

On the roles of false belief and recalcitrant fear in anorexia nervosa

Somogy Varga & Asbjørn Steglich-Petersen

Penultimate draft, forthcoming in *Mind & Language*

Abstract: The DSM-5 highlights two essential psychological features of anorexia nervosa (AN): recalcitrant fear of gaining weight and body image disturbance. Prominent accounts grant false beliefs about body weight and shape a central role in the explanation of AN behavior. In this article, we propose a stronger emphasis on recalcitrant fear. We show that such fear can explain AN behavior without the intermediary of a false belief, and thus without the associated explanatory burdens and conceptual difficulties. We illustrate how shifting the emphasis from false belief to recalcitrant fear can supplement a number of different non-doxastic models of AN.

Keywords: anorexia nervosa, false beliefs, insight, irrationality, recalcitrant emotions

1. Introduction

Anorexia nervosa (AN) is a condition generally characterized by the persistent restriction of food intake resulting in body weight that is significantly below a normal level for age, sex, developmental trajectory, and physical health (American Psychiatric Association (APA), 2013, pp. 339–341). AN is difficult to treat (Vitousek et al., 1998) and it is associated with the highest mortality rate of all psychiatric disorders (Smink et al., 2013). While its pathogenesis is poorly understood, *The Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) points to two essential psychological features of AN that are assumed to play key roles in its development and maintenance (APA, 2013, pp. 339–341).

First, individuals with AN typically exhibit an intense fear of becoming fat or gaining weight. *The International Classification of Diseases 11th Revision* (ICD-11) speaks of “extreme fear of weight gain” (WHO, 2019, 6B80).¹ Moreover, this intense emotional state is described as a *recalcitrant fear*, in the sense that the fear does not recede when individuals with AN are confronted with evidence that they are far from a body weight that would warrant it. As the DSM-5 explains, “this intense fear of becoming fat is usually not alleviated by weight loss. In fact, concern about weight gain may increase even as weight falls” (APA, 2013, p. 340).

¹ In more detailed fashion, the ICD-10 spoke of a “dread of fatness and flabbiness of body contour” (WHO, 2004, F50.0).

Second, the DSM-5 highlights what the literature refers to as a “body image disturbance”. This is usually conceived as having both perceptual and doxastic components, with special emphasis on *false beliefs* related to body weight and shape (e.g., “I am fat”; “I am overweight”) that are unresponsive to counterevidence and typically fail to recede upon losing weight.

Of these two psychological features, the literature has focused more on the body image disturbance, and in large part, implicitly or explicitly, lets an attribution of false beliefs about body weight and shape play a central explanatory role with respect to the characteristic AN behavior.² According to what we will call the “false belief model”, AN behavior should essentially be explained as the result of AN patients acting on a desire to attain an ideal body weight combined with a false and irrational belief that one significantly exceeds that weight. However, while simple and intuitive, this model is hampered by serious conceptual and empirical downsides: it is unclear how such false beliefs are formed and maintained in spite of the strong evidence to the contrary; it leaves AN behavior in patients without false beliefs unexplained; and it makes it puzzling how AN patients could have genuine insight into their own condition.

In this article, we argue that these difficulties should lead us to abandon the strong emphasis on false belief in our understanding of AN and AN behavior. While some AN patients likely do harbor such false beliefs, we should attribute them to patients more sparingly, and assign them a less prominent explanatory role when we do. Instead, we propose a stronger explanatory emphasis on the first of the two central psychological features mentioned above: recalcitrant fear of being or becoming overweight. As we will show, such fear can explain AN behavior without the intermediary of a false belief, and thus without the explanatory burdens associated with attributing such a belief; it can explain why some AN patients do develop false beliefs; and it makes better sense of how AN patients could be insightful about their own condition. In addition to this, we will argue, it can help explain the occurrence of abnormal body-shape experiences, and it is better placed to explain the nuances of AN behavior.

Clarifying this issue is important not only because a potential risk factor in the maintenance of restrictive eating behaviors continues to elude our understanding (Konstantakopoulos et al., 2012; Metral et al., 2014), but also because the nature and extent to which AN involves false belief and delusion have important implications for the effectiveness and justifiability of certain

² See, for example, Vitousek (1996) for a review. Some models are arguably exceptions to this tendency, for example: the cognitive-interpersonal maintenance model (Schmidt & Treasure, 2006); the habit model (Walsh, 2013); and the control model (Fairburn et al., 1999). See Section 5 for a discussion.

treatments. In severe cases, delusion can be associated with incompetence and hence figure as a potential reason to justify overruling the decision of patients to refuse potentially lifesaving food (for a discussion, see Draper, 2000).

We proceed as follows. In Section 2, we sketch the false belief model and draw attention to three serious difficulties hampering it. In Section 3, we introduce our alternative model of AN emphasizing recalcitrant fear, which explains AN behavior without the intermediary of a false belief. In Section 4, we show that this model not only avoids the difficulties associated with the false belief model, but has additional explanatory advantages. In Section 5, we discuss how our model supplements rather than competes with existing non-doxastic accounts of AN, before concluding in Section 6.

2. The false belief model

As mentioned above, AN is classically associated with false belief about body size, and most models of AN place a strong emphasis on false belief in the explanation of AN behavior. Although there are significant differences between such models, they agree in explaining AN behavior as the result of AN patients acting on a false belief about their body size in order to further a desire not to be overweight.

Of these two explanatory factors, the desire appears at least initially less problematic, and is certainly common in the general population, which may be one of the reasons why the literature has focused on the false beliefs about body weight and shape in explaining AN behavior. The literature describes these beliefs as displaying a wide range of intensity and persistence, and they are comprehended as occurring along a spectrum, with delusional beliefs occupying the extreme end. Although influential researchers have early on highlighted “delusional denial of thinness” as a general feature (Bruch, 1973), the beliefs in question are today mostly described as irrational “overvalued ideas” (for a critique, see Rahman et al., 2019), which, according to the DSM-5, are “unreasonable and sustained beliefs maintained with less than delusional intensity” (APA, 2013, p. 826). Correspondingly, researchers distinguish between AN “with good insight”, “with poor insight” and “with psychotic features”, depending on the degree to which patients understand the irrationality of the beliefs (Phillips et al., 1995). On such a view, insight is dimensional, ranging from delusional, to having overvalued ideas, to being completely insightful (Hartmann et al., 2013). This leads us to the difficulties that we see for the false belief model.

2.1 Explaining the formation and maintenance of false beliefs

The first serious difficulty for the false belief model is explaining how it is possible for AN patients to form and maintain such false beliefs in spite of the strong counterevidence they are continually exposed to, such as testimony from acquaintances and health care workers, or the use of a scale to determine their weight. This is especially puzzling in individuals who themselves recognize that these beliefs are irrational, but it stands in need of explanation even in individuals without insight. We will illustrate the seriousness of this task by discussing Gadsby's (2020) recent proposal to explain the false beliefs by reference to two distinct factors: abnormal body experiences and a desire driven bias.

