Toward a Better Understanding of the Development of Overweight: A Study of Eating Behavior in the Natural Environment using Ecological Momentary Assessment

A Thesis

Submitted to the Faculty

of

Drexel University

by

John Graham Thomas

in partial fulfillment of the

requirements for the degree

of

Doctor of Philosophy

May 2009

Acknowledgements

I would like to thank my mentor, Dr. Michael Lowe, whose years of patient training prepared me for this independent research project. I must also thank my other committee members, Dr. Kelly Allison, Dr. Ross Crosby, Dr. David DeMatteo, Dr. Evan Forman for their thoughtful guidance during the development and culmination of this research project. I am especially grateful to Dr. Crosby, for his introduction to the theory and technique of Ecological Momentary Assessment. This project would not have reached fruition with the help of Sapna Doshi, who helped execute the study protocol, and Lauren Lubarski, who provided vital administrative support on multiple occasions.

Table of Contents

List of Tables

List of Figures

Abstract

Toward a Better Understanding of the Development of Overweight:
A Study of Eating Behavior in the Natural Environment using
Ecological Momentary Assessment
John Graham Thomas
Michael R. Lowe, Ph.D.

Obesogenic eating behavior is driven by a combination of person-specific factors (e.g., individual differences in physiology and attitudes towards food & eating) and environmental factors (e.g., type and amount of foods available). This study used Ecological Momentary Assessment (EMA) via palmtop computers to collect real-time information about participants' environment, attitudes, and eating patterns to predict overeating (i.e., greater than usual intake during routine meals/snacks, and eating outside of a participant's normal routine) that could lead to weight gain. The EMA was completed by 43 women of normal weight, who denied any history of an eating disorder. Participants carried a palmtop computer for 7-10 days, which prompted them six times daily to answer questions about eating episodes, including the number of high-calorie foods in the environment, episodes of overeating, and fluctuations in restraint (as measured by efforts to eat less often, less portions, and fewer fattening foods). On average, 29.8% of eating episodes were characterized by self-identified overeating. Hierarchical linear models showed that BMI interacted with the number of high-calorie foods available in the environment to predict the occurrence of overeating (p = .035). Specifically, for individuals with a higher BMI, the probability of overeating was low in the absence of high-calorie foods, but quickly increased as the number of high-calorie foods available increased. For all participants, on days in which overeating occurred,

dietary restraint was significantly higher after overeating than before. A significant non-linear trend in dietary restraint was observed (p = .019), such that restraint began to increase gradually in the hours prior to self-identified overeating, and accelerated as the episode of overeating approached. Restraint reached a peak several hours after overeating. Reminiscent of Schachter's early work, the eating behavior of heavier individuals is susceptible to environmental cues. Also, dietary restraint appears to have a complex relationship with overeating in that it is unclear whether restraint leads to or results from bouts of over-consumption. In sum, healthy weight control may be facilitated by limiting high-calorie foods in the immediate environment, and by encouraging healthy dietary restraint.

CHAPTER 1: INTRODUCTION

1.1 Understanding Obesity: The Role of Excessive Energy Intake

Global rates of overweight and obesity[1] are rising (Flegal, Carroll, Ogden, & Johnson, 2002). As of 2005, a majority of adult Americans were overweight or obese. One result of this upward trend in overweight is an unprecedented increase in the number of people suffering from weight-related illness such as type 2 diabetes (Rejeski et al., 2006), cardiovascular disease (Evangelista & Miller, 2006), and certain cancers (Ferrante, Chen, Crabtree, & Wartenberg, 2007). In addition to an increased likelihood of disease, overweight is associated with shorter lifespans (Greenberg, 2006) and reproductive difficulties (Sarwer, Allison, Gibbons, Markowitz, & Nelson, 2006).

It is incumbent upon researchers and health professionals to understand the factors that are responsible for the development and maintenance of overweight and obesity. At a basic physiological level, excess adipose tissue is the result of excessive energy intake, insufficient energy expenditure, or a combination of the two (James, 2002). While a lack of physical activity may be partly responsible for the trend in increasing body weights, the human body has a very limited ability to vary its energy expenditure (Jebb, 2002). For example, running a marathon increases energy expenditure for the day by only about 50%, and spending the day in bed reduces energy expenditure by only 20%. By contrast, there is huge potential for variability in food intake. It is easy to triple normal daily energy intake by eating large quantities of energy dense food, and it is possible to reduce energy intake to as few as 400 k/cal per day by fasting, as is common during a very-low-calorie diet (Foster et al., 1992). These observations have

[1] The world health organization (WHO) defines overweight as a BMI between 25 and 30 kg/m^2, and obesity as a BMI > 30 kg/m^2.

been cited as evidence suggesting that the current obesity epidemic is driven more by problems with energy intake than energy expenditure (Jeffrey & Utter, 2003).

Overweight people are known to eat more than is required to maintain a healthy weight, and choose unhealthy foods that are high in energy density (i.e., calories per unit volume) but low in nutritional value (Mela, 2001). Furthermore, evidence suggests that these eating behaviors are a cause of overweight, and not simply a consequence of weight gain (Mela, 2001). Early research suggested that these obesity-related eating behaviors were caused by dietary restraint (i.e., attempts to restrict eating in the face of biological pressures towards weight gain; Nisbett, 1972). More recent research suggests that excessive energy intake is more likely the result of an interaction between basic appetitive systems and the current food environment (Lowe & Levine, 2005).

Lowe and Levine (2005) reviewed evidence that the eating behavior of humans and animals is driven by two appetitive systems. The first system is homeostatic. It is activated by energy deficit, and results in sensations of hunger. The second system, which is activated in the presence of palatable food, is labeled "hedonic." Lowe and Levine (2005) argue that activation of the homeostatic system is not responsible for the majority of eating episodes in developed countries because individuals in these countries learn to anticipate hunger and eat preemptively to avoid it. Given that the homeostatic system is rarely active, the majority of eating is thought to be driven by the hedonic system. As most highly palatable foods are also high in energy density, frequent activation of the hedonic system may result in excessive consumption of high-calorie foods, which contributes to weight gain.

