The Thin End of the Wedge?: The Moral Puzzle of Anorexia Nervosa

Aleksy Tarasenko-Struc
Seton Hall University, South Orange, NJ

Forthcoming in The Journal of Medicine and Philosophy. *Please cite published version.*

Abstract: The practice of force-feeding dangerously malnourished patients with anorexia nervosa (AN) raises a puzzle for clinical ethics. Force-feeding AN patients may seem justified to save their lives and to help them recover from a debilitating pathological condition. Yet clinical ethics seems committed to a robust anti-paternalism principle, on which it is normally wrong to force treatment on decisionally capacitated patients for their own good. Thus, routinely force-feeding AN patients seems to constitutes an unjustifiable exception to a well-established principle of clinical decision-making. I examine three attempts to solve the puzzle and argue that, individually or taken together, they cannot justify force-feeding those AN patients for whom this intervention would be potentially effective at enabling recovery. I conclude that no such justification is currently available. A solution to the moral puzzle of AN may come from a reevaluation of the anti-paternalism principle, a deeper clinical understanding of the psychology of AN, or even a reconceptualization of decisional capacity.

Key words: anorexia nervosa, forced treatment, autonomy, paternalism

I. A PUZZLE FOR CLINICAL ETHICS: THE FORCE-FEEDING OF A.N. PATIENTS

Anorexia nervosa (AN) is mysterious. It is characterized by an intense fear of gaining weight and a perception of oneself as overweight—regardless of how much one, in fact, weighs—along with a strong drive to severely restrict one's diet that leads, eventually, to extremely low body weight. The goal of weight loss is pursued with ruthless zeal, often at the expense of the patient's other projects, relationships, health, and even life. The etiology of AN is currently not well understood.

AN is classified as a psychiatric illness in the *DSM-5* (American Psychiatric Association [APA] 2013: 338–45). Serious medical complications from prolonged AN include those associated with starvation: reduced brain volume, weakening of the heart (e.g., bradycardia, hypotension, mitral valve prolapse leading to potentially fatal arrhythmias), electrolyte

imbalance (possibly eventuating in cramps, epileptic attacks, respiratory paralysis, or cardiac arrest), and multiple organ failure. While a significant number of anorexic patients die of such complications, many lose their lives by suicide. The death rate for AN is between 0.71 and 12.8, making it the most dangerous psychiatric illness. Yet a majority of anorexic patients achieve remission within five years (APA 2013: 342). For AN patients admitted to a hospital for severe malnourishment, force-feeding—e.g., via nasogastric tube or intravenously, perhaps under physical restraint—is a common means of enabling further treatment with a view to recovery.[1]

The policy of force-feeding dangerously malnourished adult AN patients raises a moral puzzle, though. It is a standard view in medicine that force-feeding patients in this condition is generally justified if, roughly, it is necessary for saving their lives and would likely enable their recovery. True, there may be cases in which a patient's AN is so severe and treatment-refractory that the chances of medical benefit are too slim to outweigh the various harms of forced treatment.[2] But those cases appear to be exceedingly rare, and some theorists (e.g., Geppert 2015; Giordano 2019: 325–26; Radden 2021: 145) even worry that AN's unusual symptomology and trajectory problematize clinical judgments of medical futility in this context. Thus, on the standard view, if an AN patient is dangerously malnourished, and force-feeding is necessary for saving her life and would likely enable her to recover from her condition, then force-feeding is presumptively justified, at least partly because it is in her own best interests.[3]

At first blush, the standard view looks to be at odds with the deeply held conviction that a person has a strong right to refuse life-sustaining treatment if she is decisionally capacitated with

[1] By 'force-feeding' I mean administering nutrition to a patient over her objection.
[2] See, for example, (Draper 2000); (Giordano 2005: 235–53), (2019: 329–30); and (Yager 2015).
[3] Advocates of the standard view include (Bruch 1980), (Dresser 1984), (Draper 2000), (Giordano 2005), (Giordano 2010), (Charland et al 2013), (Hope et al 2013), (Westmoreland et al. 2017), (Giordano 2019), (Hawkins 2021), and (Szmukler 2021).

respect to that refusal. This right isn't normally undermined by the fact that she will thereby bring irreversible, needless harm upon herself. Call this the *anti-paternalism principle*. The anti-paternalism principle reflects the priority enjoyed by the principle of respect for autonomy, which is typically thought to override or curtail the principle of beneficence. Consider, for instance, the old chestnut in bioethics that an adult Jehovah's Witness (JW) patient with (relevant)[4] decisional capacity has a right to refuse blood products even if he will die as a result. Yet if some AN patients retain decisional capacity (as they seem to, by standard criteria), it is wrong to force-feed them even to save their lives. In that case, force-feeding these AN patients seems to be just as wrong as it would be to forcibly administer blood products to a JW patient.[5] So, the worry is that the practice of force-feeding dangerously malnourished AN patients constitutes an unjustifiable exception to a well-established principle of clinical decision-making.

On the other hand, it is hard to resist the conclusion that clinicians are not only permitted but sometimes even morally required to force treatment on AN patients, as per the standard view. Unlike the treatment preferences of a JW patient, we may worry that those typical of AN are not expressive of an alternative system of personal values, religious worldview, or culturally specific way of life. Rather, the AN patient's refusal of care is due to a pathological condition known for impeding the autonomous capacities of those in its grip. It is, therefore, far from straightforward to identify her refusal of treatment as *her* decision. In this connection, it is to the point that many former AN patients express retrospective gratitude for the coercive life-saving measures to which they were subjected, as disturbing as their ordeal was to them (Tan, Hope, and Stewart 2003).

[4] That is, with respect to the decision to refuse life-sustaining nutrition. Going forward, I will drop the '(relevant)' qualifier from claims about AN patients' decisional capacity/incapacity, so the qualification will be implicit.

[5] Thus, I do not discuss the different moral considerations in play depending on whether the AN patient in question is force-fed on a trial basis or a long-term basis. That is because my primary interest in this paper is with the prior ethical question of whether force-feeding, even if effective in enabling recovery, is—in principle—permissible.

My goal is to determine to what extent this familiar puzzle—the *moral puzzle of AN*—admits of a solution. I examine three proposed solutions so as to reveal their limitations. The first is that force-feeding AN patients is justified because they lack (relevant) decisional capacity, so the anti-paternalism principle is not violated. The second is that force-feeding constitutes justified soft or weak paternalism—thus, that it is in harmony with the principle. And the third is that force-feeding even decisionally capacitated AN patients constitutes justified hard or strong paternalism—hence, that it is an exception to, or curtailment of, the anti-paternalism principle. I argue that these proposals, individually or taken together, cannot justify force-feeding those AN patients for whom this intervention would be potentially effective at enabling recovery. Hence, my project's contribution is largely negative. By engaging with research on AN patients' decisional (in)capacity alongside theoretical work on the scope of justified paternalism in the context of AN, we see that the moral puzzle of AN turns out to be even tougher to solve than it may seem.[6] A solution to the puzzle may only come from some quite radical development in our thinking, such as a reevaluation of the anti-paternalism principle, a deeper clinical understanding of the psychology of AN, or even a reconceptualization of decisional capacity.

