the hand position" (2010, p. 673). According to this notion, then, the body schema consists of two *kinds* of information regarding the body: short-term and long-term (de Vignemont 2010).

In terms of the internal models themselves, however, how many exist and how exactly they operate is still an open question within motor control research (Pickering and Clark 2014; Wolpert and Kawato 1998). As will become clear, our argument below does not depend on *how* such models make use of this content, only *that* they do. That is, the explanatory importance of the body schema rests on the *kinds* of body information computationally required for successful motor control, and not the specific architecture of the cognitive mechanisms that exploit such information.

3.2 Anorexia nervosa and the body schema

The notion of the body schema has its roots within cognitive neuropsychology, being used to explain several different dysfunctions related to somatosensation and motor control (Haggard and Wolpert 2005; Paillard 1999). More recently, however, it has come to play an important role in an exciting body of research in the emerging subfield of cognitive neuro*psychiatry*, which applies the methods of cognitive neuropsychology to understanding disorders of higher-level cognition, such as belief formation (Halligan and David 2001). Specifically, it has been employed to explain a number of behavioural peculiarities and problematic beliefs associated with anorexia nervosa (Gadsby 2017a, b; Keizer et al. 2013; Metral et al. 2014).

Whilst ostensibly categorized as an eating disorder—whereby the central symptom is a problematic relationship with food—anorexia has long been suspected to involve cognitive dysfunction of some kind (Bruch 1962). Operating under this assumption, there is an incredibly fecund field of cognitive neuropsychiatry devoted to uncovering the cognitive and neurological differences between patients and neurotypical subjects (Kanakam and Treasure 2013).

Whilst it has long been thought that anorexia involves distortion of the body image (Smeets 1997), a more recent branch of research focuses on uncovering evidence that patients' *body schemas* are also distorted. Indeed, this represents one of the most promising branches of research into the body schema. The prerequisite knowledge for such research arises from studies done by ecological psychologists in the 1980s. Warren and Wang (1987) first discovered that when passing through apertures (for example, a doorway), neurotypical subjects turn their shoulders at roughly the same ratio of shoulder-to-aperture width—what's known as *the critical point*. Similarly, it was found that neurotypical subjects conform to a shared ratio when *assessing* their ability to pass through apertures without turning their shoulders. Given these known ratios, cognitive neuropsychiatrists devised a number of behavioural tasks to test the body schemas of anorexia patients. These tasks can be divided into two categories: movement and action assessment.

Though a number of movement-based studies have been conducted (Keizer et al. 2013; Metral et al. 2014; see also: Engel and Keizer 2017), we will focus on the first of these experiments, in order to give a more detailed summary of its findings. In this initial study, participants were required to walk through different sized apertures whilst completing a distractor task. Movement kinematics were recorded using a camera and reflective markers. For each participant, their critical point was calculated "based on the largest aperture width for which participants (turned their shoulders) at least two times … divided by the participants' shoulder width" (4). The experimenters discovered differences in critical points between the healthy control and anorexia groups: anorexia patients displayed significantly *higher* critical points than the usual standard; in other words, their movement kinematics were equivalent to subjects with much broader shoulders. The researchers interpreted this as evidence that the patients' body schemas were *distorted*, representing their bodies as wider than reality (Keizer et al. 2013, p. 5).

Are there any possible alternative explanations of this data? For example, is it possible that anorexia patients simply *misperceive* aperture width, or perhaps exhibit

some form of non-specific kinematic abnormality? Under these alternatives, there would be no need to posit the existence of an impairment in the body schema itself.

There are further aspects of the study which rule out these alternative hypotheses. Beyond measuring movement kinematics, Keizer and colleagues also included a behavioural task aimed at measuring the spatial content of participants' body representations. Specifically, "participants were instructed to draw a vertical line on a whiteboard which represented the width of their shoulders" (Keizer et al. 2013, p. 4). What they discovered was that "if (anorexia) patients' shoulders were as wide as they estimated them to be (in this task), they would perform equal to (healthy controls) on body-scaled action" (*ibid*. 6). That is, anorexia patients were moving their bodies with the same dynamics as healthy controls, albeit *as if* they were as wide as they estimated them to be.

This finding is inconsistent with the hypothesis that patients exhibit non-specific kinematic abnormalities or simply misestimate aperture width. Neither of these can account for the finding that the width at which patients start to turn can be *predicted* by the outcome of the shoulder-width estimate task. However, this evidence *can* be accounted for by assuming that the shoulder-width estimate task is a measure of the *same information* used to determine critical point and therefore that differences in this information are what cause differences in critical point—as the experimenters themselves conclude (*ibid*.).[2]

Considering the relationship between critical points and the body schema illuminates this evidence. Specifically, whilst critical points are *measured* as the ratio between shoulder width and aperture width, these critical points are *determined* not by the body, but by the body schema (Gadsby 2017a). As per de Vignemont's model, body schema information is fed to the inverse model, determining motor trajectory (via the calculation of motor commands), which in turn determines critical point. Because the motor system computes critical points from body schema information, changes in this information explain critical point differences.

3.3 Action assessment

In order to further buttress the explanatory relevance of long-term body schema information to this research, we will briefly discuss another group of experiments, antecedent to these movement-based tasks. These studies required participants (anorexia patients and healthy controls) to *assess* their ability to pass through different sized apertures without turning their shoulders. Again, whilst healthy controls roughly conformed to the ratio discovered by Warren and Wang, anorexia patients'

[2] An anonymous reviewer pointed out the possibility that these computations might rely on "hard-coded" values, which presumably don't count as representations as they don't track the body. However, this possibility is at odds with knowledge of movement dynamics. Our motor commands are consistently body size appropriate: as we transition from childhood to adulthood and our bodies grow, our motor commands reflect this change in size. Similarly so for more sudden changes in body size (fat and muscle fluctuation, loss of limbs, etc). And, of course, motor commands are altered when the dimensions of tools are incorporated into the spatial content of the body schema (Gadsby 2017c, pp. 22–23). This evidence discounts the possibility that motor command computation relies on "hard-coded" size values, rather than a body schema representation which *tracks* the size of the body (and other action-relevant effectors such as tools).

critical points were much higher—they estimated their ability to pass through apertures as if their bodies were wider than reality (Guardia et al. 2010, 2012; Metral et al. 2014; see also: Engel and Keizer 2017). As expected, this bias in action assessment was not present when patients were asked to assess an experimenter's ability to pass through apertures (Guardia et al. 2012).

Although traditionally associated with movement, research suggests that the body schema also underlies action assessment (Schwoebel and Coslett 2005). This falls in line with the dominant psychological theory of action assessment, whereby in order to assess our ability to act, we take our motor circuitry offline, *simulating* the actions themselves (Jeannerod 2001; for review, see Declerck 2015, pp. 2–5). According to this theory, both motor control and motor simulation are driven by the same motor circuitry—motor circuitry which relies on the body schema. Again, the positing of long-term body schema information is central in accounting for why anorexia patients show differences in movement kinematics *and* action assessment: this information represents the body as wider than reality, leading to downstream effects in both these domains.

3.4 Summary

The notion of the body schema—an unconscious sensori-motor representation of the body used for action—has played a prominent role within the field of cognitive neuropsychology. Furthermore, this structure has come to play an indispensable explanatory role within the cognitive neuropsychiatric approach to understanding anorexia. Whilst cognitive neuroscience has traditionally conceptualised this structure in representational terms, there is, as of yet, no robust philosophical grounding of this assumption; this is the task we turn to now.

4 The body schema's representational credentials

We have introduced the notion of a body schema, involving two kinds of information about the body—long-term spatial characteristics and short-term information regarding current body position—which are used for motor control and action assessment. We have also discussed the central role that this notion plays within cognitive neuropsychiatry—in particular, in explaining evidence of abnormal movement kinematics and action assessment in anorexia patients. Whilst this research paints the body schema as fundamentally representational, it mustn't simply be taken at its word. Indeed, a central insight of the anti-representationalist movement is that the structures which cognitive scientists conventionally refer to as representations often do not properly fit the label (Ramsey 2007). As such, in this section we turn to validating the representational status of the body schema. First, we show that it satisfies the JDC. Specifically, its functional role is to *stand in* for something to which the brain's sensorimotor systems have restricted access. Second, we show that the notion of *content* (understood as satisfaction conditions) is indispensable to its explanatory role within the research we have drawn on. As such, the body schema passes both conditions stipulated to be jointly sufficient for representational status.

4.1 The body schema and the job description challenge

Consider the introduced distinction between two kinds of body schema information: short-term and long-term. This is a somewhat classic distinction in the body representation literature, arising due to the different ways in which we sensorily track the state of our body (O'Shaughnessy 1980). There are a whole host of afferent receptors which monitor information such as tendon tension, joint angle and muscle stretch (de Vignemont 2014, p. 990). These receptors have evolved to supply very specific information regarding current body position—information which is *consistently available*. As such, the internal states which carry this information are in constant causal connection (via the afferent receptors) with the states of affairs themselves (e.g. muscle tension, joint angle etc.). The information derived from these receptors therefore needn't be *stored* over long periods of time, instead, it is consistently replaced by new incoming signals (de Vignemont 2010, p. 672).

In contrast, there are no afferent receptors which deliver size information regarding the spatial structure of the body (i.e. size and shape) (Proske and Gandevia 2012). We rely on vision the most to gain information about our body size, but of course our body is rarely in the visual field, especially when we are attending to our environments during goal-directed action. As such, size information regarding the body is *absent* to the sensory systems. Consequently, this information must be stored and slowly updated over long periods of time (de Vignemont 2014; Gadsby 2018).

Ramsey's functional analysis of cognitive structures is relevant to this distinction between short-term and long-term body information. Internal states which carry short-term information are in constant causal contact with the afferent sensory receptors that deliver it. Given this unbroken chain of causality—from internal property to external state of affairs—states like these are best thought of as *reliable causal mediators;* they facilitate action via reliably responding to external conditions (in this case, immediate properties of body parts like joints, muscles and tendons). As Ramsey (2007) argues, however, being a reliable responder is not sufficient for representational status. If it were, this would entail an overly permissive account of representation, as there are many seemingly non-representational internal states which fulfil this same systemic role (e.g. the immune system's response to infections) (2007, p. 125). For the sake of argument, we grant that receptors do not answer the JDC. As such, we also grant that internal states which carry short-term body information are *not* representational.

On the other hand, internal states which carry long-term information about spatial parameters are a different affair. Specifically, there is no unbroken causal link from these internal states to the relevant external states of affairs. Whilst the states are *updated* via occasional sensory input, this input is not *present* to the sensory system whenever the states are in use. As such, it is plausible to think that a property of the internal states themselves must *stand in* for these spatial parameters. In the research we have drawn upon, this is exactly the functional role assigned to the long-term body schema. It is emphatically not a mere reliable causal mediator: there is no unbroken internal-state/sensory-system/external-state loop at play.

Further, not only does the body schema stand in for spatial characteristics of the body to which the brain has restricted access, but recognising this status accounts for the previously discussed evidence. Long-term body schema information enables

successful aperture passing because the motor system uses it as a proxy for shoulder-width, in order to calculate body-appropriate motor commands (de Vignemont 2010, p. 673). Further, as per the simulation hypothesis, the same motor circuitry is used for action assessment (via *offline* motor simulation). As such, this same information is also used to passively assess aperture passability. Without the body schema standing in for shoulder width in this manner, it is difficult to see how subjects would be able to successfully navigate through apertures *or* estimate their ability to do so. Thus it isn't simply that the body schema stands-in for shoulder-width, but that this standing-in relationship is *exploited* by the motor system to enable successful navigation and action assessment.

In the same way that structural features of a map function as stand-ins for physical features of the world, the evidence suggests that long-term body schema information functions as a stand-in for structural characteristics of the body. When a structure functions as a stand-in, or surrogate, for a target which the system lacks direct access, it is serving a representational role, therefore passing the JDC (Ramsey 2007, pp. 87–99, pp. 199–201). As Haugeland puts it, "(t)hat which stands in for something else in this way is a representation; that which it stands in for is its content; and its standing in for that content is representing it" (1991, p. 62; from Ramsey 2007, p. 87) or, as Ramsey (2007, p. 199) himself states (in reference to a hypothetical car navigation system), "to say that an area of the groove functions as a 'stand in' for a segment of the track is just to say that an area of the groove is playing a representational role".

The fact that the motor system uses the body schema as a surrogate in this way also provides undeniable explanatory purchase in terms of the aforementioned behavioural evidence. That the motor system uses the body schema as a surrogate explains why neurotypical subjects turn their shoulders and assess the need to turn their shoulders at consistent critical points (despite lacking sensory access to their own shoulder width). Further, the fact that anorexia patients deviate from this standard in ways that can be predicted using body size estimate tasks can be understood by assuming that their body schema information is inaccurate and that the tasks are a measure of this inaccurate information (see below).

To summarise, then, not only does long-term body schema information play a fundamentally representational role—standing in for characteristics of the body to which the brain's sensorimotor systems have restricted access—this representational role is integral to the operation of such systems. Through reference to this role we can explain the full range of behavioural evidence related to aperture passing.

4.2 The explanatory role of content

If we are right, a subject's body schema genuinely functions as a *stand-in* or *proxy* for structural characteristics of her body in a way that can be exploited to guide action and facilitate action assessment. As such, it passes Ramsey's job description challenge: it resembles the functional profile of prototypical *public representations*—that is, representations that enable us to coordinate our behaviour with domains to which we have restricted access by *standing in* for such domains (Ramsey 2007). For some, this is sufficient to ascribe representational status to a structure (Isaac 2013; Williams

2018). After all, its systemic role within the motor system is to re-present information to which the system does not have direct access.

However, we noted in Sect. 2 that many philosophers require something more. They require that the relevant states are *contentful*—at a minimum, that they possess *satisfaction conditions* of some kind. For example, this requirement is central to Hutto and Myin's (2013) claim that much of cognition should be viewed as non-representational on the grounds that we have no solution to what they call the "hard problem of content". Specifically, they argue that the kinds of covariational relationships that cognitive scientists rely upon to characterise the operation of cognitive mechanisms are insufficient to explain how such systems acquire determinate satisfaction conditions—how they come to be *answerable* to how things are:

> … anything that deserves to be called content has special properties—e.g., truth, reference, implication—that make it logically distinct from, and not reducible to, mere covariance relations holding between states of affairs. Though covariance is surely scientifically respectable, it isn't able to do the required work of explaining content (Hutto and Myin 2013, p. 67).

This kind of anti-representationalist argument is similar in character to Ramsey's. However, whereas Ramsey's argument targets the *functional role* of "receptor" states, Hutto and Myin's argument is broadly *metaphysical*: namely, that one cannot *reduce* content to co-variational relations of the sort countenanced by natural science. If this is right—if the hard problem of content is as insuperable for sub-personal cognitive processes as Hutto and Myin claim—one might take this as evidence against the representational status of the body schema.

Whilst we don't propose to wade into the enormous literature on content determination and naturalistic psychosemantics here, we will note two points that we think strongly undermine this non-representationalist challenge in the current context.

First, as noted above, the body schema is *not* a "receptor"-style state whose function is to reliable covary with bodily states. In fact, just the opposite is the case: its function is to stand in for structural characteristics of the body to which the brain does not have reliable access. Unlike mere detectors, it is not obvious how to describe the behaviour of such cognitive structures without appealing to the concept of representation (Williams and Colling 2017). If one wants to argue for a non-representational analysis of the body schema and its broader functional role, then, one must offer a non-representational description of the research. For the reasons outlined above (see also Sect. 5), we do not think that this is possible. As such, if the body schema genuinely functions in the manner we have suggested, this gives us good prima facie reason to suppose that it is contentful. As Ramsey (2007, pp. 30–31) puts it:

> if theorists can develop accounts of cognitive processes that posit representations in a way that reveals just how the posit plays a representational role in the system, then the explanation of content can wait … from the standpoint of psychological theory development, the need for an account of content-grounding is not so urgent.

In fact, it is possible to go further than this. The second—and more important—point is that satisfaction conditions *themselves* play an indispensable explanatory role in the

research we have outlined. That is, it is not just that the body schema is genuinely used as a stand-in for structural characteristics of the body (i.e. its shape and size). Central to the research outlined above is the fact that in patients with anorexia the body schema gets something *wrong*. Specifically, we saw above (Sects. 3.2, 3.3) that the best explanation for these patients' movement kinematics is that their body schema information is *inaccurate*. If this is right, however, the body schema must be contentful: not only does it possess veridicality conditions that can fail to be satisfied, but this failure of satisfaction is positively central to the research that we have drawn upon.

To quote Ramsey once more,

> The significance of a representational theory of the mind stems in large measure from the different elements that are associated with representational states as ordinarily understood. For example, when theorists posit inner representations, they typically assume that they now have an important way to explain how the system can fail to behave appropriately. *It is now possible to explain faulty behavior as sometimes stemming from false representations of the world* (Ramsey 2007, p. 12, our emphasis).

As should be clear from the research introduced above, the role of the body schema conforms to this analysis. By construing the body schema in representational terms, researchers in cognitive neuropsychiatry can explain the anomalous behaviour evidenced in anorexia nervosa in terms of *inaccuracy*. Not only does this inaccuracy explain differences in critical point, the *measurement* of this inaccurate content (via body size estimate task) is also integral to the research discussed, as it allows researchers to predict the movement kinematics of those with distorted body schemas (see Sect. 3.2).

Advocates of the "hard problem of content" might claim that taking content as given in this research is somehow illegitimate or insufficiently "naturalistic"—that to substantiate this research we must show how to explain the content of the body schema in terms of "naturalistically" kosher non-content (cf. Hutto and Myin 2013; Hutto and Satne 2015). This is not a version of naturalism that can be sustained, however. If our best science of some capacity indispensably requires the postulation of contentful stand-ins for bodily states, we should be realists about such representations in the only sense of "realism" that could matter in the representation wars (Rescorla 2013). Specifically, we understand these representation wars to concern the representational status of our best *science* of cognition (see Sect. 2), and we understand philosophical naturalism to involve a deference to science in determining how the world works (Quine 1960). As such, given that representational content seems to feature indispensably in the research we have drawn on, whether the body schema is contentful should depend on the scientific value of this research, not on the outcome of metaphysical theories of content (cf. Ramsey 2007; Rescorla 2013).

Of course, as noted above, our argument here will not go through if anti-representationalists can provide an alternative explanation of the judgemental and behavioural anomalies exhibited in anorexia as outlined in Sect. 3—an explanation that makes no reference to content. We know of no such explanations, however, and—-

given the systemic role of the body schema articulated above—we are not optimistic that any could be provided.[3]

4.3 Summary

According to the research from cognitive neuropsychiatry we have outlined, the body schema stands in for structural characteristics of the body to which the brain has restricted access, enabling fluid motor control and action assessment. As such, the body schema passes the JDC: just as one might use scribbles of ink on a napkin as a proxy for the spatial layout of a city to navigate one's way to a conference, the brains' sensorimotor systems exploit the body schema as a surrogate for the shape and size of the body to guide effective action and action assessment. Further, as outlined in Sect. 3, the content of the body schema—specifically, its inaccuracy in the case of anorexia—has been central to research within the field of cognitive neuropsychiatry. Pending an explanation of the data that avoids reference to content, we should thus assume that it is indeed contentful. The body schema thus satisfies both conditions stipulated as jointly sufficient for representational status and, as such, should be considered a genuine representation.

5 Basic cognition and the body schema

We have argued that research into the development and maintenance of anorexia nervosa rests on an explanatorily robust concept of representation: the *body schema*. So what? Even the most intransigent anti-representationalists usually concede that *some* internal structures are rightly considered representations. What's so special about the body schema?

In this section, we argue that this seemingly recondite area of research is of central importance to the "representation wars". To see this, it will be helpful to briefly outline two obvious implications of the foregoing argument.

First, our arguments concerning the representational status of the body schema are not restricted to the domain of psychopathology. A central methodological practice in cognitive neuropsychology exploits evidence concerning *breakdowns* in cognitive functioning to reveal the mechanisms underlying cognitive processes in the neurotypical population more generally (Coltheart 2001). As such, if anorexia involves a pathological case of misrepresentation, this gives us strong evidence to believe that the relevant domain of cognition (see below) involves *veridical* representation in non-pathological cases.

Relatedly, although the research we have focused on predominantly concerns aperture-passability, its implications extend far beyond this narrow domain. Specif-

[3] As an anonymous reviewer points out, given that our argument is one of abduction—i.e. that the *best explanation* of the evidence posits satisfaction conditions—in order to counter this argument anti-representationalists would need to provide not just an alternative non-representational explanation but a better one, relative to some standards of explanatory power (e.g. simplicity, coherence, predictive power, consilience with other scientific research, etc.).

ically, they extend to the domains of environmental interaction and affordance perception more generally—the domain of "basic cognition" supposed to be most hospitable to a non-representationalist treatment (Hutto and Myin 2013; Gallagher 2017; Chemero 2009). This bears significant relevance to the "representation wars" of recent cognitive science (see Sect. 2).[4] In what follows, we briefly outline some prominent anti-representational accounts of two of these paradigmatically *basic* domains (environmental interaction and affordance perception) and show how, specifically, the body schema undermines such accounts.

5.1 Environmental interaction

Anti-representationalists generally consider environmental interaction to be a pervasive and integral aspect of our cognitive economy. Radical enactivists, for example, claim not only that there is "no prospect of understanding minds without reference to interactions between organisms and their environments" (Myin and Hutto 2015, 4) but indeed that "*most* of what humans do and experience is best understood in terms of dynamically unfolding interactions with the environment" (*ibid*., 61, our emphasis). This naturally includes the assumption that environmental interaction is a fundamentally *representation-free* process (see below).