Gadsby's account adopts an approach from research on delusions that explicates delusional beliefs by highlighting abnormal experiences that are taken to be involved in their acquisition. Given the fact that many individuals with AN report experiencing their bodies as overly "large" or "fat", defined as "the somatic sensation of being overweight that does not entirely correlate with one's actual weight" (Mehak & Racine, 2020, p. 1400), it is initially plausible to think that these experiences could cause and maintain the relevant body size beliefs.

Some indirect evidence comes from research on body image disturbances that uses body size estimation (BSE) tasks, showing that while individuals with AN do not exhibit general problems with size perception, they tend to overestimate their body size (Cash & Deagle III, 1997; for reviews, see Smeets et al., 1999; Gardner & Brown, 2014). As the BSE tasks solicit participants to estimate the size of their body without visual reference to their body (e.g., looking in a mirror), completing them requires consulting a visual, mental image of their body, which is inferred to represent their body as larger than reality. Such conscious mental imagery of an oversized body likely involves experiencing the body as oversized (Gadsby, 2017a).

While BSE studies focus on visual modality, others indicate that body image disturbance in individuals with AN affects multiple modalities and is comprehensive, including enlarged body schemas that affect motor control and motor simulation abilities (Kaizer, 2014).³ For example, when completing tasks such as passing through apertures, compared to controls, individuals with AN rotate their shoulders for relatively larger door widths, suggesting that their body schemas present their body size as larger than their actual size (Beckmann et al., 2021; Keizer et

³ The body schema builds on proprioceptive information from the muscles, joints, and skin and comprises sensorimotor representations of the body. It is functionally separable from the body image (e.g., during the rubber hand illusion) and is involved in posture and movement (de Vignemont et al., 2021).

al., 2013; Metral et al., 2014). The effect also occurs when individuals with AN mentally simulate walking through apertures, but it disappears in conditions where the task is to mentally simulate another person walking through an aperture (Guardia et al., 2012).

While it is plausible to think that these experiences could help ground and reinforce false beliefs about body size,⁴ Gadsby recognizes that it remains puzzling why such beliefs remain unresponsive to the strong and unambiguous counterevidence that patients are typically exposed to. In order to explain this, Gadsby (2020) proposes that a “(pathologically) strong desire to neither be overweight nor be seen as overweight” (p. 622) constitutes a second factor, leading to motivational biases in the acquisition and interpretation of evidence.

First, individuals engage in biased evidence gathering like repeated body checking that selectively focuses on body parts that are perceived as “fat”. While this raises the likelihood of finding supporting evidence for the belief “I am fat”, individuals with AN also tend to engage more in comparing themselves to others than controls, and they compare themselves more with individuals who they believe have thinner bodies (Blechert et al., 2009; Green et al., 2009), resulting in more “they are thinner than me” judgments.

Second, individuals with AN exhibit bias when it comes to the interpretation of ambiguous information related to their own body size. For example, they tend to interpret comments from others like “You look healthy” as evidence for the belief that “I am fat” (Jackman et al., 1995), and tend to discount or explain away counterevidence. Espeset et al. (2011) note that participants sometimes reject strong evidence about their low weight from a scale, and offer alternative, implausible explanations like the scale being manipulated or broken, or that they have an uncommonly light bone structure.

Gadsby (2020) argues that such biases arise from the undesirability of harboring particular false beliefs. While we usually seek to acquire and maintain true beliefs and avoid false ones, not all false beliefs are equally undesirable to hold. Depending to a large extent on one’s psychological makeup (e.g., desires, values, risk aversiveness, and commitments), there are certain “costly” false beliefs that agents have a stronger desire to avoid than others. For example, falsely believing that one is vulnerable to liver disease may be costly in the sense that it causes both psychological distress and costly behaviors (seeking medical treatment, performing

⁴ Indeed, in more recent work, Gadsby (2022a) argues that beliefs formed on the basis of proprioceptive misperceptions of bodily boundaries could, in some cases, render false beliefs about body size rational. It is unclear, however, to what extent Gadsby takes this point to address the explanatory issue at hand. Even if the false beliefs can in some cases be characterized as rational, we take it that, in most cases, there will still be a question as to why the beliefs remain unresponsive to counterevidence.

intensive exercise, and changing dietary habits) that one had otherwise not prioritized. However, many would regard falsely believing that one is *not* vulnerable to liver disease as even more costly, likely leading to a missed opportunity to implement potentially life-saving changes.

According to Gadsby, in individuals with AN, when the cost disparity reaches a significant level, the individual's treatment of evidence becomes biased toward minimizing costly errors, which can promote the adoption and maintenance of certain false beliefs. This account draws inspiration from literature suggesting the existence of a “twisted” form of self-deception (i.e., acquiring false beliefs due to motivational influences), which involves deceiving oneself into believing something one desires to be false (Mele, 1999, 2001). Applied to the case of AN, holding certain false beliefs (e.g., “I am fat”) is likely costly, as it is associated with serious psychological distress, especially in light of the strong preoccupation with body size and strong desire to be thin and to be perceived as thin. However, it can be offset by the costs associated with holding the opposite false belief (e.g., “I am thin”), because such an error will most likely reduce the force of the motivation to exercise and to decrease food intake. For someone with such a strong desire for thinness, holding the false belief “I am fat” could be more advantageous, because it ensures that one will never let one’s guard down and become overweight. The following diagram summarizes Gadsby's two-factor version of the false belief model:

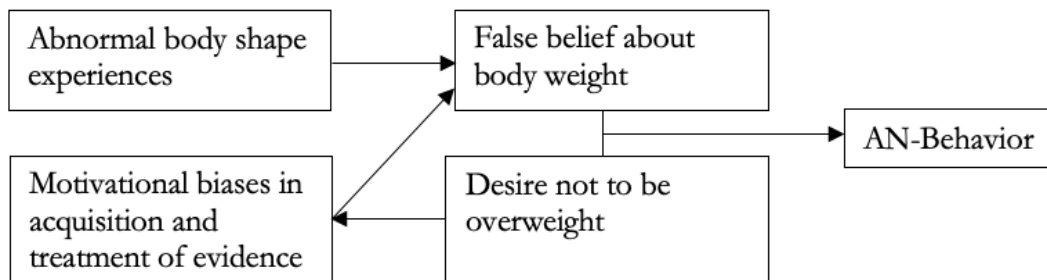


Figure 1: The two-factor false belief model.