II. THE DECISIONAL INCAPACITY ACCOUNT

The first and most intuitive proposal for solving the puzzle is to maintain that AN patients lack decisional capacity to refuse nutrition. The anti-paternalism principle would not rule out force-

[6] In presenting my topic as the search for a solution to a puzzle, I run the risk of trivializing a condition that has tragically claimed the lives of many young people. Some readers may worry that I am thereby treating AN as primarily posing a problem for thought as opposed to a problem for practice—and for real human individuals. Other readers may take me to be indulging in an extended (and callous) bout of devil's advocate. I take these concerns seriously. Still, there is a tension in our moral thinking about the care of AN patients, which places our concern for patients' welfare in contention with our commitment to respecting their autonomy, while challenging both our anti-paternalistic convictions and our conception of decisional capacity. Until we can resolve this tension, a satisfactory justification for force-feeding AN patients—and forcing treatment in other contexts—may remain unavailable.

feeding them to save their lives, then. JW patients, by contrast, generally have decisional capacity to refuse blood products, making it wrongly paternalistic to force treatment on them.[7]

To see the plausibility of this view—the *Decisional Incapacity Account*—consider the case of K:

> K is a 20-year-old woman previously diagnosed with obsessive-compulsive disorder and depression who has been fixated on her food intake for two years. When she left home for college, she began to exercise daily and to complain of intense nausea during meals. Eventually she felt a visceral urge to vomit whenever she ate. Diagnostics and consultations with numerous gastroenterologists revealed no gastric condition, and her symptoms were ultimately revealed to be psychogenic—a result of AN. By the time of her most recent ED admission, she weighed 85 pounds and barely ate at all. She refused to believe her AN diagnosis and insisted that her digestive system was compromised. K thus declined food and a feeding tube, believing that they would exacerbate the problem.

K lacks decisional capacity by the criteria of the dominant approach to capacity assessment: the *four abilities model* (Grisso and Appelbaum 1998). On this model, a patient is decisionally capacitated (with respect to some treatment decision) if and only if she demonstrates the following abilities: (1) expressing a choice, (2) understanding, (3) appreciation, and (4) reasoning.[8] K is clearly capable of communicating her decision(s) to the care team. And we can

[7] For a dissenting voice on this topic, see (Martin 2007). As I will assume for the sake of argument that JW patients aren't decisionally incapacitated in virtue of their religious beliefs, I will set Martin's nuanced discussion aside.

[8] While it is relatively uncontroversial what the ability to express a choice is, there is far less agreement concerning the character of the other abilities. But I'll stipulate that understanding consists, roughly, in comprehension of the relevant medical facts; appreciation involves, at least, believing these facts to be true of oneself; and reasoning is tantamount to drawing (deductive and inductive) inferences from the facts believed to further claims (e.g., about how different options may affect what one values), in a way that's governed by elementary norms of consistency.

imagine that she can understand her diagnosis, her treatment options, and their consequences, as well as reason about the implications of these options for what she cares about and values. As she is deluded about the nature of her condition, however, she fails to meet the appreciation criterion, which requires not only comprehending a diagnosis but also believing it of oneself.

More generally, not only is it clear that malnutrition may adversely affect a person's decision-making abilities (by, e.g., impairing her concentration and making her more prone to unclear thinking), there is also some evidence that AN patients in particular tend to exhibit a reduced ability to appreciate relevant information by applying it to themselves. In two separate qualitative studies, Jacinta Tan and colleagues (2003) (2006) found that some AN patients indicated problems with appreciation, although in each case the authors admitted that the majority of patients interviewed met the cognitive benchmarks of decisional capacity (as measured by the MacArthur Competence Assessment Tool for Treatment [MacCAT-T]). In a larger and more recent study, moreover, Isis Elzakkers and colleagues (2016) discovered that 34.3% of participants in the study (= 24/70) had impaired decision-making ability, with lower mean BMI predictive of poorer decision-making and deficits in appreciation especially salient.[9]

It is well known, though, that not all AN patients are in the grip of factual delusions about their condition, as the studies just cited themselves indicate. Unlike K, some AN patients believe that they have AN and even *want* to continue to have it. They regard a thin figure as supremely and intrinsically valuable and derive meaning from the pursuit of this ideal, perhaps also seeing it as the pinnacle of self-control and autonomy. By the standard criteria for decisional capacity, these AN patients qualify as decisionally capacitated. Nor can we convict these AN patients of

[9] Cf. (Sherri Turrell et al 2010), which interviewed 35 adolescents with AN and found deficits in their ability to reason (compared with the control group). Because this study investigates the decision-making abilities of adolescent AN patients, however, its application to the population of adult AN patients is unclear.

decisional incapacity *just* on the grounds that they have a psychiatric disorder. Clinical diagnosis is generally taken to be inadmissible as a basis for a determination of incapacity, and people with mental illness may still retain decisional capacity.

A more effective and more common strategy, though, is to turn this argument against the Decisional Incapacity Account on its head. If patients who report valuing their AN turn out to have decisional capacity by the standard criteria, well, so much the worse for those criteria. These patients are clearly incapacitated, so we should revise or expand the existing criteria for decisional capacity. At least three alternative approaches to decisional capacity have been proposed in the literature.

The Authenticity Approach

The first approach is, in effect, to add an authenticity criterion to the Grisso-Appelbaum model. Jacinta Tan and colleagues (2006) argue that some AN patients lack decisional capacity because they hold "pathological values": roughly, values that inform their treatment decision, yet that are not properly their own but rather are caused by a psychological disorder. Likewise, Scott Kim (2016) contends that one component of decisional capacity is "the ability to value"—that is, the ability to make treatment decisions in light of one's authentic values, rather than fringe motivational elements that contradict those values and that are attributable to a psychological disorder.[10] George Szmukler (2018: 101–21; 2021) has defended a similar criterion as well.

This view—the *authenticity approach*—faces familiar challenges.[11] For one, not all values that originate in a psychological disorder are clearly decisionally incapacitating. For example, obsessive-compulsive disorder may prompt a person to see the value of cleanliness and

[10] Notably, Kim claims that assessing capacity in AN patients is bound to be challenging; see (Kim 2016: 200).
[11] For criticism, see (Whiting 2009) and (Charland and Hawkins 2020).

rigor, yet valuing these things need not be a mark of decisional incapacity. But a deeper issue is that a person's authentic identity can be *founded* on her disorder, and the resulting values are not obviously incapacitating, either. This point can be illustrated by another clinical vignette:

> E is a 34-year-old woman diagnosed with AN in early adulthood. Ever since adolescence, she has idolized thin women like her mother. She saw her mother as not only the epitome of beauty and grace but also as possessed of an enviable self-mastery. Achieving and maintaining a thin body figure soon became the most important thing in the world to E. She gradually gave up activities that she valued—tennis, travel—and her relationships with friends and family eventually frayed, particularly after her parents' death. E now devotes herself single-mindedly to the pursuit of thinness. Told by her primary care physician that she must pursue psychotherapy to resolve her AN or she will die within a year, she responds, "I'm willing to take that risk: I can't imagine my life without it."