Yet as we have just shown, environmental interactions are *not* necessarily unmediated by representation: our interactions with apertures are driven by representations of the subject's own body size. Further, as just noted, this insight is not restricted to the recondite domain of aperture passing. If we are right—if the body schema performs the systemic role that we have outlined in this paper—then a considerable portion of our environmental interactions must be similarly mediated by representations of body size.

This argument is substantiated by introspection and commonsense. Consider the broad range of goal-directed movements we make in a day: switching on the office light, turning on a computer monitor, reaching for a coffee cup, bringing it to our mouth for a sip, and so on. None of the motor commands for these actions can be calculated without relying on the spatial content of the relevant body parameters. For example, the motor commands necessary for reaching for coffee cups will markedly differ between subjects with different arm lengths. Indeed, it is difficult to see how *any* goal-directed interaction with one's environment could proceed without this relevant information. The spatial diversity of the objects in our environments, coupled with the spatial diversity of our action-relevant effectors (e.g. arms, legs, hands), positively requires that successful environmental interaction must take head of the dimensions of the body.

It is worth taking a moment to explore an influential argument *in favour* of the non-representational credentials of environmental interaction, in order to see how, exactly,

[4] Whilst different groups of anti-representationalists emphasize different activities, basic cognition is generally characterized as consisting of our capacities for "online" sensorimotor engagement with the environment e.g. learning, skilled action, environmental interaction, action-oriented perception (Dreyfus 2002; Gallagher 2017; Hutto and Myin 2013). This contrasts against "higher-order" capacities—such as language, thought, memory, planning etc.—which are generally regarded as "representation-hungry" (Clark and Toribio 1994).

the body schema undermines them. Enactivists fortify the non-representational credentials of this process via the notion of *dynamic couplings* (Gallagher 2017; Hutto and Myin 2013, p. 6). In contrast to linear relations holding between inputs and outputs, dynamic couplings consist of causal loops between organism and environment (Clark 1997, pp. 163–167). These loops are importantly *recurrent*, involving ongoing reciprocal influence, linking sensory input, motor output, and environment. Such looping relationships, it is argued, are not usefully characterised as involving *stand-ins* for those parts of the outer world the organism responds to (Van Gelder 1995). As such, the notion of representation never comes into consideration.

However, the body schema is exactly such a stand-in. Instead of an unbroken recurrent causal loop between subject and world, then, an intermediary step within this process *can* be identified—namely one which involves inner surrogates standing in for body size. The dynamic coupling is thus broken by a decidedly internal, decouplable surrogate; not of the world, but of the body.

5.2 Affordance perception

Whilst many anti-representationalists take environmental interactions as key to explaining the broader scope of cognition, others focus more closely on the notion of affordance perception. Gibson originally defined affordances as the actions provided or furnished by one's environment (1979, p. 127), of which passing through apertures is a prime example (Warren and Wang 1987). Perception of affordances was specifically contrasted with conventional, representational theories of perception (Richardson et al. 2008). As such, they became a popular focus point for many anti-representationalist accounts of cognition (Anderson 2014; Chemero 2009). Chemero, for example, claims that Gibson's ecological psychology—along with its principle that "perception is of affordances"—provides a "non-representational background theory" for the radical embodied approach to cognition (pp. 98–99). Echoing the popularity of affordance perception amongst non-representationalists, Zipoli Caiani writes, "conceiving affordances as dispositional properties allows for a lawful account of how agents perceive and interact with action-related properties of the environment, *without relying on mental representations* and computations" (2017, p. 665, our emphasis).

Yet knowledge of the cognitive and neurological systems which enable affordance processing has grown in leaps and bounds since Gibson first introduced the term. Today, we are in a significantly enlightened position, where the dominant assumption is that affordance processing is enabled through simulation of the actions themselves (Garbarini and Adenzato 2004; Declerck 2013, 2015; Zipoli Caiani 2017). As discussed, this same assumption is crucial in understanding the evidence from anorexia research—whereby it is suggested a distorted body schema, used by the motor system, figures in both action and affordance perception tasks, explaining why differences in critical point are evident in both.

Again, simple introspection supports the relevance of the spatial characteristics of the body schema to a considerable number of affordances. To see this, consider a defining claim of ecological psychologists: that affordances are relational, implicating not just properties of the environment but properties of the agent (Chemero 2003). In

particular, the spatial properties of an agent's body play a strong role in determining the affordances of an environment: a foot-stool is perceived as a seat for a child but not an adult *because* of their differences in size (Heft 1989, p. 3). As such, body size is a pervasive factor in determining affordances.

In the case of affordance perception, its non-representational credentials are borrowed from the notion of *direct perception*. Anti-representationalists claim that temporally invariant patterns in the sensory input itself are sufficient to specify affordances. As such, reconstructive—that is, re-presentational—perception of these environmental affordances is unnecessary: organisms can directly "resonate" to this environmentally embedded information (Anderson 2014; Chemero 2009; Gibson 1979; Noë 2004).

Irrespective of the legitimacy of this approach for cases of ordinary exteroceptive perception, however, we can now see why it fails in the case under discussion. Even if one thinks that some properties of the environment and agent *can* be directly tracked, such as the width of apertures (through vision) or the angle of joints (through proprioception), the spatial parameters of the body *cannot*—humans (and other animals) simply did not evolve afferent receptors dedicated to tracking the spatial parameters of the body. We can—and do—perceive these parameters (largely through vision), but this access is patchy at best, occurring sporadically and over long periods of time (de Vignemont 2014).

As such, factors like shoulder width, arm length, or the size of one's grip aperture are not dimensions one can simply *pick up* or *resonate to*, yet they are precisely the kinds of dimensions that determine affordances. Again, this point is buttressed by the evidence from anorexia research. Anorexia patients do not have a breakdown in their exteroceptive perceptual abilities, resulting in some kind of dysfunctional pickup of environmental variants pertaining to shoulder or aperture width (however that would work); rather, they *internally* misrepresent their own body size (Smeets 1997; Smeets et al. 1999; Spitoni et al. 2015; Keizer et al. 2011, 2012; Metral and Keizer 2017).

5.3 Summary

Whilst anti-representationalists tend to think of basic cognitive processes such as environmental interaction and affordance perception as representation-free, we have shown this assumption to be false. Cognition over such domains is heavily reliant on long-term body schema information regarding the spatial characteristics of the body. Furthermore, this information satisfies the anti-representationalists' own desiderata for being considered genuinely representational, putting considerable strain on their position within the representation wars.

6 Conclusion

We have drawn on research in cognitive neuropsychiatry concerning the malfunctioning of the body schema in patients with anorexia, and have argued that this body schema is robustly representational—that it passes Ramsey's job description challenge, and that its content is central to its explanatory role within such research. Further, we have

argued that this seemingly recondite area of research is of central relevance to the heated representation wars that have arisen in cognitive science in recent decades: it substantiates the view that representational cognition underlies environmental interaction and affordance perception in neurotypical subjects, the domains which are assumed to be most hospitable to purely representation-free explanations.

If our argument is along the right lines, it strongly undermines recent prophecies of a radical non-representational shift in cognitive theorizing (Hutto and Myin 2013; Ramsey 2007). Nevertheless, we acknowledge the limits of our argument. We do not claim—and have not demonstrated—that a non-representational interpretation of the evidence we have discussed in this paper is *impossible*, and we welcome attempts to accommodate this evidence in strictly representation-free terms. Any such non-representational treatment would need to account for the systematic differences between patients with anorexia and neurotypical subjects in the aperture passing-tasks we have outlined, as well as the ability to predict these differences using the outcomes of body size estimate tasks without appealing to an internal representation of the subjects' shoulder-width. Again, we do not claim that this cannot be done, but—for the reasons enumerated in Sect. 4—it seems unlikely.

Cognitive neuropsychology has enjoyed great success in understanding the mechanisms underlying our cognitive capacities by exploring cases in which they malfunction. We hope that the current paper has illustrated how fruitful this strategy can be with respect to questions about the existence and extent of cognitive representation. Specifically, the extensive and growing body of literature on psychopathology provides a rich and compelling cache of empirical data to explore in answering such questions, precisely because such research so often revolves around the notion of *misrepresentation*—whether through false belief, false perception, or otherwise. Misrepresentation has long been considered the key to understanding representation in general (Dretske 1986). If this is right, understanding may be gained by conducting debates surrounding representation in contexts where cognitive systems appear to get something *wrong*—to be in error relative to how things are in the world. We have exploited this strategy to reveal something important: the prospects of a global non-representationalism about basic cognition seem dim. Pathological misrepresentation of the body in patients with anorexia helps to reveal the extent to which affordance perception and environmental interactions are underpinned by a very specific kind of representation—a mirror not of nature, but of the subject's own body.

Acknowledgements We would like to thank the Cognition and Philosophy Lab at Monash University and two anonymous reviewers for their feedback. This research was supported by an Australian Government Research Training Program (RTP) Scholarship (S.G.) and the Arts and Humanities Research Council (D.W).

References

Alsmith, A. J. T., & de Vignemont, F. (2012). Embodying the mind and representing the body. *Review of Philosophy and Psychology, 3*(1), 1–13.

Anderson, M. (2014). *After phrenology*. Cambridge, MA: MIT Press.

Barrett, L. (2011). *Beyond the brain: How the body shapes the way we think*. New Jersey: Princeton University Press.

Bechtel, W. (2008). *Mental mechanisms*. Hoboken: Taylor and Francis.

Bermúdez, J. (2010). *Cognitive science*. Cambridge: Cambridge University Press.
Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence, 47*(1–3), 139–159. https://doi.org/10.1016/0004-3702(91)90053-m.
Bruch, H. (1962). Perceptual and conceptual disturbances in anorexia nervosa. *Psychosomatic Medicine, 24,* 187–194.
Chemero, A. (2003). An outline of a theory of affordances. *Ecological Psychology, 15*(2), 181–195.
Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, MA: MIT Press.
Clark, A. (1997). *Being there: Putting brain, body and world together again*. Cambridge, MA: MIT Press.
Clark, A. (2015). Predicting peace: The end of the representation wars—A reply to Michael Madary. In T. Metzinger & J. M. Windt (Eds.), *Open MIND: 7(R)*. Frankfurt am Main: MIND Group. https://doi.org/10.15502/9783958570979.
Clark, A., & Toribio, J. (1994). Doing without representing? *Synthese, 101*(3), 401–431.
Coltheart, M. (2001). Assumptions and methods in cognitive neuropsychology. In B. Rapp (Ed.), *The handbook of cognitive neuropsychology: What deficits reveal about the human mind* (pp. 3–21). New York, NY: Psychology Press.
de Vignemont, F. (2010). Body schema and body image—Pros and cons. *Neuropsychologia, 48*(3), 669–680.
de Vignemont, F. (2014). A multimodal conception of bodily awareness. *Mind, 123*(492), 989–1020.
Declerck, G. (2013). Why motor simulation cannot explain affordance perception. *Adaptive Behavior, 21*(4), 286–298.
Declerck, G. (2015). How we remember what we can do. *Socioaffective Neuroscience and Psychology, 5,* 24807.
Dretske, F. (1986). Misrepresentation. In R. Bogdan (Ed.), *Belief: Form, content, and function* (pp. 17–36). Oxford: Oxford University Press.
Dreyfus, H. L. (2002). Intelligence without representation–Merleau-Ponty's critique of mental representation the relevance of phenomenology to scientific explanation. *Phenomenology and the Cognitive Sciences, 1*(4), 367–383.
Engel, A., Friston, K., & Kragic, D. (2015). *The pragmatic turn* (1st ed.). Cambridge: MIT Press.
Engel, M. M., & Keizer, A. (2017). Body representation disturbances in visual perception and affordance perception persist in eating disorder patients after completing treatment. *Scientific Reports, 7*(1), 16184.
Fodor, J. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind*. Cambridge, MA: MIT Press.
Gadsby, S. (2017a). Anorexia nervosa and oversized experiences. *Philosophical Psychology, 30*(5), 594–615.
Gadsby, S. (2017b). Explaining body size beliefs in anorexia. *Cognitive Neuropsychiatry, 22*(6), 495–507.
Gadsby, S. (2017c). Distorted body representations in anorexia nervosa. *Consciousness and Cognition, 51,* 17–33.
Gadsby, S. (2018). How are the spatial characteristics of the body represented? A reply to Pitron & de Vignemont. *Consciousness and Cognition, 62*, 163–168.
Gallagher, S. (2017). *Rethinking the mind*. New York: Oxford University Press.
Gallagher, S., & Cole, J. (1995). Body image and body schema in a deafferented subject. *Journal of Mind and Behaviour, 16,* 369–389.
Garbarini, F., & Adenzato, M. (2004). At the root of embodied cognition: Cognitive science meets neurophysiology. *Brain and Cognition, 56*(1), 100–106.
Garzón, F. (2008). Towards a general theory of antirepresentationalism. *The British Journal for the Philosophy of Science, 59*(3), 259–292. https://doi.org/10.1093/bjps/axl007.
Gibson, J. (1979). *The ecological approach to visual perception* (1st ed.). Boston: Houghton Mifflin.
Grush, R. (1997). The architecture of representation. *Philosophical Psychology, 10*(1), 5–23.
Guardia, D., Conversy, L., Jardri, R., Lafargue, G., Thomas, P., Dodin, V., et al. (2012). Imagining one's own and someone else's body actions: Dissociation in anorexia nervosa. *PLoS ONE, 7*(8), e43241.
Guardia, D., Lafargue, G., Thomas, P., Dodin, V., Cottencin, O., & Luyat, M. (2010). Anticipation of body-scaled action is modified in anorexia nervosa. *Neuropsychologia, 48*(13), 3961–3966.
Haggard, P., & Wolpert, D. M. (2005). Disorders of body scheme. In H. J. Freund, M. Jeannerod, M. Hallett & R. Leiguarda (Eds.), *Higher-order motor disorders: From neuroanatomy and neurobiology to clinical neurology* (pp. 261–271).Oxford: Oxford University Press.
Halligan, P. W., & David, A. S. (2001). Cognitive neuropsychiatry: Towards a scientific psychopathology. *Nature Reviews Neuroscience, 2*(3), 209–215.

Haugeland, J. (1991). 'Representational genera. In W. Ramsey, S. Stich, & D. Rumelhart (Eds.), *Philosophy and connectionist theory* (pp. 61–89). Hillsdale, NJ: Lawrence Erlbaum.
Head, H., & Holmes, G. (1911). Sensory disturbances from cerebral lesions. *Brain, 34*(2–3), 102–254.
Heft, H. (1989). Affordances and the body: An intentional analysis of Gibson's ecological approach to visual perception. *Journal for the Theory of Social Behavior, 19,* 1–30.
Hutto, D., & Myin, E. (2013). *Radicalizing enactivism*. Cambridge, MA: MIT Press.
Hutto, D., & Satne, G. (2015). The natural origins of content. *Philosophia, 43*(3), 521–536. https://doi.org/10.1007/s11406-015-9644-0.
Isaac, A. (2013). Objective similarity and mental representation. *Australasian Journal of Philosophy, 91*(4), 683–704. https://doi.org/10.1080/00048402.2012.728233.
Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *NeuroImage, 14*(103), 109.
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.
Keizer, A., Smeets, M., Dijkerman, H., Uzunbajakau, S., Van Elburg, A., & Postma, A. (2013). Too fat to fit through the door: First evidence for disturbed body-scaled action in anorexia nervosa during locomotion. *PLoS ONE, 8*(5), e64602.
Keizer, A., Smeets, M., Dijkerman, H., Van den Hout, M., Klugkist, I., Van Elburg, A., et al. (2011). Tactile body image disturbances in anorexia nervosa. *Psychiatry Research, 190*(1), 115–120.
Keizer, A., Smeets, M., Dijkerman, H., Van Elburg, A., & Postma, A. (2012). Aberrant somatosensory perception in anorexia nervosa. *Psychiatry Research, 200,* 530–537.
Metral, M., Guardia, D., Bauwens, I., Guerraz, M., Lafargue, G., Cottencin, O., et al. (2014). Painfully thin but locked inside a fatter body: Abnormalities in both anticipation and execution of action in anorexia nervosa. *BMC Research Notes, 7*(1), 707.
Myin, E., & Hutto, D. D. (2015). REC: Just radical enough. *Studies in Logic, Grammar and Rhetoric, 41*(1), 61–71.
Noë, A. (2004). *Action in perception*. Cambridge, MA: MIT Press.
O'Shaughnessy, B. (1980). *The will: Dual aspect theory*. Cambridge: Cambridge University Press.
Paillard, J. (1999). Body schema and body image: A double dissociation in deafferented patients. In G. N. Gantchev, S. Mori & J. Massion (Eds.), *Motor control, today and tomorrow* (pp. 197–214). Sofia: Academic Publishing House.
Pickering, M. J., & Clark, A. (2014). Getting ahead: Forward models and their place in cognitive architecture. *Trends in Cognitive Sciences, 18*(9), 451–456.
Proske, U., & Gandevia, S. (2012). The proprioceptive senses: Their roles in signaling body shape, body position and movement, and muscle force. *Physiological Reviews, 92*(4), 1651–1697.
Quine, W. (1960). *Word and object* (1st ed.). Cambridge: Technology Press of the Massachusetts Institute of Technology.
Ramsey, W. (2007). *Representation reconsidered*. Cambridge: Cambridge University Press.
Rescorla, M. (2013). Bayesian perceptual psychology. In M. Matthen (Ed.), *Oxford handbook of the philosophy of perception*. Oxford, NY: Oxford University Press.
Richardson, M., Shockley, K., Fajen, B. R., Riley, M. A., & Turvey, M. (2008). Ecological psychology: Six principles for an embodied–embedded approach to behaviour. In P. Calvo & T. Gomila (Eds.), *Handbook of cognitive science: An embodied approach* (pp. 161–187). New York: Elsevier.
Rosenberg, A. (2015). The genealogy of content or the future of an illusion. *Philosophia, 43*(3), 537–547.
Ryle, G. (1949). *The concept of mind*. Chicago: University of Chicago Press.
Schwoebel, J., & Coslett, H. B. (2005). Evidence for multiple, distinct representations of the human body. *Journal of Cognitive Neuroscience, 17*, 543–553.
Shapiro, L. (2011). *Embodied cognition*. London: Routledge.
Smeets, M. (1997). The rise and fall of body size estimation research in anorexia nervosa: A review and reconceptualization. *European Eating Disorders Review, 5*(2), 75–95.
Smeets, M., Ingleby, J., Hoek, H., & Panhuysen, G. (1999). Body size perception in anorexia nervosa: A signal detection approach. *Journal of Psychosomatic Research, 46,* 465–477.
Spitoni, G., Serino, A., Cotugno, A., Mancini, F., Antonucci, G., & Pizzamiglio, L. (2015). The two dimensions of the body representation in women suffering from anorexia nervosa. *Psychiatry Research, 230*(2), 181–188.
Van Gelder, T. (1995). What might cognition be, if not computation? *Journal of Philosophy, 92*(7), 345–381. https://doi.org/10.2307/2941061.

Von Eckardt, B. (2012). The representational theory of mind. In K. Frankish & W. Ramsey (Eds.), *The Cambridge handbook of cognitive science* (1st ed., pp. 29–50). Cambridge: Cambridge University Press.

Warren, W., & Wang, S. (1987). Visual guidance of walking through apertures: Body-scaled information for affordances. *Journal of Experimental Psychology. Human Perception and Performance, 13*(3), 371–383.

Williams, D. (2017). Predictive processing and the representation wars. *Minds and Machines*. https://doi.org/10.1007/s11023-017-9441-6.

Williams, D. (2018). Pragmatism and the predictive mind. *Phenomenology and the Cognitive Sciences*. https://doi.org/10.1007/s11097-017-9556-5.

Williams, D., & Colling, L. J. (2017). From symbols to icons: The return of resemblance in the cognitive neuroscience revolution. *Synthese*. https://doi.org/10.1007/s11229-017-1578-6.

Wittgenstein, L., & Anscombe, G. (1953). *Philosophical investigations* (1st ed.). Oxford: Blackwell.

Wolpert, D., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience, 3*(Suppl), 1212–1217.

Wolpert, D. M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks, 11*(7), 1317–1329.

Zipoli Caiani, S. (2017). When the affordances disappear: Dynamical and computational explanations of optic ataxia. *Theory and Psychology, 27*(5), 663–682.

Linking Text Between Chapters 5 and 6

In the previous chapter, we addressed an issue of great importance to the chapters that preceded it, namely, the legitimacy of positing representations to explain the mind. We argued that the role of the (long-term) body schema in explaining difference in movement and affordance processing associated with anorexia nervosa justifies the use of mental representation, illustrating how representation plays an indispensable role in accounting for some of our most basic mental capacities. This matters for how we approach explaining mental disorders and the unusual beliefs associated with them. Many disorders involve unusual beliefs about the body, for example, that one's limb belongs to someone else (Cogliano et al., 2012), one has two heads (Ames, 1984), one's body is infested by parasites (Prakash et al., 2012), or one's body is dead (Young et al., 1994). Contrary to arguments from philosophers, the previous chapter illustrates the benefit of looking to the brain's representational mechanisms in order to explain this category of body-related unusual beliefs.