The effect of the food environment was described by Peters (2003), who illustrated how environmental changes in the composition and availability of food may be responsible for much of the increase in the prevalence of obesity over the past three decades in the United States. The four changes in the food environment that are most responsible for the obesity epidemic appear to be increases in portion size, energy density, dietary variety, and food availability and palatability. These factors are reviewed in greater detail later.

While the evidence suggests that the hedonic appetitive system and the obesogenic food environment both contribute to weight gain independently, it is their interaction that is most troubling. Lowe and Levine (2005) argue that the hedonic system is constantly activated by our current food environment because of the easy availability of highly palatable foods in large portions. The result is a near constant drive to eat the foods that are most likely to cause weight gain. Unless the urge to eat palatable foods is somehow denied, the result is excessive energy intake. Over time, elevated intake leads to overweight.

1.2 Obesity Treatment and Prevention: Interventions to Limit Excessive Energy Intake

Researchers have attempted to take what is known about the causes of excessive energy and apply the knowledge to interventions aimed at weight control. Efforts to limit excessive energy intake are at the heart of most empirically validated treatments for weight loss and weight maintenance (Wadden, Butryn, & Wilson, 2007). The so-called lifestyle approach to weight control is the current gold standard (Wadden et al., 2007). Lifestyle modification programs have multiple components, in which participants must

invest considerable time and energy if they are to be successful. Some aspects of lifestyle modification programs, such as instructions to limit portion size and eat low-calorie foods, are proximally related to reducing the number of calories consumed. Other aspects, such as learning to regulate emotions and finding peer support, are more distally related to curbing excessive intake. Nevertheless, the end goal is usually better control over eating (or increased physical activity).

Lifestyle interventions lasting 16 to 24 weeks typically produce an average weight loss of about 10% of starting weight (Wadden et al., 2007). This amount of weight loss is associated with significant improvements in physical and psychological health (Wadden et al., 2007). However, lifestyle programs have been criticized for several reasons. First, a weight loss of 10% would not put most individuals back into the healthy weight range. Second, these interventions are successful only for patients who are highly compliant (Wadden et al., 1995). The percentage of compliant participants in most behavioral weight loss studies ranges from 30% to 60% of those who attend at least one treatment session (Kaplan & Atkins, 1987). Worse still, this low range may be an overestimate, as additional participants often fail to attend even a single session. However, rates of dropout before the initiation of treatment are rarely reported. The low rates of compliance may be related to any number of factors including the large investment in time and energy that is required, unrealistic expectations for weight loss, or ambivalence about the need for weight loss. Regardless of the reason, only a minority of participants in lifestyle programs will ever be successful in achieving a meaningful weight loss.

Unfortunately, even those patients who are able to lose weight typically regain the lost weight within 5 years (Elfhag & Rossner, 2005). There are very few patients who

maintain a significant weight loss over the long term. In fact, these patients are so rare that a national organization, the National Weight Control Registry, was founded to track and study these individuals (Daeninck & Miller, 2006). Evidently, long-term maintenance of a weight loss is not achievable for most individuals, at least given existing intervention technology.

The evidence reviewed above indicates that once overweight is achieved, it is very unlikely that an individual will lose an amount of weight that will result in significant health benefits, much less return to the normal weight range. Furthermore, even if an individual is able to lose weight, the loss is almost always short lived. It is beyond the scope of the current investigation to review all of the available weight loss treatments. Pharmacological treatments generally have similar or slightly lower efficacy than lifestyle programs (Chaput, St-Pierre, & Tremblay, 2007). Surgical interventions are costly and carry risks of mortality and morbidity that limit their usefulness (Powell, Calvin, & Calvin, 2007).Thus, weight loss treatment is unlikely to ever result in a long-term solution to the obesity epidemic.

Given the limitations of weight loss treatment, many researchers are now turning their efforts towards prevention. Compared to weight loss treatments, interventions for weight gain prevention are in their infancy. In order to create effective preventive interventions, it is essential to understand what causes the primary risk factor for the development of overweight (i.e. excessive energy intake). As alluded to previously, the study of excessive energy intake has produced several competing theories. A recently developed theory that includes a focus on appetitive systems and the food environment has shown much promise (Lowe & Levine, 2005). However, this theory is new and needs

additional testing to confirm its usefulness as a potential starting point for weight gain prevention interventions.

1.3 Goals of the Present Study: A Better Understanding of Excessive Energy Intake

The goal of the current study is to collect information that may be used to test theories of eating behavior related to excessive energy intake, and to refine these theories where they are found to be inadequate. The design will incorporate measures of trait-like characteristics such as restrained eating and sensitivity to the food environment, as well as a state-like measures of restraint and the obesogenicity of the food environment. Data will be collected using ecological momentary assessment (EMA), which allows participants to be measured throughout the day, over the course of several days, using palm top computers. This method of assessment overcomes many of the limitations of more traditional instruments such as questionnaires (Smyth et al., 2001). The following sections provide further information on the theories of eating to be tested, and the methods that will be used to test them.

1.4 Introduction to Restraint Theory

The development and maintenance of excessive energy intake is often conceptualized within a theoretical framework known as restraint theory. The emphasis on restrained eating is partially attributable to Nisbett (1972), who conducted a variety of studies to understand why overweight persons display unhealthy eating behaviors. More specifically, he sought to understand why overweight people often ate more than those of normal weight in selected contexts. Nisbett's work led him to develop set-point theory to explain the behavior of obese persons. His work formed the foundation of what was to become a field replete with a variety of theories of eating behavior, as well as research

findings to support them. The most well-known restraint theories are reviewed in this section, in chronological order of their development.

1.5 Early Restraint Theory

1.5.1 *The "Internal-External" Theory of Obesity and Nisbett's Set Point Theory*

Early in the study of eating behaviors, Nisbett and his colleague Schachter (e.g., Schachter, 1968) noticed several differences between normal weight and obese persons. For example, compared to normal weight participants, obese participants eat more in response to negative emotions and the presence of highly palatable food. Schachter reasoned that the differences between obese and normal participants were the result of obese participants' reliance on external (i.e., environmental) cues for eating, and a corresponding de-emphasis on internal cues of hunger and satiety, which normal weight persons relied on to a much greater extent. A variety of studies supported this conclusion. Subsequently, Schachter developed the internal-external theory of obesity, which suggests that obese individuals' reliance on external cues is an important *cause* of their overweight (Schachter, 1968).