Not only does E endorse her AN, it is, ultimately, woven into the fabric of her personality.[12] Moreover, there appears to be no inconsistency between her deeper, authentic values and her AN-generated values. For even before she was motivated to myopically value the pursuit of a thin figure, thinness appealed to her prior to and, to some extent, even independently of her AN. Perhaps her AN even crystallized values that would have organized her pursuits anyway. Thus, E's valuing thinness should count as an authentic rather than a pathological value—and *a fortiori*

[12] Indeed, AN is classified an egosyntonic disorder, meaning that those afflicted with it tend to value it and its effects; for discussion, see, e.g., (Gregertsen, Mandy, and Serpell 2017). As far as I am aware, it is not currently known how many AN patients embrace their condition as fully as E. That said, there is ample clinical evidence that AN patients tend to see their condition in a positive light, and sometimes to experience it as interwoven with their identity, all of which presents an obstacle to recovery. See also (Nordbø et al 2006) and (Voswinkel et al 2021).

she should qualify as retaining the ability to value—which would entail that she is decisionally capacitated. The anti-paternalism principle then prohibits force-feeding E even to save her life.

The Passion Approach

A second strategy—the *passion approach*—is due to Louis Charland and colleagues (2013). They claim that a person may be decisionally incapacitated due to the influence of a *passion* on her orientation to the world—roughly, an abnormal affective syndrome that co-opts their intellectual and volitional capacities in the service of an obsessional fixed idea. And they canvass some evidence that AN is just such a passion, and a particularly deadly one at that. If AN is a passion, in this sense, then it may undermine autonomy and thus decisional capacity. They write:

> It is not clear that a person driven by a passion is acting autonomously. The behavior and decisions (including to refuse treatment) are not the result of careful and rational reflection on life-goals (that is the paradigm of autonomous behavior). The decisions (e.g., to refuse treatment) may be rationalized in ways that meet the standard criteria for capacity, but be made as a result of the affective motivating power of the passion. (Charland et al 2013: 363)

We may worry, however, that this approach overgeneralizes, counting too many patients as decisionally incapacitated. For by the authors' (2013: 355) own admission, quite a few affective conditions qualify as passions, including jealousy, ambition, fear, and romantic love. These states of mind can all become sites of obsession that increasingly dominate a person's psyche and impede rational reflection. Yet not all passions are decisionally incapacitating, and it is

unclear how to distinguish those that are from those that are not. Until we can, we should be cautious about concluding, as Charland and colleagues (2013: 363) do, that "[t]o help the person to be free from the grip of the passion and attain a greater degree of autonomy over her life, it may be necessary to" force treatment on her. Imagine a patient so single-mindedly dedicated to football that he considers foot/leg amputations worse than death. As bizarre as this value is, and although it prevents him from full rationality, his holding it is not necessarily incapacitating. If he refuses a medically necessary foot amputation on these grounds, then, his clinicians are not permitted to override his will, not even to save his life—or to restore him to full autonomy.

The Welfarist Approach

A third and final strategy that I will consider—the *welfarist approach*—has been recently defended by Jennifer Hawkins (2016, forthcoming). Hawkins proposes that the standard four abilities are necessary for decisional capacity but insufficient. In particular, a person with these abilities is decisionally incapacitated if two further conditions are met. First, there is "good evidence… that the patient is making a serious, prudential mistake here and now" *from his own perspective*—that is, he is needlessly and perhaps irrevocably sacrificing what he cares about (or what he will eventually come to care about). Second, the patient is "known to have a condition that, in turn, is known to make those who have it more likely than ordinary to make prudential mistakes" (Hawkins forthcoming: 6–7). Because some AN patients meet these two conditions, force-feeding them turns out to be permissible when it is necessary for saving their lives.

Hawkins's novel view requires a more extensive treatment than I can provide here. But, as a first pass, we can note that her additional conditions for decisional incapacity seem too wide, with concerning implications for the justifiability of paternalistic interventions in patients' lives.

Take, for example, a patient with nicotine dependence who resolutely and repeatedly refuses to give up smoking when his physician suggests prescribing a nicotine nasal spray to help him break his unhealthy habit. This patient is making a serious prudential mistake, due to a clinical condition—nicotine dependance—which has been known to make its bearers more likely than ordinary patients to make such mistakes. Yet he does not seem to lack decisional capacity to refuse the nasal spray, nor does he seem to lose his right against paternalistic interference with this medical decision.

In fact, Hawkins's additional conditions for decisional incapacity do not even require that the patient be making a serious prudential mistake *in virtue of* a clinical condition known for rendering people more disposed than usual to making *that very kind* of prudential mistake. So, her view seems committed to counting even more patients as decisionally incapacitated.

Suppose, for instance, that our nicotine-dependent patient has been diagnosed with stage I lung cancer and refuses chemotherapy because his mother underwent it after her own cancer diagnosis and it was unsuccessful. As his prognosis is good, this patient is making a serious prudential mistake in refusing. He also has a clinical condition that disposes people who have it to making prudential mistakes more than usual. So, he meets Hawkins's two conditions. It seems to follow that the nicotine-dependent patient lacks decisional capacity *to refuse chemotherapy*, even though his refusal is not causally related in any significant way to his nicotine dependence.

In summary, three approaches to revising the standard criteria for decisional capacity—the authenticity approach, the passion approach, and the welfarist approach—all have serious issues. In drawing this conclusion, however, I want to emphasize that the project of

reformulating criteria for decisional capacity may still be warranted. But, in light of the issues that I have identified, we must look elsewhere for a solution to the moral puzzle of AN.[13]

III. THE SOFT PATERNALISM ACCOUNT

According to the Soft Paternalism Account, force-feeding AN patients constitutes *soft paternalism*: roughly, interfering with someone's freedom for her own good either (a) when her prospective action is substantially nonvoluntary—i.e., caused by some controlling influence or based on lack of (relevant) empirical knowledge—or (b) to verify that she is, in fact, acting voluntarily. Force-feeding might instead qualify as *weak paternalism*, on this view: restricting someone's freedom for her own good when her prospective choice of means would defeat her own ends. This view exploits the fact that acts of soft/weak paternalism are far easier to justify when they are necessary for saving a person's life. In preventing someone from (potentially) nonvoluntarily bringing severe harm to herself, we are, in effect, acting in accordance with her own will, her deeper values and preferences. Thus, the Soft Paternalism Account rests on the idea that AN patients refuse life-sustaining nutrition against their own will, so coercively treating them does not run afoul of the anti-paternalism principle: it's in harmony with the principle.

Does force-feeding an AN patient amount to a relatively benign form of paternalism? There is one kind of situation in which it surely would constitute soft paternalism. Suppose that the patient's malnourishment is emergent, she continues to refuse nutrition, yet no strong evidence of decisional capacity on her part (or perhaps of her assigning overriding value to the

[13] I have not discussed value-laden conceptions of decisional capacity, on which a person is decisionally capacitated only if she holds (or assigns particular weight to) certain substantive values. See, for example, (Craigie 2011) and (Holroyd 2012: 156–60) for defenses of such views in connection with AN. Briefly, although I have no knockdown argument against these views, I believe that criteria for decisional capacity should be informed by a commitment to liberal neutrality, which is incompatible with a requirement to give particular weight to certain substantive values.

maintenance of a thin figure) is available to her care team. In that case, force-feeding might be justified because it allows her clinicians to determine whether she appreciates the risks of her persistent refusal and is willing to accept those risks—hence, it is soft paternalism (Bunch 2023).