While the previous chapters focused on developing and expanding an account of false beliefs in eating disorders, in the next two chapters I turn to a distinct phenomenon: imposter syndrome. In the imposter syndrome, successful, intelligent, and capable individuals believe that their success is due to luck and fear that they will be exposed as imposters. There are some important similarities between the conditions. For example, both involve a failure of self-knowledge: in the case of eating disorders, about body size; in the case of imposter syndrome, about one's abilities. Similarly, both groups go to considerable lengths to overcome their self-perceived shortcomings. While people with eating disorders engage in severe weight loss behaviours, people with imposter syndrome work many times harder than necessary (Clance, 1985), causing themselves significant mental anguish and exhaustion in the process (Hutchins et al., 2018). As I will show, there are also important similarities in the causes of these conditions.

In chapter 1, I argued that the reasoning biases associated with eating disorders can be explained with reference to the "Friedrich-Trope-Liberman" account of hypothesis

testing. According to this account, people test hypotheses (such as whether they have reached their ideal size), in ways that minimise costly errors: false beliefs with negative consequences. Specifically, people seek out, attend to, and interpret evidence in ways that avoid costly errors. Within this framework, I argued that people with eating disorders avoid evidence in favour of the proposition "I am my ideal size", because falsely believing that proposition is represented as particularly costly. I also suggested that there is a motivational benefit to believing that one has not yet met or surpassed their ideal size: it ensures never slipping into complacency, allowing oneself to become overweight. This motivational element is consistent with reports from people with eating disorders who admit to selectively focusing on their "fat" body parts, in order to motivate themselves.

This suggestion presents an important idea: sometimes we are biased into holding negative beliefs about ourselves—e.g. "I am not thin enough"—because of the motivational benefits of doing so. It is understandable that such a phenomenon would occur in eating disorders: these individuals are incredibly motivated towards weight loss and go to extreme lengths to achieve their body size goals.

Note an important point regarding this argument, the mechanisms associated with biased reasoning in eating disorders are not dysfunctional. Unlike distorted body representations, these mechanisms are operating as they do in the neurotypical population—in ways that are sensitive to the idiosyncratic concerns of the individual. What causes biased reasoning are the exceedingly strong desires involved in eating disorders: a desire to achieve and maintain a thin body, no matter the cost. This strong desire, in conjunction with the usual mechanisms of hypothesis testing, causes the observed bias. This account bears an important prediction: other groups who hold similarly strong desires should exhibit similar reasoning biases, albeit in domains related to the content of their own desires. Imposter syndrome represents a way in which to test this prediction.

The following chapter evaluates the possibility that just as people with eating disorders are motivated by their desire to be thin, people with imposter syndrome are motivated by

their desire to succeed. In order to gain this benefit, I argue, both groups are biased towards believing something negative about themselves. As promised, this account draws some previously underappreciated links between these two seemingly dissimilar conditions.

AUSTRALASIAN JOURNAL OF PHILOSOPHY
https://doi.org/10.1080/00048402.2021.1874445

Imposter Syndrome and Self-Deception

Stephen Gadsby

Monash University

ABSTRACT

Many intelligent, capable, and successful individuals believe that their success is due to luck, and fear that they will someday be exposed as imposters. A puzzling feature of this phenomenon, commonly referred to as *imposter syndrome*, is that these same individuals treat evidence in ways that maintain their false beliefs and debilitating fears: they ignore and misattribute evidence of their own abilities, while readily accepting evidence in favour of their inadequacy. I propose a novel account of imposter syndrome as an instance of self-deception, whereby biased evidence treatment is driven by the motivational benefit of negative self-appraisal. This account illuminates a number of interconnected philosophical and scientific puzzles related to the explanation, definition, and value of imposter syndrome.

ARTICLE HISTORY Received 17 May 2020; Revised 13 November 2020

KEYWORDS imposter syndrome; imposter phenomenon; self-deception; rationality; motivated belief; belief-based utility

1. What Is Imposter Syndrome?

> Sarah is a philosophy student in a highly ranked Ph.D. program, supervised by a well-regarded professor. She received excellent grades throughout her undergraduate degree, which, along with glowing references from her teachers, earned her a spot in the prestigious programme. Sarah's supervisor regularly praises her work, and audience members respond positively to her talks. Despite this, Sarah considers herself to be less intelligent than the other graduate students. While they clearly have what it takes to be in the programme and to move into permanent positions, Sarah believes that her current success is largely due to luck.
>
> This situation causes considerable mental anguish for Sarah, who fears that her inadequacy will be exposed. She works incredibly hard, obsessing over the minor details of her written prose and presentation style. She will not be convinced of her own merit. In receiving written feedback from her supervisor, she hurriedly skims the positive comments, seeking out the few that indicate shortfalls. While reflecting on the feedback that she has received on her presentations, Sarah remembers few of the (mostly) positive comments but can vividly recall the tone of a derisive audience member who rebutted her during question time, many months ago.

The above vignette might strike readers as familiar. Perhaps you recognise this behaviour in someone you know; perhaps you even recognise it in yourself. This constellation of behaviours and attitudes is referred to colloquially as *imposter syndrome*, a widely observed phenomenon affecting individuals from all walks of life—most notably, college students and career academics [Langford and Clance 1993].

There is much that we do not know about imposter syndrome. Although the concept was introduced in the 1970s [Clance and Imes 1978], it has struggled to attract scientific attention [Bravata et al. 2019], and, until recently, has escaped the interest of philosophers [Hawley 2019a, 2019b; Paul 2019; Slank 2019]. Nevertheless, there are a number of intriguing and important philosophical issues related to the phenomenon.

First and foremost, the condition is epistemically puzzling: why does Sarah engage in faulty reasoning concerning her own abilities, especially when it causes her such mental anguish? Imposter syndrome appears to undermine the assumption that humans are, by and large, epistemically and pragmatically rational—that we reason in ways that track truth and promote well-being [Bortolotti 2014]. We are thus faced with the task of explaining why people with imposter syndrome deviate so significantly from these norms of rationality.

There are also definitional issues to resolve. Philosophers have made an excellent start on outlining potential necessary and sufficient conditions for imposter syndrome [Hawley 2019b; Paul 2019] and on drawing parallels between imposter syndrome and other apparent instances of irrationality [Hawley 2019a], yet there is still much to be done. For example, an outstanding question is that of how to categorise the phenomenon: among the myriad of known mental states, traits, and maladies, where does it fit?

Finally, there is the issue of value. Among scientists and philosophers, the condition is seen as overwhelmingly negative (in value). Yet none have questioned whether imposter syndrome might provide some (positive) value to those who exhibit it.

In this paper, I will make progress on the task of explaining, defining, and evaluating imposter syndrome. I will provide a novel account of the phenomenon, whereby people with imposter syndrome are motivated to downplay their own ability due to the motivational benefits of doing so. Consequently, imposter syndrome qualifies as an instance of self-deception. A corollary of this position is that imposter syndrome cannot be understood without considering the value that it holds for some individuals.

To be clear, I do not intend to provide an account of imposter syndrome that is true for all of those who exhibit it. One point of consensus from the scientific literature is that the category of imposter syndrome is heterogeneous: different individuals exhibit it for different reasons—psychological, developmental, and socio-cultural—and, for any particular individual, their condition might need to be explained with reference to multiple factors [Clance and Imes 1978; Harvey and Katz 1985; Sakulku and Alexander 2011; Leonhardt et al. 2017]. I will simply argue that, within the concept of imposter syndrome, we must make room for a self-deceptive variant.

2. Defining Imposter Syndrome

Imposter syndrome is not a syndrome in the common sense of the term: it involves neither disease nor disorder [Clance 1985: 23]. Clance and Imes [1978] were careful in giving it the more neutral label ‘imposter phenomenon’, in order to avoid this association. To have imposter syndrome is simply to exhibit a set of related emotions, attitudes, and behaviours. Here, following the most common definition in the psychological literature, I assume imposter syndrome to involve three features—affective, doxastic, and behavioural [Harvey and Katz 1985].

2.1 The Affective and Doxastic Features

The affective feature involves a fear of being discovered and exposed as an imposter—one who does not belong in the role that they occupy and who does not deserve the success that they achieve. As Leonhardt et al. [2017: 7] say, 'these individuals describe their greatest fear as being exposed as incompetent, less intelligent and thus as a phony.' This fear leads to a host of negative psychological consequences, such as anxiety, stress, and emotional exhaustion [ibid.].

Fear of being discovered is related to the doxastic feature of the condition, standardly characterised as a belief in one's own inadequacy [Clark et al. 2014]. Specifically, those who suffer from imposter syndrome believe that they lack *ability*. Such beliefs are not only false—they are just as able as, if not more able than, their peers—but unjustified, as people with imposter syndrome are exposed to significant evidence in favour of their own ability (see below).

A crucial point to note is that people with imposter syndrome do not hold negative beliefs regarding all of their role-relevant attributes. Many believe that they are charming, personable and, importantly, that they are hard workers [Hawley 2019b: 205]. Thus, imposter syndrome is not a case of domain general negative self-evaluation: inadequacy beliefs only refer to certain kinds of abilities (see section 5.2). Commonly (especially in the academic context), inadequacy beliefs refer to lacking intelligence, but the relevant properties can vary between different contexts and cases. For example, in her first-person account of imposter syndrome, Olberding [2018] describes believing that she 'lacked the standard cultural and class equipage of academe'.[1]

Some discussions of these inadequacy beliefs assume that they refer to 'innate' or 'fixed' abilities [Kumar and Jagacinski 2006; Paul 2019; Slank 2019]. While this is true much of the time, a better specification is that the abilities are difficult to acquire. Cultural equipage is not innate or fixed, but it is difficult to gain without a particular kind of upbringing.

2.2 The Behavioural Feature

People with imposter syndrome are exposed to significant evidence in favour of their abilities: positive test scores, awards, accolades from colleagues and, of course, the simple fact that they are in the position they are in (while others are not) all speak to the fact that they are suitably able. The most puzzling feature of the condition is this evidence's failure to exert the epistemic effect that one would expect. The solution to this puzzle lies in the behavioural feature of imposter syndrome.

Imposter syndrome is defined by a range of biases in the treatment of evidence related to the individual's abilities: 'Imposters … dismiss praise, derogate the accuracy of positive evaluations, and engage in other behaviors that insulate them against information that would validate their competence and worth' [Leary et al. 2000: 727]. The most notable of these biases pertains to the interpretation of evidence. Sarah regularly misattributes her own success, as do others with imposter syndrome [Clance and OToole 1987: 52]:

[1] Precisely what cultural and class equipage entails is difficult to pin down; it seems to point to a constellation of properties, such as the way one speaks (e.g. accent and vocabulary), the clothes one wears, and the knowledge one holds (e.g. who composed 'Ode to Joy') [ibid.].

> They attributed their successes to hard work, luck, knowing the right people, being in the right place at the right time, or to their interpersonal assets such as charm and the ability to relate well, rather than to ability or competence … These subjects were ingenious at negating objective external evidence that indicated they were indeed very bright.

Negative interpretation of success is accompanied by an inordinate focus on, and acceptance of, markers of failure. [ibid.].

On the standard account of imposter syndrome, explaining biased evidence treatment is crucial to explaining the condition. Fear of being exposed as an imposter stems from an unjustified belief in inadequacy, and this unjustified belief is adopted and maintained due to a biased treatment of the available evidence. To explain imposter syndrome, then, we must explain this behavioural feature.

3. Self-Deception

I will argue that some instances of imposter syndrome ought to be classified as self-deception. What is self-deception? Consider a common example:

> Jessica's teenage son has begun to act strangely. He comes home late at night with bloodshot eyes, smelling of marijuana. Despite evidence that her son might be smoking marijuana, Jessica will not accept it. She brushes off her husband's attempts to convince her of this truth. Upon finding various paraphernalia in his room, she concludes that these must have been left by a friend.

A key feature of self-deception is present in this example—biased treatment of the available evidence. In brushing off her husband's arguments and explaining away the paraphernalia, Jessica is contributing to the maintenance of her own false belief. According to the standard understanding of self-deception, her own desire for her son to be drug-free causes her to engage in these biased practises.

Consider another variant of self-deception:

> Chris loves his wife but believes that she is having an affair. Although she has excellent work-related reasons for arriving home late occasionally and no prior record of infidelity, Chris takes his wife's working late as evidence of an affair. He carefully attends to her social plans, interpreting slightly anomalous events as planned rendezvous with her lover.

Here the standard understanding falls apart: Chris does not desire that his wife is having an affair; in fact, he desires the opposite. Philosophers refer to this kind of self-deception as twisted [Mele 2001], dreadful [Van Leeuwen 2007], and negative [Funkhouser 2019], and there is a broad consensus that a complete account of self-deception ought to accommodate both these forms [Funkhouser 2019; Mele 2001]. I will argue that Chris and Sarah are alike in that each suffers from twisted self-deception.

3.1 Defining Self-Deception

Here, I assume a definition of self-deception as involving three conditions [Mele 2001: 120]. First, the self-deceived believe a proposition that is unwarranted by the available evidence.[2,3] Second, this unwarranted belief is maintained through biased evidence

[2] Note that 'available evidence' refers not to the evidence that an individual possesses, but to the evidence that is 'easily available' to them [Lynch 2012: 441]. This ensures that those who possess belief-warranting evidence in virtue of their biased evidence gathering (e.g. actively seeking out supporting evidence) might still qualify.

[3] I put aside the requirement that those who are self-deceived engage in behaviour that suggests that, at some level, they are aware that the relevant belief is false [Schwitzgebel 2002]. Philosophers generally accept that

treatment. This comes in many different forms [Funkhouser 2019: 116], the most common of which relate to: search—seeking out supporting (and avoiding contradictory) evidence—sampling—selectively attending to, and recalling from memory, supporting evidence—and interpretation—positively interpreting supporting evidence, while rationalising away contradictory evidence.

People with imposter syndrome satisfy both of these conditions: they believe in their own inadequacy (despite this belief being unwarranted by the available evidence), and they maintain this belief through biased evidence treatment—most notably, biased interpretation. Whether imposter syndrome qualifies as self-deception depends on its satisfaction of a third condition, related to the cause of these biases. Specifically, self-deception requires that the relevant biases are motivated—namely, driven by desires, emotions, or incentives of some form [Funkhouser 2019: 14].

One philosophically vexed aspect of self-deception involves characterising the precise form of motivational bias at play. Accounts of self-deception are divided into two broad camps over the issue—intentionalists and motivationalists. Modelling their accounts on interpersonal forms of deception, intentionalists claim that the self-deceived intentionally deceive themselves [Bermúdez 2000]. In contrast, motivationalists claim that, while self-deception is motivationally driven, it is unintentional [Mele 2001]. I will return to this distinction in section 5.4.

4. Belief-Based Utility and Evidence Treatment

In this paper, I will draw heavily from an account of self-deception from the economics literature, a guiding assumption of which is that 'people derive utility not only from possessions and experiences, but also from beliefs' [Golman et al. 2017: 128]. In their work, Bénabou and Tirole [2016] emphasize two categories of this belief-based utility. The first category is affective: beliefs about ourselves and our prospects have a 'direct and powerful affective impact' [ibid.: 143]. Seeing ourselves and our prospects in a positive light is satisfying (utility), while seeing ourselves and our prospects in a negative light causes sadness and worry (disutility). The second category is motivational: 'confidence in one's ability and chances of success (or those of teammates) can be a powerful motivator to pursue difficult long-term goals and persevere through adversity' [Bénabou 2015: 6]. For example, the belief that one is almost at the end of a difficult task motivates one to persevere.

What are the effects of belief-based utility? Crucially, they are not direct: we do not simply revise our beliefs on consideration of the relevant utility. Instead, belief-based utility biases our evidence treatment practises—in the realms of evidence search, sampling, and interpretation—such that our (future) selves are guided towards beliefs with higher utility [Brocas and Carrillo 2000; Sharot and Sunstein 2020]. According to this framework, the biased evidence treatment associated with self-deception stems from attempts to maximize belief-based utility.

Consider how this framework applies to the aforementioned examples of self-deception. We can explain Jessica's behaviour by assuming that she is biased away from the affective disutility (anxiety) associated with believing that her beloved son

there are some cases of self-deception where the self-deceived show all of the markers of belief in the proposition, with no indication that they believe the opposite [Mele 2001; Lynch 2012; Funkhouser 2019]. Thus, I will focus on this more straightforward notion of self-deception.

is smoking marijuana (see Barnes [1997]). In Chris's case, believing that his wife is unfaithful may be perceived as providing a certain motivational utility. As Mele (borrowing from Pears [1984]) puts it [2001: 95]:

> the value of his jealousy-inspired belief in his wife's infidelity lies in its capacity in combination with his desire for her fidelity, to lead him to take steps to reduce the chance that she will have affairs by, for example, increasing his vigilance.

This excessive vigilance may be perceived as necessary to maintain or control the relationship, even though (paradoxically) it might drive the wife away [ibid.: 101].

It is important to note that economic accounts of self-deception do not assume that the self-deceived are aware of their bias. Indeed, as Bénabou and Tirole write [2016: 147],

> the process of manipulating one's own attention, memory, or awareness must not be too transparent. There must be some opaqueness as to what exactly one is failing to update to, some ambiguity as to why certain actions are taken or not taken.

This framework is thus consistent with recent accounts of self-deception that argue for the empirical plausibility of unconscious self-deceptive processes [Funkhouser and Barrett 2016].

5. Imposter Syndrome: Maladaptive or Motivational?

Some psychologists suggest that imposter syndrome stems from problematic parenting styles or family dynamics. As Langford and Clance [1993: 497] write, 'people who experience impostor feelings are likely to come from families in which support for the individual is lacking, communications and behaviors are controlled by rules, and considerable conflict is present' (see also Clance and Imes [1978] and Thompson [2004]). Accordingly, 'To truly understand the Impostor Phenomenon … it's essential to start at the beginning—with the Impostor's family' [Clance 1985: 465]. On these accounts, imposter syndrome is neither intentional nor motivated, as those who exhibit it are simply acting out maladaptive reasoning styles that were internalised during childhood [ibid.: 55–62]. In contrast, I argue that imposter syndrome, in many cases, is driven by the benefits of believing in one's own inadequacy. In order to understand this form of imposter syndrome, we must consider the utility of inadequacy beliefs, along with the factors (psychological and situational) that determine that utility.

Consider the utility associated with believing that one lacks ability, as individuals with imposter syndrome do. There is obvious disutility involved, in terms of the negative affect that stems from such a belief. As Bénabou and Tirole [2016: 143] note, holding negative beliefs about ourselves (that we lack ability) and our prospects (that we are likely to fail) causes distress and anxiety. Thus, all things being equal, we should be biased away from such beliefs. This is precisely what the psychological literature suggests, as there is a general tendency towards positive evaluation of own abilities [Taylor and Brown 1988]

In the case of imposter syndrome, this negative utility is especially potent. People with imposter syndrome do not simply believe that they lack ability; they believe that they lack the ability that is crucial for succeeding in a greatly valued role. Not only do they suffer from this belief, they suffer from the resulting fear of their inadequacy being discovered and exposed. When considered in light of affective disutility,

the biased evidence treatment associated with imposter syndrome appears maladaptive, undermining these individuals' own basic desire for happiness. When faced with explaining behavioural patterns that work against someone's own interests, psychologists commonly look to the person's history. Thus, it is unsurprising that many psychologists have assumed a relationship between imposter syndrome and childhood experiences.

In contrast to these accounts, I argue that there is an overlooked benefit of the condition. Specifically, I argue that beliefs in low ability bestow a motivational benefit, one which is particularly attractive for those who wish to succeed in contexts where the pathway to success is both challenging and opaque. In what follows, I discuss a number of preconditions that (jointly) cause a belief in low ability to provide high motivational utility. First, the individual strongly desires to succeed: this ensures that any belief that aids in success will provide high utility. Second, the individual believes that the domain is one where considerable effort is required to succeed: this ensures that beliefs that motivate effort provide high utility. Third, the domain is one where the individual believes that effort can substitute ability, such that a belief that one lacks ability will have a motivating (rather than demotivating) effect. Finally, the domain is one where accurate belief provides low utility, rendering self-deception more advantageous.

As I will show, these conditions are commonly found in contexts where imposter syndrome thrives. Following the emphasis in the scientific literature, I will develop this argument with reference to the academic context. Nevertheless, I contend that such conditions will hold in other contexts where imposter syndrome is prevalent.

5.1 Desire for Success where Considerable Effort is Necessary

The first precondition is that people with self-deceptive imposter syndrome strongly desire to succeed. This claim is uncontroversial. Descriptions of imposter syndrome from the scientific literature regularly describe these individuals' 'strong desire … to be the very best among their peers' [Schubert and Bowker 2019: 749]. The presence of such a desire is also consistent with evidence that imposter syndrome highly correlates with success driven traits such as fear of failure and perfectionism [Clance and OToole 1987; Kumar and Jagacinski 2006; Sakulku and Alexander 2011].

The second precondition is that people with self-deceptive imposter syndrome believe that, natural abilities notwithstanding, succeeding requires considerable effort. This ensures that beliefs that motivate effort will provide high utility. Considerable effort is necessary in contexts where competition is high and succeeding is difficult, such as those where imposter syndrome is commonly found. Consider the context of academia. Succeeding in this domain is exceedingly difficult: research suggests that only 3.5% of Ph.D. students secure a permanent position at a university and, of those, only 12% become professors [Taylor et al 2010: 14]. This ensures that even those who are bright and talented do not succeed. In challenging domains such as these, substantial effort is required, which people with self-deceptive imposter syndrome are aware of.

