The Enigmatic Persistence of Anorexia Nervosa

B. Timothy Walsh, M.D.

**Objective:** In this review, based on recent advances in cognitive neuroscience, the author presents a formulation in which the marked persistence of anorexia nervosa can be usefully understood as a well-ingrained maladaptive habit.

**Method:** The author reviewed the relevant literature on the development and course of anorexia nervosa and interpreted critical features in light of developments in cognitive neuroscience.

**Results:** Anorexia nervosa is a well-characterized disorder with remarkable persistence both across history and among affected individuals. Food restriction, the salient behavioral feature of the disorder, often begins innocently but gradually takes on a life of its own. Over time, it becomes highly entrenched and resistant to change through either psychological or pharmacological treatment. Cognitive neuroscience has described two related but distinct processes that underlie the acquisition of new patterns of behavior, namely, action-outcome and stimulus-response learning. It is likely that both processes are engaged in the development of anorexia nervosa and that stimulus-response learning (that is, habit formation) is critical to the persistence of the dieting behavior.

**Conclusions:** The formulation of the dieting behavior characteristic of anorexia nervosa as a well-entrenched habit provides a basis for understanding the striking persistence of this disorder. This model helps explain the resistance of anorexia nervosa to interventions that have established efficacy in related disorders and implies that addressing the dieting behavior is critical, especially early in the course of the illness, before it has become ingrained.

(Al J Psychiatry 2013; 170:477–484)

---

Anorexia nervosa is a serious psychiatric disorder with remarkable persistence, both over centuries and among many individuals who develop it. Anorexia nervosa occurs primarily in cultures where food is plentiful and where, for some reason, special value is attached to dieting and weight loss. Its salient features—a relentless pursuit of thinness accompanied by an intense fear of weight gain despite being significantly underweight—have remained constant over time. The first known description of the disorder in the medical literature was provided by Richard Morton in 1689, and it has been speculated that some medieval saints may have suffered with anorexia nervosa (3, 4). In the late 19th century, Sir William Gull coined the term “anorexia nervosa” in his description of three young women who would clearly meet DSM-IV and the proposed DSM-5 criteria for the disorder (5).

The course of anorexia nervosa is also often remarkably persistent. Although the range of outcomes is wide, anorexia nervosa is one of the deadliest of psychiatric disorders, with an estimated mortality rate at least five times higher than that expected (6). Among the few factors associated with a more favorable outcome are adolescent onset (in contrast to adult onset) and a shorter duration of illness (7). The positive significance of an early adolescent onset of anorexia nervosa is in contrast to mood and psychotic disorders, in which early adolescent onset is associated with a more refractory course (8, 9).

Although a number of risk factors have been identified, including being female, being an adolescent, and having an obsessional style, they are not particularly specific, and they account for only a small fraction of the variation in occurrence (10). The overwhelming majority of adolescent girls, even those who tend to be obsessional, do not develop anorexia nervosa. Similarly, while dieting behavior is almost universal among adolescent girls in the United States, it develops into anorexia nervosa in very few of them.
THE ENIGMATIC PERSISTENCE OF ANOREXIA NERVOSA

few. Most cases of anorexia nervosa begin unremarkably. Many young women begin to diet in hopes of losing weight and gaining attractiveness and self-esteem, often in response to one of the ordinary challenges of adolescence, such as beginning to date or leaving home for camp or school. In the weeks or months that follow, among the few who develop anorexia nervosa, the dieting takes on a life of its own and evolves into an unrelenting pursuit that becomes the individual’s primary focus. Early intervention appears to interrupt the increasingly restrictive dieting and growing social isolation. But in many individuals, dieting and weight loss dominate mental life for years and severely impair physical health and social and occupational development.