Otherwise, though, force-feeding an AN patient counts as justifiable soft paternalism only if one of two conditions is met. First, her refusal of life-sustaining nutrition is based on false or incomplete information that, if corrected or supplemented, would motivate her to accept care. Second, her refusal is the upshot of a controlling influence that undercuts voluntary choice. Furthermore, even if neither condition is met, force-feeding an AN patient might instead constitute weak paternalism if her treatment refusal is self-defeating, in that the probable results of that refusal would prevent her from obtaining an outcome that she values even more highly.

Do AN patients meet any of these conditions? I consider three answers to this question.

False Empirical Belief

First, it might be thought that AN patients refuse treatment on the basis of false empirical beliefs about their body. They persist in believing that they are too fat when they are, in fact, utterly emaciated. Furthermore, they seem to believe that they are too fat regardless of how thin they become. For some AN patients, there may be no such thing as having a body that is *too* thin.

One tempting reply to this view is to insist that AN patients shouldn't be interpreted as expressing a straightforward empirical belief about their bodies. Rather, they are expressing a value-laden belief—one that is informed by the overridingly high value that they place on a thin figure. Their beliefs are thus unlike that of a person who picks up a glass of gasoline, intending to drink it, because he falsely believes that it is full of gin. The matter is complicated, however.

There is, in fact, some empirical support for the hypothesis that AN patients harbor false beliefs about their body size/shape. Stephen Gadsby (2017) (2020) (2022) has argued that AN patients regularly form false beliefs about their bodies as a result of misperceiving their bodily boundaries, and that these beliefs play a substantial causal role in sustaining the psychiatric condition. Gadsby's account is complex, and I cannot adequately summarize it here. But even if it holds of some individuals who suffer from AN, there is reason to question its generality.

As Somogy Varga and Asbjørn Steglich-Petersen (2023: 6–7) point out, some AN patients seem to disavow the belief that they are fat, suggesting that these patients might be expressing fear of being/becoming fat instead.[14] Because the concept of fatness (as AN patients use it) is saturated with fatphobic attitudes, I believe that a partial version of the tempting reply is in order: in claiming that they are fat, some AN patients may be expressing value-laden beliefs, not neutral empirical beliefs. If so, the patients in question need not hold false empirical beliefs about their bodies at all, and, absent an alternative explanation, coercively treating these AN patients will not constitute soft or weak paternalism in the first place.

Compulsion

Second, AN patients might be in the grip of a compulsion: a motivational force that is impossible for a person to resist and that effectively removes her freedom to choose. Amanda Evans (2021:484–86), for instance, has recently canvassed empirical evidence that AN originates in a process of pathological habit-formation or through increasing sensitization to disorder-specific cues that mirrors the genesis of addiction.[15] And in a qualitative study by Lauren Godier and

[14] AN patients may also suspend judgment about whether they are, in fact, fat.

[15] It should be noted that Evans's (2021) discussion is oriented to empirical explanation, not to justification. So, the view outlined in the rest of this subsection should not be attributed to Evans, who does not discuss paternalism at all.

Jessica Park (2015), over a third of AN patients interviewed reported that they experienced compulsivity as central to AN, with over two thirds avowing that they felt that they had diminished control over their food restriction.[16] It may therefore be thought that AN patients' treatment refusal is the result of a compulsion that undermines voluntary choice.

There is much to be said for this account. It helps identify one respect in which some AN patients are alienated from themselves, which might reduce their autonomy or even undermine the voluntariness of their refusal of life-sustaining nutrition. Consider the following case:

> J is a 23-year-old man who has struggled with AN since he began swimming competitively three years ago. He regularly experiences his condition as an alien force that he can resist only with great resolve, a persistent threat to the people and things he cares about—including life itself. Sometimes, though, he copes with his condition by convincing himself that a slim figure is his heart's desire. He can maintain the ruse only for so long, luckily. Although he realizes that he needs treatment, he finds it exceedingly difficult to adhere to a care plan, fearing that he is not strong enough to continue.

Because J is unwillingly afflicted with AN, force-feeding him would amount to soft or weak paternalism, and it would be justified if it brought his behavior in line with his deeper values.

Do all or even many AN patients refuse life-sustaining nutrition under the influence of a compulsion, though? We may think not. For one, a nonnegligible proportion of AN patients do seek treatment voluntarily, so it cannot be that all such patients are in the grip of a motivational

[16] Notably, a genetic link between AN and obsessive-compulsive disorder has also been discovered. See (Yilmaz et al 2020).

force that is literally irresistible.[17] Perhaps, then, the claim is only that the motivational force associated with AN is *difficult*, but not impossible, to resist. In that case, however, it seems implausible that AN patients' food restriction behavior fails to be voluntary after all.

Nor does it seem right to think of force-feeding certain AN patients as soft or weak paternalism, both of which involve—in effect—interfering with an individual's liberty so as to prevent her from acting against her deeper values. Recall E, whose sense of self is founded on a robust identification with her desire for a thin figure. Unlike J, E doesn't bear AN unwillingly at all. On the contrary: she suffers AN *willingly*, even if her behavior is strongly influenced by her clinical condition. She wants to have AN and even reflectively endorses it. So, disturbingly, starvation would be in line with her deeper values, while force-feeding her would contravene them. And because AN is an egosyntonic condition, E's relation to her condition is hardly exceptional.

More generally, when people act on desires that are so strong as to be difficult to resist (such as those associated with substance use disorder), their autonomy may thereby be compromised to some extent. But their actions may still be voluntary, particularly if their capacity to value is intact and the action is in line with their deeply held, reflectively endorsed values. The burden of proof is on those who insist that E's refusal of nutrition is not voluntary.

Post Hoc Rationalization

Third, AN patients might be motivated to refuse life-sustaining nutrition by a kind of powerful, affectively induced illusion—albeit not, perhaps, by an ordinary false empirical belief.

[17] See (Coffino et al 2019). Here my point parallels Hanna Pickard's claim that people with substance use disorder do not generally behave compulsively but use substances because they see doing so as a valuable part of their identity. See, e.g., (Pickard 2021).

In a well-known study of AN, Tony Hope and colleagues (2013) hypothesized, based on extensive interviews with AN patients, that such patients harbor two conflicting mental states. They believe that they are extremely, dangerously thin. Yet they also experience an intense fear or anxiety that tells them that they are "fat", and this emotion leads them to refuse life-sustaining nutrition. The conflict between these two mental states also leads to profound cognitive dissonance and ambivalence, however. When asked by clinicians to give reasons for their refusal, AN patients therefore feel compelled to fabricate a post hoc rationalization for their alleged choice, claiming that they would prefer to die rather than to put on weight.