Overall, the two-factor account aims to show that in individuals with AN, the (pathological) desire not to be overweight leads to biases and a “twisted” form of self-deception culminating in the acquisition of the false belief “I am fat”. The strong desire to maintain thinness renders holding the false belief “I am fat” less costly than the error of holding the opposite false belief. It constitutes a second factor required to explain the acquisition and maintenance of the false belief.

Gadsby's two-factor account offers a rich and helpful perspective on AN behavior, and it rightly insists that certain bodily experiences in individuals with AN could help cause and maintain the false belief "I am fat". But it also faces some difficulties. The success of the two-factor account depends on the truth of the claim that in individuals with AN, self-deception and the acquisition of the false belief "I am fat" constitutes the less costly alternative. But it bears noting that the cost associated with the acquisition of this belief is quite high, as it will surely increase the psychological distress caused by the dread of fatness. To compare, if somebody with arachnophobia comes to acquire the belief that he is in close proximity to a spider, then the penetrating fear that already haunts him can be expected to deepen the distress. But then, the explanation does not sit well with most accounts of self-deception, which, quite plausibly, maintain that the advantage of such false beliefs is typically the reduction of unease and anxiety (Barnes, 1997, p. 46).

Gadsby could reply that the increased emotional distress is outweighed by the costs associated with holding the opposite false belief (e.g., "I am thin"), which likely reduces the motivation to decrease food intake and raises the probability of gaining weight. But this line of reply is only truly convincing if one can reasonably expect that holding the false belief reliably leads to this result. This, however, is far from certain. Coming to falsely believe the proposition "I am fat" could easily lead to resignation or to some kind of acceptance of one's body size, which may lead to the opposite result: It may reduce the motivation to decrease food intake and raise the probability of gaining weight. In addition, if the aim of the belief acquisition is to decrease food intake and lower the probability of gaining weight, then it seems that a less costly alternative is possible: *withholding judgment* about one's body might be an equally effective measure to ensure that one remains vigilant about one's weight. So why not opt for this alternative? Some theorists argue that withholding of judgment is subject to a higher degree of voluntary control than belief in the kind of cases under discussion. For example, according to Schroeder (2012), one can decide to withhold judgment if one deems the consequences of acting on a false belief sufficiently grave. Lastly, it bears mentioning that motivating the kind of self-deception described by Gadsby seems to require a rather intricate level of instrumental reasoning, as well as very specific expectations about the effects of holding the relevant false belief for one's own behavior. This in itself speaks against attributing this kind of self-deception to AN patients on the scale required for Gadsby's account to be sufficiently general.

2.2 AN behavior without false beliefs

While the above points are not in themselves decisive, they do illustrate that explaining the formation and maintenance of false beliefs about body size is far from trivial. But there is a more immediate problem with the false belief model: Not all AN patients seem to harbor false beliefs about their body size, yet still display the characteristic AN behavior. Such behavior will go unexplained by the model.

At least some AN patients report a more complicated constellation of doxastic states and indicate that their AN behavior persists in the absence of believing that they are overweight. Here are two examples:

“Deep down I know that I cannot be fat. But I believe that your brain can deceive you in many ways. I cannot trust my senses, because they keep telling me that I’m fat” (unnamed patient, quoted in Kaizer, 2014, p. 15).

“I actually understand that it can’t be quite so, because I do know the size of my clothes, so it can’t actually be true. And it’s rather confusing and sometimes frustrating because I actually know that it’s impossible that I could be that fat, because then these clothes wouldn’t fit me. But at the same time it doesn’t make sense with how I see myself” (Johanna, BMI 16, quoted in Espeset et al., 2011, p. 184).⁵

In these examples, individuals with AN report not only distrusting their (proprioceptive) experiences, but also that they *know* or *understand* that they could not be fat. Since they know or understand that they are not fat, they do not hold the belief that they are fat.⁶ But then, given the prominent explanatory role that belief plays in the false belief model, it seems unable to offer a satisfactory explanation of the relevant behavioral patterns. It leaves AN behavior unexplained in those individuals with AN that do not harbor false beliefs.

The existence of such cases does not exclude, of course, that false beliefs can play a strong role in causing AN behavior in those patients that *do* harbor such beliefs, or in patients that experience fluctuating belief states, moving back and forth between believing and not believing

⁵ Others cite perceptual counterevidence: “I did not think I was too fat; quite the contrary. There was a skeleton peering at me from behind the looking glass” (Halban, 2009, p. 184; cited in Bowden, 2012).

⁶ While it may be possible to believe a proposition that one simultaneously believes to be false, it is not possible to know a proposition one believes to be false, since such a belief would defeat the knowledge. This kind of defeat is known as “doxastic defeat” (Lackey, 1999). For a recent explanation, see also Taylor (2022).

that they are overweight. But that would imply that AN behavior should be explained differently, depending on whether or not the individual patient has false beliefs about her weight. We take it that it would be preferable for explanatory models to explain AN behavior in substantially the same way across all patients, thus motivating the search for a model that does not place a heavy explanatory burden on false beliefs.

Our point here relies on the common general assumption that the strongest explanatory models explain the target phenomenon in substantially the same way across all instances (see, e.g., Keas, 2018). Of course, it may turn out that such a uniform model of AN behavior cannot be provided due to significant heterogeneity in the population of individuals with AN. In that case, questions might arise as to whether AN should be split into several categories. Still, we take it that if a unified explanatory model can be found, such a model would be preferable.⁷

2.3 A puzzle about insight

Related to the above, in many cases, individuals with AN recognize that their thoughts or ideas about their body size are not merely false, but irrational in the sense that they are adopted without appropriate evidence and/or remain unresponsive to counterevidence. Correspondingly, as mentioned, researchers distinguish between AN “with good insight”, “with poor insight”, and “with psychotic features” depending on the degree to which the patient is aware of the irrationality of these ideas (Phillips et al., 1995). But how is such insight possible if the insight targets a belief of one’s own? We will bring this question into sharper focus by drawing on an analogous puzzle about insight in obsessive compulsive disorder (OCD) recently described by Taylor (2022) and further discussed in our own previous work, which we adapt to AN below (Steglich- Petersen & Varga, 2022).

We may distinguish between two kinds of insight. First, insight might amount to recognizing that the contents of one’s thoughts about body size are *false*. This is called “world-directed” insight, because it involves recognizing that something is false as a matter of worldly fact.

⁷ Gadsby (2017b) notes the heterogeneity amongst AN patients, and that many do not hold the relevant false beliefs about their body size. Consequently, he embraces a “piecemeal approach” and maintains that in the population that harbors these false beliefs, understanding how they arise will help solve a piece of the puzzle. Embracing such a piecemeal approach would certainly alleviate some of the difficulties we have pointed to. Nonetheless, even in cases in which AN patients hold the relevant false beliefs, it cannot be readily assumed that the beliefs will constitute the best candidates for explanation, and the search for a more unified explanation would thus still be warranted.