Efforts at improving outcomes through the development of specific treatment interventions have proven frustrating. One of the few significant advances in recent years is the articulation and promulgation of a family-based treatment approach developed at the Maudsley Hospital in London several decades ago (11, 12). In stark contrast to previous approaches, this method specifically charges the parents of the affected individual with a major role in intervention. Controlled studies support the utility of this approach (13), but it is focused on adolescents with anorexia nervosa, who are those most readily treated. Anorexia nervosa in adults, including those only in their 20s, is generally much more refractory to intervention. Treatments of established efficacy in related disorders, such as bulimia nervosa and mood and anxiety disorders, including cognitive-behavioral therapy and selective serotonin reuptake inhibitors, are surprisingly ineffective (14). Older individuals with anorexia nervosa can be helped to restore normal weight in structured treatment programs, and they may exhibit substantial improvement in both physical and psychological symptoms, but relapse rates after discharge are high (15). Thus, despite better recognition and presumably more informed treatment approaches, the outcome of anorexia nervosa, particularly for adults, has probably not improved substantially in the past 50 years (7).

These salient facts pose a challenge to our understanding of this disorder. How is it that children and adolescents in generally good health develop an illness that is often chronic, severe, and possibly deadly but that may, especially if treated early, fully remit without any long-term consequences? In this review I extend an earlier discussion (16) and address one important facet of this enigma, namely, the remarkable persistence of anorexia nervosa once it has become established.

Eating Behavior and the Persistence of Anorexia Nervosa

Individuals with anorexia nervosa achieve and maintain a significantly low body weight primarily by dieting and, in a significant fraction of individuals, by engaging in excessive physical activity. (Although in this review I focus on the dieting behavior for the sake of clarity of exposition, the discussion can readily be extended to include increased physical activity as well.)

Among the striking characteristics of anorexia nervosa is the homogeneity of disturbances in eating behavior across affected individuals. Individuals with the disorder uniformly restrict their caloric intake by consuming a limited variety of foods that are low in energy density. Once established, such behavior—which has been documented in laboratory settings—is highly resistant to change, persisting even after normal weight has been restored in intensive treatment programs through consumption of a varied diet typical of individuals without eating disorders (17–20). Recent work by our group strongly suggests that the persistence of such behavior is a major contributor to relapse (15, 21, 22).

Therefore, an improved understanding of why it is so difficult for individuals with anorexia nervosa to alter their eating behavior is critical to improving our understanding of the development of the disorder and to improving treatment interventions. The formulation I present here attempts to explain how such dieting behavior becomes established and persistent. Its central element is that this behavior has become habitual and is grounded in the neural mechanisms that underlie habit formation and persistence. This formulation is based on a substantial body of animal and human research, much of which is focused on the neural mechanisms underlying substance abuse.

Diетing Behavior in Anorexia Nervosa Is Habitual

Graybiel (2) provided a useful working definition of habits from the perspective of cognitive neuroscience: habits are behaviors that are not innate, that are engaged in repeatedly and become fixed, that appear to occur without conscious effort, and that are elicited by a variety of stimuli. The dieting behavior characteristic of anorexia nervosa fulfills this definition. Although there are hints that some individuals who develop anorexia nervosa may have been picky eaters as children (23), there are no indications that eating behavior or rate of weight gain are abnormal prior to the onset of illness. In order to lead to significant weight loss, the dieting behavior must be engaged in repeatedly, over weeks to months, and the studies of eating behavior cited above demonstrate that it becomes markedly fixed (18, 20). Individuals with anorexia nervosa appear to diet without obvious attempts to contemplate the specific elements of eating behavior; conversely, attempts to consume a more varied and calorically dense diet require substantial effort (24). The characteristic food choices are elicited by a variety of stimuli, including the presentation of foods at mealtime (17, 18), and by emotional stress (25, 26).
How Habits Develop

Substantial research over the past several decades has articulated two related but distinct processes responsible for the acquisition and persistence of behavior that is not innate: action-outcome and stimulus-response learning (2, 27–29). Much of this research has been carried out using laboratory animals, especially rodents, but the key findings have also been established in humans.

In action-outcome learning, also referred to as instrumental or operant conditioning, an animal detects that some action, such as pushing a lever, leads to a reward, such as the presentation of food, and the likelihood of the animal’s engaging in the behavior is sensitive to the reward value of the outcome. Hence, if at some point the behavior no longer leads to the receipt of the reward, its frequency will diminish, and eventually it will occur only sporadically. This form of learning is critical to the acquisition of new behaviors, and it involves specific neural pathways, including the amygdala, the ventral striatum (the nucleus accumbens), and the orbitofrontal cortex.