This account is the most plausible of those considered so far. Not only does it appear to capture the fraught experience that many AN patients have of their condition, it also suggests that they may be acting against their better judgment in refusing treatment—and under the spell of a delusion, no less. So, even if they are not (strictly speaking) decisionally incapacitated, force-feeding them to save their lives might still be justified as soft or weak paternalism.[18] Furthermore, if AN patients are typically as ambivalent as Hope and colleagues (2013) theorize, in that they exhibit radically inconsistent or shifting treatment preferences, force-feeding may still be justifiable on similar grounds even if their refusal of care is voluntary and informed.

I believe, however, that this account presents an inconclusive case for the claim that AN patients in general refuse nutrition involuntarily on the basis of an affective delusion. For one, we might doubt the idea that because these patients intensely fear gaining weight, they don't really prefer to die rather than to put on weight. Suppose that I have an intense fear of needles. Then I may so strongly prefer not to be stuck with one that I would be willing to risk great bodily

[18] Interestingly, because patients in this situation appear to hold no (relevant) false empirical beliefs, they should count as decisionally capacitated, at least by standard criteria. See (Hawkins and Charland 2020) for discussion. If so, however, they may still be candidates for soft or weak paternalistic interventions.

harm to avoid this prospect. Why wouldn't the analogous claim hold of AN patients? It is true, of course, that a person who is *unwillingly* moved by her intense fear of needles may not be acting voluntarily. But someone who sees her fear of needles as fitting, who thinks that needles are so horrible that it would be very bad to be stuck with one, and who is *willingly* moved by that fear can rightly be attributed a preference to avoid needles at all costs, it seems.[19]

We may also doubt that the account articulated by Hope and colleagues (2013) applies to AN patients generally, no matter how well it fits the authors' interview subjects.[20] By their own admission (Hope et al 2013: 25), all of their interviewees were people who had already accepted treatment for AN, so it is unsurprising that they would be deeply conflicted about nutrition. But then we cannot responsibly infer, from that data, that other AN patients act against their better judgment on the basis of similar delusions. Indeed, the account may not fit patients who, like E, appear to more wholeheartedly embrace the value of maintaining a slim figure, in which case force-feeding cannot amount to soft paternalism, any more than my strong, reflectively endorsed preference not to encounter needles amounts to soft paternalism. And to the extent that AN patients accord overridingly high value to this goal, their refusal of nutrition is not self-defeating. If so, though, it cannot count as weak paternalism, either: E seems to be pursuing what she most values!

Thus, arguments purporting to show that force-feeding AN patients constitutes justified soft or weak paternalism appear to be severely restricted in scope, at best. While some such

[19] I thank an anonymous reviewer for this example.

[20] Compare Cragie's (2011: 331–33)'s view that there is reason to believe that AN patients generally lack decisional capacity because they (1) regret having acted on their AN-based values later, (2) (in some cases) report cognitive dissonance with respect to their food refusal, and (3) exhibit low emotional arousal during decision-making. Here, too, I'm not sure how much these data show about AN patients generally. Regret over having acted on one's values shows that people can be ambivalent or that they could come to adopt very different and perhaps incompatible values (Pugh 2020: 229–33), as does a tendency toward cognitive dissonance vis à vis conflicting values. I'm also skeptical that the fact that a person is relatively unemotional while making decisions provides significant evidence that these decisions fail to meet a baseline of autonomous decision-making.

patients no doubt act on the basis of inadequate information or out of a compulsion that undermines voluntary choice, others appear not to suffer any factual delusions and to pursue thinness with far less ambivalence, making their treatment refusals voluntary. Nor do patients who refuse nutrition because they endorse their AN seem to be pursuing their goals in a clearly self-defeating way. Force-feeding some AN patients—like E—will therefore fail to qualify as soft or weak paternalism at all; hence, doing so is in conflict with the anti-paternalism principle.

IV. THE HARD PATERNALISM ACCOUNT

I will examine one final proposal, defended by Jennifer Hawkins (2021) and more extensively by Simona Giordano (2005) (2010) (2019). Proponents of this view—the Hard Paternalism Account—endorse the anti-paternalism principle and affirm that some AN patients may indeed have decisional capacity. They also acknowledge that force-feeding such patients constitutes *hard paternalism* (roughly, restricting someone's freedom for her own good even when her prospective action would be substantially voluntary and undertaken in light of relevant empirical knowledge) or *strong paternalism* (roughly, restricting someone's freedom for her own good when the goal of her prospective action is mistaken, bad, or potentially harmful to her).[21] They insist, however, that force-feeding is justified in the case of some AN patients, because the character of AN weakens the otherwise powerful moral objection to the practice.

Why, then, might the standard obligation to respect a decisionally capacitated patient's refusal of (even life-sustaining) care be suspended or diminished when that patient has AN? Hawkins and Giordano offer strikingly different answers to this question. Their views raise broader issues concerning the relationship between the principle of respect for autonomy and the

[21] The distinction between hard and strong paternalism doesn't matter to my argument against this view. Accordingly, going forward, I will largely not distinguish these forms of paternalism.

principle of beneficence, and, ultimately, concerning the relationship between the value of free choice and the value of well-being. I evaluate each view in turn, starting with Hawkins's.

Autonomy as Constituent of Well-Being

Hawkins's (2021) utilitarian account has origins in the work of John Stuart Mill (2003: 95). On this view, the master value is that of well-being, and autonomy is valuable as a component of it, in that is typically—but not always—better for people to be able to act on the basis of their conception of the good, without interference from others. Call this picture *autonomy-as-constituent*.

It follows, per Hawkins, that paternalistic interventions are justified when there is strong evidence that a patient is making a serious prudential mistake, as when some AN patients refuse nutrition. For, in such cases, allowing a person to act on the basis of his own conception of the good is not, in fact, conducive to his well-being. Force-feeding and other coercive interventions may then be justified, on utilitarian grounds, as they have the effect of "salvaging a great deal of welfare (subjectively construed) at relatively minor costs to overall lifetime freedom" (Hawkins 2021: 157).

Hawkins's view invites two worries. First, the same rationale seems to favor paternalistic interventions in plenty of other cases that seem intuitively to be ruled out by the anti-paternalism principle. In the earlier case of the JW patient, it does not seem to accord with the spirit of this principle to forcibly administer a blood transfusion while giving him the following justification: "Yes, we're violating your (decisionally capacitated) wishes and contravening your moral beliefs. But, look, it's for your own good—you'll be dead without it—and, besides, we're still

leaving you with a sufficient amount of freedom. Once you leave the hospital, you can act on your religious convictions—strange as they may seem to us—as much as you please."

Second, I doubt that Hawkins's view captures what we find morally objectionable in paternalistically overriding a (decisionally capacitated) patient's wishes. When we refrain from treating someone over objection, even though his decision is self-destructive, we generally do not—and should not—calculate the costs and benefits to him of curbing his freedom to make that decision, so as to determine whether it really is in his best interests to be free to act in that way. Rather, we should, in Daniel Groll's (2012) terminology, treat the patient's will as "substantively decisive"—as giving us exclusionary reasons to act on his will even, sometimes, at the expense of his welfare. This attitude is, arguably, part and parcel of respecting the patient as a person.