World-directed insight: Where S has a thought m with content p, good insight into m can take the form of recognizing that p is false.

For example, an individual with AN plagued by the thought that she is overweight might recognize that she is not in fact overweight. This can amount to *knowing* that the thought is false, and it is quite natural to describe AN patients with good insight as having such knowledge. Indeed, to have “insight” or “recognize” imply what we would normally consider knowledge, since these mental states are both *factive* (i.e., true) and *rational* (i.e., held on a reasonable basis).

Rather than being a matter of recognizing that one's thoughts about body size are false, the second kind of insight is a matter of recognizing that they are *irrational*, for example, in the sense of conflicting with one's evidence. Since rationality depends on one's own state of mind, rather than on the world, Taylor calls this “self-directed” insight.

Self-directed insight: Where S has a thought m with content p, good insight into m can take the form of recognizing that it is irrational to have m.

For example, an AN patient plagued by the thought that she is overweight might recognize that this thought is *irrational*. Such recognition can amount to knowing that the thought is irrational.

The two kinds of insight typically go together, but do not ascribe the same property to the target state. A thought can be false without being irrational (e.g., if it is based on good reasons, but happens to be false), and it can be irrational without being false (e.g., if it is not based on good reasons, but turns out to be true).⁸

The insight puzzle arises when we try to describe the first-order thoughts that the two kinds of insight take as their objects: None of the most natural characterizations of them can accommodate both world-directed and self-directed insight. Candidate characterizations can be divided into two groups: beliefs (or some other kind of doxastic states) and non-doxastic states. The thesis that AN thoughts about body size are simply beliefs is in many ways a natural characterization, and it would also accommodate the possibility of self-directed insight, since it is possible for beliefs to be irrational and hence to be the objects of self-directed insight. And if AN thoughts were beliefs, they often *would* be irrational because they lack appropriate

⁸ The distinction between falsity and irrationality is standard in epistemology. For a discussion, see Cohen and Comesana (2023).

evidential grounding. However, if understood as beliefs, AN thoughts could not be the objects of *world-directed* insight, which consists in knowing that the content of the target thought is false. If AN thoughts are beliefs that some content *p* is true, the possession of such beliefs would make it impossible to know that *p* is false, since one cannot know a content to be false that one believes to be true (see fn. 4). Hence, insofar as world-directed insight is possible, AN thoughts cannot be beliefs.

These issues might lead us to instead interpret AN thoughts as a different kind of doxastic state, such as *doubting* or *suspending judgment* about one's weight. But such interpretations fall prey to the same problem: While they may be capable of being irrational and thus figure as objects of self-directed insight, they cannot be the objects of world-directed insight; they are incompatible with knowledge, since one cannot know that one is not overweight while doubting or suspending judgment that one is not overweight.

The above considerations speak against interpreting AN thoughts as doxastic states, such as false beliefs. It should be noted however, that it is equally puzzling how non-doxastic states could be the objects of both kinds of insight. Suppose, for example, that AN thoughts should be interpreted as mere *aliefs* or *imaginings*. Both of these states allow for world-directed insight: Aliefs are a type of mental states introduced by Tamar Gendler (2008a) to explain cases where one is disposed to think, feel, and act in ways that are contrary to one's beliefs. For example, when stepping on to the Grand Canyon Skywalk, I might be disposed to think of and feel afraid of falling, even while believing that the walkway is safe, in which case I can be said to "alieve" that it is *not* safe. As Taylor (2022) notes, this is compatible with knowing that the walkway is safe. Similarly, if imagining is simply a matter of having a mental image or representation of some scenario or proposition, then imagining *p* is also compatible with knowing that *p* is false. But while aliefs and imaginings allow world-directed insight into their falsity, they do not allow self-directed insight into their irrationality, since these states cannot be irrational. Aliefs are defined as *arational* states (Gendler, 2008b, p. 557), and while imaginings may be unpleasant or irrelevant, they cannot in themselves be irrational. So, for non-doxastic accounts of AN, such as our own, a puzzle remains about self-directed insight. We return to address that below.

3. The explanatory potential of recalcitrant fear

Above, we have pointed out three difficulties associated with the false belief model. While we do not take these difficulties to settle the matter, we think that they are sufficiently serious to

motivate the search for a model that does not rely on attributing false beliefs about weight to individuals with AN. We now turn to show that an account instead stressing the role of recalcitrant fear can explain AN behavior without having to face the difficulties discussed above.

We take our cue from Michael Strober (2004), who speculates that weight aversion and compulsive dieting in AN express a heritable tendency for fear-based learning and phobic avoidance, effects which likely interact with personality traits that accelerate emotional conditioning. Moreover, Strober holds that a property of this overexpressed fear-based learning is that weight phobia is encapsulated from cognitive control, accounting for the unusually protracted and often chronic course of the illness. He further hypothesizes that underlying the conditioning of weight-related fear are abnormalities in limbic structures known to be involved in emotion-driven behavior, specifically the propensity to fear and to rapidly acquire conditioned fear behavior (Strober, 2004, p. 507).

What is striking about Strober's proposal in the present context, is that AN behavior is understood as a direct expression of fear, without the intermediary of false belief, and that this fear is not itself understood as caused or sustained by false beliefs, but rather as evading cognitive control by what will often be a realistic understanding of one's own condition. Below, we provide an elaboration and motivation of this basic idea. For this, we draw on the philosophical literature on recalcitrant emotions, which characteristically endure in spite of standing in a tension or conflict with one's beliefs (D'Arms & Jacobson, 2003; Döring, 2015).

3.1 Recalcitrant emotion

The literature on recalcitrant emotions has focused in large part on accounting for their *irrationality* (for a more detailed account, see Steglich-Petersen & Varga 2022). When it comes to this, it is important to steer clear of the problems facing both *judgmentalism* and alternative *neojudgmentalist* accounts of emotions. On judgmentalist accounts of emotions (e.g., Lyons, 1980; Nussbaum, 2001), emotions should, at least in part, be identified with evaluative judgments or beliefs, which means that when we undergo recalcitrant emotions, we are irrational by virtue of holding inconsistent beliefs. However, this account attributes too much irrationality to those harboring recalcitrant emotions (Helm, 2001), and requires violating the principle of charity (Greenspan, 1988; D'Arms & Jacobson, 2003, pp. 129–130). The neojudgmentalist account does not face this problem, as it associates emotions with a cognitive aspect (e.g., “evaluative construal”) that is weaker than judgment or belief (e.g., Brady, 2009, 2013; de Sousa, 1991; Roberts, 2003). Nonetheless, as Helm (2001) illustrates by drawing an

analogy with recalcitrant perceptions in visual illusions, the problem is that such evaluative construals may not impute enough irrationality to recalcitrant emotion.