Later, Nisbett (1972) proposed that the differences in obese and normal weight eating behavior are the result of obese persons holding their weight below its biologically appropriate level, or "set point." He suggested that each individual is endowed with a certain number of fat cells, and that the body tends to regulate eating behavior such that these fat cells maintain a certain degree of fullness. Subsequently, persons who are endowed with a greater than average compliment of fat cells would be predisposed to overweight. However, Nisbett recognized that overweight is not well tolerated in industrialized societies. There is a well recognized pressure, for women especially, to

have a lean body. Thus, Nisbett imagined that some obese individuals would be sufficiently motivated by a desire for thinness that they would successfully restrict their intake to a degree that would reduce their weight well below their biological set point. It was presumed that the reward of possessing a more socially acceptable body would be sufficient to maintain caloric restriction in the face of biological pressures towards weight gain. However, a consequence of weight suppression, and the subsequent tension between biological and psychological mechanisms over eating, was assumed to be an increased sensitivity to the taste, smell and sight of food, which would sometimes lead to episodes of overeating.

1.5.2 *Herman and Polivy's Boundary Model*

Herman and Polivy took Nisbett's work one step further. They argued that normal weight individuals who suppress their weight would show some of the same abnormal eating behaviors that were observed in studies of the obese (Herman & Polivy, 1975). Herman and Polivy developed the Restraint Scale (RS) to identify normal weight individuals who restrict their eating out of a desire to avoid biological pressures toward weight gain (Herman & Polivy, 1984). A series of well known studies showed that restrained eaters who were identified by the RS exhibited counter-regulatory eating in response to a preload (Herman & Polivy, 1980). The typical pre-load study utilized a 2x2 design in which restrained and unrestrained eaters were randomly assigned to consume a high-calorie preload such as a milkshake, or no preload, before participating in a taste test of some highly palatable food, usually consisting of ice cream or cookies. The amount of food consumed during the taste test was surreptitiously recorded. An interaction between preload and restraint status was typically observed (Herman & Polivy, 1980, 1984).

When not given a preload, restrained eaters tended to eat less during the taste than unrestrained eaters during the taste-test. When given a preload, unrestrained eaters would compensate for the extra calories by eating less during the taste test. In contrast, the eating of restrained eaters would become disinhibited, such that they ate more during the taste-test when given a preload than when not. In essence, consuming high-calorie foods was thought to cause the restrained eaters to lose control of their eating, which resulted in overeating.

There are a multitude of variations on the preload paradigm. A high-calorie preload is not the only stimulus that has been shown to provoke disinhibited eating in restrained eaters. For example, the same effect is observed with a small amount of "forbidden" food (Knight & Boland, 1989), manipulations involving negative affect (Herman & Polivy, 1980), and cognitive tasks requiring concentration (Ward & Mann, 2000). In each case, restrained eaters eat less than unrestrained eaters in the absence of an experimental manipulation, but eat more following one.

Herman and Polivy developed the boundary model to explain the eating behaviors of restrained eaters (Herman & Polivy, 1984). The theory suggests that normal eating behavior of unrestrained eaters is controlled primarily by sensations of hunger and satiety. Hunger is thought to be stimulated in response to insufficient caloric intake resulting in energy deficit. Once a sufficient amount of food has been eaten, the body triggers the end of a meal by producing sensations of satiety. In times of relative energy balance, in the absence of sensations of hunger or satiety, eating may be simulated by psychological or environmental factors such as the presence of highly palatable food or other people eating nearby. As can be seen in Figure 1, the "zone of biological

indifference" is the term used to describe situations in which psychological, rather than physiological, factors most influence eating.

The eating behavior of restrained eaters is thought to be governed by a similar, but modified, system. More specifically, restrained eaters are thought to impose an artificial cognitive diet boundary just above their hunger boundary in the zone of biological indifference. This diet boundary represents a cognitive rule governing eating, such as an amount of food or a calorie limit that must not be exceeded. In the absence of a disinhibiting stimulus, a restrained eater eats not to satiety, but to this cognitive diet boundary. Consequently, restrained eaters experience hunger more often than unrestrained eaters, which eventually leads to a habituation to sensations of hunger among restrained eaters. Additionally, restrained eaters' cognitive diet boundary is easily disrupted, as shown by preload studies. Overeating is thought to result from a combination of chronic hunger and disruption of the cognitive diet boundary. Repeated episodes of overeating eventually erode restrained eaters' sensation of satiety so that it, too, becomes less influential. The end result of restrained eating is the widening of the zone of biological indifference. Thus, the effect of external (i.e., environmental and psychological) cues comes to influence eating behavior more strongly than internal cues of hunger and satiety. In this way, a cycle of restrained eating and overeating develops in which each iteration of the cycle increases the likelihood of later excessive energy intake.

1.5.3 *Correlates and Consequences of Restrained Eating*

As mentioned above, research on restrained eating is of special interest because Herman and Polivy have suggested that restraint contributes to weight gain and obesity (Polivy & Herman, 1983) and is a major cause of eating disorders (Polivy & Herman, 1985). At first glance, the empirical evidence seems to support these conclusions (Lowe, 1993). However, as restraint theory has evolved, our understanding of the relationships between restraint and these undesirable outcomes has become more sophisticated.

Given that the RS is purportedly associated with both efforts at caloric restriction and a propensity toward overeating, it is not surprising that researchers have found a variety of relationships with weight and obesity status. Cross-sectional studies have found positive correlations between the RS and measures of weight, BMI, and percentage overweight (for a review, see Lowe & Thomas, in press). The results of prospective studies are more mixed. One study reported that the RS predicted weight gain among adult women, but not men, over a one year period, when the relationship was analyzed in a multiple linear regression including other physiological, demographic, and activity variables (R. C. Klesges, Isbell, & Klesges, 1992). However, the RS failed to prospectively predict changes in body weight in three studies involving college students (R. Klesges, Klem, Epkins, & Klesges, 1991; Lowe et al., 2006; Tiggemann, 1994). Notably, no studies have reported that the RS predicts weight loss.