Autonomy as Independent Constraint on the Promotion of Well-Being

Let us turn to Giordano's view. She, too, holds that force-feeding an AN patient with decisional capacity may be justified despite our commitment to the anti-paternalism principle. That is because, Giordano (2005: 237) claims, our obligation to honor the refusals of treatment made by decisionally capacitated adults is only a *prima facie* obligation, and "the principle of respect for patient autonomy, in the case of anorexia nervosa, is weaker than it is in other cases."

While the language of *prima facie* obligation may suggest that respect for autonomy is one duty among others, Giordano seems, rather, to be claiming that autonomy functions as an *independent constraint* on the promotion of the patient's well-being. Call this picture *autonomy-as-independent-constraint*.[22] Thus, the view that she expresses appears to be a deontological

[22] In (Giordano 2005: 56), for instance, she claims that "autonomy is a primary value" that "is more important than welfare," arguing that a person's autonomous decisions ground a constraint on other people's efforts to promote his welfare—a constraint whose force is independent of the value of his welfare. This constraint takes the form of a

view, not a utilitarian one. On this view, a decisionally capacitated patient's treatment decisions/wishes ground a strong presumptive obligation to provide her only with treatment that does not conflict with those decisions/wishes—*an obligation whose force is independent of the value of his well-being*. This obligation is not absolute. In some cases, it might be defeated by some other obligation, such as an obligation of beneficence. But those cases are exceptional, and a very high justificatory bar must be cleared for the constraint in question to be overridden.

Which considerations weaken the principle of respect for autonomy, in Giordano's view? She cites three. First, in general, the AN patient's "death is avoidable, and the condition is reversible" (Giordano 2005: 249). Second, the patient may yet live many happy years if she survives and recovers from her condition, and she may even be retrospectively grateful for the coercive treatment to which she was subjected (Giordano 2019: 327). Third, the patient's loved ones will suffer greatly if she dies, perhaps particularly if her death was avoidable (Giordano 2005: 249–50). Taken together, these facts license "exceptionalism" in the treatment of AN patients: "partial derogation from the usual principles of ethical decision making"—primarily, that is, from the normally overriding principle of respect for autonomy (2019: 316).

The problem, however, is that these three factors are present in other clinical contexts. The death of a JW patient after he refuses a life-sustaining blood transfusion is also avoidable—indeed, even more easily so than a dangerously malnourished AN patient's. We can imagine, too, that his loved ones will be devastated if he dies. And some JW patients' attitudes toward accepting blood products may be just as reversible as some AN patients' attitudes toward food. For people do leave the JW faith, for various reasons—say, because they take their lifestyle or sexual orientation to be forbidden by their religion. True, the reversal is not instantaneous. But

prima facie obligation to defer to the individual's autonomous decisions even, sometimes, at the expense of his welfare.

the same is true, typically, of AN: it may take years to escape the specter of self-starvation. So, if we believe that it is nevertheless wrong to forcibly administer a blood transfusion on this patient, then it seems that we are committed to the view that force-feeding AN patients is also wrong.[23]

Or consider a patient who has undertaken a hunger strike. He does not wish to die, he says. But he is willing to risk death—even to die—in order to send a political message about his society's indifference to climate change, which, he feels, poses a threat to humanity's survival. We can also imagine that Giordano's three factors are present in his case as well. His death is avoidable, his family will suffer greatly if he dies, and because self-starvation is immensely difficult, his intention to continue the hunger strike may not be ironclad and may be reversed. But, again, it is far from obvious that it is right to force-feed the conscientious hunger striker.

Rejecting Liberal Individualism?

Perhaps a far more radical move is needed to justify hard/strong paternalism. According to Jennifer Radden (2021), because AN patients often do have decisional capacity by ordinary criteria, force-feeding them will always violate what I have called the anti-paternalism principle. Yet this principle is grounded in a broader conception of the person that we have reason to reject, she claims. In rejecting it, we may find a basis for justifying the force-feeding of AN patients.

Injunctions against medical paternalism in this context are, Radden (2021: 147) claims, rooted in the *liberal, individualistic model of persons*. On this picture, persons are self-contained individuals in pursuit of goals that they autonomously determine according to their own conception of the good; accordingly, they should be as free from others' interference in that endeavor as possible. As against this picture, feminist theorists have of late recommended a

[23] Giordano (2010: 150–51) accepts the conclusion that Jehovah's Witnesses, too, should be forcibly treated if they gave up the relevant religious convictions after treatment and expressed retrospective gratitude for the treatment.

relational model of persons, on which persons are interdependent with others both causally and constitutively. Individuals depend on others causally in that they become persons through the actions of others (e.g., caregivers), and they depend on others constitutively in that their actions, values, and identities are essentially bound up with the actions, values, and identities of others.

Per Radden, the normative upshot of the relational model of persons is, in effect, that decision-making authority is shared between patients and those with whom they have significant relationships. And this idea, she claims, does (in some cases) make space for a different justification of force-feeding AN patients. Because "the patient is not a unitary individual but part of a larger social whole" (per the relational model), "the person's interests and perspectives are no longer exclusively privileged, and she is recognized as only one among the others comprising her identity." In that case, however, "she cannot remain the sole object of medical attention, and decisions over nutrition may not seem to be hers alone to make" (2021: 147–8). That is, force-feeding AN patients—in deference to her loved ones' wishes—may be justified.[24]

Even if we accept the relational model of persons, it's hard to see how Radden's normative conclusion follows. The fact that my identity is entwined with yours or that my actions (indeed, my very ability to act) substantially depend on your actions doesn't obviously imply that you are authorized to force treatment on me, even for my own benefit. More to the point, though, decisionally capacitated adults are not normally liable to have their treatment decisions about their own body vetoed by well-intentioned loved ones. The appeal of this idea is not confined to those who assent to liberal individualism. It seems to follow from the eminently

[24] It is unclear how strong Radden's view is meant to be. On a strong version, the patient's wishes and the wishes of her loved ones carry *equal* moral weight with respect to the decision of whether to pursue a particular medical intervention. On a moderate version, by contrast, the patient's wishes carry *more* weight than the wishes of her loved ones, but her loved ones are permitted to veto her treatment decisions if (in their view) those decisions are harmful to her. Fortunately, we need not determine which of these views Radden endorses. Both are dubious, as I will argue.

plausible conviction that every person's body is, morally speaking, their own private domain, as well as the idea that we have strong rights against others that they not invade our bodies.

Radden's view implies that a progressive JW patient who consents to receiving a life-sustaining blood transfusion may have his treatment decision overridden by his more conservative family members. Her view also implies that a young ex-Catholic woman with a life-threatening pregnancy who seeks an abortion may have her treatment decision overruled by her loved ones, who believe that the sin of abortion damages the soul worse than death. These implications are difficult to accept, to put it mildly, even if one rejects liberal individualism.

Perhaps there is some way that Radden's view can avoid these implications while still justifying the force-feeding of AN patients. But if so, force-feeding must be justified at least partly in virtue of the special character of AN, yet Radden's discussion makes it unclear how it is distinctive. In any case, there are strong reasons not to discard liberal individualism wholesale. Certain aspects of this view are well worth preserving, such as its emphasis on value neutrality in assessments of decisional capacity and its prioritization of liberty from others' interference. Indeed, rejection of the latter idea commits Radden's view to such unacceptable implications.