While new behaviors are typically acquired through action-outcome learning, if they are repeated and continue to lead to the acquisition of rewards, the behaviors will eventually become relatively insensitive to the receipt of the reward. This process, referred to as stimulus-response learning, is the basis for habit formation. The key neural structures involved are neuroanatomically close to, but distinct from, those required for action-outcome learning: the dorsolateral striatum (the caudate/putamen) and the dorsolateral prefrontal cortex.

A critical element in stimulus-response learning is overtraining—the repetition of the behavior. A popular term referring to this phenomenon in athletic pursuits is the development of “muscle memory,” acquired, for example, by repeatedly practicing a baseball swing. “Muscle memory” is a misnomer, as muscles have no capacity for memory; it is the brain that (hopefully) learns a productive swing cued by the appearance of a baseball.

In the model presented here, which is based on these principles of learning, it is hypothesized that the dieting behaviors of individuals with anorexia nervosa begin as goal-directed actions that lead to weight loss, which is highly rewarding (action-outcome learning). Over time, the dieting behaviors are engaged in persistently and repeatedly and thereby become overtrained and habitual (stimulus-response learning). Once they have become established as habitual, the behaviors are highly resistant to change and are a critical element in the persistence of the disorder. The next sections describe critical elements in this learning process.

Dieting Behavior Is Strongly Reinforced

By definition, individuals with anorexia nervosa diet effectively enough to lose weight (or, if still growing, to fail to gain weight as expected). In the current Western culture, successful weight loss is an often wished for and encouraged goal that is rarely achieved. Therefore, the initial weight loss of individuals with anorexia nervosa provides evidence of impressive self-control and personal accomplishment, leading to enhanced self-esteem. In addition, many individuals who develop anorexia nervosa describe receiving compliments at the beginning of their successful efforts at weight loss.

Such feelings about the emotional significance of weight loss are well known to clinicians experienced in the care of individuals with anorexia nervosa and have been documented in numerous descriptions of the disorder (30–32). Bruch (33), writing in the latter half of the 20th century, provided some of the most compelling accounts of the symptoms and internal motivations of individuals with anorexia nervosa. Describing the perspective of one individual, Bruch wrote, “That she could be so thin gave her a sense of pride, power, and accomplishment” (p. 2). Another individual “explained that losing weight was giving her power, that each pound lost was like a treasure that added to her power” (pp. 4–5). Garfinkel and Garner (34), in their important volume on anorexia nervosa, noted, “While anorexia nervosa patients have an overriding drive to be thin, they do not offer any explanation; they merely say they will feel better the thinner they are” (p. 3). Thus, the dieting behavior leads to weight loss, which is highly rewarding and is thereby positively reinforced.

Individuals with anorexia nervosa frequently report that dieting behavior also helps them cope with negative affect, and recent studies have begun to document this phenomenon experimentally. Engel et al. (25), using ecological momentary assessment, found that increased levels of negative affect were followed by restriction of food intake. Wildes et al. (26) found that induction of negative affect in a laboratory setting led to an increase in self-reported eating disorder symptoms, although intake of a test meal was not altered. These observations suggest that in individuals with anorexia nervosa, dieting behavior may be triggered by negative affect and serves as a method of modulating emotional distress. The relief of negative affect thus likely constitutes an additional, negative reinforcer of dieting behavior. In other words, dieting behavior is rewarded both by leading to weight loss and enhanced self-esteem and by helping individuals cope with emotional difficulties.

Dieting Itself Becomes Rewarding

It is likely that in anorexia nervosa, dieting behavior itself becomes rewarding through a phenomenon called “conditioned reinforcement” (35). Conditioned reinforcement occurs when a stimulus or situation is regularly paired with an established reward. Such phenomena are viewed as a critical element in the development of the complex set of cues that, over time, become associated
with craving for drugs of abuse and of the rituals associated with drug taking themselves becoming rewarding (36, 37). Since in anorexia nervosa dieting behavior is paired with a potent reward—weight loss—it is likely that, over time, dieting behavior itself takes on rewarding properties. Such a phenomenon helps to explain the distinction between “good” (low-calorie) and “bad” (high-calorie) foods described with remarkable consistency by individuals with anorexia nervosa and the difficulties they experience in making even minor changes in their dieting practices.