For our purposes, we adopt Michael Brady's (2009, 2013) neo-judgmentalist account, which comprehends emotions as responses to states of affairs in the world that are of significance to us. They involve a number of processes that mobilize motivational and cognitive resources. For example, emotions fine-tune attention and help filter information, enabling us to tackle large amounts of competing environmental stimuli (Jenkins & Oatley, 1996). They enhance stimulus detection, focus attention on the relevant stimuli, and increase perceptual sensitivity to them. For example, fear focuses attention on threatening stimuli, makes it difficult to disengage, and assigns a lower priority to the processing of neutral information (Najmi et al., 2012).

Brady (2009) argues that emotions involve *inclinations* to assent to and act on evaluative construals. For example, when fear activates our cognitive resources, this leads to increased attentional focus on stimuli that confirm an evaluative construal of the situation as dangerous. This, in turn, inclines us to assent to the situation being dangerous and to act accordingly. Of course, it is possible to resist assenting and acting in ways that the particular construal would dictate. As Brady argues, this creates the kind of tension that characterizes recalcitrant emotional experience: One is *inclined* to act on and assent to an evaluative construal, but assents instead to an opposing construal.

Brady's account explains the irrationality of recalcitrant emotion without ascribing inconsistent beliefs or judgments. On his account, recalcitrant emotions are irrational because they involve *significant practical and cognitive costs*, and because they *violate a substantive epistemic norm* (Brady, 2009, pp. 426–429). First, recalcitrant emotion involves mobilizing motivational resources to prepare for action that the subject does not believe is required. This is a waste of motivational resources and it likely distracts the agent from preparing to act in ways that would assist the pursuit of other goals. For example, if one believes that the environment is perfectly safe, then remaining increasingly attentive to signs of danger is a waste of resources. In addition, in order to pursue her goals, the agent needs to activate additional resources to counteract the attentional bias and the inclination to believe.

Second, being inclined to assent to a construal in light of a conflicting evaluative belief can be seen as epistemically irrational in several ways. It is epistemically irrational in virtue of *what* construal it inclines the subject to assent to. This is because recalcitrant emotion inclines one to assent to a construal that one regards as false. Also, it is epistemically irrational in virtue of *how*

it inclines the subject to assent. This is because the inclination is not based on what the subject regards as good reasons. In fact, the agent will take these reasons to support her considered judgment that conflicts with the emotion. Finally, it is epistemically irrational in that it predisposes the agent to search for reasons that would justify her emotional experience.

3.2 The recalcitrant fear model of AN

With this in place, we are in a position to sketch our recalcitrant fear model of AN. Just as the false belief model, our model acknowledges that experiential body image disturbances might play a role in AN behavior. However, instead of emphasizing false beliefs, it stresses the role of recalcitrant fear, which we take to be able to explain AN behavior without the intermediary of a false belief, and thus without having to confront the difficulties faced by the false belief models. The following diagram summarizes the recalcitrant fear model:

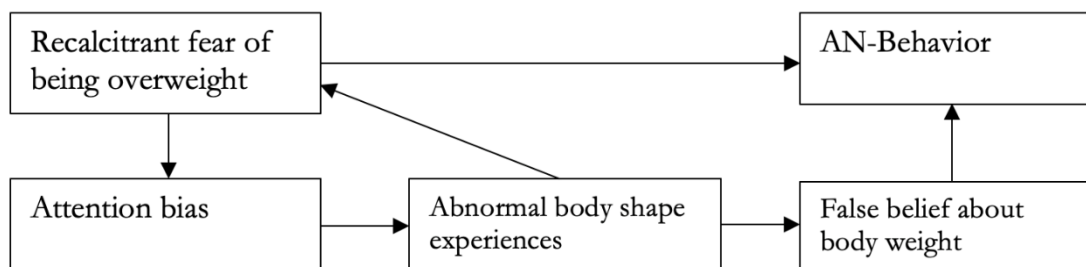


Figure 2: Recalcitrant fear model.

According to this model, recalcitrant fear can explain AN behavior along two pathways. First, along a *primary* and *direct* pathway, the model maintains that recalcitrant fear can cause AN behavior without requiring mediating beliefs. This is in line with the above understanding of recalcitrant emotions as in part involving inclinations to act in ways that are not in line with the subject's beliefs. When these emotion-driven inclinations overwhelm the subject's cognitive control, they can thus give rise to behavior in the absence of the relevant belief. In support of the direct pathway, note that we often appeal to emotions to explain behavior without citing beliefs and desires. In fact, it is now relatively common among researchers to think that emotions

can display independent motivational force, and replace or supplement belief- desire pairs (see e.g., Döring, 2003; Scarantino, 2017; Yip, 2022).⁹

Second, because we accept the view that emotions adjust attentional mechanisms leading to perceptual changes that assist the mobilization of motivational and cognitive resources, the model can also accommodate that the recalcitrant fear can lead to attentional biases and more abnormal experiences of the body, which, in turn, can cause false beliefs and thus create a *secondary* and *indirect* pathway to AN behavior. Displaying some resemblance to the false belief model, this pathway allows false beliefs to play a role in the explanation of AN behavior. But given that the indirect pathway is not the only option leading to AN behavior, our model accommodates that some individuals with AN do not harbor such false beliefs. Crucially, our model stops short of assigning false beliefs a prominent explanatory role and, in our view, the indirect pathway and the ascription of false beliefs may be reserved to the relatively rare cases of AN without insight.

4. Advantages of the recalcitrant fear model

We now turn to show that our model faces none of the issues that we have highlighted in our discussion of the false belief model. Moreover, we argue that it displays additional explanatory advantages.

First, by attributing false beliefs about body size to AN patients much more sparingly, our account does not face the task of explaining how such false beliefs could be formed and maintained despite the presence of strong counterevidence. It was puzzling how such beliefs could arise on the scale supposed by the false belief model; but it is not puzzling that they occasionally result from the attentional biases caused by recalcitrant emotion in the way outlined above. Our account therefore does not need to posit a complicated “twisted” form of self-deception, on which one deceives oneself into believing something one desires to be false, or suppose that the false belief “I am fat” is less costly than the opposite false belief.