In regards to the development of disordered eating, Herman and Polivy have suggested that restrained eaters' frequent episodes of overeating may eventually translate into binge eating episodes, which are followed by redoubled efforts at caloric restriction, and the use of radical weight control techniques such as vomiting and laxative use

(Polivy & Herman, 1985). The end result is the familiar bulimic cycle of binge eating, compensatory behavior, and caloric restriction. Prospective studies confirm that high scores on the RS predict the development of binge eating (Stice, Killen, Hayward, & Taylor, 1998) and bulimic pathology (Killen et al., 1994; Killen et al., 1996). In a sample of 967 adolescent girls who where followed over a four-year period, Killen et al. (1994) found that girls who developed bulimic symptoms had greater scores on both the Concern for Dieting (CD) and Weight Fluctuation (WF) subscales of the RS at baseline, compared to girls who remained asymptomatic. In a similar study of 543 female high school students, Stice et al. (1998) reported that RS scores at baseline predicted onset of objective binge eating, subjective binge eating, and purging.

Some researchers have called for a moratorium on dieting because of the evidence linking restrained eating with overweight and eating disorders (e.g., Polivy & Herman, 1984; Brownell & Rodin, 1994). However, evidence has begun to accumulate that casts doubt on the validity of this recommendation. As we shall see, the link between restrained eating and undesirable outcomes such as weight gain and eating disorders is more complex than was previously thought.

1.6 Criticisms of Early Restraint Theory

1.6.1 *The Definition of Restraint*

One of the most problematic aspects of restraint theory is the fact that Herman and Polivy and other researchers treat the terms “restrained eating” and “dieting” as synonymous. Dieting is usually thought of as an attempt to lose weight by eating less food, thereby creating a state of energy deficit in the body. However, the majority of research on the relationship between restraint and actual caloric intake finds no evidence

that restrained eaters consume fewer calories than their unrestrained counterparts. For example, Laessle, Tuschl, Kotthaus, and Pirke (1989) failed to find a correlation between RS and mean caloric intake over a seven day period in a sample of 60 normal weight women ($r = -.04$). Similarly, De Castro (1995) found no relationship between total caloric intake and RS over a seven day period in a sample of 201 male and 157 female adult participants. In a study by French, Jeffery, and Wing (1994), RS score was not related to caloric intake over a sixth month period, as measured by the Block Food Frequency Questionnaire (FFQ; Block et al., 1986).

Notably, Laessle et al. (1989) and De Castro (1995) measured caloric intake via self-report food diaries. However, self-reported dietary intake has poor validity in general (Bandini, Schoeller, Dyr, & Dietz, 1990; Lichtman et al., 1992; Livingstone, Prentice, & Strain, 1990; Prentice et al., 1986), but especially among overweight samples (Lichtman et al. 1992; Prentice et al., 1986) and restrained eaters (Bathalon et al., 2000; Bingham et al., 1995). Both of these groups tend to underreport food intake to a significantly greater degree than unrestrained normal weight individuals. Thus, null findings between the RS and food intake may mask a trend towards greater consumption among restrained eaters.

Laboratory studies of dietary restraint have yielded mixed results. While preload studies typically find the disinhibition effect among restrained eaters, a minority have not (Ouwens, vanStrien, & vanderStaak, 2003). Furthermore, Stice et al. (2004) conducted a series of studies involving the unobtrusive measurement of food intake and found no significant correlations between intake and the RS. Additionally, Stice, Cooper, Schoeller, Tappe and Lowe (in press) conducted studies using doubly-labeled water and a measure of food purchases over a 3 month period, which indicated that restrained eaters

do not consume less than unrestrained eaters. Doubly-labeled water is water in which the hydrogen and oxygen have been replaced with uncommon isotopes of these elements (Schoeller & van Santen, 1982). The metabolic rate (i.e., calories expended per day) of a participant may be measured by administering doubly-labeled water, and then later measuring excretion of deuterium and the O-18 isotope. When combined with objectively measured weight, this test can be used to very accurately measure food intake.

One of the early signs that something major was amiss with restraint theory were studies of the relationship between restraint and dieting. One early study showed that most restrained eaters are not on a diet to lose weight (Lowe, Whitlow, & Bellwoar, 1991). Rather, they express *concern* over their current weight, and a history of weight fluctuation, but are not currently eating less than unrestrained eaters (as described above). Furthermore, Lowe, et al. (1991) conducted a preload study in which restrained eaters were divided into two groups based on whether they reported that they were currently on a diet to lose weight. If the imposition of a cognitive diet boundary was the cause of overeating as Herman and Polivy had suggested, then one would expect the restrained dieters to be most susceptible to disinhibited eating during the taste test. However, the opposite relationship was found. Restrained dieters were protected from disinhibition, whereas non-dieters were not. A similar pattern of results was found by Hetherington and Rolls (1991).

Taken together, studies of naturalistic and laboratory eating behavior seem to indicate that restrained eaters are not eating less than unrestrained eaters - either in the short term or the long term. Additionally, most restrained eaters are not currently on a diet to lose weight, and contrary to the predictions of restraint theory, current dieting is

protective against disinhibited eating. Thus, “restraint” and “dieting” appear to be different constructs associated with different effects on behavior.

1.6.2 *What does the Restraint Scale Measure?*

If the RS is not a measure of hypocaloric dieting, then the question arises as to what *is* it assessing? Researchers have remarked that the RS seems to be measuring several constructs simultaneously. Van Strien (1999) has called the RS a measure of “failed dieting” because it includes questions related to weight concern and overeating. In a series of confirmatory factor analyses, Van Strien et al. (2007) found that a model in which the restraint scale loaded on three factors representing dieting, overeating, and body image concerns was a superior fit to a model in which dietary restraint loaded only on the dieting factor. During the development of the RS, Herman and Polivy identified two subscales called Concern for Dieting and Weight Fluctuation. Exploratory factor analyses tend to support this conceptualization of the RS, but item-level confirmatory factor analyses tend to find that neither a one-factor solution, nor a two-factor solution fit the data particularly well. The results of the confirmatory factor analysis lend further credence to the idea that the RS is multifactorial, and that there is a poor understanding of all the constructs that it measures. Nevertheless, Herman and Polivy maintain that the construct of restraint is best conceptualized as incorporating both efforts towards restriction and periodic overeating. Subsequently, they encourage the use of the total RS score rather than its components.