**Rewards Are More Rewarding During Adolescence**

Most cases of anorexia nervosa begin during adolescence, a time of dramatic psychological and physical development. During the teens and into the 20s, substantial cortical development contributes importantly to emotional and cognitive maturation—for example, to the ability to postpone the receipt of rewards (38). Many psychiatric disorders emerge during this phase of life, undoubtedly reflecting a complex interplay of environmental, social, and internal factors (39). Adolescents are disproportionately sensitive to the receipt of reward (40). For example, the response of the ventral striatum (a site well established in the assessment of the potential reward value of stimuli) to rewarding stimuli has been shown to be greater in adolescents than in children and adults, suggesting that the same reward is assessed as being more rewarding at this stage of development (41). The attraction of rewarding stimuli, coupled with the relative immaturity of inhibitory neural systems, has been suggested as an important factor underlying the vulnerability of adolescents to developing substance abuse (42), and it may similarly contribute to the vulnerability to anorexia nervosa during this stage of development.

**The Dieting Behavior Becomes Habitual**

The preceding sections have made the case that at the onset of anorexia nervosa, weight loss is highly rewarding. Thus, the acquisition of dieting behavior is a classic example of action/outcome learning. Other factors lead to these behaviors becoming habitual.

First, in order to lead to significant weight loss, dieting behavior must be engaged in persistently and repeatedly. The behavior becomes highly practiced and overtrained, thereby becoming more entrenched. This sequence is the classic technique producing habit learning (2).

Second, the rewards of dieting behavior are intermittent. Daily weight loss is not inevitable, and the receipt of social accolades for becoming thinner occurs only occasionally. Skinner’s early work on instrumental conditioning established that an intermittent and unanticipated pattern of reinforcement encourages establishment of habitual behavior (43, 44).

Third, the onset of anorexia nervosa typically begins during a period of stress. Most of the stressors are those routinely encountered during adolescence, but they are nonetheless emotionally compelling at the time they are experienced. Both rodent and human studies have demonstrated that behaviors learned during a period of stress are more likely to become habitual than those learned at other times (45, 46). Finally, by definition, anorexia nervosa involves significant weight loss. There is good evidence that substantial weight loss enhances the tendency to develop compulsive patterns of behavior. The Minnesota starvation study conducted during World War II (47) documented the development of irrational compulsive behaviors in previously healthy subjects during a period of starvation. Conversely, in anorexia nervosa, weight gain is associated with improvement in symptoms of obsessiveness (48). Furthermore, a substantial body of work by Carr et al. (49, 50) has established that weight loss enhances the rewarding effects of drugs of abuse and up-regulates mechanisms of synaptic plasticity in the nucleus accumbens. These findings suggest a mechanistic basis for the development of particularly well-entrenched habits, such as resistance of cocaine-conditioned behavior to extinction in food-restricted rodents (51).

**Relationship to Other Models**

Many models have been suggested to account for the development and persistence of anorexia nervosa. Strober (52) hypothesized that individuals who develop anorexia nervosa have an underlying propensity to anxiety and to fear-learning and a greater resistance to its extinction than unaffected individuals. Like the model presented here, Strober’s model is based on relatively recent advances in our understanding of how behavior is acquired and is shaped by internal and external stimuli. It suggests that negative emotion, specifically anxiety, about weight gain leads to enduring patterns of dieting. The model presented here similarly suggests that dieting and weight loss become learned techniques to relieve anxiety and other negative emotions but that this development is only one of several factors contributing to such behaviors becoming strongly rewarding. Both models suggest that once acquired, dieting behavior in anorexia nervosa is remarkably resistant to extinction, but the formulation I present proposes that this resistance, to a substantial degree, reflects the establishment of well-entrained stimulus-response learning—that is, habits. As hypothesized by Strober, there may also be an innate resistance to the extinction of learned behaviors.

Other recent models have focused on the roles of emotional regulation (53–55) and of cognitive and cognitive-behavioral disturbances (56, 57). While generally acknowledging that
weight loss and dieting are highly reinforcing for individuals with anorexia nervosa and are useful in responding to negative emotion, these models do not grapple in detail with precisely how, in terms of fundamental processes of learning, such behaviors become so rewarding and persistent.