One could worry, of course, that the recalcitrant fear model does not fare much better than the false belief model in this respect: Is it not equally puzzling how recalcitrant fear can arise and persist in spite of counterevidence? We do not think so. It is standard in evolutionary psychology to comprehend fear reactions as modular evolutionary adaptations that aim to solve problems

⁹ For example, it is natural to cite someone's anger to explain why he slammed the door. While some think that a number of actions resist belief-desire explanation, others argue that the motivational force is best comprehended as generating and modulating desires, which would render it consistent with the belief-desire analysis (see, e.g., Yip, 2022).

faced by our hunter-gatherer ancestors (Cosmides & Tooby, 2000; for a critical discussion, see de Sousa, 2008). Fear prepares us to react, and once it is activated, it functions relatively, but not entirely, encapsulated from reasoning (Majeed, 2022). Thus, recalcitrance can at least in part be explained by the modular nature of emotions (Griffiths, 1990). There are some harmless mismatches between some of the evolved emotional responses and our contemporary environments. For example, low-intensity recalcitrant fears that do not seriously interfere with our lives are very common (e.g., when standing on the edge of a cliff or seeing a snake). However, certain fears are harmful and maladaptive, and best seen as dysregulated responses to generally non-threatening environmental factors (Buss, 2019). Some of them are directed towards objects and situations that seem relevant from a phylogenetic perspective (e.g., snakes, dirtiness) whereas others (e.g., fat) make more sense from an ontogenetic perspective (Öhman & Mineka, 2001). The latter type does clearly not figure among the threats in our evolutionary history, but can rather be seen as involving an overexpression of fear-based learning and phobic avoidance (Quirk & Gehlert, 2003; Strober, 2004). On such a view, recalcitrance is not particularly puzzling: Socially reinforced unease with weight together with a propensity to fear conditioning jointly contribute to the prolonged fears that are resistant to extinction and encapsulated from cognitive control.

A second advantage of our model is its wider scope. Due to its strong emphasis on false beliefs in explanations of AN behavior, the false belief model is not able to explain AN behavior in AN patients that do *not* harbor false beliefs. In contrast, our model can make sense of how AN behavior, fueled by recalcitrant fear along the direct pathway, can persist in the absence of the false belief.

Third, our account solves the puzzle about insight. The challenge was, recall, that while beliefs cannot be the object of world-directed insight, since one cannot know what one believes to be false, the most obvious candidates for non-doxastic interpretations of the relevant state (e.g., aliefs or states of imagination) cannot be the object of self-directed insight, since these states cannot be irrational. This problem is solved by understanding the object of insight as a state of recalcitrant fear. Unlike aliefs and imaginations, emotions are clearly capable of being rational and irrational,¹⁰ and are hence also capable of being objects of self-directed insight. And it is also relatively clear what world-directed insight would amount to when directed at an emotional state, such as fear. Above, we understood emotions as involving an evaluative construal of the

¹⁰ One might object that fear at least requires that one regards the feared scenario as less than fully certain not to obtain.

states or situations that the emotions are directed at. World-directed insight into the “falsity” of an emotion would then amount to knowing that the evaluative construal associated with that emotion is false. So, world-directed insight into one’s fear of x would amount to knowing that x is not in fact dangerous. Importantly, this combination of attitudes seems possible, both psychologically and epistemically. For example, psychologically speaking, I might fear flying even though I believe flying to be safe (that is just what it means for the fear to be recalcitrant). Still, emotions of this sort do not defeat the fact that one’s belief constitutes knowledge. For example, fear of flying does not defeat knowledge that flying is safe.¹¹

This leaves the question of how world-directed insight is possible for AN patients that *do* harbor the false belief that they are overweight. According to the above reasoning, such a false belief would make it impossible for them to know that they are not in fact overweight, and thus impossible for them to have world-directed insight. But this, we take it, is an acceptable result, especially if we are right that it is appropriate to ascribe false beliefs to much fewer AN patients than commonly supposed.

So, the recalcitrant fear model avoids the difficulties faced by the false belief model. In addition to this, we want to highlight two further explanatory advantages. First, note that while the false belief model relies on abnormal body experiences to explain the formation and maintenance of false beliefs, it leaves those very experiences unexplained. An advantage of the present model is that the focus on the emotion enables a more unified explanation and casts at least some light on the experiential aspects. To see how, it is important to note first that the current evidence is inconclusive as to the exact nature of the experiential body image disturbance. In particular, the evidence is inconclusive as to whether or not such disturbances can involve *visual misperceptions* of one’s own body shape and size (Gadsby, 2022b). Alternative suggestions compatible with the evidence include the hypothesis that subjects simply apply harsher standards when reporting on what are in fact accurate visual representations of themselves; and the hypothesis that the disturbances should be understood as a matter of attending selectively to parts or aspects of one’s body that one finds the most unattractive, even if the visual representation of the body is in fact accurate. Biased attention can be extremely powerful and can easily cause subjects to report “seeing” themselves as overweight.

In lieu of decisive experimental data, there are reasons to prefer interpreting the experiential body image disturbance as a matter of attention bias rather than visual mis- perception. First,

¹¹ This might be true but it does not pose a problem to our view, as knowing p is compatible with being less than fully certain that p. For a prominent example of this fallibilist understanding, see Williamson (2000).

while the existence of attention biases is well documented, the latter interpretation would imply the truth of the more controversial thesis that visual experience can be cognitively penetrated. Second, the interpretation in terms of attention bias is the simpler of the two interpretations in the basic sense that, whereas AN is *already* known to be associated with attention bias (Tuschen-Caffier et al., 2015), it is *not* at present known to be associated with visual misperception. An explanation in terms of attention bias would thus not commit us to any new cognitive malfunctions, beyond those already known to be at work.

If the experiential body image disturbance in AN should indeed be understood as a matter of attention bias, the recalcitrant emotion model suggests a natural explanation of such bias. As mentioned above, it is uncontroversial that emotions contribute to filtering information, by directing our attention to emotionally valenced stimuli that are processed faster than neutral stimuli. So, the recalcitrant fear model may help explain the experiential aspects of AN.

This brings us to the second additional explanatory advantage: An explanation of AN behavior in terms of recalcitrant fear, rather than belief-desire pairs, makes better sense of the nuances of AN behavior. In particular, it makes better sense of the strongly ritualistic and rule-focused aspect of AN behavior, especially around eating and exercise. Common eating rituals in AN include eating very slowly or quickly; cutting food into small pieces; preventing foods from touching each other; using specific cutlery or china; chewing food a particular number of times; eating foods in a particular order; and more (Glasofer et al., 2016, p. 226). What is striking about such rituals is that they do not stand in a realistic causal relationship to the desired outcome of losing weight, or at least not a causal relationship that is strong enough to motivate the intense preoccupation with them. If the rituals were to be explained by relevant belief-desire pairs, then we would have to attribute not only false beliefs about body weight to AN patients, but also false beliefs about the effectiveness of the ritualistic behaviors with respect to the desire to lose weight. This, we take it, is not a very plausible upshot of that model. A more plausible explanation emerges if we understand the ritualistic behavior as (maladaptive) strategies to quench or control emotional distress. On this model, the AN patients in many cases understand that the rituals will not help them lose weight. But AN patients are in the painful situation that their intense fear of being overweight does not respond to their cognitive understanding of the

situation, and to quench or control the fear, they thus engage in behaviors that to outsiders will seem excessive or ineffectual.^{12 13}

5. Relationship to non-doxastic accounts

As noted above, ours is not the only model of AN that downplays the role of false belief. For example, the cognitive-interpersonal maintenance model (Schmidt & Treasure, 2006), the habit model (Walsh, 2013), and the control model (Fairburn et al., 1999), all seem to share this feature, which raises the question of how our model relates to them. Without going into too much detail, it seems to us that emphasizing recalcitrant fear in many ways supplements, rather than runs counter to, the general ideas that animate these alternative models.