Despite the influence of Herman and Polivy, there is a growing sentiment that the causes, correlates, and consequences of restraint can only be understood when restraint itself is properly operationalized. Subsequently, researchers have attempted to identify

the constructs of Herman and Polivly's restraint scale that are responsible for its relationship with overweight and eating disordered behavior. The results include new measures of eating behavior and new ways of thinking about risk factors for weight gain and eating disordered behaviors, including overeating.

1.7 New Measures of Restraint and Recent Trends in Restraint Research

1.7.1 *The Development of the Three Factor Eating Questionnaire and the Dutch Eating Behavior Questionnaire*

Stunkard and Messick (1985) were concerned with two aspects of the RS. First, the RS contains at least two sources of variance, which make it difficult to account for the relationships between the RS and its correlates. Of special concern was the observation that the RS appeared to be confounded with overweight because of the weight fluctuation items, which seemed to artificially inflate the RS score of obese persons. Second, they noted that overweight restrained eaters did not necessarily show evidence of disinhibition, as did normal weight restrained eaters. This weakened the link between RS restraint and risk for overweight.

To address these limitations, Stunkard and Messick (1985) created the Three Factor Eating Questionnaire (TFEQ; also known as the Eating Inventory). The TFEQ consists of three psychometrically distinct subscales for Cognitive Restraint, Disinhibition, and Hunger. Van Strien, Frijters, Bergers, and Defares (1986), responding to the same concerns as Stunkard and Messick, created the Dutch Eating Behavior Questionnaire (DEBQ). Their measure also contains three psychometrically distinct subscales, corresponding to Restraint, External Eating, and Emotional Eating. The two restraint subscales of these measures are highly correlated (Lowe & Thomas, in press),

which may be partially explained by the fact that they both borrowed items from Pudel's measure of latent obesity (Pudel, Metzdorrf, & Oetting, 1975).

The restraint scales of the TFEQ (i.e., the TFEQ-R) and the DEBQ (i.e., the DEBQ-R), represent a major improvement in the assessment of restrained eating because they eliminate two confounds - between restraint and overeating, and between restrained eating and overweight - that characterize the RS. In support of this claim are findings (reviewed in Lowe and Thomas, in press) that the TFEQ-R and DEBQ-R have weak or non-existent relationships with the other subscales of the TFEQ and DEBQ that tap different types of excessive eating (stemming from disinhibition, hunger, negative emotions and external food stimuli). Studies on the relationship between the TFEQ-R, DEBQ-R and body size are mixed. Some studies find a positive correlation between body size and these measures of restraint, while others do not (Lowe and Thomas, in press). One might expect to find a correlation between body size and restraint, even in the absence of psychometric confounds, simply because individuals with larger body sizes may be more likely to be motivated to restrain their eating in an attempt to limit or reverse their level of overweight.

The TFEQ-R and DEBQ-R have an advantage that, unlike the RS, many studies of normal weight individuals find that the TFEQ-R and DEBQ-R are negatively related to food intake (Lowe & Thomas, in press). Also, participants in behavioral weight loss trials routine show increases in TFEQ-R as they lose weight (Björvell, Rössner, & Stunkard, 1986; Foster et al., 1998) These findings tend to support the view that the TFEQ-R and DEBQ-R may do a better job of assessing actual caloric restriction than the RS. However, none of the three restraint scales identifies restrained eaters who are in a state of energy

deficit (i.e., successful dieters). Stice et al. (in press, 2004) used doubly labeled water and unobtrusive measures of eating in laboratory and naturalistic settings to show that restrained eaters (as measured by the TFEQ-R and DEBQ-R) consume at least as much as unrestrained eaters. One study showed that individuals scoring highly on the TFEQ-R had a relatively high intake (Laessle et al., 1989). By way of a possible explanation for the failure to find a link between restraint and food consumption, Gorman et al. (1993) found that items in the TFEQ-R related to cognitive aspects of restraint were frequently endorsed, whereas items related to actual behavioral restraint were endorsed at very low rates. Multidimensional scaling revealed that the TFEQ-R could be conceptualized as a continuum, with the frequently endorsed cognitive restraint items at the low end, and the infrequently endorsed behavioral items at the high end. In essence, most restrained eaters apparently think about dieting and would like to restrict their intake, but few take the behavioral steps to make it a reality. Thus, while the TFEQ-R and DEBQ-R are improvements over the RS in that these two scales are related to food intake, none of the three measures can be thought of as a measure of "dieting."

One of the intended uses of "pure" measures of restraint such as the TFEQ-R and DEBQ-R was to determine the degree to which the restraint component of the RS was responsible for its ability to predict overeating. Notably, restrained eaters identified by the TFEQ-R and the DEBQ-R do not generally show disinhibited eating during preload studies (Lowe & Maycock, 1988; Rogers & Hill, 1989; Tuschl, Laessle, Platte, & Pirke, 1990; Westerterp, Nicolson, Boots, Mordant, & Westerterp, 1988; Westerterp-Plantenga, Wouters, & ten Hoor, 1991). More commonly, a tendency towards disinhibited eating (as measured by the Disinhibition subscale of the TFEQ, for example) is a better predictor of

overeating (e.g., Ouwens et al., 2003; van Strien, Cleven, & Schippers, 2000). If the restraint component of the RS were responsible for disinhibited eating, one would assume that the relationship between restraint and disinhibited eating would be stronger for "purer" measures of restraint (i.e., the DEBQ-R and TFEQ-R). In fact, the opposite is true.

There is additional evidence from the newer restraint scales that restraint is not causing overeating. Whereas one study found the RS to be positively associated with binge eating (Wardle, 1980), the TFEQ-R was found to have a weak negative relationship with binge eating in two studies (Lowe & Caputo, 1991; Marcus, Wing, & Lamparski, 1985) If one is willing to assume that the TFEQ-R and DEBQ-R are "pure" measures of restraint, then these findings suggest that components of the RS *other* than restraint are responsible for the overeating of restrained eaters.