5.2 Effort and Ability are Substitutes

The third precondition is the belief that effort can substitute ability. Bénabou and Tirole [2002: 873] point out that, in most contexts, ability and effort are

complementary: they combine multiplicatively to determine outcomes. Due to this, belief in our own ability has a positive motivational effect. In scenarios where ability and effort are complementary, if one believed that one has little or no ability, then even a significant amount of effort would provide one with little return. If I believe that I am entirely hopeless at chess then—fearing a short, embarrassing, and unfulfilling game—I might not bother playing at all. Believing that I am somewhat capable will prompt me to play, under the assumption that, with a little effort, I might win (or at least be competitive). Thus, in many cases, positive belief in one's own abilities is motivational. Because those beliefs provide this utility, we are biased towards them.

Nevertheless, in some contexts, negative self-appraisal has a positive motivational effect [ibid.: 904]:

> A student preparing for exams may … discount his previous good grades as attributable to luck or lack of difficulty. A young researcher may understate the value of his prior achievements, compared with what will be required to obtain tenure. A dieting person who lost a moderate amount of weight may decide that he 'looks fatter than ever,' no matter what others or the scale may say.

In such scenarios, we exaggerate the difficulty of the task at hand in order to steer ourselves away from 'coasting' or 'slacking off'. This phenomenon appears in a number of different contexts. For example, Norem and Cantor provide evidence that, when facing tasks with a high possibility of failure, many students dwell on how unprepared they are 'in order to get [themselves] to work harder' [1986: 1213]. Similarly, evidence suggests that people with eating disorders visually attend to the parts of their bodies that appear 'fatter', in order to motivate their own dietary restraint [Gadsby 2020: 616].

In certain contexts, negative self-appraisal provides motivational utility by signalling to ourselves the need for effort. When negative self-appraisal is related to ability, motivational benefit will only occur if the individual believes that effort and ability are potential substitutes [Bénabou and Tirole 2016: 145].[4] When we believe that effort can substitute for ability, negative appraisal of our own ability can be motivationally beneficial. There is good reason to think that people with imposter syndrome believe that this is possible. After all, they devote considerable time and effort towards achieving their goals, despite believing that they lack ability. In fact, such efforts are assumed to result from these beliefs [Yaffe 2020: 1].[5] This intense diligence and hard work also pays off, in terms of 'excellent performance and approval from authorities', precisely the forms of compensation required to succeed [Clance and Imes 1978: 244].

The effort that these individuals devote cannot be explained exclusively through reference to a desire for success. Desire does not drive effort when one believes that one lacks the necessary ability, and that ability cannot be substituted for by effort. No matter how strongly I desire to be a fashion model, I know that I am simply not blessed with the right genetics: no amount of make-up, hair styling, or gruelling gym sessions will compensate for that fact. I do not bother with these activities

[4] For example, this condition often holds in domains where the rewards are of a pass-fail nature, such as 'graduating from school, making a sale, being hired or fired (tenure, partnership), proposing marriage, etc.' [Bénabou and Tirole 2002: 905]. As anyone who has underachieved throughout high school will recognise, believing that you are terrible at a subject but needing to pass it in order to graduate is a wonderful motivator for studying.
[5] In fact, some psychologists suggest that the motivational effect of inadequacy beliefs is two-fold: first, people with imposter syndrome work harder to overcome their lack of ability; second, they work harder so as not to have their lack of ability discovered [Clance and Imes 1978: 244].

because I know that they cannot compensate for my genetic shortfalls. This explains why, despite believing that they lack ability, people with imposter syndrome do not simply give up; those who believe that they lack ability, but that they can make up for it with effort, persist in pursuing their goals.

While those with imposter syndrome believe that they lack attributes that would make succeeding easier—intelligence, brilliance, talent, cultural equipage—they do not hold negative beliefs about all of their attributes: they commonly believe themselves to be hard workers [Hawley 2019b: 205]. Believing that they did not possess the capacity to work hard would undermine the motivational benefit of believing that the task at hand is difficult to achieve. This helps to define the inadequacy beliefs associated with self-deceptive imposter syndrome: those who exhibit it believe that they lack the kinds of abilities that can be substituted for with effort.

5.3 Knowledge of Ability Provides Low Utility

The final precondition for self-deceptive imposter syndrome is that knowledge of one's own ability provides low utility. In most contexts, knowledge of our abilities is useful for achieving our goals because it aids in planning. If my goal is to make it into the local basketball team, then an accurate assessment of my basketball skills will help. For example, if I know that my offensive game is unparallelled, but that I cannot shoot free throws, then I know what I need to work on. I can focus my efforts on improving my free throw ability, up to a level at which the local team will allow me to join. I will not waste resources (time and effort) practising offensive skills, because I know that this strategy will not pay off.

However, in some domains, knowledge of our abilities is not so useful. This is the case when the pathway to success is particularly opaque. Consider, again, the academic context. What it takes to succeed in this domain is overwhelmingly difficult to discern, as is evident in the following anecdote from a junior academic [Anonymous Academic 2018]:

> I visited the university's HR advisor and asked her—naively—what I would have to do to eventually be appointed professor. Her answer was frank: 'I have no idea'. My group leader, a professor, couldn't help either and told me that 'there is no clearly defined path that will get you there'. The university agreed to investigate the issue, and later introduced a set of criteria to define what was expected of academics at different levels. But those criteria were set so high that it was impossible for anybody to achieve them—including the professors themselves.

One might respond that, while there might not be a precisely laid-out pathway to success in academia, most academics still possess a (rough) idea of what is required —publication in leading journals, citations, respect from colleagues, and the like. The important point, however, is that it is unclear which specific activities are those to which one should dedicate time to in order to reap such benefits; if it were otherwise, our hard drives and notebooks would not be so full of half-finished projects and abandoned ideas.

While, in most scenarios, knowledge of one's abilities is useful for planning, this is only the case when the necessary actions are known (if I know what I need to do, I know whether it is the kind of thing of which I am capable). In situations where the pathway to success is opaque, accurate knowledge of one's own abilities does not grant this advantage; even perfect self-knowledge cannot overcome not knowing

what needs to be done. This point is crucial. As economic models of self-deception stress, when there is no utility provided by holding accurate beliefs, alternative sources of utility play a stronger role in biasing evidence treatment [Brunnermeier and Parker 2005].[6]

5.4 Summary

Self-deceptive imposter syndrome involves a strong desire to succeed in a challenging domain, where the pathway to success is opaque. This ensures that people are willing to exert considerable effort to succeed and, given the difficulty of succeeding, this is precisely what is required. It also ensures that knowledge of one's own abilities is less useful for achieving this goal, as it does not aid in planning. In these domains, those who believe that ability can be substituted for with effort derive a motivational benefit from believing that they lack ability. For such individuals, this additional motivation is regarded as advantageous, as it brings them closer to the success that they so strongly desire. For individuals who satisfy these conditions, the motivational utility provided by a belief in lacking ability will outweigh the relevant affective disutility. According to economic models of self-deception, such individuals will seek out, attend to, remember, and interpret evidence in ways that reinforce this belief [Bénabou and Tirole 2016].

According to the proposed account, some who suffer from imposter syndrome do so for the same reason that Jessica believes that her son is not smoking marijuana and that Chris believes that his wife is having an affair—namely, because the utility provided by those beliefs outweighs their disutility. Consequently—and in contrast to the family dynamics account—imposter syndrome is explained with reference to belief-based utility and the biases that it produces. Not only does this explain the biased way in which people with imposter syndrome interpret evidence, it predicts that they will exhibit other forms of bias. For example, people with self-deceptive imposter syndrome might exhibit bias in the realms of evidence search and sampling, whereby they will predominately seek out, attend to, and recall evidence that supports their inadequacy beliefs (as described in the example of Sarah).

While this proposal provides one explanation for the behavioural element of imposter syndrome, it also illustrates how these practices qualify as motivated, since these individuals' own desire to succeed biases them towards the belief that they are inadequate, due to its motivational utility. Thus, the account illustrates how those who exhibit imposter syndrome might qualify as self-deceived, according to motivationalist accounts.

It is also consistent with stricter accounts of self-deception. For example, some philosophers hold that self-deception must be not only motivated, but motivated by a desire to believe the relevant proposition [Nelkin 2002; Funkhouser 2005]. Crucially, these accounts do not insist that the self-deceived are aware of, or control, this desire. As Nelkin [2002: 395] writes, 'the desire to believe that p is true need not be conscious. It is likely not to be actively contemplated at the time during which the

[6] Of course, this does not entail that knowledge of one's own abilities provides no utility whatsoever; it might provide many sources of utility (for example, when it is positive, it will cause happiness). The important point, however, is that, in contexts where imposter syndrome is prevalent, knowledge of one's own abilities does not make succeeding easier.

belief that p is formed, and it might even be difficult for an agent to recognize.' According to the proposed account of self-deceptive imposter syndrome, those who exhibit it unconsciously desire to believe in their own inadequacy, due to its motivational benefits.

Is this proposal consistent with intentionalism? Do those with self-deceptive imposter syndrome intend to deceive themselves into believing that they are inadequate? This depends on what one considers to be entailed by intention. The biasing of our evidence treatment practices certainly appears purposive, as it guides our future selves towards beliefs with high utility (see Funkhouser [2019: 62]). If intentionalists are willing to concede that self-deception can occur as a relatively automatic and unconscious process (unlike most forms of intention), then the proposed model is consistent with intentionalism. Some intentionalists are willing to concede this possibility [Bermúdez 2000], but, at that point, one might argue that the line between motivationalism and intentionalism has become too blurry (see Funkhouser [2019: 67]).

I will not take a stand on which account of self-deception is correct, or is most befitting of the above proposal. There are surprising difficulties in teasing apart different accounts of self-deception, and many philosophers point to the same processes as either being intentional or (non-intentionally) motivated [ibid.]. To adequately adjudicate the different issues at play here is beyond the scope of this paper.

I have argued for one particular form of self-deceptive imposter syndrome in this paper, but I leave open the possibility of alternative forms. While I have focused on the motivational utility derived from negative self-appraisal, there might be other, as-of-yet, underappreciated forms of utility provided by such beliefs [Sharot and Sunstein 2020]. There might even be forms of self-deceptive imposter syndrome that are consistent with family dynamic accounts. Childhood experiences might directly result in the adoption of maladaptive reasoning styles, but they might also operate via the mechanisms of motivated reasoning, through instilling a desire to believe in one's own inadequacy. According to such an account, the relevant motivational biases would still need to be explained with reference to these experiences, rather than to belief-based utility alone.[7]

Before concluding, I will briefly address a challenge sometimes pitched against belief-based utility accounts of self-deception [Pinker 2011].[8] The challenge, applied here, is to explain why people who desire success (in challenging and opaque domains) do not simply modify their behaviour, so as to exert as much effort as possible. Why would they self-deceive when this simpler route is available? In order to support my account, I need not answer this question. As discussed, research shows that some students focus on their lack of preparedness before exams and that some people with eating disorders focus on the 'fat' parts of their bodies. Crucially, both these groups claim that they engage in these practices in order to gain motivational benefits. So, people do engage in biased evidence treatment in order to motivate themselves. My argument is simply that self-deceptive imposter syndrome is another instance of this phenomenon; I need not also explain why the phenomenon occurs.

[7] This would constitute a form of self-deception that is importantly different from the more commonly discussed varieties, introduced in section 3. Thanks to an anonymous reviewer for pointing out this possibility.
[8] Thanks to Eric Funkhouser for bringing this argument to my attention.

6. Conclusion

In this paper, I made progress on the task of explaining, defining, and evaluating imposter syndrome. I provided a novel explanation of a sub-class of imposter syndrome, as driven by the motivational utility provided by negative self-appraisal, within contexts where certain preconditions hold. This contrasts with psychological accounts that assume that the condition is maladaptive, resulting from problematic family dynamics. It also contrasts with contemporary philosophical accounts, which emphasize the role of epistemic obstacles, arising from social factors, as driving imposter syndrome [Hawley 2019b; Paul 2019; Slank 2019]. Nevertheless, I do not claim to contradict these accounts. As I have emphasized, imposter syndrome is a heterogenous phenomenon: different individuals exhibit it for different reasons, and different factors might drive it for the same individual. The proposed account opens the door for future research into how these different accounts connect and conflict.

How we explain the cause of imposter syndrome bears on how we define it. My argument shows that, within the category of imposter syndrome, we must make room for a self-deceptive variant. This identifies an unrecognised, yet remarkably widespread example of self-deception, one of which many philosophers and psychologists have first-hand experience. In turn, this opens the door for the use of first-hand experience as well as the literature on imposter syndrome to inform philosophical debates regarding self-deception, helping to resolve some of the many disputes present in this literature.

One question for future research pertains to the prevalence of self-deceptive imposter syndrome: what proportion of those who suffer from imposter syndrome are self-deceived, compared to, say, those who are acting out maladaptive behavioural patterns learnt during childhood? While this is an open question, I note that factors related to parenting style and family dynamics have only been found to correlate weakly with measures of imposter syndrome [Sakulku and Alexander 2011: 80–2]. In contrast, some preconditions that I have stipulated as driving self-deceptive imposter syndrome (for example, perfectionism) highly correlate [ibid.: 86]. Given these findings, it might be the case that self-deceptive imposter syndrome is as common as, if not more so than, forms of imposter syndrome that stem from problematic childhood experiences.

Finally, the proposed account introduces an important consideration for how we attempt to treat imposter syndrome. If imposter syndrome is underpinned by motivational bias, then motivations ought to be a target for treatment. This could be approached psychologically—by working with sufferers of imposter syndrome to reflect on and re-assess the utility provided by inadequacy beliefs—or situationally—by working to change the conditions that determine that utility.

However, we might also question whether imposter syndrome necessarily ought to be treated. In the psychological and philosophical literature, this point is undisputed: while researchers disagree on how imposter syndrome should be treated, they all assume that it should be. Yet, on the proposed account, there is significant value to the condition.[9] Perhaps, then, for those facing situations where the pathway to success is both challenging and opaque, but also considerably desirable, imposter syndrome is not something that should be treated. Perhaps such individuals should not

[9] This is not to say that imposter syndrome is, overall, epistemically good, only that it has underappreciated benefits (see Bortolotti [2015]).

seek to rid themselves of the relevant beliefs, but rather should mitigate the beliefs' affective disutility. Such a possibility is broached by Olberding [2018], who suggests that, rather than trying to reject the imposter label, we simply embrace it, stripping it of the associated low self-esteem and the fear of being exposed.[10] After all, being a hard worker who does not rely on natural ability to succeed is an identity that one ought to embrace, rather than one of which to be ashamed. Perhaps those with self-deceptive imposter syndrome should strive for the state in which Olberding now finds herself: 'I sometimes still feel a fraud in academic environments, but neither do I mind it much.'[11]

Disclosure Statement

No potential conflict of interest was reported by the author(s).

Funding

This research was supported by an Australian Government Research Training Program (RTP) Scholarship.

ORCID

Stephen Gadsby http://orcid.org/0000-0002-2302-844X

References

Anonymous Academic 2018. The Path to Professorship Is Long, Rocky and Confusing. Why? *The Guardian*. URL = https://www.theguardian.com/higher-education-network/2018/aug/24/the-path-to-professorship-is-long-rocky-and-confusing-why

Barnes, A. 1997. *Seeing through Self-Deception*, Cambridge: Cambridge University Press.

Bénabou, R. 2015. The Economics of Motivated Beliefs, *Revue d'économie politique* 125/5: 665–85.

Bénabou, R. and J. Tirole 2002. Self-Confidence and Personal Motivation, *The Quarterly Journal of Economics* 117/3: 871–915.

Bénabou, R. and J. Tirole 2016. Mindful Economics: The Production, Consumption, and Value of Beliefs, *Journal of Economic Perspectives* 30/3: 141–64.

Bermúdez, J.L. 2000. Self-Deception, Intentions, and Contradictory Beliefs, *Analysis* 60/4: 309–19.

Bortolotti, L. 2014. *Irrationality*, Cambridge: Polity Press.

Bortolotti, L. 2015. The Epistemic Innocence of Motivated Delusions, *Consciousness and Cognition* 33: 490–9.

Bravata, D., S.A. Watts, A.L. Keefer, D.K. Madhusudhan, K.T. Taylor, D.M. Clark, R.S. Nelson, K.O. Cokley, and H.K. Hagg 2019. Prevalence, Predictors, and Treatment of Impostor Syndrome: A Systematic Review, *Journal of General Internal Medicine* 35/4: 1252–75.

[10] If psychologists are correct in saying that some of the motivational utility of imposter syndrome stems from the fear of being discovered (see note 5), then this strategy might reduce motivational utility. Nevertheless, this may be an acceptable trade off, given the decrease in affective disutility.

[11] I would like to thank Simon van Baal, Tim Bayne, Andrew Corcoran, Manja Engel, Eric Funkhouser, Jakob Hohwy, Alex Kiefer, Andy McKilliam, Dan Williams, Iwan Williams, Jennifer Windt, and two anonymous reviewers for helpful discussion and/or written feedback on the manuscript.

Brocas, I. and J.D. Carrillo 2000. The Value of Information when Preferences are Dynamically Inconsistent, *European Economic Review* 44/4–6: 1104–15.
Brunnermeier, M.K. and J.A. Parker 2005. Optimal Expectations, *American Economic Review* 95/4: 1092–118.
Clance, P.R. 1985. *The Impostor Phenomenon: Overcoming the Fear that Haunts your Success*, Atlanta, GA: Peachtree [kindle version].
Clance, P.R. and S.A. Imes 1978. The Imposter Phenomenon in High Achieving Women: Dynamics and Therapeutic Intervention, *Psychotherapy: Theory, Research and Practice* 15/3: 241–7.
Clance, P.R. and M.A O'Toole 1987. The Imposter Phenomenon: An Internal Barrier to Empowerment and Achievement, *Women and Therapy* 6/3: 51–64.
Clark, M., K. Vardeman, and S. Barba 2014. Perceived Inadequacy: A Study of the Imposter Phenomenon among College and Research Librarians, *College and Research Libraries* 75/3: 255–71.
Funkhouser, E. 2005. Do the Self-Deceived Get What They Want? *Pacific Philosophical Quarterly* 86/3: 295–312.
Funkhouser, E. 2019. *Self-Deception*, Abingdon: Routledge.
Funkhouser, E. and D. Barrett 2016. Robust, Unconscious Self-Deception: Strategic and Flexible, *Philosophical Psychology* 29/5: 682–96.
Gadsby, S. 2020. Self-Deception and the Second Factor: How Desire Causes Delusion in Anorexia Nervosa, *Erkenntnis* 85/3: 609–26.
Golman, R., D. Hagmann, and G. Loewenstein 2017. Information Avoidance, *Journal of Economic Literature* 55/1: 96–135.
Harvey, J.C. and C. Katz 1985. *If I'm so Successful, Why do I Feel Like a Fake? The Impostor Phenomenon*, New York: St. Martin's Press.
Hawley, K. 2019a. Conspiracy Theories, Impostor Syndrome, and Distrust, *Philosophical Studies* 176/4: 969–80.
Hawley, K. 2019b. What Is Impostor Syndrome? *Proceedings of the Aristotelian Society Supp. Vol.* 93/1: 203–26.
Kumar, S. and C.M Jagacinski 2006. Imposters Have Goals Too: The Imposter Phenomenon and its Relationship to Achievement Goal Theory, *Personality and Individual Differences* 40/1: 147–57.
Langford, J. and P.R. Clance 1993. The Imposter Phenomenon: Recent Research Findings Regarding Dynamics, Personality and Family Patterns and their Implications for Treatment, *Psychotherapy: Theory, Research, Practice, Training*, 30/3: 495–501.
Leary, M.R., K.M. Patton, A.E. Orlando, and W. Wagoner Funk 2000. The Impostor Phenomenon: Self-Perceptions, Reflected Appraisals, and Interpersonal Strategies, *Journal of Personality* 68/4: 725–56.
Leonhardt, M., M.N. Bechtoldt, and S. Rohrmann 2017. All Impostors Aren't Alike–Differentiating the Impostor Phenomenon, *Frontiers in Psychology* 8: 1505.
Lynch, K. 2012. On the 'Tension' Inherent in Self-Deception, *Philosophical Psychology* 25/3: 433–50.
Mele, A.R. 2001. *Self-Deception Unmasked*, Princeton: Princeton University Press.
Nelkin, D.K. 2002. Self–Deception, Motivation, and the Desire to Believe, *Pacific Philosophical Quarterly* 83/4: 384–406.
Norem, J.K. and N. Cantor 1986. Defensive Pessimism: Harnessing Anxiety as Motivation, *Journal of Personality and Social Psychology* 51/6: 1208–17.
Olberding, A. 2018. The Outsider, *Aeon* URL = https://aeon.co/essays/how-useful-is-impostor-syndrome-in-academia
Paul, S. K. 2019. What Should 'Impostor Syndrome' Be? *Proceedings of the Aristotelian Society Supp. Vol.* 93/1: 227–45.
Pears, D.F. 1984. *Motivated Irrationality*, Oxford: Clarendon Press.
Pinker, S. 2011. Representations and Decision Rules in the Theory of Self-Deception, *Behavioral and Brain Sciences* 34/1: 35–7.
Sakulku, J. and J. Alexander 2011. The Impostor Phenomenon, *The Journal of Behavioral Science* 6/1: 75–97.
Schubert, N. and A. Bowker 2019. Examining the Impostor Phenomenon in Relation to Self-Esteem Level and Self-Esteem Instability, *Current Psychology* 38/3: 749–55.
Schwitzgebel, E. 2002. A Phenomenal, Dispositional Account of Belief, *Noûs* 36/2: 249–75.
Sharot, T. and C.R. Sunstein 2020. How People Decide What They Want to Know, *Nature Human Behaviour* 4: 14–19.