I conclude that the Hard Paternalism Account faces powerful challenges, and that we should be skeptical of arguments purporting to show that force-feeding AN patients constitutes justified hard/strong paternalism. These arguments have unacceptable moral implications for the proper care of other patients, such as hunger strikers and Jehovah's Witnesses. (Or at least their moral implications contravene the spirit of the anti-paternalism principle.) The problem is only compounded, moreover, if we look to alternatives to liberal individualism as a basis for a justification for coercively treating AN patients. We should therefore either reject the Hard Paternalism Account or reconsider our allegiance to the anti-paternalism principle.

V. CONCLUSION: NOT SOLVING—BUT SHRINKING—THE PUZZLE?

Time to take stock. I have been arguing that three proposed solutions to the moral puzzle of AN are vulnerable to powerful objections. These views tend to problematically overgeneralize beyond cases of AN, licensing paternalistic interventions in contexts where a patient plausibly falls under the protection of the anti-paternalism principle. As stated, then, none of the proposals discussed, individually or taken together, articulates a moral justification for force-feeding AN patients for whom this intervention would be potentially effective at enabling recovery, although some of these accounts do identify considerations that justify coercively treating some such patients. So, in routinely force-feeding such patients, we may be treating them differently than other patients whose treatment refusals are based on apparently irrational beliefs.

My conclusion is conditional in two noteworthy respects. First, it follows only if AN patients often have decisional capacity, which, by standard criteria, they sometimes appear to retain. Second, it follows only if we accept the anti-paternalism principle, and a strong formulation of it at that: that a decisionally capacitated patient's treatment decisions/wishes function as a robust, independent constraint on the promotion of her well-being, giving others exclusionary reasons not to overrule her will even to prevent her from harming herself.

For all I have said, we may have reason to reject the principle itself or the strong reading in particular. I have assumed that it is true, but I have provided no argument for it. Discarding or revising the principle may be preferable to concluding that force-feeding some AN patients is wrong. My point is only that the price is endorsing forcible blood transfusions for JW patients and force-feedings for conscientious hunger strikers, among many other acts of medical paternalism. We may also doubt that AN patients often have decisional capacity even if they

meet standard criteria. But then, as I argued, it is unclear how to supplement these criteria without radically expanding our conception of which patients are decisionally incapacitated.

We still seem to be ensnared in our original dilemma. If anything, the foregoing arguments may only have entangled us in it further. Having conveyed the contours (and complexity) of the moral puzzle of AN, though, I want now to strike a more optimistic note.

It bears emphasizing that some approaches examined above have yielded a presumptive justification for force-feeding certain AN patients. Some—like K—are decisionally incapacitated, so they are not fully protected from forced treatment (the Decisional Incapacity Account), while others—like J—are appropriate candidates for soft or weak paternalism (the Soft Paternalism Account). The hard case, again, is that of E: given that she endorses and identifies with her AN yet retains decisional capacity (per the four-abilities model), force-feeding her does not appear to be justified. Coercively treating E therefore constitutes an exceptional violation of the normally overriding principle of respect for autonomy.

In other contexts, and particularly outside of psychiatric settings, clinicians are willing to allow patients to refuse life-sustaining treatment, even—sometimes—on the basis of imprudent and substantively irrational beliefs. We do not generally force chemotherapy on people who unreasonably believe that homeopathic therapies would be more effective, nor do we normally force intubation on people who refuse it out of fear because their loved one died after having been intubated.[25] It's the exceptional nature of coercively treating AN patients despite their retaining decisional capacity that is morally troubling. In any case, even if the moral puzzle of AN has not been solved, exactly, it has at least been greatly shrunk.

[25] I am grateful to an anonymous reviewer for suggesting this example and for pushing me to clarify this aspect of my concern.

Merely diminishing the moral puzzle of AN will probably strike many as unsatisfactory. But it is important to remember that therapeutic interventions remain an option for AN patients like E. If she chooses to remain hospitalized for medical management of her starvation-related symptoms, there is hope that she may be persuaded to voluntarily accede to therapy. Such therapy should be oriented to reconstituting her identity on a different, less destructive basis.

Although it has largely been theoretical, my discussion suggests some directions for clinical practice. Negatively, it suggests that force-feeding ought not to enjoy the default status in the care of dangerously malnourished AN patients that it sometimes does today. More positively, however, my discussion indicates that greater effort should be made to justify its use in particular cases, that priority should be given to prompting the patients in question to voluntarily accept psychotherapy (and, more generally, to interventions that engage the patient's autonomous capacities rather than overriding them), and that care providers must try to understand the value that AN holds for these patients and what role it may play in the constitution of their identity.

In closing, I would like to gesture at a potentially fruitful avenues for justifying the coercive treatment of patients who endorse and identify with their AN. It might turn out that patients like E are, in fact, more like J in their psychology—deeply ambivalent about their food restriction, at best, and in the grip of self-deception when refusing care. The danger of this approach is that it's all too easy to force treatment on people on the grounds that they really want it, deep down, but convince themselves otherwise. Nevertheless, the weight of clinical evidence may favor this interpretation of self-endorsed AN. More insight into its etiology is needed.

Otherwise, we will need a new conception of decisional capacity that fits patients like E. That view should give due weight to at least some tenets of liberal individualism. Decisionally capacitated patients have a strong right to act against their own best interests without coercive

interference from concerned others, and patients should not be deemed incapacitated just in virtue of their substantive values, their prudential irrationality, or their psychiatric diagnosis. In the end, it is hard for most of us to shake the conviction that AN behavior—even that of E—falls short of some baseline of rational or autonomous well-functioning in decision-making that, when met, grounds a robust right to make decisions that are bad only for oneself. AN behavior of this kind challenges us to rethink our conception of autonomy and its relation to decisional capacity, as well as how irrational or heteronymous a person must be for that right to be overridden or lost.

ACKNOWLEDGMENTS

I presented versions of this essay at a meeting of the Psychiatry and Psychotherapy Affinity Group at the 2023 American Association for Bioethics and Humanities Conference, the Washington University Medical Ethics Student Society, and the International Philosophy of Medicine Roundtable. I'd like to thank audiences at these venues for stimulating discussion. I'm particularly grateful to John Han, Garson Leder, Brian Reese, and two anonymous referees for specific challenges and suggestions.

REFERENCES

American Psychiatric Association, ed. 2013. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, D.C.: American Psychiatric Association.

Bruch, H. 1980. *The Golden Cage: The Enigma of Anorexia Nervosa*. London: Open Books.

Bunch, L. 2023. Some treatments over objection, not others, are ethically justifiable when managing anorexia nervosa. *American Journal of Bioethics* 23(8): 99–101.

Charland, L., T. Hope, A. Stewart, and J. Tan. 2013. Anorexia nervosa as a passion. *Philosophy, Psychiatry, & Psychology* 20(4): 353–65.