1.7.2 *Lowe's Three Factor Model of Dieting*

Given the evidence cited above, Lowe (1993) concluded that the overeating seen in restrained eaters was not the result of cognitive restraint or naturalistic dietary practices. Rather, he suggested that three factors, including Frequency of Dieting and Overeating, Current Dieting, and Weight Suppression, best accounted for the behavior of restrained eaters and dieters. While each factor was thought to influence overeating independently, one of the advantages of the model was its ability to explain complex behaviors though interactions of the three factors.

Frequent episodes of dieting and overeating, which restrained eaters show to a greater degree than unrestrained eaters (Lowe, 1993), were thought to contribute to future overeating eating partly because of the erosion of physiological sensations of hunger and

satiety, and the subsequent reliance on external cues to govern eating behavior. Additionally, evidence suggests that repeated weight loss and regain, as is typical in obese individuals, may result in metabolic changes that increase the likelihood of future weight gain (Lowe, 1993).

Current dieting was thought to protect against overeating, largely because of evidence that dieters are not susceptible to overeating, as are restrained non-dieters. Also, by separating current dieting from a history of dieting and overeating, Lowe (1993) emphasized the important difference between restraint and weight loss dieting.

The study of weight suppression, defined as the difference between an individual's highest weight and his or her current weight, was in its infancy at the time. Early studies (see Lowe, 1993) suggested that weight suppressors were successful dieters who had maintained a weight loss. One study reported that the average length of weight suppression was 20 months. Weight suppressors were thought to have maintained their weight loss because of appetitive changes that protected them from the rewarding aspects of delicious food. This hypothesis was supported by studies finding that weight suppressors had a reduced liking for sweet foods. Whatever the source of their ability to maintain a weight loss, weight suppressors appeared to be largely immune to disinhibition in preload studies.

1.7.3 *Recent Trends in Restraint Research*

In their early research, Herman and Polivy conceptualized restraint as an attempt to resist biological pressures toward overweight (Herman & Polivy, 1975). As the study of eating disorders became more popular in the early 1980s, Herman and Polivy began describing restraint as resulting from an unhealthy desire to attain an unrealistic feminine

ideal for physical appearance, that was propagated in large part by the national media (Polivy & Herman, 1985). This conceptualization of restraint, and the studies linking with restraint to undesirable outcomes such as weight gain, binge eating, and eating disorders, resulted in the recommendation that dieting should be avoided. In his description of his Three Factor Model, Lowe (1993) did not make a point of distinguishing between a desire to become very thin or a desire to avoid overweight as motivations for dieting behavior. However, he notes that his Frequency of Dieting factor is compatible with the idea that repeated attempts at dieting may be related to undesirable outcomes.

Since the early 1980's, the United States and other developed countries have become overwhelmed by an obesity epidemic (Flegal et al., 2002). The number of overweight and obese individuals is far greater than the number of individuals suffering from an eating disorder (Lowe & Levine, 2005). As such, Brownell and Rodin (1994) point out that "[i]t may be important to separate dieting in individuals who are close to normal weight from dieting in those who are heavier. Valid arguments that dieting can be pathological in the former group have been used to discourage treatment for the latter group" (p. 787). Lowe and Levine (2005) argue that if dieting in normal weight individuals is driven by a desire for thinness, then relinquishing dieting would have little or no effect on body weight. However, among overweight individuals, or normal weight individuals who are prone to weight gain, abandoning efforts at caloric restriction would likely result in unhealthy weight gain.

In support of their argument that restraint should be encouraged among individuals prone to unhealthy weight gain, Lowe and Levine (2005) suggest that

restrained eating is not to be blamed for excessive energy intake. Instead, these authors propose that restraint is associated with excessive energy intake because the two phenomenon are the result of some other factor or factors that are causing both restraint *and* overeating. In fact, restraint may simply be a response to the weight gain that usually results from excessive energy intake.

Lowe and Levine (2005) and Lowe and Butryn (2007) suggest that one of the factors contributing to overeating and subsequent restraint is an interaction between a predisposition to consume energy rich foods, and the current food environment. Evolutionary accounts of eating regulation indicate that there are two appetitive systems that stimulate eating. The purpose of both of these systems is to ensure that the body possesses sufficient energy for normal functioning. The homeostatic system responds to energy deficit by stimulating hunger. In an environment replete with easily obtained energy-rich foods, the homeostatic system would be enough to ensure survival. However, humans, like most other animals, did not evolve in such an environment. Rather, early human ancestors likely lived in an environment where obtaining food was difficult, and periods of famine were common. Thus, it was advantageous to anticipate hunger by "stocking up" on high-calorie food when it was available, even in the absence of any energy deficit. The motivation to consume high-calorie foods did not, and does not, come from the homeostatic appetitive system that produces hunger. Rather, there is evidence that a second system, labeled the hedonic system, drives consumption of high-calorie foods by linking the consumption of these foods to basic neural reward systems. In other words, there is reason to believe that the consumption of high-calorie foods is intrinsically rewarding for humans and other animals.

Lowe and Levine (2005) and Lowe and Butryn (2007) review evidence from animal experiments that supports the theory of dual appetitive motivations. For example, there appear to be separate neural peptides associated with the two systems. Orexigenic agent neuropeptide Y (NPY), which is manufactured by rats in a state of energy deficit, significantly increases behaviors aimed at food acquisition. Thus, NPY may be a neurochemical mediator of the homeostatic system. In contrast, opioid peptides appear to be involved in the rewarding aspects of eating. Unlike NPY, opioid peptides decrease during periods of energy deficit. However, opioid peptides are produced in greater quantity in response to a high fat/sucrose diet. Furthermore, NPY increases intake of food generally, whereas opioid peptides selectively increase intake of highly palatable foods. The fact that separate neurochemical mediators exist to drive food in energy deficit, and in the presence of palatable foods, strengthens the argument for dual appetitive systems.