Slank, S. 2019. Rethinking the Imposter Phenomenon, *Ethical Theory and Moral Practice* 22/1: 205–18.
Taylor, M., B. Martin, and J. Wilsdon 2010. The Scientific Century: Securing our Future Prosperity, *The Royal Society*. URL = https://royalsociety.org/topics-policy/publications/2010/scientific-century
Taylor, S.E. and J.D. Brown 1988. Illusion and Well-Being: A Social Psychological Perspective on Mental Health, *Psychological Bulletin* 103/2: 193–210.
Thompson, T. 2004. Failure-Avoidance: Parenting, the Achievement Environment of the Home and Strategies for Reduction, *Learning and Instruction* 14/1: 3–26.
Van Leeuwen, D.S.N. 2007. The Product of Self-Deception, *Erkenntnis* 67/3: 419–37.
Yaffe, Y. 2020. Does Self-Esteem Mediate the Association between Parenting Styles and Imposter Feelings among Female Education Students? *Personality and Individual Differences* 156: 109789.

Linking Text Between Chapters 6 and 7

In the previous chapter, I discussed the phenomenon of imposter syndrome, along with the fears, beliefs, and reasoning biases it involves. I developed an account of the condition within a framework of motivated reasoning derived from behavioural economics. On this account, many who exhibit imposter syndrome negatively evaluate themselves due to the motivational benefits of doing so. Accordingly, imposter syndrome qualifies as a widespread and theoretically important instance of self-deception, one that exhibits important similarities to eating disorders, in terms of the motivational biases involved.

This thesis started with the supposition that people with imposter syndrome reason in ways that undermine their own basic desire for happiness, causing themselves considerable stress and anxiety, for no apparent gain. Such reasoning violates norms of pragmatic rationality, which dictate that we reason in ways that promote our own interests and well-being. In contrast, my account vindicates the pragmatic rationality of imposter syndrome, by illustrating the value of negative self-appraisal in certain contexts and demonstrating how these individuals behave in ways that do advance their own interests—specifically, their interest in succeeding.

The first step in further pursuing this account of self-deceptive imposter syndrome is to establish ways in which these reasoning biases can be studied. In the next and final chapter of this thesis, I present the results of a pilot study aimed at doing precisely this. In this study, postgraduate students were required to solve a number of reasoning problems and rate their own performance. We found that participants who scored higher on a measure of imposter syndrome were more negative in their self-evaluations. Interestingly, we also found that when controlling for depression and low self-esteem, imposter syndrome was only predictive of one particular form of negative self-evaluation: evaluation compared to others. This represents a critical feature of the interdisciplinary approach I advocate for: employing methods from the sciences to test hypotheses and generate new data to inform theoretical and philosophical models.

7. Biased Performance Evaluation in the Imposter Phenomenon: Content, Cause, and Authenticity

Abstract

Background: The imposter phenomenon involves a bias towards negative evaluation of one's own performances. For example, people high in the imposter phenomenon often expect themselves to fail exams. In this study, we explore the relationship between imposter phenomenon and biased performance evaluation. First, we addressed the possibility that the negative evaluations associated with the imposter phenomenon are feigned for social benefit. Second, we assessed whether they are better explained with reference to other constructs, namely, depression and low-self-esteem. Finally, we tested the hypothesis that these evaluations are driven by biases in the search for evidence, such that those high in imposter phenomenon seek out more negative feedback regarding their own performance (leading to more negative judgments).

Method: Using the online platform Prolific, we recruited a sample consisting of graduate students living in the United Kingdom, United States, and Europe (n = 163). Participants were required to solve a number of reasoning problems and report their confidence in their performance during the test. After receiving some (self-selected) feedback on their performance, they estimated their objective performance (how many problems that they correctly solved) and comparative performance (what percent of the sample that they outperformed).

Results: Participants high in imposter phenomenon evaluated their performance more negatively. This occurred in an anonymous online setting, with no social incentive to feigning such behaviour, contradicting claims that negative self-evaluation associated with IP is feigned for social benefits. When controlling for the effect of depression and low self-esteem, the imposter phenomenon was only significantly predictive of one particular form of negative performance evaluation: comparison to others. Finally, contrary to our prediction, we found no significant relationship between imposter phenomenon and a bias in selecting feedback.

Conclusions: Our results suggest that individuals high in imposter phenomenon exhibit a bias towards negative performance evaluation, in an anonymous online setting where

there is no social incentive to pretend. When controlling for depression and low self-esteem, the relationship between imposter phenomenon and negative performance evaluation holds only for a particular form of evaluation: comparison to others.

1. Introduction

In the imposter phenomenon (IP), successful and intelligent people believe that they are incompetent and fear being exposed as such (Clance & Imes, 1978). While the phenomenon is most commonly associated with, and researched in, the academic setting , it affects people from a wide variety of professions (Clark et al., 2014; Fried-Buchalter, 1997). In fact, some estimates suggest that up to 70% of individuals have experienced IP at some point in their life (Cozzarelli & Major, 1990, p. 1).

A key feature of IP is a bias towards negative evaluation of one's own performances. People high in IP expect themselves to perform poorly in upcoming exams, and judge themselves as having performed poorly, after completing exams (Cozzarelli & Major, 1990). This bias is considered to be a crucial feature of, and driving force behind, IP (Clance, 1985). By judging themselves to have performed poorly, individuals high in IP reinforce their belief in their own incompetence; by expecting themselves to perform poorly, they exacerbate their fears of being exposed as imposters.

One important issue regarding this relationship relates to distinctions in the content of performance evaluations. Previous research has focused on the difference between *prospective* evaluation (how well one judges that they will perform, in an upcoming task) and *reflective* evaluation (how well one judges that they performed, after completing a task) (Cozzarelli & Major, 1990). However, an underexplored form of bias involves *ongoing* evaluations (how one judges their performance, at the time of performing). Another important distinction is between *objective* performance evaluations (how well one judges that they performed) and *comparative* performance evaluation (how well one judges that they performed, compared to their peers). In order to better understand the relationship between IP and biased performance evaluation, we must understand the specific content of the bias.

Another important question relates to the cause of the bias. A recent theoretical account of IP provides one potential explanation. On this account, IP qualifies as a form of motivated reasoning, wherein people high in IP are motivated to believe that they lack ability (for example, that they are unintelligent), because this belief motivates them to work harder ("Because I am not intelligent, I will have to work harder in order to succeed") (Gadsby, forthcoming). A consequence of this account is that people high in IP negatively evaluate their own performances in order to maintain these negative (but motivationally beneficial) beliefs.

The literature on motivated reasoning suggests a number of strategies that motivated reasoners use to bias themselves (Bénabou & Tirole, 2002). One such strategy involves searching for evidence in an asymmetric manner. For example, in a recent study participants were motivated to believe that they had performed well in a general knowledge test (Solda et al., 2019). After completing the test, they were allowed to select questions for which they would receive feedback (whether they answered correctly or not). Participants who were motivated to believe that they had performed well selected more questions that they had answered correctly. In other words, they chose to receive more positive feedback, biasing themselves towards (over)confidence in their own performance (see also: Smith et al., 2017). The motivated reasoning account of IP suggests that people high in IP will exhibit the opposite pattern of bias: because they are motivated to form negative beliefs about their own abilities, they will actively seek out more negative evidence regarding their performances (Gadsby, forthcoming).

While the motivated reasoning account of IP offers some useful predictions about the cause of negative performance evaluations in IP, it also faces some challenges. One challenge relates to the authenticity of negative performance evaluations. Some researchers suggest that people high in IP might not genuinely believe that they are incompetent and judge themselves to have performed poorly. Instead, they claim that the negative performance evaluations associated with IP are a social strategy (Leary et al., 2000; McElwee & Yurak, 2007), conducted for interpersonal benefits, such as lowering others' expectations, conveying a sense of modesty, protecting one's image in the face of failure, and eliciting encouragement and support from others (Leary et al., 2000, p. 752).

While Leary et al. leave open the possibility that "true imposters"—who genuinely believe that they are inadequate and expect themselves to fail—exist, they nevertheless claim that "people who score high on measures of [IP] are strongly motivated by self-presentational concerns" (p. 753). On this account, IP does not qualify as a form of motivated reasoning, as the relevant beliefs and reasoning biases are feigned.

Another challenge to the motivated reasoning account of IP pertains to the relationship between IP and associated traits, namely, low self-esteem and depression. IP highly correlates with low self-esteem and depression (Chrisman et al., 1995; Cozzarelli & Major, 1990; McGregor et al., 2008; Thompson et al., 1998), both of which are associated with a bias towards negative self-evaluation (Campbell & Fairey, 1985; Kovacs & Beck, 1978). Consequently, people high in IP may engage in negative performance evaluations in virtue of being depressed or low in self-esteem. If this was the case, the biases associated with IP would not qualify as a form of motivated reasoning, as they are driven by other, domain general, traits. More generally, this would suggest that the IP construct does not offer explanatory benefits regarding biased performance evaluation, over and above these more well-known and well-researched conditions.

This study investigated the relationship between IP and performance evaluation bias, in light of the foregoing considerations. To do so, we adapted an experimental paradigm from the motivated reasoning literature, wherein participants were required to solve a set of reasoning problems and evaluate their own performance (Solda et al., 2019). This study had four aims. The first aim was to explore the relationship between IP and different forms of performance evaluation bias. After solving each problem, participants were required to rate their confidence in the accuracy of their solution, indicating a form of ongoing performance evaluation. We also measured two forms of retrospective performance evaluation: objective and comparative. After receiving some self-selected feedback on their performance (see below), participants were required to estimate their total accuracy on the set of problems (objective performance evaluation), as well as what percent of other participants that they outperformed (comparative performance evaluation).

The second aim was to test the aforementioned prediction of the motivated reasoning account of IP: that people high in IP will exhibit a bias towards searching for more negative feedback regarding their own performances. To test this hypothesis, we allowed participants to select between one and four of their solutions to receive feedback on, testing whether those higher in IP would show a bias towards receiving feedback for problems that they were less confident of (i.e. solutions that would more likely be false, constituting negative feedback).

The third aim was to address the possibility that the biased performance evaluation associated with IP is feigned. To address this, we conducted the experiment in an anonymous online setting, thereby removing any possibility for social gain by feigning negative performance evaluation. The fourth and final aim was to test the comparative influence of depression and low self-esteem in driving negative performance evaluations. To do so, we measured participants' levels of depression and self-esteem and measured the effect of these factors on performance evaluation bias.

We constructed three hypotheses. First, contrary to the claim that the negative performance evaluation associated with IP is feigned for social benefit, we predicted that IP would be associated with a bias towards negative performance evaluation, in this anonymous online setting. We further predicted that this bias would hold in all three forms of performance evaluation (ongoing, objective, and comparative evaluation). Second, we predicted that, in line with the motivated reasoning account, IP would be associated with a search bias, such that those high in IP would be biased towards selecting more negative feedback regarding their performance. Third, we predicted that the relationship between IP and these four forms of bias, (ongoing, objective, and comparative evaluation, as well as search) would persist when statistically controlling for the influence of depression and low self-esteem, contradicting the possibility that performance evaluation bias can be better explained with reference to these distinct constructs.

2. Method

2.1. Ethics

This study was approved by Monash University Human Research Ethics Committee (MUHREC Project ID: 25939). Participants were informed about the design and purpose of the study and provided informed consent before taking part.

2.2. Participants

201 participants were recruited online, through the platform Prolific (https://www.prolific.co/). We recruited participants between the ages of 18 and 65, who were currently residing in United Kingdom, United States, and continental Europe, and were currently enrolled in graduate studies (e.g. MA/MSc/MPhil/PhD). All participants were fluent in English and had no current or prior diagnosis of a neurological or psychiatric condition.

Participants were compensated a base rate of £1.70 for their time. They were also given a bonus, designed to incentivise effort. For each participant, 2 of their answers were randomly selected and they were paid £1 for each correct answer. Including bonuses, the total payment, per participant, fell between £1.70 and £3.70.

Participants (5) were excluded when they failed one of two measures of effort, either scoring less than 3 out of 16 questions correctly or failing a question designed to test effort ("What is the sixth month of the year"). Participants (31) were also excluded for failing one of two measures of general interest in their score on the reasoning task, either for choosing not to see their final score and the set of answers (13) or selecting 4 answers in a row (e.g. 1, 2, 3 ,4) when selecting feedback (18). We reasoned that those who showed no interest in the outcome of the task would thoughtlessly select questions to receive feedback for, undermining the data on search bias. The final sample included 163 participants (93 females; 70 males; *M* age = 25.4; SD: 4.4).

2.3. Material and Procedure

Data was collected through the Qualtrics platform (http://www.qualtrics.com/).

After providing informed consent, participants filled out a demographic questionnaire, completed the main task, and then completed the three questionnaires (CIPS, RSS, BDI-II, provided in that order). They were then given the opportunity to return directly to Prolific or view their total score and the complete set of answers.

2.3.1. Main Task

In the main task, participants were required to complete a set of 16 reasoning problems ("designed to test your intelligence"). These problems were taken from the international cognitive reasoning ability resource (ICAR) (Condon & Revelle, 2014) and the test of figural analogies (TFA) (Blum et al., 2016). These included verbal reasoning, letter and number series, three-dimensional rotation, matrix reasoning, and figural analogy problems (for two examples, see figures 1 and 2).

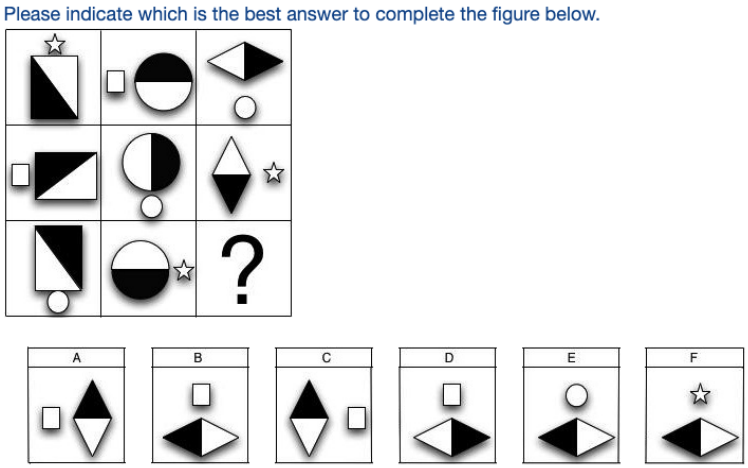


Figure 1. Example problem (matrix reasoning).

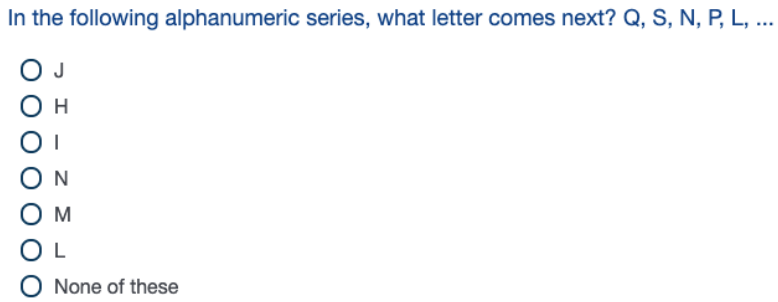


Figure 2. Example problem (letter and number series).

Participants were allowed two minutes to solve each problem. After providing a solution, they were required to report their confidence in it ("How confident are you that you answered correctly?") on a scale from 0 ("not confident at all") to 100 ("completely confident").

After solving the entire problem set, participants were shown each of the problems, along with the confidence that they reported in their solutions. They were asked to select between one and four of the problems to receive feedback for: "You may now check whether some of your answers were correct. You will only be informed about whether each answer was correct, you will not be given the answers to the questions".

After receiving feedback for the problems that they selected, they were asked to evaluate their performance. First, they estimated how many problems (0-16) that they had solved correctly. Second, they estimated the percent of other participants (graduate students in the UK, US, and Europe) that they outperformed (0%-100%).

2.3.2 Questionnaires

After completing the main task, participants answered three questionnaires. In order to measure IP, we employed the Clance Imposter Phenomenon Scale (CIPS) (Clance, 1985). Although several measures of IP exist, the CIPS is the most commonly employed and exhibits the strongest validity and reliability (a = .91; Holmes et al., 1993).

To measure self-esteem, we employed the Rosenberg Self-esteem Scale (Rosenberg, 1965) (RSS), a widely used measure of global self-esteem (a = .94; Kolligian & Sternberg, 1991) which has been found to correlate with the IP, as measured by the CIPS (Sonnak & Towell, 2001).

To measure depression, we used the Beck Depression Inventory II (a = .93, Beck, Steer, & Brown, 1996) (BDI-II), a commonly used scale to measure depression, which has also been found to correlate with IP, as measured by the CIPS (McGregor et al., 2008).

2.4. Data Preparation and Analysis

2.4.1. Data Preparation

CIPS, BDI-II, and RSS scores were calculated by summing the responses to each statement of the respective questionnaires. Participants' total accuracy was recorded on the reasoning problems, as well as their ongoing reported confidence for each of their solutions, and their retrospective performance evaluations. Based on these scores, four dependent variables were calculated:

Ongoing evaluation bias was calculated by subtracting a participant's score (%) from the mean of their (ongoing) confidence ratings. Negative values represent a negative bias (being less confident in one's performance than is warranted), while positive values represent a positive bias (being more confident than is warranted);

Objective evaluation bias was calculated by subtracting participants' actual score from their estimated score. Negative values indicate a negative bias (underestimating one's own score), while positive values indicate a positive bias (overestimating it);

Comparative evaluation bias was calculated by subtracting the percent of the sample a participant actually outperformed from the percent that they estimated themselves to have outperformed. Negative values indicate a negative bias (underestimating one's own comparative performance), while positive values indicate a positive bias (overestimating it);

Search bias was calculated by subtracting the mean confidence of the participants' searched answers from the overall mean confidence. A negative value represents a negative bias (participants choosing feedback for answers that they were less confident of, i.e. they believed were more likely to be inaccurate), while a positive value represents a positive bias (participants choosing feedback for answers that they were more confident of, i.e. they believed were more likely to be accurate).

2.4.2 Analysis

All analysis was conducted using IBM SPSS Statistics (27.0.0.0). To test our first hypothesis, that IP predicts negative performance evaluation in an anonymous setting, we conducted three independent linear regressions, using Ongoing evaluation bias, Objective evaluation bias, and Comparative evaluation bias as dependent variables and, in each case, CIPS as a predictor. To test our second hypothesis, that IP predicts a negative bias in the search for feedback, we conducted an additional linear regression, using Search bias as a dependent variable and CIPS as predictor. For each of the regressions, assumptions were tested according to Field (2009). One-directional hypotheses were specified (CIPS predicting negative bias), so one-tailed p-values were used for regression coefficients.

In order to further explore the relationship between IP and performance evaluation bias, we conducted additional between-group analyses, comparing participants who scored high in IP against those who scored low. Following the approach of previous studies (Ferrari & Thompson, 2006), we split the data set into two groups, based on their CIPS scores: the top-third of participants (high imposter group; n = 54; range = 70–91; M = 79.37, SD = 6.67) and the bottom-third (low imposter group; n = 54; range = 26–58; M = 49.91, SD = 6.67). These ranges are consistent with the suggested CIPS cut-off between "impostors" and "non-impostors" (61) (Holmes et al., 1993). To test for a between-group difference, we conducted independent sample t-tests, using each of the dependent variables: Ongoing evaluation bias, Objective evaluation bias, Comparative evaluation bias, and Search bias. Normality was assessed with the Kolmogorov–Smirnov test. In both the low-IP and high-IP groups, objective evaluation bias was significantly non-normal, low-IP: $D(54) = .158$, $p = .002$; high-IP: $D(54) = .160$, p. = .001. In the

high-IP group, search bias was also significantly non-normal, D(54) = .138, p. = .012. However, in each case the extent of kurtosis and skewness fell within an acceptable range (±2) (George & Mallery, 2010). Homogeneity of variance was assessed with Levene's test. One-directional hypotheses were specified (the high-IP group exhibiting more negative evaluations than the low-IP group), so one-tailed p-values were used.

Our third hypothesis was that IP would predict negative performance evaluation and negative search bias when controlling for self-esteem and depression. First, in order to replicate previous findings and confirm the relationship between IP, depression, and self-esteem in our sample, we performed Pearson correlations using the CIPS, BDI-II, and RSS scales as dependent variables. Hypotheses provided by previous findings were one-directional (CIPs positively correlating with depression and negatively correlating with self-esteem), so one-tailed p values were used. For each of the questionnaire outcomes, the assumption of normality was tested. The BDI-II score was found to be significantly non-normal (according to the Kolmogorov–Smirnov test), D(163) = .118, p < .001. However, both skewness (.954) and kurtosis (.610) fell within an acceptable range (±2) (George & Mallery, 2010).

Next, we conducted four mixed linear regressions with CIPS, RSS, and BDI-II scores as forced entry predictors, and Ongoing evaluation bias, Objective evaluation bias, Comparative evaluation bias, and Search bias as dependent variables. For each of the regressions, assumptions were tested according to Field (2009). One-directional hypotheses were specified (CIPS predicting negative bias), so one-tailed p-values were used for regression coefficients.