According to the control model, an extreme need to control eating is a key feature of AN, which in turn is strengthened by the tendency in Western societies to link self-worth to body shape and weight. The dieting behavior is reinforced positively through a sense of success and negatively through fear of gaining weight (Fairburn et al., 1999, pp. 2–3). Our model can accommodate the idea that the fear of becoming overweight is linked to a fear of losing control. Moreover, as the need for control is a natural expression of fear, stressing recalcitrant emotion seems like a natural corollary to the control account.

Walsh's habit model is primarily an attempt to understand the persistence of AN behavior, which it comprehends as an ingrained maladaptive habit, grounded in neural mechanisms that enable the formation and persistence of habits (Walsh, 2013, p. 479). The food restricting behavior itself becomes rewarding (conditioned reinforcement), and at least in certain stages of the disorder, individuals with AN do not experience their behavior as habitual and cue driven (see Evans, 2021). Walsh (2013, p. 480) also notes that dieting behavior becomes a learned technique to control negative emotions, which fits well with our model: behavior engaged in to quench distress from recalcitrant fear can be a part of the explanation of how self-enforcing and entrenched patterns of habitual behavior develop.

Finally, the cognitive-interpersonal maintenance model stresses that AN often becomes highly valued by people with the condition and proposes various factors to explain the mainte-

¹² In Steglich-Petersen and Varga (2022), we propose a similar explanation of compulsive behavior in OCD. We take the large comorbidity between AN and OCD (Mandelli et al. (2020) estimate that 44% of AN patients have a lifetime-comorbidity with OCD) to support this similarity of explanation.

¹³ In Steglich-Petersen and Varga (n.d.), we explore a further explanatory advance of the recalcitrant fear model over the false belief model, namely its superior ability to explain body checking, which is one of the most common problematic behaviors associated with AN.

nance of this stance. Drawing on an evolutionary approach to psychopathology, Schmidt and Treasure (2006, p. 345) maintain that characteristic AN behavior has a defensive function aiming at the reduction of social threat and dealing with intense negative emotions that, according to the authors, are often triggered by close relationships. Again, it is easy to see how a recalcitrant fear of being overweight could motivate patients to value a set of behavioral patterns and habits that help them quench and control the distress of experiencing this fear.

In sum, moving emphasis from false belief to recalcitrant fear as the basic driver in AN can supplement a number of different non-doxastic models of AN.

6. Conclusion

Body image disturbance and recalcitrant fear are regarded as essential psychological features of AN that play key roles in its development and maintenance. Focused more on the body image disturbance, many explanatory attempts of AN behavior grant a central role to false beliefs about body weight. We offered reasons for thinking that we should abandon the strong explanatory emphasis on false belief. Instead, we proposed a stronger explanatory emphasis on recalcitrant fear and argued that our recalcitrant fear model not only sidesteps the explanatory burdens and conceptual difficulties associated with false belief model, but also displays additional advantages.

References

- American Psychiatric Association (Ed.). (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. American Psychiatric Association.
- Barnes, A. (1997). *Seeing through self-deception*. Cambridge University Press.
- Beckmann, N., Baumann, P., Herpertz, S., Trojan, J., & Diers, M. (2021). How the unconscious mind controls body movements: Body schema distortion in anorexia nervosa. *International Journal of Eating Disorders*, 54(4), 578–586.
- Bowden, H. (2012). A phenomenological study of anorexia nervosa. *Philosophy, Psychiatry, & Psychology*, 19(3), 227–241.
- Brady, S. B. (2009). The irrationality of recalcitrant emotions. *Philosophical Studies*, 145(3), 413–430.
- Brady, S. B. (2013). *Emotional insight: The epistemic role of emotional experience*. Oxford University Press.
- Buss, D. M. (2019). *Evolutionary psychology: The new science of the mind*. Routledge.
- Cash, T. F., & Deagle, E. A., III. (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.
- Cohen, S., & Comesana, J. (2023). Rationality and truth. In J. Dutant & F. Dorsch (Eds.), *The new evil demon problem*. Oxford University Press.
- Cosmides, L., & Tooby, J. (2000). Evolutionary psychology and the emotions. *Handbook of emotions*, 2(2), 91–115.
- D'Arms, J., & Jacobson, D. (2003). The significance of recalcitrant emotion. *Royal Institute of Philosophy Supplement*, 52,

127–145.