An array of CNS disturbances have been described in anorexia nervosa that may affect the processing of a range of stimuli, including ones linked to reward (58–60), but the models based on these findings do not articulate how such disturbances specifically lead to the persistence of dieting behavior.

The model proposed here does not directly challenge or contradict these other proposals. Rather, it places them in a broader context by suggesting specific mechanisms through which weight loss and dieting behavior become intensely rewarding and, eventually, habitual. Once acquired, these learned behaviors may well be utilized to cope with negative emotion and be sustained by innate or acquired biological disturbances affecting the CNS.

Implications

The proposed model has a number of implications. The hypothesis that the core behaviors characteristic of anorexia nervosa are habitual helps account for the established observation that the disorder is much more readily treated in younger individuals with shorter durations of illness (7); in such individuals, complete recovery is a routine outcome. With a shorter duration of illness, the individual has had less time to overtrain the dieting behavior, and the establishment of dieting behavior as rewarding in its own right is less secure.

This formulation also supports the recommendation that interventions focusing on interrupting the behavior and restoring weight to normal are crucial in achieving full recovery. The mechanisms described suggest that weight loss and dieting behavior have acquired reinforcing value and are likely to persist unless directly addressed. It is notable that the first phase in family-based treatment, an intervention of documented utility for younger patients, emphasizes the need for parents to actively refeed their child, that is, to interrupt the dieting behavior (12). It may also be useful to help patients recognize that dieting behavior has become largely automatic and may be triggered by a wide range of external and internal cues, including negative affective states. For patients with long-standing anorexia nervosa, it may be important to enhance current interventions by adapting techniques employed in habit reversal therapy that have been found to be useful in addressing disorders such as trichotillomania and Tourette’s disorder (61).

The suggested neural basis underlying the development of persistent dieting behavior in anorexia nervosa is fundamentally identical to that proposed to account for the development of persistent abuse of inherently rewarding substances such as cocaine (2). These neurobiological similarities provide a compelling explanation for the striking clinical similarities between the behavior of individuals with substance use disorders and that of individuals with anorexia nervosa. In both disorders, individuals become intensely focused on the desired result (acquisition of the abused substance or weight loss, respectively) to the exclusion of occupational and social activities, may engage in deceitful behavior to attain it (stealing money to buy drugs or lying about whether or what they have eaten), and sacrifice physical health in pursuit of the rewarding goal. Indeed, Overeaters Anonymous specifically includes individuals with anorexia nervosa among those who may be helped by its 12-step program (62). But the model proposed here in no way suggests that individuals with anorexia nervosa are “addicted” to a substance or a food. Rather, the basis of the similarity is that the same neural structures are engaged in the acquisition of dieting behavior as habitual and in the chronic abuse of rewarding substances.

Indirectly, the proposed model may offer at least a partial explanation for the surprising relative ineffectiveness of interventions such as cognitive-behavioral therapy, interpersonal therapy, and antidepressants for anorexia nervosa (63–65). Such interventions are, in general, quite effective in treating affective disorders and related eating disorders, such as bulimia nervosa. However, there is no reason to believe that they possess substantial utility in addressing well-entrenched habits of behavior. Similar considerations may help explain the relative ineffectiveness of such treatments for substance abuse.

The proposed model provides an explanation for the occurrence of anorexia nervosa over many centuries. The critical element is that dieting and weight loss are valued and therefore capable of providing a potent subjective reward. In current Western society, dieting and being thin are equated with health and, especially for women, with attractiveness. In the middle ages, other rewards of food restriction were prominent, including as a route to asceticism and spiritual purity (4).

The proposed model suggests that distinct neural circuits are involved in the acquisition and perpetuation of the dieting behavior characteristic of anorexia nervosa. In the early stages of illness, as weight loss and dieting behavior are recognized and “coded” as rewarding, the amygdala, the ventral striatum, and the orbitofrontal cortex are engaged. Over time, however, as the dieting behavior becomes habitual, although these centers remain involved, the dorsal striatum and the dorsolateral prefrontal cortex play a more critical role. Given the increasing sophistication of experimental paradigms using functional MRI to identify the relative contributions of different neural circuits in supporting decisions, including about what to eat, it should be possible to examine these hypotheses in individuals restricting their food intake,
ranging from healthy dieters to individuals with chronic and severe anorexia nervosa (66).