3. Results

3.1. Imposter Phenomenon, Performance Evaluation Bias, and Search Bias

Linear regressions discovered statistically significant relationships between CIPS and Ongoing evaluation bias (R = .219, R^2 = .048, F(1,162) = 8.108, p = .006); CIPS and Objective evaluation bias (R = .216, R^2 = .047, F(1,162) = 7.864, p = .006); CIPS and Comparative evaluation bias (R = .205, R^2 = .042, F(1,162) = 7.086, p = .009). No

significant relationship was found between CIPS and Search bias (R = .025, R^2 = .001, F(1,162) = .099, p = .754) (statistics, see Table 1). This means that those who scored higher on IP were more negative in their performance evaluations, both during the task, and afterwards; and both objectively and when comparing themselves to other participants.

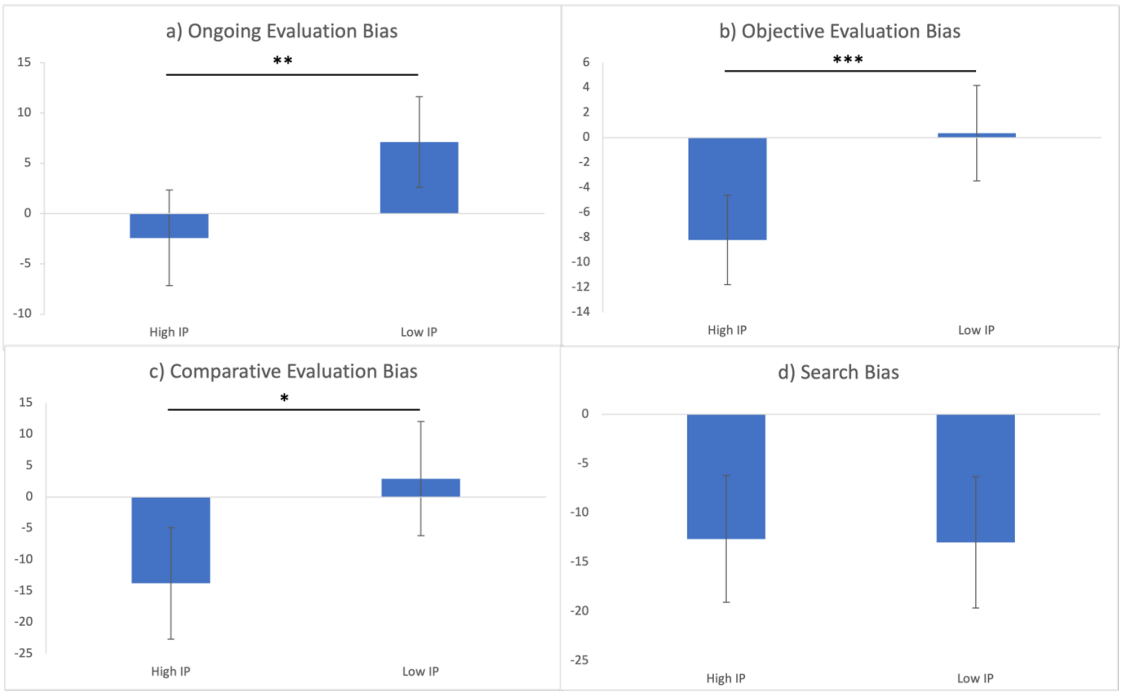
Table 1: Four linear regressions with CIPS as predictor.

Outcome	**B**	**SE B**	***B***	**p**
Ongoing Evaluation Bias	**-.273**	**.096**	**-.219**	**.003**
Objective Evaluation Bias	**-.229**	**.082**	**-.216**	**.003**
Comparative Evaluation Bias	**-.499**	**.188**	**-.205**	**.005**
Search Bias	-.042	.135	-.025	.377

Note. Unstandardized coefficient B, standard error of B, standardised coefficient β, and p-values for regression coefficients.

3.2. Between-Group Analysis

Four independent sample t-tests discovered the following: there was a significant difference in Ongoing evaluation bias between the high-IP (M = -2.42, SD = 17.45) and low-IP (M = 7.13, SD = 16.49) groups; t(106) = 2.92, p = .002, d = 0.563 (figure 1a). There was a significant difference in Objective evaluation bias between the high-IP (M = -8.22, SD = 13.12) and low-IP (M = 0.35, SD = 14.02) groups; t(106) = 3.28, p = .001, d = 0.631 (figure 1b). There was a significant difference in Comparative evaluation bias between the high-IP (M = -13.77, SD = 32.61) and low-IP (M = 2.93, SD = 33.34) groups; t(106) = 2.63, p =.05, d = 0.506 (figure 1c). There was no significant difference in Search bias between the high-IP (M = -12.64, SD = 24.40) and low-IP (M = -13, SD = 23.6) groups; t(106) = 2.92, p = .469, d = 0.015 (figure 1d). This means that participants in the high-IP group were significantly more negative in their ongoing, comparative, and objective evaluations, than those in the low-IP group.



Figures 3a, 3b, 3c, 3d: Between-group comparison of Ongoing evaluation bias, Objective evaluation bias, Comparative evaluation bias, and Search bias. Error bars represent 95% CI. * = p < .05; ** = p < .01; *** p < .001

3.3. Controlling for Self-Esteem and Depression

According to the results of the Pearson correlations, CIPS positively correlated with BDI (r = .57, p < .001) and negatively correlated with SE (r = -.67, p. <.001). This means that those who scored higher on IP also scored higher in depression and lower on self-esteem.

Four multiple linear regressions with forced entry of the three predictors (CIPS, RSS, BDI-II) discovered: Ongoing evaluation bias (R = .298, R^2 = .089, F(1,162) = 5.158, p = .002); Objective evaluation bias (R = .246, R^2 = .060, F(1,162) = 3.4, p = .019); Comparative evaluation bias (R = .229, R^2 = .052, F(1,162) = 2.929, p = .035); Search bias (R = .142, R^2 = .020, F(1,162) = 1.096, p = .353) (for more information, see tables 2-5). When controlling for the influence of self-esteem and depression, IP was no longer a significant predictor of Objective evaluation bias or Ongoing evaluation bias, though it remained a significant predictor of Comparative evaluation bias (see: tables 1, 2, and 3).

However, depression was a significant predictor of Ongoing evaluation bias, when controlling for self-esteem and IP (see table 1).

Tables 2-5: Multiple Linear Regressions, with CIPS, RSS, and BDI-II as predictors.

Ongoing Evaluation Bias	**B**	**SE B**	***B***	**p**
(Constant)	26.355	17.153		.063
CIPS	-.146	.130	-.117	.132
RSS	-.261	.387	-.081	.251
BDI-II	**-.510**	**.200**	**-.275**	**.006**

Objective Evaluation Bias	**B**	**SE B**	***B***	**p**
(Constant)	9.521	14.819		.261
CIPS	-.146	.112	-.137	.099
RSS	-.017	.334	-.006	.480
BDI-II	-.229	.172	-.145	.093

Comparative Evaluation Bias	**B**	**SE B**	***B***	**p**
(Constant)	65.658	34.147		.028
CIPS	**-.569**	**.259**	**-.234**	**.015***
RSS	-.909	.770	-.144	.120
BDI-II	-.435	.397	-.120	.138

Search Bias	**B**	**SE B**	***B***	**p**
(Constant)	-9.555	24.473		.349
CIPS	-.190	.185	-.111	.154
RSS	.112	.552	.025	.420
BDI-II	.462	.285	.181	.054

Tables 2-5. Unstandardized coefficient B, standard error of B, standardised coefficient β, and p-values for regression coefficients.

4. Discussion

In order to make progress on the issues surrounding the relationship between IP and biased performance evaluation, experimental paradigms are needed to measure different forms of this bias. This study presents a new experimental paradigm to accomplish this. In an online setting, participants were required to solve a set of reasoning problems, while evaluating each of their solutions, and, after receiving some self-selected feedback regarding the accuracy of their solutions, estimate their own performance.

We constructed the following hypotheses. First, that in an anonymous online setting—where there was no social benefit to feigning negativity—participants high in IP would exhibit a bias towards negative performance evaluation, in each of the three forms (Ongoing evaluation, Objective evaluation, Comparative evaluation). Second, that IP would additionally be associated with a negative Search bias, consisting of a tendency towards selecting feedback for answers that they felt less confident in (i.e., problems that they were more likely to have solved incorrectly). Third, that the relationship between IP and these four forms of bias would hold while statistically controlling for depression and self-esteem.

The study confirmed our first hypothesis, as IP predicted a bias towards negative estimates in each of the three forms of performance evaluation (Ongoing, Objective, and Comparative). This finding was further confirmed by additional between-group analysis, comparing the top and bottom third of the sample, wherein participants who exhibited high IP were significantly more negative in their performance evaluation, in each of the three forms.

Our findings did not confirm the second hypothesis, that IP would be associated with a negative search bias. IP was not a significant predictor of Search bias and there was no significant between-group difference in search bias between high IP and low IP groups. Rather, both groups chose to receive feedback on questions that they were comparatively less confident in.

Regarding the third hypothesis, that IP would predict performance evaluation bias in each of the four forms, while controlling for the influence of depression and low self-esteem, we made three discoveries. First, depression, but not IP or self-esteem, significantly predicted bias in ongoing performance evaluation. Second, none of the measures individually predicted bias in objective performance evaluation or search bias. Third, IP, but not self-esteem or depression, significantly predicted bias in comparative performance evaluation.

These results illuminate a number of issues regarding the relationship between IP and biased performance evaluation. Our findings are consistent with the claim that those high in IP exhibit bias in how they evaluate their own performances. In our study, participants higher in IP were less confident in the accuracy of their solutions, while performing (Ongoing performance bias), and they were more negative in their retrospective evaluation of their performance, both objectively and comparatively. Performance evaluations in this study were made completely anonymously, provided to a researcher who the participants had never met and would never meet; there was thus no social value for participants in pretending to negatively evaluate themselves. Contrary to previous suggestions that many, if not most, individuals who score highly in measures of IP engage in feigned negative performance evaluation for social advantage, our results suggest that these biases are still present in contexts where there is no social advantage to such behaviour.

That said, we did not employ social setting as a manipulation in this study, and so our results are not quantitatively informative of the role of social setting in driving these factors. It might be that while some people who score high in IP feign negative evaluations, others do so genuinely, and the presence of this latter group was sufficient to

drive the observed results (Leonhardt et al., 2017). In any case, our findings suggest that social gains are not the only causally relevant feature, and that anonymous online studies that remove social setting are a useful tool for exploring the association between IP and bias in performance evaluation.

Regarding the comparative influence of low self-esteem and depression on performance evaluation bias, our findings are also illuminating and important. First, they suggest that researchers interested in IP should consider the effect that low self-esteem and depression have on driving different forms of performance evaluation bias. For example, while those high in IP might be less confident in their performances while taking a test, this may be better explained with reference to these individuals' higher levels of depression.

The finding that IP predicts Comparative performance evaluation bias, when controlling for self-esteem and depression, is also important. Previous research on performance evaluation bias in IP has focused on objective evaluations of performance (e.g. obtaining a certain grade) (Cozarelli & Major 1990). However, classic accounts of the imposter phenomenon emphasize its socially comparative nature. Those high in IP specifically downplay their abilities related to their peers, insisting that they are less capable and intelligent than them (thus meriting the "imposter" label) (Clance, 1985). Our findings are consistent with this characterisation and suggest that while comparisons may not be feigned for social benefit, they are of a socially comparative nature. Consequently, comparative evaluations are a more appropriate target for understanding the specific forms of biased performance evaluation associated with IP.

Our results did not support the hypothesis that negative performance evaluation in IP is maintained by a bias in searching for evidence. While participants high in IP were more negative in their performance evaluations, they did not select more negative feedback than other participants. Negative performance evaluation may, instead, be underpinned by distinct strategies. For example, instead of seeking out negative feedback, those high in IP may exhibit biased memory recall, disproportionately remembering problems that they could not solve (Bénabou & Tirole, 2002, p. 149). They might also be more biased in

their interpretation of the feedback that they received, allowing negative feedback (incorrect solutions) to more strongly influence their judgments about performance (Lord et al., 1979). Finally, bias in the search for feedback may still be associated with IP, though the present study was unable to discover it.

In summary, our results showed that individuals high in IP exhibit a bias towards negative performance evaluation, in an anonymous online setting. Further, when controlling for depression and low self-esteem, the relationship between IP and negative performance evaluation holds only in a particular form of evaluation: comparison to others, suggesting that this form of evaluation is the most appropriate target for future research.

Conclusion

1. Thesis Summary

I began this thesis with a characterisation of eating disorders and imposter syndrome as involving philosophically and scientifically puzzling examples of unusual belief. Philosophically, the beliefs appear to severely violate norms of rationality; scientifically, as they are not associated with any clear form of cognitive or neurological dysfunction. Over the course of the thesis, I illuminated these philosophical and scientific puzzles, while demonstrating the productivity of a resolutely and exhaustively interdisciplinary approach towards understanding unusual beliefs.

The interdisciplinary approach I advocated for throughout the thesis involves two specific principles. The first is that philosophers should cast their nets far and wide in terms of the disciplines from which to draw, acting as facilitators of knowledge between disparate bodies of literature. The second is that philosophers should employ methods and tools usually reserved for other researchers (such as experimental paradigms), in order to more effectively contribute to our understanding of unusual belief. Over the course of the thesis, I applied both of these principles. I drew from a broad variety of disciplines—philosophy, psychology, cognitive neuroscience, cognitive neuropsychiatry, and behavioural economics—and combined different methods—conceptual analysis, theoretical psychology, and empirical research—to generate new ways of understanding these eating disorders and imposter syndrome, along with their philosophical and scientific implications.

Chapter 1 introduced an empiricist account of the false body size beliefs found in eating disorders. On this account, these beliefs are grounded and reinforced by misleading experiences of body size. This account corresponds with a general two-factor approach, wherein unusual experiences provide the content of unusual beliefs, while other factors contribute to their maintenance. I argued that the second factor in this model consisted of biases in the way in which evidence about body size is treated, and I provided a

unifying explanation for these biases, within the Friedrich-Trope-Liberman model of hypothesis testing.

This chapter provided the foundations of a model that was expanded, tested, applied, and enriched throughout the following chapters. This process began in chapter 2, where I amended the model with an additional form of experience: direct proprioceptive misperception of bodily boundaries. This chapter also drew out the philosophical consequences of this addition, in terms of vindicating the epistemic rationality of eating disorders. I argued that, contrary to the assumptions of many philosophers, people with eating disorders are not severely epistemically irrational in the way they form and maintain beliefs about their bodies.

Both chapters 1 and 2 addressed the philosophically and scientifically puzzling aspects of false body size beliefs in eating disorders. While such beliefs appear scientifically perplexing—in virtue of their lack of clear neurological or cognitive aetiology—I illustrated that the explanatory approaches usually reserved for noticeably dysfunctional beliefs can gain explanatory traction in this domain. This presents an important methodological suggestion: by carefully scrutinizing the available scientific literature, researchers can identify potential belief-grounding experiences associated with disorders that do not involve obvious dysfunction. Additionally, the arguments presented in these chapters suggest that, contrary to their appearances, these beliefs do not represent extreme violations of norms of epistemic rationality. Rather, there is sufficient evidence to justify the relevant beliefs, though this evidence is not of a form that is obvious, or accessible, to third parties.

These chapters also illuminated the pragmatic rationality of these beliefs. Specifically, chapter 1 illustrated that, contrary to appearances, the relevant beliefs do play a role in advancing these individuals' desires, as negative appraisal of body size aids in achieving and maintaining thinness. This account illuminates and enriches our understanding of the pragmatic rationality of such beliefs and suggests that researchers should pay careful attention not only to the unusual experiences associated with mental disorders, but to unusual value systems as well.

The following chapters aimed at extending and enriching this empiricist account, through addressing two additional forms of misleading body size experience associated with eating disorders: visual misperception and tactile misperception. In line with the interdisciplinary focus of the thesis, chapters 3 and 4 demonstrated concrete methods through which to gather data and make progress on a key assumption of all empiricist models: that those who hold unusual beliefs perceive the world in unusual ways.

Chapter 3 assessed the claim that people with eating disorders visually misperceive their own body size. If this claim is true, it has crucial implications for the proposed empiricist account, as well as our broader understanding of visual perception. While first-person reports support the claim, behavioural research has failed to confirm it. This is due to methodological issues that face the paradigms used to test the claim. I outlined a number of ways to overcome these issues in order to make progress on the question of whether eating disorders involve visual misperception.

Chapter 4 tested the hypothesis that people with eating disorders misperceive touch. Again, this claim bears importance for the proposed empiricist model, as it suggests yet another misleading experience of body size that these individuals suffer from: tactile misperception of body size. By manipulating a previously employed paradigm, we uncovered evidence that the differences in estimating tactile distances associated with anorexia nervosa may not stem from perceptual differences, as has been assumed. Consequently, it appears unlikely that people with eating disorders misperceive their bodies through touch.

Up to this point, the thesis presented a novel, robust, and illuminating account of the false beliefs associated with eating disorders. According to this model, people with eating disorders misperceive their own body size, in part, due to exhibiting distorted mental representations of their own bodies. Nevertheless, despite its empirical plausibility, such a model would draw the suspicions of a substantial number of philosophers. These philosophers challenge the legitimacy of mental representations, insisting that they should not feature in explanations of psychological phenomena. Chapter 5 addressed this

philosophical objection, pointing out the explanatorily indispensable role that mental representations—specifically, representations of the body—play in explaining the kinematic and affordance processing differences associated with anorexia nervosa. This suggests a fruitful way in which to make progress on philosophical debates regarding mental representation, namely, by careful paying attention to the scientific models that include them and the explanatory benefits that they gain by doing so.

While the previous chapters focused on developing and expanding an account of false body size beliefs in eating disorders, the next two chapters discussed a (seemingly) distinct phenomenon: imposter syndrome. Nevertheless, in the course of these chapters, I illustrated a number of important similarities in the aetiologies of these two conditions. Chapter 6 argued for a new account of imposter syndrome, according to which it qualifies as a form of self-deception. On this account, people with imposter syndrome downplay evidence of their own abilities due to the motivational benefits of doing so. This motivational benefit is perceived as necessary for succeeding in particularly challenging domains. This provides a new account of the cause of imposter syndrome and introduces the condition as a philosophically important instance of self-deception.

Chapter 7 introduced a new experimental paradigm, adapted from the literature on self-deception, designed to further explore the biased reasoning associated with imposter syndrome. In this paradigm, post graduate students were required to solve a number of online reasoning problems and estimate their own performance. We discovered that participants who scored higher on a measure of imposter syndrome were more negative in their performance evaluation. Additionally, we discovered that, when controlling for depression and low self-esteem, our measure of imposter syndrome was only predictive of one particular form of negative self-evaluation: evaluation compared to others.

Once again, these two chapters reinforce the importance of interdisciplinarity in understanding unusual beliefs. The philosophical framework presented in chapter 6 rested on a number of assumptions and produced a number of predictions. By testing the assumptions and predictions of this framework, we were able to progress, enrich, and

extend the proposed model, as well as uncover new insights, for example, those related to the comparative nature of self-evaluation bias in imposter syndrome.

These chapters also uncovered a meaningful similarity between eating disorders and imposter syndrome: both conditions appear to be, at least partly, maintained by motivated reasoning, driven by the benefits of holding the relevant beliefs. This suggests yet another fruitful way to approach the understanding and explanation of unusual and seemingly irrational beliefs: by considering the benefits that holding those beliefs might bestow. According to two influential frameworks of reasoning—from the social psychology and behavioural economics literatures, respectively—costs and benefits strongly determine the way in which we seek out, attend to, and interpret evidence. This suggests that just as unusual experiences may be the cause of many unusual beliefs, so too might the benefits associated with holding those beliefs (see: Bortolotti, 2020). To (once again) echo a point from McKay and colleagues (2017, p. 393): motives are powerful doxastic forces, a point that researchers would do well to appreciate.

In sum, the presented chapters suggested that there are many subtle but important pathways to unusual belief. In order to recognise these less obvious pathways, however, we must adopt a resolutely and exhaustively interdisciplinary approach, seeking out insights regarding belief formation from disparate literatures and bringing these insights to bear on the relevant beliefs. In what follows, I finish the thesis by drawing out some important implications of the foregoing chapters and pointing towards future directions for research.

2. Implications

There are a number of implications of the work presented here. While I have touched on some of these, at different points throughout thesis, I will highlight some of the most significant ones here, so as to clarify the unique contribution of this thesis and the future directions it offers.

The proposed empiricist account has important implications for the explanation and treatment of eating disorders. While biases in evidence treatment are already recognised as a critical feature of these conditions and a worthy target of treatment, there is no clear, unifying framework to account for these biases (Fairburn et al., 1999; Shafran et al., 2004). Clinicians attempt to address biases in evidence treatment directly, by persuading their clients to be more objective in the way in which they assess their own body size (Delinsky & Wilson, 2006; Shafran et al., 2007). However, such approaches produce generally unpromising results (Murray et al., 2019; van den Berg et al., 2019). Accordingly, there is a need for new approaches towards the explanation and treatment of these biases.

In contrast to current approaches, I argued for a unifying explanation for the biases in evidence treatment associated with eating disorders, as driven by the costs and benefits associated with holding certain beliefs about body size. This suggests that a fresh approach must be taken towards treating such biases, namely, by addressing those costs and benefits. For example, by working with their clients to recognise and reconsider the costs and benefits of holding such beliefs, clinicians are likely to have more success in adjusting the way these individuals seek out, attend to, and interpret evidence regarding their own body size.