Charland, L. and J. Hawkins. 2020. Decision-making capacity. *The Stanford Encyclopedia of Philosophy* [Online]. Available: https://plato.stanford.edu/entries/decision-capacity/

Coffino, J., T. Udo, and C. Grilo. 2019. Rates of help-seeking in US adults with lifetime DSM-5 eating disorders: Prevalence across diagnoses and differences by sex and ethnicity/race. *Mayo Clinic Proceedings* 94(8): 1415–426.

Craigie, J. 2011. Competence, practical rationality, and what a patient values. *Bioethics* 25(6): 326–33.

Draper, H. 2000. Anorexia nervosa and respecting a refusal of life-prolonging therapy: A limited justification. *Bioethics* 14(2): 120–33.

Dresser, R. 1984. Legal and policy considerations in treatment of anorexia nervosa patients. *International Journal of Eating Disorders* 3(4): 43–51.

Elzakkers, I., U. Danner, H. Hoek, and A. van Elburg. 2016. Mental capacity to consent to treatment in anorexia nervosa: Explorative study. *BJPsych Open* 2(2): 147–53.

Evans, A. 2023. Anorexia nervosa: Illusion in the sense of agency. *Mind and Language* 38: 480–94.

Gadsby, S. 2017. Anorexia nervosa and oversized experiences. *Philosophical Psychology* 30(5): 594–615.

————. 2020. Self-deception and the second factor: How desire causes delusion in anorexia nervosa. *Erkenntnis* 85(3): 609–26.

————. 2022. The rationality of eating disorders. *Mind and Language* 38(3): 732–49.

Geppert, C. 2015. Futility in chronic anorexia nervosa: A concept whose time has not yet come. *The American Journal of Bioethics* 15(7): 34–43.

Giordano, S. 2005. *Understanding Eating Disorders: Conceptual and Ethical Issues in the Treatment of Anorexia and Bulimia Nervosa*. Oxford: Oxford University Press.

————. 2010. Anorexia and refusal of life-saving treatment: The moral place of competence, suffering, and the family. *Philosophy, Psychiatry, and Psychology* 17(2): 143–54.

————. 2019. Anorexia nervosa: A case for exceptionalism in ethical decision making. *Philosophy, Psychiatry, and Psychology* 26(4): 315–31.

Godier, L., and J. Park. 2015. Does compulsive behavior in anorexia nervosa resemble an addiction? A qualitative investigation. *Frontiers in Psychology* 20(6): 1–12.

Gregertsen, E., W. Mandy, and L. Serpell. 2017. The egosyntonic nature of anorexia: An impediment to recovery in anorexia nervosa treatment. *Frontiers in Psychology* 8: 1–9.

Grisso, T. and P. Appelbaum. 1998. *The Assessment of Decision-Making Capacity: A Guide for Physicians and Other Health Professionals*. Oxford: Oxford University Press.

Groll, D. 2012. Paternalism, respect, and the will. *Ethics* 122(4): 692–720.

Hawkins, J. 2021. Why even a liberal can justify limited paternalistic intervention in anorexia nervosa. *Philosophy, Psychiatry, and Psychology* 28(2): 155–58.

————. Forthcoming. Affect, values and problems assessing decision-making capacity. *American Journal of Bioethics:*1-–12.

Holroyd, J. 2012. Clarifying capacity: Values and reasons. In *Autonomy and Mental Disorder*, ed. Lubomira Radoilka, 145–69. Oxford: Oxford University Press.

Hope, T., A. Stewart, and J. McMillan, 2013. Agency, ambivalence and authenticity: The many ways in which anorexia nervosa can affect autonomy. *International Journal of Law in Context* 9(1): 20–36.

Kim, S. 2016. The place of ability to value in the evaluation of decision-making capacity. In *Philosophy and Psychiatry: Problems, Intersections and New Perspectives*, ed. Daniel D. Mosely and Gary Gala, 189-203. New York: Routledge.

Martin, A. 2007. Tales publicly allowed: Competence, capacity, and religious belief. *Hastings Center Report* 37(1): 33–40.

Mill, J.S. 2003. *Utilitarianism and On Liberty*, ed. M. Warnock. Oxford: Blackwell Publishing.

Norbø, R., E. Espeset, K. Gulliksen, F. Skårderud, and A. Holte. The meaning of self-starvation: Qualitative study of patients' perception of anorexia nervosa. *International Journal of Eating Disorders* 39(7): 556–64.

Pickard, H. 2021. Addiction and the self. *Noûs* 55(4): 737–61.

Pugh, J. 2020. *Autonomy, Rationality, and Contemporary Bioethics*. Oxford: Oxford University Press.

Radden, J. 2021. Food refusal, anorexia, and soft paternalism: What's at stake?'. *Philosophy, Psychiatry, and Psychology* 28(2): 141–50.

Szmukler, G. 2018. *Men in White Coats: Treatment under Coercion*. Oxford: Oxford University Press.

————. 2021. Anorexia nervosa, lack of "coherence" with deeply held beliefs and values, and involuntary treatment. *Philosophy, Psychiatry, and Psychology* 28(2): 151–54.

Tan, J., T. Hope, and A. Stewart. 2003. Competence to refuse treatment in anorexia nervosa. *International Journal of Law and Psychiatry* 26: 697–707

Tan, J., T. Hope, A. Stewart, and R. Fitzpatrick. 2006. Competence to make treatment decisions in anorexia nervosa: Thinking processes and values. *Philosophy, Psychology, and Psychiatry* 13(4): 267–82.

Turrel, S., M. Peterson-Badali, and D. Katzmann. 2011. Consent to treatment in adolescents with anorexia nervosa. *International Journal of Eating Disorders* 44(8): 703–7.

Varga, S. and A. Steglich-Petersen. 2023. On the roles of false belief and recalcitrant fear in anorexia nervosa. *Mind and language* [Online first].

Westmoreland, P., C. Johnson, M. Stafford, R. Martinez, and P. Mehler. 2017. Involuntary treatment of patients with life-threatening anorexia nervosa. *Journal of the American Academy of Psychiatry Law* 45: 419–25.

Voswinkel, M., C. Rijkers, J. van Delden, and A. van Elburg. 2019. Externalizing your eating disorder: A qualitative interview study. *International Journal of Eating Disorders* 9(128): 1–12.

Whiting, D. 2009. Does decision-making capacity require the absence of pathological values?. *Philosophy, Psychiatry, and Psychology* 16(4): 341–44.

Yager, J. 2015. The futility of arguing about medical futility in anorexia nervosa: The question is how would you handle highly specific circumstances?. *The American Journal of Bioethics* 15(7): 47-50.

Yilmaz, Z., M. Halvorsen, J. Bryois, D. Yu, L. Thornton, S. Zerwas, N. Micali, R. Moessner, C. Burton, G. Zai, L. Erdman, M. Kas, P. Arnold, L. Davis, J. Knowles, G. Breen, J. Scharf, G. Nestadt, C. Mathews, C. Bulik, M. Mattheisen, J. Crowley, and the Eating Disorders Working Group of the Psychiatric Genomics Consortium, Tourette Syndrome/Obsessive–Compulsive Disorder Working Group of the Psychiatric Genomics Consortium. 2020. Examination of the shared genetic basis of anorexia nervosa and obsessive-compulsive disorder. *Molecular Psychiatry* 25(9): 2036–46.