The taste preferences of rats are mostly innate. Thus, the foods that are most rewarding are "hardwired." In humans, taste preferences are largely learned through the association of ingested food and the delivery of metabolized energy during digestion. In other words, humans learn taste preferences for the foods that contain the most energy. Apparently, the association between taste and energy content becomes sufficiently strong, such that taste, and not energy content, eventually becomes the primary reinforcer for eating high-calorie foods. Furthermore, delicious foods keep their motivational properties regardless of current energy balance. These phenomena, in concert with the hedonic appetitive system, seem to be driving much of the eating behavior of humans (Lowe & Butryn, 2007; Lowe & Levine, 2005). For example, consumption of a highly palatable food is greater than consumption of a less palatable food, even when they are equal in

energy density. Some good-tasting foods, such as sugarless gum and diet soda, are repeatedly consumed despite a lack of calories. Perhaps most importantly, we also know that the mere presence of delicious food often results in eating, even during periods of energy excess (e.g., following the consumption of a satisfying meal, in the short-term, and obesity in the long-term).

In the environment of early human ancestors, the hedonic appetitive system would result in occasional excessive energy intake, which was adaptive because it created fat stores that could be called on in times of famine. The current food environment is much different. Instead of a scarcity of high-calorie foods, there is an abundance of foods rich in fat and sugar.

The four changes in the food environment that are most responsible for the obesity epidemic appear to be increases in portion size; energy density; dietary variety; and food availability and palatability. Portion sizes in developed countries generally, and in the United States in particular, have increased dramatically in the last few decades (Smiciklas-Wright, Mitchell, Mickle, Goldman, & Cook, 2003; Hill & Peters, 1998). Increased portion sizes contribute to weight gain because they cause increased food intake: people generally eat a greater volume of food, and subsequently more calories, when given larger portions (Rolls, 2003).

Holding palatability constant, the weight (or, in some studies, volume) of food people eat on a daily basis remains fairly constant regardless of the macronutrient composition or energy density of the food (Yao & Roberts, 2001). Because food weight or volume is a primary determinant of satiety, increased consumption of energy dense foods, such as those high in fat, can promote weight gain (Kral & Rolls, 2004). As energy

density increases, intake volume does not naturally decrease in order to compensate for the added calories, which can lead to passive over-consumption (Blundell & Macdiarmid, 1997).

Meals that are composed of a variety of foods lead to greater energy intake when compared to calorie-matched meals that are composed of only one food type (Raynor & Epstein, 2001). In laboratory studies, when participants are provided with the same meal for several days, they eat less than participants who are given a variety of meals (Meiselman, De Graff, & Lesher, 2000; Zandstra, De Graff, Van Trijp, 2000). Thus the vast array of food choices in supermarkets, restaurants, and homes may be leading to higher energy intake.

Finally, a *variety* of energy dense, palatable foods are omnipresent in the current environment. Palatable food has become an unavoidable stimulus, which is problematic because simply increasing food visibility increases consumption (Wansink, 2004). As reviewed by Yeomans, Blundell, and Leshem (2004), there are many indications that regular exposure to palatable food in and of itself stimulates hunger. Many, if not most, eating situations that people encounter in the current food environment involve combinations of factors that can increase food intake (e.g., restaurant meals involve a wide variety of highly palatable and energy dense foods served in large portions). In sum, many aspects of the current food environment (i.e., the structure, composition and availability of food) appear to contribute to the development of overweight.

When the hedonic appetitive system responds to these changes in the food environment, the result is a nearly constant drive to consume the large portions of delicious and calorie rich foods that are almost always available. Thus, the hedonic

appetitive system, which was a boon to human ancestors, now appears to be a significant contributor to the development of unhealthy overweight. Furthermore, restraint is a desirable characteristic in normal weight individuals, if one assumes that it is usually employed to prevent excessive energy intake and not to create negative energy balance. Given that the majority of humans in developed countries live in essentially identical food environments in regards to portion size, energy density, variety, and availability, then variations in the strength of the hedonic appetitive system will influence the degree to which an individual is susceptible to excessive energy intake leading to overweight.

1.7.4 *The Power of Food Scale*

Lowe et al. (2007) developed the Power of Food Scale (FPS) to measure hedonic hunger. As the PFS was designed to tap a construct thought to contribute to overeating, it is reassuring that the measure correlates with the overeating subscales of the TFEQ (i.e., Disinhibition and Hunger) and DEBQ (i.e., Emotional Eating and External Eating). Furthermore, the PFS is not correlated with "pure" measures of restraint such as the TFEQ-R and DEBQ-R, which suggests that the PFS is not just another redundant measure of restraint. However, the possibility remains that restraint may moderate the expression of the appetitive system that is ostensibly measured by the PFS. In a combined sample of obese and normal weight individuals, PFS scores were significantly higher in the obese group, as would be expected if PFS contributes to the development of overweight (Lowe et al., 2007).

1.8 Tying it All Together: Hypotheses Tested in the Current Study

The theory of hedonic hunger (Lowe & Butryn, 2007; Lowe and Levine, 2005) suggests that (a) individuals with a high degree of appetitive responsiveness are the most

susceptible to excessive energy intake leading to weight gain, (b) restraint is a reaction toward this predisposition to excessive intake, and (c) restraint may interact with appetitive responsiveness to determine food intake. These hypotheses have never been tested in a study of naturalistic eating behavior. Neither has a study been done to compare the predictions of the theory of hedonic hunger (Lowe & Butryn, 2007; Lowe and Levine, 2005) with those of traditional theories of restraint, which suggest that restraint is a direct cause of overeating behavior. The current investigation proposed to accomplish both of these goals. The specific hypotheses of the current study were:

Hypothesis 1a: The PFS will be positively related to the probability that an eating episode will be characterized by greater-than-usual food intake (GTUI) during measurement of eating behavior in a naturalistic setting.

Hypothesis 1 b: During an eating episode, the number of highly palatable foods available will be positively related to the likelihood of greater-than-usual food intake during the measurement of eating behavior in a naturalistic setting.

Hypothesis 1 c: The number of highly palatable foods available will moderate the relationship between PFS and frequency of greater-than-usual food intake, such that the positive relationship between the PFS and likelihood of greater-than-usual food intake will be greater as the number of highly palatable foods available increases.