It is possible that the key elements of the proposed model might be extended to help us understand the persistence of maladaptive behaviors in other psychiatric conditions. If a behavioral pattern develops and is repetitively engaged in at a vulnerable and malleable time of life, such as during adolescence, and is initially associated with a powerful reward, such as relief from an intensely negative emotion, it may become habitual, especially if, at least intermittently, it continues to be associated with the receipt of a reward. The application of this notion is straightforward if the reward is obvious, such as taking a drug of abuse or consuming highly palatable food in bulimia nervosa or binge eating disorder. But perhaps this sequence also plays a role in the development and persistence of maladaptive interpersonal relationships and risky behaviors.

Summary

The developmental and psychosocial context in which anorexia nervosa develops sets the stage for weight loss and, over time, the characteristic dieting behaviors themselves to become intensely rewarding. Several additional factors lead to dieting behavior becoming a well-established habit and very resistant to change. This formulation thereby provides a possible insight into the impressive persistence of anorexia nervosa once it has developed. However, other factors, both innate and acquired, undoubtedly also contribute to its chronicity, including biological, psychological, and environmental influences that reduce emotional and behavioral flexibility.

It is also important to emphasize that this proposal addresses only one facet of a complex disorder. For example, it does not directly address the difficult but critical issue of vulnerability: Why do only a very small number of individuals who begin to diet during adolescence go on to develop anorexia nervosa? The proposed model raises the possibility that the among the risk factors is an innate ability to establish habitual behaviors or, conversely, a difficulty in extinguishing well-practiced behaviors. An association between anorexia nervosa and both obsessive-compulsive disorder and obsessive-compulsive personality disorder is well documented, and there is good evidence that the symptoms of individuals with anorexia nervosa extend beyond issues related to eating or weight (10, 67). Relevant to the formulation proposed here, a recent study focusing on obsessive-compulsive disorder found evidence that affected individuals, relative to healthy comparison subjects, rely more on habitual patterns of response and are less able to adjust their behavior to changing contingencies (68). Vulnerability may also be increased by individuals’ difficulties in managing some of the emotional challenges of adolescence and by environmental influences, such as the perceived advantages of strict dieting. Presumably, such innate and external factors set the stage for the development and persistence of relentless dieting.

The proposed model focuses on the development and persistence of the salient behavioral feature of anorexia nervosa—severe restriction of food intake—and does not specifically address other important features of the disorder, such as the emergence of binge eating, a phenomenon that, over time, occurs in approximately one-half of those with anorexia nervosa (69). It is important to recognize that despite the occurrence of episodes of binge eating, often followed by purging, these individuals continue to restrict their food intake in order to maintain their low weight (70). Since such individuals are prone to engage as well in impulsive behaviors other than binge eating, it seems reasonable to hypothesize that these phenomena are a manifestation of difficulties with inhibiting responses to naturally rewarding stimuli, including palatable food, after long periods of starvation. Such phenomena were well documented among the men who participated in the Minnesota starvation study (71).

Each of the elements underlying the proposed model has a basis in fact, but the resulting formulation requires elaboration and testing. The propensity of individuals with anorexia nervosa to develop and extinguish habits when underweight and after short-term and long-term recovery should be examined. The activity of neural centers involved in the recognition of rewards and the development of habits is certainly relevant. It may be useful to examine such phenomena in animal models. Finally, the articulation and testing of treatment interventions based on the hypothesized mechanisms would be of potential heuristic importance and clinical utility.

References


43. Humphreys LG: The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions. J Exp Psychol 1939; 25:141–158


Clinical Guidance: Anorexia Nervosa Viewed as Habitual Behavior

Many patients with anorexia nervosa report that dieting helps them cope with negative affect, but once the behavior is paired with a desired outcome, it may become rewarding itself. This formulation of anorexia nervosa by Walsh is consistent with findings that adolescent onset and shorter duration of illness are among the few factors associated with better outcome. Cognitive-behavioral therapy, interpersonal therapy, and antidepressants are ineffective, whereas interrupting the behavior and restoring normal weight are considered crucial. Younger patients often respond to family-based treatment, in which parents refeed the child to interrupt the dieting behavior. Patients with long-standing anorexia nervosa may be helped by adding habit-reversal techniques to current interventions.