While my account introduces this novel approach towards treating reasoning biases in eating disorders, it also suggests that such an approach will only be so effective in correcting the relevant beliefs. This is because even if people with eating disorders were to treat evidence in an unbiased manner, they would still suffer from consistent, misleading experiences of body size which, in many cases, would be sufficient for causing them to adopt and maintain their false beliefs (as illustrated in chapter 2). Consequently, I suggest that targeting these misleading experiences of body size is an indispensable component of effective treatment.

Several strands of research suggest that the misleading experiences of body size associated with eating disorders can be targeted and adjusted. According to the proposed model, most of these experiences—i.e. false self-other comparisons, affordance

misperception, and proprioceptive misperception—stem from distorted body representations. In order to treat eating disorders and correct the relevant beliefs, clinicians must address these distorted body representations, by adjusting their distorted content and ensuring that those adjustments last (Gadsby, 2019b). There are a number of ways to approach this. Some research suggests that virtual reality paradigms can reduce body representation distortion, in lasting ways (Keizer et al., 2016). Other techniques involve providing clients with eating disorders with multi-modal sensory feedback regarding their own body size (Keizer et al., 2019). The proposed account vindicates the usefulness of these techniques; only by addressing this body representation distortion, will we be able to redress oversized experiences, removing the evidence that grounds and reinforces false body size beliefs. In short then, this thesis contributes both a clarification and confirmation of the best treatment practises for these harmful disorders.

The proposed model contributes clear and valuable pathways for future research. For example, one important task is to understand how body representations become distorted and why this distortion persists (Gadsby, 2017b). While this question has not been our focus here, it is an important piece of the puzzle in order to develop and extend the proposed model. Additionally, future research should focus on whether the forms of oversized experience posited by this model are indeed genuine. Chapters 3 and 4 illustrated some ways in which to approach this task, namely, by carefully (re)designing and employing behavioural tasks aimed to measure misperception. In doing so, researchers can ensure that differences in the behaviour of participants with eating disorders stem from misperception, as opposed to alternative factors such as attitudes, judgments, and demand characteristics.

This thesis also contributed novel arguments regarding the cause and categorisation of imposter syndrome. Rather than exclusively stemming from problematic childhood experiences—a common assumption in the psychological literature (Clance & Imes, 1978; Thompson, 2004)—I argued that many who suffer from imposter syndrome are motivated reasoners, who bias themselves towards negative self-evaluation due to the motivational benefits of doing so. As suggested for eating disorders, to address the relevant biases, we should address the relevant costs and benefits associated with holding

those beliefs. This could be approached at an individual level, by prompting these individuals to reconsider the costs and benefits of their inadequacy beliefs, or at the environmental level, by addressing the circumstances that determine those costs and benefits. Differently from the case of eating disorders, however, we might also question whether the relevant biases necessarily need addressing. Perhaps for some who exhibit mild and minimally harmful forms of imposter syndrome, downplaying their own abilities to give themselves a motivational boost is an acceptable and sustainable strategy, one that they should be allowed to pursue so long as it helps them to achieve their goals.

Eating disorders and imposter syndrome cause considerable distress and suffering for many who suffer from them, as well as for many who love and care for these sufferers. My hope is that this thesis will contribute to a deeper understanding of these conditions and the unusual beliefs associated with them and emphasise the decisive role that resolutely and exhaustively interdisciplinary research can play in contributing to this effort.

References

Aardoom, J. J., Dingemans, A. E., Op't Landt, M. C. S., & Van Furth, E. F. (2012). Norms and discriminative validity of the Eating Disorder Examination Questionnaire (EDE-Q). *Eating behaviors, 13*(4), 305-309.

Adler, J. E. (2002). Akratic believing? *Philosophical Studies, 110*(1), 1-27.

Adler, J. E. (2006). *Belief's own ethics*. MIT Press.

Alleva, J., Jansen, A., Martijn, C., Schepers, J., & Nederkoorn, C. (2013). Get your own mirror. Investigating how strict eating disordered women are in judging the bodies of other eating disordered women. *Appetite, 68*, 98-104.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Association.

Ames, D. (1984). Self shooting of a phantom head. *The British Journal of Psychiatry, 145*(2), 193-194.

Arcelus, J., Mitchell, A. J., Wales, J., & Nielsen, S. (2011). Mortality rates in patients with anorexia nervosa and other eating disorders: a meta-analysis of 36 studies. *Archives of general psychiatry, 68*(7), 724-731.

Badawy, R. L., Gazdag, B. A., Bentley, J. R., & Brouer, R. L. (2018). Are all impostors created equal? Exploring gender differences in the impostor phenomenon-performance link. *Personality and Individual differences, 131*, 156-163.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Bayne, T., & Pacherie, E. (2005). In defence of the doxastic conception of delusions. *Mind & Language, 20*(2), 163-188.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory Manual (2nd ed.)*. San Antonio, TX: The Psychological Corporation Harcourt Brace & Company.

Beckmann, N., Baumann, P., Herpertz, S., Trojan, J., & Diers, M. (2020). How the unconscious mind controls body movements: Body schema distortion in anorexia nervosa. *International Journal of Eating Disorders*.

Bénabou, R., & Tirole, J. (2002). Self-confidence and personal motivation. *The quarterly journal of economics, 117*(3), 871-915.

Ben-Tovim, D. I., Walker, M. K., Murray, H., & Chin, G. (1990). Body size estimates: Body image or body attitude measures? *International Journal of Eating Disorders, 9*(1), 57-67.

Bermúdez, J. L. (2001). Normativity and rationality in delusional psychiatric disorders. *Mind & Language, 16*(5), 493-457.

Blum, D., Holling, H., Galibert, M. S., & Forthmann, B. (2016). Task difficulty prediction of figural analogies. *Intelligence, 56*, 72-81.

Bortolotti, L. (2010). *Delusions and other irrational beliefs*. Oxford University Press.

Bortolotti, L. (2014). *Irrationality*. John Wiley & Sons.

Bortolotti, L. (2020). *The Epistemic Innocence of Irrational Beliefs*. Oxford University Press.

Bourget, D., & Whitehurst, L. (2004). Capgras syndrome: a review of the neurophysiological correlates and presenting clinical features in cases involving physical violence. *The Canadian Journal of Psychiatry, 49*(11), 719-725.

Bravata, D. M., Watts, S. A., Keefer, A. L., Madhusudhan, D. K., Taylor, K. T., Clark, D. M., Nelson, R. S., Cokley, K. O., & Hagg, H. K. (2019). Prevalence, Predictors, and Treatment of Impostor Syndrome: a Systematic Review. *Journal of General Internal Medicine*, 1-24.

Broks, P. (2020). *Unholy anorexia*. Aeon. https://aeon.co/essays/anorexia-nervosa-is-a-battle-site-between-body-and-soul

Brooks, K. R., Mond, J., Mitchison, D., Stevenson, R. J., Challinor, K. L., & Stephen, I. D. (2019). Looking at the figures: visual adaptation as a mechanism for body-size and-shape misperception. *Perspectives on Psychological Science, 15*(1), 133-149.

Brooks, K. R., Mond, J. M., Stevenson, R. J., & Stephen, I. D. (2016). Body image distortion and exposure to extreme body types: contingent adaptation and cross adaptation for self and other. *Frontiers in neuroscience, 10*, 334.

Brown, J. (2020). Epistemically blameworthy belief. *Philosophical Studies*(177), 3595–3614.

Bruch, H. (1974). *Eating disorders. Obesity, anorexia nervosa, and the person within*. Routledge & Kegan Paul.

Calugi, S., El Ghoch, M., Conti, M., & Dalle Grave, R. (2018). Preoccupation with shape or weight, fear of weight gain, feeling fat and treatment outcomes in patients with anorexia nervosa: A longitudinal study. *Behaviour research and therapy, 105*, 63-68.

Campbell, J. D., & Fairey, P. (1985). Effects of self-esteem, hypothetical explanations, and verbalization of expectancies on future performance. *Journal of personality and social psychology, 48*(5), 1097.

Carter, J., Blackmore, E., Sutandar-Pinnock, K., & Woodside, D. (2004). Relapse in anorexia nervosa: a survival analysis. *Psychological medicine, 34*(4), 671-679.

Cash, T., & Deagle, E. (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), 107-126.

Cazzato, V., & Sacchetti, S. (2020, October 1). Why looking in the mirror is so hard for people with eating disorders. The Conversation. https://theconversation.com/why-looking-in-the-mirror-is-so-hard-for-people-with-eating-disorders-144758

Chislenko, E. (2016). Moore's Paradox and Akratic Belief. *Philosophy and Phenomenological Research, 92*(3), 669-690.

Chrisman, S. M., Pieper, W., Clance, P. R., Holland, C., & Glickauf-Hughes, C. (1995). Validation of the Clance imposter phenomenon scale. *Journal of personality assessment, 65*(3), 456-467.

Clance, P. R. (1985). *The impostor phenomenon: Overcoming the fear that haunts your success.* Peachtree Pub Ltd.

Clance, P. R., & Imes, S. A. (1978). The imposter phenomenon in high achieving women: Dynamics and therapeutic intervention. *Psychotherapy: Theory, Research & Practice, 15*(3), 241.

Clark, M., Vardeman, K. K., & Barba, S. E. (2014). Perceived inadequacy: A study of the imposter phenomenon among college and research librarians. *75*(3), 255–271.

Cogliano, R., Crisci, C., Conson, M., Grossi, D., & Trojano, L. (2012). Chronic somatoparaphrenia: A follow-up study on two clinical cases. *Cortex, 48*(6), 758-767.

Coltheart, M. (2007). The 33rd Sir Frederick Bartlett lecture cognitive neuropsychiatry and delusional belief. *Quarterly Journal of Experimental Psychology, 60*(8), 1041-1062.

Coltheart, M., Menzies, P., & Sutton, J. (2010). Abductive inference and delusional belief. *Cognitive neuropsychiatry, 15*(1-3), 261-287.

Condon, D. M., & Revelle, W. (2014). The International Cognitive Ability Resource: Development and initial validation of a public-domain measure. *Intelligence, 43*, 52-64.

Cornelissen, K. K., McCarty, K., Cornelissen, P. L., & Tovée, M. J. (2017). Body size estimation in women with anorexia nervosa and healthy controls using 3D avatars. *Scientific reports, 7*(1), 1-15.

Cornelissen, K. K., Widdrington, H., McCarty, K., Pollet, T. V., Tovée, M. J., & Cornelissen, P. L. (2019). Are attitudinal and perceptual body image the same or different? Evidence from high-level adaptation. *Body image, 31*, 35-47.

Cozzarelli, C., & Major, B. (1990). Exploring the validity of the impostor phenomenon. *Journal of social and clinical psychology, 9*(4), 401-417.

Craig, J. C., & Johnson, K. O. (2000). The two-point threshold: Not a measure of tactile spatial resolution. *Current Directions in Psychological Science*, 9(1), 29-32.

Currie, G., & Jureidini, J. (2001). Delusion, Rationality, Empathy: Commentary on Martin Davies et al. *Philosophy, Psychiatry, & Psychology, 8*(2), 159-162.

D'Amour, S., & Harris, L. R. (2017). Perceived face size in healthy adults. *PLOS one, 12*(5), e0177349.

Delinsky, S. S., & Wilson, G. T. (2006). Mirror exposure for the treatment of body image disturbance. *International Journal of Eating Disorders, 39*(2), 108-116.

De Vignemont, F. (2010). Body schema and body image—Pros and cons. *Neuropsychologia*, 48(3), 669-680.

Draper, H. (2000). Anorexia nervosa and respecting a refusal of life-prolonging therapy: a limited justification. *Bioethics, 14*(2), 120-133.

Ellis, H. D., & Szulecka, T. K. (1996). The disguised lover: a case of Frégoli delusion. In Peter Halligan & J. Marshall (Eds.), *Method in Madness: Case Studies in Cognitive Neuropsychiatry* (pp. 39-50).

Ellis, H. D., & Young, A. W. (1990). Accounting for delusional misidentifications. *The British Journal of Psychiatry, 157*(2), 239-248.

Engel, M. M., van Denderen, K., Bakker, A. R., Corcoran, A. W., Keizer, A., & Dijkerman, H. C. (2020). Anorexia nervosa and the size-weight illusion: No evidence of impaired visual-haptic object integration. *PloS one*, 15(8), e0237421.

Engel, M. M., & Keizer, A. (2017). Body representation disturbances in visual perception and affordance perception persist in eating disorder patients after completing treatment. *Scientific reports, 7*(1), 1-9.

Espeset, E. M., Gulliksen, K. S., Nordbø, R. H., Skårderud, F., & Holte, A. (2012). Fluctuations of body images in anorexia nervosa: patients' perception of contextual triggers. *Clinical psychology & psychotherapy, 19*(6), 518-530.

Espeset, E. M., Nordbø, R. H., Gulliksen, K. S., Skårderud, F., Geller, J., & Holte, A. (2011). The concept of body image disturbance in anorexia nervosa: an empirical inquiry utilizing patients' subjective experiences. *Eating disorders, 19*(2), 175-193.

Fairburn, C. G. (2008). *Cognitive behavior therapy and eating disorders*. Guilford Press.

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders, 16*(4), 363-370.

Fairburn, C. G., & Cooper, Z. (2007). Thinking afresh about the classification of eating disorders. *The International journal of eating disorders, 40*(S3), S107.

Fairburn, C. G., & Harrison, P. J. (2003). Eating disorders. *Lancet (London, England), 361*(9355), 407-416.

Fairburn, C. G., Shafran, R., & Cooper, Z. (1999). A cognitive behavioural theory of anorexia nervosa. *Behaviour research and therapy, 37*(1), 1-13.

Farrell, C., Lee, M., & Shafran, R. (2005). Assessment of body size estimation: A review. *European Eating Disorders Review: The Professional Journal of the Eating Disorders Association, 13*(2), 75-88.

Fassino, S., Pierò, A., Gramaglia, C., & Abbate-Daga, G. (2004). Clinical, psychopathological and personality correlates of interoceptive awareness in anorexia nervosa, bulimia nervosa and obesity. *Psychopathology, 37*(4), 168-174.

Feenstra, S., Begeny, C. T., Ryan, M. K., Rink, F. A., Stoker, J. I., & Jordan, J. (2020). Contextualizing the Impostor "Syndrome". *Frontiers in psychology, 11*, 3206.

Ferrari, J. R., & Thompson, T. (2006). Impostor fears: Links with self-presentational concerns and self-handicapping behaviours. *Personality and Individual differences, 40*(2), 341-352.

Field, A. (2009). *Discovering statistics using IBM SPSS statistics*. Sage.

Firestone, C. (2013). On the origin and status of the "El Greco fallacy". *Perception, 42*(6), 672-674.

Firestone, C., & Scholl, B. J. (2014). "Top-down" effects where none should be found: The El Greco fallacy in perception research. *Psychological science, 25*(1), 38-46.

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *Behavioral and Brain Sciences, 39*, 1-72.

Fox, J., & Weisberg, S. (2018a). *An R companion to applied regression*. Sage publications.

Fox, J., & Weisberg, S. (2018b). Visualizing fit and lack of fit in complex regression models with predictor effect plots and partial residuals. *Journal of Statistical Software, 87*(9), 1-27.

Frankish, K. (2009). Delusions: A two-level framework. *Psychiatry as cognitive neuroscience: Philosophical perspectives*, 269-284.

Fried-Buchalter, S. (1997). Fear of success, fear of failure, and the imposter phenomenon among male and female marketing managers. *Sex roles, 37*(11-12), 847-859.

Friedman, J. (2013). Suspended judgment. *Philosophical Studies, 162*(2), 165-181.

Funkhouser, E. (2019). *Self-Deception*. Taylor & Francis.

Gadsby, S. (2017a). Anorexia nervosa and oversized experiences. *Philosophical Psychology, 30*(5), 594-615.

Gadsby, S. (2017b). Distorted body representations in anorexia nervosa. *Consciousness and cognition, 51*, 17-33.

Gadsby, S. (2017c). Explaining body size beliefs in anorexia. *Cognitive neuropsychiatry, 22*(6), 495-507.

Gadsby, S. (2018). How are the spatial characteristics of the body represented? A reply to Pitron & De Vignemont. *Consciousness and cognition*, *62*, 163-168.

Gadsby, S. (2019a). Body Representations and Cognitive Ontology: Drawing the boundaries of the body image. *Consciousness and cognition, 74*, 102772.

Gadsby, S. (2019b). Manipulating body representations with virtual reality: Clinical implications for anorexia nervosa. *Philosophical Psychology, 32*(6), 898-922.

Gadsby, S. (2020). Self-deception and the second factor: How desire causes delusion in Anorexia nervosa. *Erkenntnis, 85*(3), 609-626.

Gadsby, S. (2021). Imposter Syndrome and Self-Deception. *Australasian Journal of Philosophy*, 1-15.

Gadsby, S., & Williams, D. (2018). Action, affordances, and anorexia: body representation and basic cognition. *Synthese, 195*(12), 5297-5317.

Ganapini, M. B. (2019). Belief's minimal rationality. *Philosophical Studies*, 1-20.

Gandevia, S. C., & Phegan, C. (1999). Perceptual distortions of the human body image produced by local anaesthesia, pain and cutaneous stimulation. *The Journal of physiology, 514*(2), 609-616.

Gardner, R. (2012). Measurement of perceptual body image. *Encyclopedia of body image and human appearance, 2*, 526-532.

Gardner, R. M., & Brown, D. L. (2014). Body size estimation in anorexia nervosa: a brief review of findings from 2003 through 2013. *Psychiatry Research, 219*(3), 407-410.

Garner, D. M., & Garfinkel, P. E. (1982). Body image in anorexia nervosa: Measurement, theory and clinical implications. *The International Journal of Psychiatry in Medicine, 11*(3), 263-284.

Garner, D. M., Garfinkel, P. E., Stancer, H. C., & Moldofsky, H. (1976). Body image disturbances in anorexia nervosa and obesity. *Psychosomatic medicine*.

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press.

George, D., & Mallery, P. (2010). *SPSS for windows step by step: A simple study guide and reference, 17.0 update*. Allyn & Bacon.

Gerrans, P. (2014). *The measure of madness: Philosophy of mind, cognitive neuroscience, and delusional thought*. MIT Press.

Gilleen, J., & David, A. S. (2005). The cognitive neuropsychiatry of delusions: from psychopathology to neuropsychology and back again. *Psychological medicine, 35*(1), 5.

Goldner, E. (1989). Treatment refusal in anorexia nervosa. *International Journal of Eating Disorders, 8*(3), 297-306.

Grisso, T., & Appelbaum, P. S. (1998). *Assessing competence to consent to treatment: A guide for physicians and other health professionals*. Oxford University Press, USA.

Guardia, D., Conversy, L., Jardri, R., Lafargue, G., Thomas, P., Dodin, V., Cottencin, O., & Luyat, M. (2012). Imagining one's own and someone else's body actions: dissociation in anorexia nervosa. *PLOS one, 7*(8).

Guardia, D., Lafargue, G., Thomas, P., Dodin, V., Cottencin, O., & Luyat, M. (2010). Anticipation of body-scaled action is modified in anorexia nervosa. *Neuropsychologia, 48*(13), 3961-3966.

Harvey, J. C., & Katz, C. (1985). *If I'm so successful, why do I feel like a fake?: The impostor phenomenon*. St. Martin's Press.

Hawkins, Jennifer and Louis C. Charland, "Decision-Making Capacity", The Stanford

Encyclopedia of Philosophy (Fall 2020 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2020/entries/decision-capacity/>.

Hawley, K. (2019). I—What Is Impostor Syndrome? Aristotelian Society Supplementary Volume,

Hohwy, J., & Rosenberg, R. (2005). Unusual experiences, reality testing and delusions of alien control. *Mind & Language, 20*(2), 141-162.

Holmes, S. W., Kertay, L., Adamson, L. B., Holland, C., & Clance, P. R. (1993). Measuring the impostor phenomenon: A comparison of Clance's IP Scale and Harvey's IP Scale. *Journal of personality assessment, 60*(1), 48-59.

Hume, D. (2000). *An enquiry concerning human understanding: A critical edition* (Vol. 3). Oxford University Press.

Hutchins, H. M., Penney, L. M., & Sublett, L. W. (2018). What imposters risk at work: Exploring imposter phenomenon, stress coping, and job outcomes. *Human Resource Development Quarterly, 29*(1), 31-48.

Ichino, A. (2020). Superstitious confabulations. *Topoi, 39*(1), 203-217.

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.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

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.

Kearns, H. (2013, November 26). Feel like a fraud? You might have imposter syndrome. The Conversation. https://theconversation.com/feel-like-a-fraud-you-might-have-imposter-syndrome-20745

Keizer, A. (2014). *Body representation disturbances in anorexia nervosa* Utrecht University].

Keizer, A., Engel, M. M., Bonekamp, J., & Van Elburg, A. (2019). Hoop training: a pilot study assessing the effectiveness of a multisensory approach to treatment of body image disturbance in anorexia nervosa. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity, 24*(5), 953-958.

Keizer, A., Smeets, M. A., Dijkerman, H. C., Uzunbajakau, S. A., van Elburg, A., & Postma, A. (2013). Too fat to fit through the door: first evidence for disturbed body-scaled action in anorexia nervosa during locomotion. *PLOS one, 8*(5).