- de Sousa, R. (1991). *The rationality of emotions*. MIT Press.
- de Sousa, R. (2008). Against emotional modularity. In L. Faucher & C. Tappolet (Eds.), *The modularity of emotions* (pp. 29–50). University of Calgary Press.
- de Vignemont, F., Pitron, V., & Alsmith, A. J. (2021). What is the body schema? In Y. Ataria, S. Tanaka, & S. Gallagher (Eds.), *Body schema and body image: New directions* (pp. 3–31). Oxford University Press.
- Döring, S. A. (2003). Explaining action by emotion. *The Philosophical Quarterly*, 53(211), 214–230.
- Döring, S. A. (2015). What's wrong with recalcitrant emotions? From irrationality to challenge of agential identity. *Dialectica*, 69(3), 381–402.
- Draper, H. (2000). Anorexia nervosa and respecting a refusal of life-prolonging therapy: a limited justification. *Bioethics*, 14(2), 120–133.
- Espeset, E., Nordbø, R. H., Gulliksen, K. S., Skarderud, 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.
- Evans, A. (2021). Anorexia nervosa: Illusion in the sense of agency. *Mind & Language*, 1–15. <https://doi.org/10.1111/mila.12385>
- Fairburn, C. G., Shafran, R., & Cooper, Z. (1999). A cognitive behavioural theory of anorexia nervosa. *Behaviour Research and Therapy*, 37(1), 1–13.
- 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. (2020). Self-deception and the second factor: How desire causes delusion in anorexia nervosa. *Erkenntnis*, 85(3), 609–626.
- Gadsby, S. (2022a). The rationality of eating disorders. *Mind & Language*, 1–18. <https://doi.org/10.1111/mila.12421>
- Gadsby, S. (2022b). Visual misperception in eating disorders. *Perception*, 50(11), 933–949.
- 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.
- Gendler, T. (2008a). Alief and belief. *Journal of Philosophy*, 105(10), 634–663.
- Gendler, T. (2008b). Alief in action (and reaction). *Mind & Language*, 23(5), 552–585.
- Green, M. A., Scott, N. A., Cross, S. E., Liao, K. Y. H., Hallengren, J. J., Davids, C. M., Carter, L. P., Kugler, D. W., Read, K. E., & Jepson, A. J. (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.
- Greenspan, P. (1988). *Emotions and reasons*. Routledge.
- Griffiths, P. (1990). Modularity, and the psychoevolutionary theory of emotion. *Biology and Philosophy*, 5(2), 175–196.
- Guardia, D., Conversy, L., Jardri, R., Lafargue, G., Thomas, P., Dodin, V., Cottencin, O., & Luyat, N. (2012). Imagining one's own and someone else's body actions: Dissociation in anorexia nervosa. *PLoS One*, 7(8), e43241.
- Halban, E. (2009). *Perfect: Anorexia and me*. Random House.
- Hartmann, A. S., Thomas, J. J., Wilson, A. C., & Wilhelm, S. (2013). Insight impairment in body image disorders: Delusionality and overvalued ideas in anorexia nervosa versus body dysmorphic disorder. *Psychiatry Research*, 210(3), 1129–1135.
- Helm, B. W. (2001). *Emotional reason: Deliberation, motivation, and the nature of value*. Cambridge University Press.
- 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.
- Kaizer. (2014). *Body representation disturbances in anorexia nervosa* (dissertation). Utrecht: Utrecht University.

- Keas, M. N. (2018). Systematizing the theoretical virtues. *Synthese*, 195(6), 2761–2793.
- Keizer, A., Smeets, M. A. M., 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., & Oulis, P. (2012). Delusional body image beliefs in eating disorders. *Psychiatry Research*, 200(2-3), 482–488.
- Lackey, J. (1999). Testimonial knowledge and transmission. *The Philosophical Quarterly*, 49, 471–490.
- Lyons, W. (1980). *Emotion*. Cambridge University Press.
- Majeed, R. (2022). Modularity and the politics of emotion categorisation. In J. Deonna, C. Tappolet, & F. Teroni (Eds.), *A tribute to Ronald de Sousa*. Retrieved from: <https://www.unige.ch/cisa/related-sites/ronald-de-sousa/>
- Mandelli, L., Stefano, D., Albert, U., De Ronchi, D., & Atti, A.-R. (2020). Rates of comorbid obsessive-compulsive disorder in eating disorders: A metaanalysis of the literature. *Journal of Affective Disorders*, 277, 927–939.
- 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.
- Mele, A. (1999). Twisted self-deception. *Philosophical Psychology*, 12(2), 117–137.
- Mele, A. (2001). *Self-deception unmasked*. Princeton University Press.
- 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. *BioMed Central Research Notes*, 7(1), 1–11.
- Najmi, S., Kuckertz, J. M., & Amir, N. (2012). Attentional impairment in anxiety: Inefficiency in expanding the scope of attention. *Depression and Anxiety*, 29(3), 243–249.
- Nussbaum, M. (2001). *Upheavals of thought*. Cambridge University Press.
- Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*, 108(3), 483–522.
- Phillips, K. A., Kim, J. M., & Hudson, J. I. (1995). Body image disturbance in body dysmorphic disorder and eating disorders: Obsessions or delusions? *Psychiatric Clinics of North America*, 18(2), 317–334.
- Quirk, G. J., & Gehlert, D. R. (2003). Inhibition of the amygdala: Key to pathological states. *Annals of the New York Academy of Sciences*, 985, 263–272.
- Rahman, T., Meloy, J. R., & Bauer, R. (2019). Extreme overvalued belief and the legacy of Carl Wernicke. *Journal of the American Academy of Psychiatry and Law*, 47(2), 180–187.
- Roberts, R. (2003). *Emotions: An essay in aid of moral psychology*. Cambridge University Press.
- Schmidt, U., & Treasure, J. (2006). Anorexia nervosa: Valued and visible. A cognitive-interpersonal maintenance model and its implications for research and practice. *British Journal of Clinical Psychology*, 45(3), 343–366.
- Schroeder, M. (2012). Stakes, withholding, and pragmatic encroachment on knowledge. *Philosophical Studies*, 160(2), 265–285.
- Smeets, M. A., Ingleby, J. D., Hoek, H. W., & Panhuysen, G. E. (1999). Body size perception in anorexia nervosa: A signal detection approach. *Journal of Psychosomatic Research*, 46(5), 465–477.
- Smink, F. R., van Hoeken, D., & Hoek, H. W. (2013). Epidemiology, course, and outcome of eating disorders. *Current Opinion in Psychiatry*, 26, 543–548.
- Steglich-Petersen, A., & Varga, S. (2022). Obsessive-compulsive disorder and recalcitrant emotion: Relocating the seat of irrationality. *Philosophical Psychology*, 1–26. <https://doi.org/10.1080/09515089.2022.2066514>
- Steglich-Petersen, A., & Varga, S. (n.d.). Body checking in anorexia nervosa: From inquiry to habit. Unpublished manuscript.
- Strober, M. (2004). Pathologic fear conditioning and anorexia nervosa: On the search for novel paradigms. *The International*

- Journal of Eating Disorders*, 35(4), 504–508.
- Taylor, E. (2022). Discordant knowing: A puzzle about insight in obsessive-compulsive disorder. *Mind & Language*, 37(1), 73–93.
- 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.
- Vitousek, K., Watson, S., & Wilson, G. T. (1998). Enhancing motivation for change in treatment-resistant eating disorders. *Clinical Psychology Review*, 18, 391–420.
- Vitousek, K. M. (1996). The current status of cognitive-behavioral models of anorexia nervosa and bulimia nervosa. In P. M. Salkovskis (Ed.), *Frontiers of cognitive therapy* (pp. 383–418). The Guilford Press.
- Walsh, B. T. (2013). The enigmatic persistence of anorexia nervosa. *American Journal of Psychiatry*, 170, 477–484.
- Williamson, T. (2000). *Knowledge and its limits*. Oxford University Press.
- World Health Organization. (2004). *ICD-10: International statistical classification of diseases and related health problems (11th ed.)*. WHO <https://apps.who.int/iris/handle/10665/42980>
- World Health Organization. (2019). *International statistical classification of diseases and related health problems (11th ed.)*. WHO Retrieved from: <https://icd.who.int/>
- Yip, B. (2022). Emotions as modulators of desire. *Philosophical Studies*, 179(3), 855–878.