Keizer, A., Smeets, M. A. M., Dijkerman, H. C., Van den Hout, M., Klugkist, I., Van Elburg, A., & Postma, A. (2011). Tactile body image disturbance in anorexia nervosa. *Psychiatry Research, 190*(1), 115-120.

Keizer, A., Smeets, M. A. M., Dijkerman, H. C., van Elburg, A., & Postma, A. (2012). Aberrant somatosensory perception in Anorexia Nervosa. *Psychiatry Research, 200*(2-3), 530-537.

Keizer, A., van Elburg, A., Helms, R., & Dijkerman, H. C. (2016). A virtual reality full body illusion improves body image disturbance in anorexia nervosa. *PLOS one, 11*(10), e0163921.

Kennett, J., Matthews, S., & Snoek, A. (2013, 2013-September-25). Pleasure and Addiction [Hypothesis and Theory]. *Frontiers in Psychiatry, 4*(117). https://doi.org/10.3389/fpsyt.2013.00117

Kilteni, K., Maselli, A., Kording, K. P., & Slater, M. (2015). Over my fake body: body ownership illusions for studying the multisensory basis of own-body perception. *Frontiers in human neuroscience, 9*, 141.

Kinnaird, E., Stewart, C., & Tchanturia, K. (2020). Interoception in Anorexia Nervosa: Exploring Associations With Alexithymia and Autistic Traits. Front. *Psychiatry, 11*, 64.

Kolligian, J., & Sternberg, R. J. (1991). Perceived Fraudulence in Young Adults: Is There an'Imposter Syndrome'? *Journal of personality assessment, 56*(2), 308-326.

Konstantakopoulos, G., Varsou, E., Dikeos, D., Ioannidi, N., Gonidakis, F., Papadimitriou, G., & Oulis, P. (2012). Delusionality of body image beliefs in eating disorders. *Psychiatry Research, 200*(2-3), 482-488.

Kovacs, M., & Beck, A. T. (1978). Maladaptive cognitive structures in depression. *American journal of Psychiatry, 135*(5), 525-533.

Leary, M. R., Patton, K. M., Orlando, A. E., & Wagoner Funk, W. (2000). The impostor phenomenon: Self-perceptions, reflected appraisals, and interpersonal strategies. *Journal of personality, 68*(4), 725-756.

Lee, S., & Kwok, K. (2005). Cross-cultural perspectives on anorexia nervosa without fat phobia. In *EDNOS: Eating disorders not otherwise specified* (pp. 218-235). Routledge.

Leonhardt, M., Bechtoldt, M. N., & Rohrmann, S. (2017). All impostors aren't alike–differentiating the impostor phenomenon. *Frontiers in psychology, 8*, 1505.

Levy, N. (2018). Showing our seams: A reply to Eric Funkhouser. *Philosophical Psychology, 31*(7), 991-1006.

Levy, N. (2019). Is conspiracy theorising irrational. *Social Epistemology Review and Reply Collective, 8*(10), 65-76.

Levy, N. (2020). Belie the belief? Prompts and default states. *Religion, Brain & Behavior, 10*(1), 35-48.

Linardon, J., Phillipou, A., Castle, D., Newton, R., Harrison, P., Cistullo, L. L., Griffiths, S., Hindle, A., & Brennan, L. (2018). Feeling fat in eating disorders: Testing the unique relationships between feeling fat and measures of disordered eating in anorexia nervosa and bulimia nervosa. *Body image, 25*, 163-167.

Lippman, C. W. (1952). Certain hallucinations peculiar to migraine. *Journal of Nervous and Mental Disease, 116*, 346–351.

Longo, M. R. (2015). Implicit and explicit body representations. *European Psychologist*, 20(1), 6–15

Longo, M. R., Azañón, E., & Haggard, P. (2010). More than skin deep: body representation beyond primary somatosensory cortex. *Neuropsychologia, 48*(3), 655-668.

Longo, M. R., & Haggard, P. (2010). An implicit body representation underlying human position sense. *Proceedings of the National Academy of Sciences, 107*(26), 11727-11732.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of personality and social psychology, 37*(11), 2098.

Matthews, E. (2000). Autonomy and the psychiatric patient. *Journal of applied philosophy, 17*(1), 59-70.

Matthews, J., Nagao, K., Ding, C., Newby, R., Kempster, P., & Hohwy, J. (2020). Raised visual contrast thresholds with intact attention and metacognition in functional motor disorder. *Cortex*.

McCloskey, M., & Chaisilprungraung, T. (2017). The value of cognitive neuropsychology: The case of vision research. *Cognitive Neuropsychology, 34*(7-8), 412-419.

McElwee, R. O. B., & Yurak, T. J. (2007). Feeling versus acting like an impostor: Real feelings of fraudulence or self-presentation? *Individual Differences Research, 5*(3), 201-220.

McFarlane, T., Urbszat, D., & Olmsted, M. P. (2011). "I Feel Fat": An experimental induction of body displacement in disordered eating. *Behaviour research and therapy, 49*(4), 289-293.

McGregor, L. N., Gee, D. E., & Posey, K. E. (2008). I feel like a fraud and it depresses me: The relation between the imposter phenomenon and depression. *Social Behavior and Personality: an international journal, 36*(1), 43-48.

McKay, R. (2012). Delusional inference. *Mind & Language, 27*(3), 330-355.

McKay, R., Langdon, R., & Coltheart, M. (2007). Models of misbelief: Integrating motivational and deficit theories of delusions. *Consciousness and cognition, 16*(4), 932-941.

McLean, B. F., Mattiske, J. K., & Balzan, R. P. (2017). Association of the jumping to conclusions and evidence integration biases with delusions in psychosis: a detailed meta-analysis. *Schizophrenia Bulletin, 43*(2), 344-354.

Mehak, A., & Racine, S. E. (2020). Understanding "feeling fat" and its underlying mechanisms: The importance of multimethod measurement. *International Journal of Eating Disorders, 53*(9), 1400-1404.

Mergen, J., Keizer, A., Koelkebeck, K., van den Heuvel, M. R., & Wagner, H. (2018). Women with Anorexia Nervosa do not show altered tactile localization compared to healthy controls. *Psychiatry research*, *267*, 446-454.

Meteyard, L., & Davies, R. A. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language, 112*, 104092.

Metral, M., Guardia, D., Bauwens, I., Guerraz, M., Lafargue, G., Cottencin, O., & Luyat, M. (2014). Painfully thin but locked inside a fatter body: abnormalities in both anticipation and execution of action in anorexia nervosa. *BMC research notes, 7*(1), 707.

Michaelian, Kourken and John Sutton, "Memory", *The Stanford Encyclopedia of Philosophy* (Summer 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2017/entries/memory/>.

Mölbert, S. C., Klein, L., Thaler, A., Mohler, B. J., Brozzo, C., Martus, P., Karnath, H.-O., 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.

Mölbert, S. C., Klein, L., Thaler, A., Mohler, B. J., Brozzo, C., Martus, P., Karnath, H.-O., 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.

Mond, J. M., Hay, P. J., Rodgers, B., Owen, C., & Beumont, P. (2004). Validity of the Eating Disorder Examination Questionnaire (EDE-Q) in screening for eating disorders in community samples. *Behaviour research and therapy, 42*(5), 551-567.

Mondraty, N., & Sachdev, P. (2011). "My Body Is My Template": Why Do People Suffering from Anorexia Nervosa See Their Bodies Differently? In V. R. Preedy, R. R. Watson, & C. R. Martin (Eds.), *Handbook of Behavior, Food and Nutrition* (pp. 3257-3270). Springer New York. https://doi.org/10.1007/978-0-387-92271-3_203

Morton, J. M. (2017). Reasoning under scarcity. *Australasian Journal of Philosophy, 95*(3), 543-559.

Moscone, A.-L., Amorim, M.-A., Le Scanff, C., & Leconte, P. (2017). A model-driven approach to studying dissociations between body size mental representations in anorexia nervosa. *Body image, 20*, 40-48.

Mountjoy, R. L., Farhall, J. F., & Rossell, S. L. (2014). A phenomenological investigation of overvalued ideas and delusions in clinical and subclinical anorexia nervosa. *Psychiatry Research, 220*(1-2), 507-512.

Murphy, R., Straebler, S., Cooper, Z., & Fairburn, C. G. (2010). Cognitive behavioral therapy for eating disorders. *Psychiatric Clinics, 33*(3), 611-627.

Murray, S. B., Quintana, D. S., Loeb, K. L., Griffiths, S., & Le Grange, D. (2019). Treatment outcomes for anorexia nervosa: a systematic review and meta-analysis of randomized controlled trials. *Psychological medicine, 49*(4), 535-544.

Niedenthal, P. M., & Wood, A. (2019). Does emotion influence visual perception? Depends on how you look at it. *Cognition and Emotion, 33*(1), 77-84.

Nordbø, R. H., Espeset, E. M., Gulliksen, K. S., Skårderud, F., Geller, J., & Holte, A. (2012). Reluctance to recover in anorexia nervosa. *European Eating Disorders Review, 20*(1), 60-67.

O'Connell, J. E., Bendall, S., Morley, E., Huang, C., & Krug, I. (2018). Delusion-like beliefs in anorexia nervosa: An interpretative phenomenological analysis. *Clinical Psychologist, 22*(3), 317-326.

Øverås, M., Kapstad, H., Brunborg, C., Landrø, N. I., & Lask, B. (2014). Memory versus perception of body size in patients with anorexia nervosa and healthy controls. *European Eating Disorders Review, 22*(2), 109-115.

Owens, D. (2002). Epistemic akrasia. *The Monist, 85*(3), 381-397.

Palmer, B. (2003). Concepts of eating disorders. *Handbook of eating disorders, 2*, 1-10.

Paul, S. K. (2019). II—What Should 'Impostor Syndrome'Be? *Aristotelian Society Supplementary Volume, 93*(1), 227-245.

Perrykkad, K., & Hohwy, J. (2020). Modelling me, modelling you: the autistic self. *Review Journal of Autism and Developmental Disorders, 7*(1), 1-31.

Peviani, V., & Bottini, G. (2018). The distorted hand metric representation serves both perception and action. *Journal of Cognitive Psychology, 30*(8), 880-893.

Phillipou, A., Mountjoy, R. L., & Rossell, S. L. (2017). Overvalued ideas or delusions in anorexia nervosa? *Australian & New Zealand Journal of Psychiatry, 51*(6), 563-564.

Pickard, H. (2020). Addiction and the self. *Noûs*, 1-25.

Prakash, J., Shashikumar, R., Bhat, P., Srivastava, K., Nath, S., & Rajendran, A. (2012). Delusional parasitosis: Worms of the mind. *Industrial psychiatry journal, 21*(1), 72.

Pitron, V., Alsmith, A., & de Vignemont, F. (2018). How do the body schema and the body image interact? *Consciousness and cognition*, *65*, 352-358.

Pitron, V., Alsmith, A., & de Vignemont, F. (2018). How do the body schema and the body image interact?. *Consciousness and cognition*, 65, 352-358.

Probst, M., Vandereycken, W., Coppenolle, H. V., & Vanderlinden, J. (1995). The Body Attitude Test for patients with an eating disorder: Psychometric characteristics of a new questionnaire. *Eating disorders, 3*(2), 133-144.

Proctor, L., & Morley, S. (1986). 'Demand Characteristics' in Body-Size Estimation in Anorexia Nervosa. *The British Journal of Psychiatry, 149*(1), 113-118.

Proske, U., & Gandevia, S. C. (2012). The proprioceptive senses: their roles in signaling body shape, body position and movement, and muscle force. *Physiological reviews, 92*(4), 1651-1697.

R Core Team. (2019). R: A language and environment for statistical computing.

Radden, J. (2010). *On delusion*. Routledge.

Raleigh, T. (2019). Suspending is believing. *Synthese*, 1-26.

Ramachandran, V. S., & Hirstein, W. (1998). The perception of phantom limbs. The DO Hebb lecture. *Brain: a journal of neurology, 121*(9), 1603-1630.

Ratcliffe, M. (2019). Grief and phantom limbs: A phenomenological comparison. *New Yearbook for Phenomenology and Phenomenological Philosophy, XVII*, 77-96.

Rettler, L. (2018). In defense of doxastic blame. *Synthese, 195*(5), 2205-2226.

Rikani, A. A., Choudhry, Z., Choudhry, A. M., Ikram, H., Asghar, M. W., Kajal, D., Waheed, A., & Mobassarah, N. J. (2013). A critique of the literature on etiology of eating disorders. *Annals of neurosciences, 20*(4), 157.

Risso, G., Martoni, R., Erzegovesi, S., Bellodi, L., & Baud-Bovy, G. (2020). Visuo-tactile shape perception in women with Anorexia Nervosa and healthy women with and without body concerns. *Neuropsychologia, 149*, 107635.

Riva, G. (2014). Out of my real body: cognitive neuroscience meets eating disorders. *Frontiers in human neuroscience*, *8*, 236.

Rock, I. (1966). *The nature of perceptual adaptation*. Basic Books.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton University Press.

RStudio Team. (2015). RStudio: integrated development for R. *RStudio, Inc., Boston, MA URL http://www. rstudio. com, 42*, 14.

Rubinstein, A. (2007). Instinctive and cognitive reasoning: A study of response times. *The Economic Journal, 117*(523), 1243-1259.

Schilder, P. (1935). *The image and appearance of the human body* (Vol. 163). Routledge.

Schwitzgebel, E. (2001). In-between believing. *The Philosophical Quarterly, 51*(202), 76-82.

Shafran, R., & Fairburn, C. G. (2002). A new ecologically valid method to assess body size estimation and body size dissatisfaction. *International Journal of Eating Disorders, 32*(4), 458-465.

Shafran, R., Fairburn, C. G., Robinson, P., & Lask, B. (2004). Body checking and its avoidance in eating disorders. *International Journal of Eating Disorders, 35*(1), 93-101.

Shafran, R., Farrell, C., Lee, M., & Fairburn, C. G. (2009). Brief cognitive behavioural therapy for extreme shape concern: An evaluation. *British Journal of Clinical Psychology, 48*(1), 79-92.

Shafran, R., Lee, M., Payne, E., & Fairburn, C. G. (2007). An experimental analysis of body checking. *Behaviour research and therapy, 45*(1), 113-121.

Slade, P. D. (1985). A review of body-image studies in anorexia nervosa and bulimia nervosa. *Journal of Psychiatric Research, 19*, 255-265.

Slade, P. D., & Russell, G. (1973). Awareness of body dimensions in anorexia nervosa: Cross-sectional and longitudinal studies. *Psychological medicine, 3*(2), 188-199.

Smeets, M. A. (1997). The rise and fall of body size estimation research in anorexia nervosa: A review and reconceptualization. *European Eating Disorders Review: The Professional Journal of the Eating Disorders Association, 5*(2), 75-95.

Smeets, M. A., Klugkist, I. G., van Rooden, S., Anema, H. A., & Postma, A. (2009). Mental body distance comparison: a tool for assessing clinical disturbances in visual body image. *Acta Psychologica, 132*(2), 157-165.

Smeets, M. A., & Panhuysen, G. E. (1995). What can be learned from body size estimation? It all depends on your theory. *Eating disorders, 3*(2), 101-114.

Smeets, M. A., Smit, F., Panhuysen, G. E., & Ingleby, J. D. (1997). The influence of methodological differences on the outcome of body size estimation studies in anorexia nervosa. *British Journal of Clinical Psychology, 36*(2), 263-277.

Smith, M. K., Trivers, R., & von Hippel, W. (2017). Self-deception facilitates interpersonal persuasion. *Journal of Economic Psychology, 63*, 93-101.

Solda, A., Ke, C., Page, L., & Von Hippel, W. (2019). Strategically delusional. *Experimental Economics*, 1-28.

Solis, C. (2019, October 10). Rebecca Goldstein Interview What Is It Like to Be a Philosopher? https://www.whatisitliketobeaphilosopher.com/rebecca-goldstein/

Sonnak, C., & Towell, T. (2001). The impostor phenomenon in British university students: Relationships between self-esteem, mental health, parental rearing style and socioeconomic status. *Personality and Individual differences, 31*(6), 863-874.

Spitoni, G. F., Serino, A., Cotugno, A., Mancini, F., Antonucci, G., & Pizzamiglio, L. (2015). The two dimensions of the body representation in women suffering from Anorexia Nervosa. *Psychiatry Research, 230*(2), 181-188.

Steinhausen, H.-C. (2002). The outcome of anorexia nervosa in the 20th century. *American journal of Psychiatry, 159*(8), 1284-1293.

Stice, E., Shaw, H., & Nemeroff, C. (1998). Dual pathway model of bulimia nervosa: Longitudinal support for dietary restraint and affect-regulation mechanisms. *Journal of social and clinical psychology, 17*(2), 129-149.

Stich, S. P. (1978). Beliefs and subdoxastic states. *Philosophy of Science, 45*(4), 499-518.

Stokes, D. (2013). Cognitive penetrability of perception. *Philosophy Compass, 8*(7), 646-663.

Stokes, D. (2014). Cognitive Penetration and the Perception of Art. *dialectica, 68*(1), 1-34.

Tan, J. O., Hope, T., Stewart, A., & Fitzpatrick, R. (2006). Competence to make treatment decisions in anorexia nervosa: thinking processes and values. *Philosophy, psychiatry, & psychology: PPP, 13*(4), 267.

Thompson, M. A., & Gray, J. J. (1995). Development and validation of a new body-image assessment scale. *Journal of personality assessment, 64*(2), 258-269.

Thompson, T. (2004). Failure–avoidance: Parenting, the achievement environment of the home and strategies for reduction. *Learning and Instruction, 14*(1), 3-26.

Thompson, T., Davis, H., & Davidson, J. (1998). Attributional and affective responses of impostors to academic success and failure outcomes. *Personality and Individual differences, 25*(2), 381-396.

Tong, J., Mao, O., & Goldreich, D. (2013). Two-point orientation discrimination versus the traditional two-point test for tactile spatial acuity assessment. *Frontiers in human neuroscience*, 7, 579.

Tosi, G., & Romano, D. (2020). The longer the reference, the shorter the legs: How response modality affects body perception. *Attention, Perception, & Psychophysics*, 1-13.

Tuschen-Caffier, B., Bender, C., Caffier, D., Klenner, K., Braks, K., & Svaldi, J. (2015). Selective visual attention during mirror exposure in anorexia and bulimia nervosa. *PLOS one, 10*(12), e0145886.

Valenzuela-Moguillansky, C. (2013). An Exploration of the Bodily Experience of Persons Suffering from Fibromyalgia. *Constructivist Foundations, 8*(3).

Van den Berg, E., Houtzager, L., de Vos, J., Daemen, I., Katsaragaki, G., Karyotaki, E., Cuijpers, P., & Dekker, J. (2019). Meta-analysis on the efficacy of psychological treatments for anorexia nervosa. *European Eating Disorders Review, 27*(4), 331-351.

Van Leeuwen, N. (2014). Religious credence is not factual belief. *Cognition, 133*(3), 698-715.

Vandereycken, W. (2006). Denial of illness in anorexia nervosa—a conceptual review: part 1 diagnostic significance and assessment. *European Eating Disorders Review: The Professional Journal of the Eating Disorders Association, 14*(5), 341-351.

Vandereycken, W., & Van Humbeeck, I. (2008). Denial and concealment of eating disorders: a retrospective survey. *European Eating Disorders Review: The Professional Journal of the Eating Disorders Association, 16*(2), 109-114.

Vitousek, K., Watson, S., & Wilson, G. T. (1998). Enhancing motivation for change in treatment-resistant eating disorders. *Clinical psychology review, 18*(4), 391-420.

Vitousek, K. M. (1996). The current status of cognitive-behavioral models of anorexia nervosa and bulimia nervosa. In P. Salkovskis (Ed.), *Frontiers of Cognitive Therapy*. The Guilford Press.

Vollberg, M. C., Gaesser, B., & Cikara, M. (2021). Activating episodic simulation increases affective empathy. *Cognition, 209*, 104558.

Waller, G., & Hodgson, S. (1996). Body image distortion in anorexia and bulimia nervosa: the role of perceived and actual control. *The Journal of nervous and mental disease, 184*(4), 213-219.

Weinstein, S. (1968). Intensive and extensive aspects of tactile sensitivity as a function of body part, sex and laterality. *The skin senses*.

Wickham, H. (2016). *ggplot2: elegant graphics for data analysis*. Springer.

Wignall, S. J., Thomas, N. A., & Nicholls, M. E. (2017). Fat or fiction? Effects of body size, eating pathology, and sex upon the body schema of an undergraduate population. *Body image, 23*, 135-145.

Williamson, D. A., White, M. A., York-Crowe, E., & Stewart, T. M. (2004). Cognitive-behavioral theories of eating disorders. *Behavior modification, 28*(6), 711-738.

Wolf, N. (2013). *The beauty myth: How images of beauty are used against women*. Random House.

Wooldridge, T. (2018). The entropic body: Primitive anxieties and secondary skin formation in anorexia nervosa. *Psychoanalytic Dialogues, 28*(2), 189-202.

Worsnip, A. (2018). The conflict of evidence and coherence. *Philosophy and Phenomenological Research, 96*(1), 3-44.

Young, A. W., Leafhead, K. M., & Szulecka, K. (1994). The capgras and cotard delusions. *Psychopathology, 27*(3-5), 226-231.

Zhang, Y., Touyz, S., Burns, B., Vartanian, L., & Abbott, M. (2014). An empirical examination of the mislabelling of fat as an emotion in sub-clinical eating disorder groups. *Journal of Eating Disorders, 2*(1), 1-1.