Hypothesis 2: Eating episodes involving greater-than-usual food intake will be followed by a period of increasing restraint, but will not be preceded by a period of increasing restraint.

Hypothesis 3: The PFS and TFEQ-R will interact to predict the probability that an eating episode will be characterized by greater-than-usual food intake. The conceptualization of

dietary restraint described by Herman and Polivy would predict that participants with high scores on both the PFS and the TFEQ-R would be most likely to experience greater-than-usual intake, whereas the conceptualization of restraint described by Lowe would predict that participants with high scores on the PFS, but low scores on the TFEQ-R, would be most likely to experience greater-than-usual intake. Thus, a specific prediction is not made regarding the nature of the relationships specified in hypothesis 3. Rather, the outcome will be used to test the two competing conceptualizations of restraint.

1.9 Ecological Momentary Assessment

1.9.1 *An Introduction to Ecological Momentary Assessment*

The bulk of the data that was used to test the hypotheses described in the previous section were collected via Ecological Momentary Assessment (EMA). This technique overcomes many of the shortcomings of more traditional methods of data collection by allowing participants to be measured in an ongoing fashion in their natural environment (Smyth et al., 2001). Early forms of EMA were expensive and difficult to implement. However, recent advances in technology have made EMA practical and affordable. Subsequently, EMA is rapidly becoming popular, especially among researchers studying health-related behaviors.

The defining feature of EMA is repeated measurements in the natural setting in the absence of an experimenter (Nazarian, Smyth, & Sliwinski, 2006). Early EMA studies provided participants with pencil-and-paper questionnaires and a cellular pager or preprogrammed timepiece that would signal them several times per day to complete a self-report assessment. More recently, researchers have adapted personal digital assistants (PDAs), such as the ubiquitous PalmPilot®, to both signal participants and record their

responses. This method is particularly advantageous because of its simplicity. It is easy for most participants to carry a PDA on their person and respond to its signals, whereas it might be more difficult to have a packet of questionnaires and a signaling device on one's person at all times.

Stone and Shiffman (1994) originally developed EMA to study patients with chronic pain. The authors were concerned with the reactive effects of the pain diaries that were typically used in research on chronic pain. Reactive effects occurred when the requirement of keeping a pain diary caused patients to attend to their pain perhaps more than they otherwise would, thus increasing the frequency and severity of pain. An early study using EMA revealed that reactive effects were much less prevalent when data were collected via EMA than with traditional pain diaries.

Since Stone and Shiffman (1994), about 60 studies using EMA have been reported in the social science literature. Most of these involve research on psychiatric and health-related outcomes such as smoking cessation, binge eating, heart disease, and bulimic behaviors. The two most represented fields of research in the EMA literature appear to be addictions and eating disorders.

In 2007 alone, there have already been over 15 studies published using EMA. The surge in the EMA technique seems to be driven primarily by the advantages it confers over more traditional types of measurement methods.

One of the biggest concerns with traditional measurement methods is the limitation of retrospective self-report (Nazarian et al., 2006). Typical self-report measures instruct participants to recall events that occurred anywhere from days to years in the past. In addition to simple forgetting, these measures are unreliable because of well

known biases in autobiographical recall. For example, there is a tendency to recall events in accordance with one's preexisting views of behavior and human nature, and for more recent events to influence the recalled content of previous events (Brown & Harris, 1978). Additionally, a person's current mood state has been demonstrated to have an effect on which events are recalled, and the content of those events (Teasdale & Fogerty, 1979). As a final example, recall is heavily influenced by salient events during the period to be remembered (Nazarian et al., 2006). For example, the experience of an exciting and happy event, such as a birthday party, in the previous month will tend to bias memory such that a person will tend to recall being happy during the weeks preceding and following the party, regardless of whether this was actually the case. Measurements collected via EMA are much less susceptible to the biases and limitations of retrospective report because participants need only recall events from the preceding minutes or hours.

Another concern of traditional measurement methods is the limitation of laboratory generalization. Most studies measure participants in a laboratory setting of some sort that is different from the participants' natural environment(s). However, being measured in such an environment, which is by definition artificial, potentially introduces bias into the measurements. There are at least two important reasons for why laboratory findings might not generalize to natural settings. First, the analogue setting may fail to recreate the naturalistic conditions under which an observed behavior normally occurs. Measurements carried out in the lab are not necessarily sensitive to changes in behavior, cognitions, or emotions that result from contact with stimuli in the natural environment. In other words, the laboratory may not capture the "flow" of events that unfolds in the natural environment. Therefore, any conclusions about relationships observed in the

laboratory must be qualified by the unknown influence of the analogue environment on the relationships that are observed.

Additionally, when making observations in the laboratory that rely on aspects of self-report, whether retrospective or contemporaneous, it may be difficult to accurately measure stimulus-response associations which the participant is unaware of. In EMA by contrast, detecting such relationships are only dependent on accurate recording of events themselves – the connections are defined purely in statistical terms. For example, an individual is only able to report that being in the presence of highly palatable food often results in an episode of overeating if they themselves have noticed the link. EMA overcomes the limitation of laboratory generalization by measuring participants in their natural environments. Furthermore, it does not require participants to be aware and report on links between events in their environments and their own reactions, as participants are measured continuously over time.

The fact that EMA data are collected in a continuous fashion confer additional benefits. Where as traditional measurement techniques are often limited in their ability to determine causality because of the scarcity of observations, cross-sectional designs, and the necessity for retrospective self-report, with EMA it is sometimes possible to explore the order of cause and effect for a recurring series of events. A major requirement of inferring causal relations is that an ostensible cause must precede its effect – the ability to temporally order such events allows stronger causal inferences to be drawn, but because the methodology is still correlational in nature one still cannot draw causal inferences. An additional benefit of EMA is the large number of observations that are collected, which increase statistical power, and subsequently the likelihood of detecting relationships if

they exist. Advanced analytic techniques such as hierarchical linear modeling (HLM) are especially well suited to the analysis of the time-series data that are obtained during EMA because of their ability to accurately model relationships over time, even in the face of missing data. The ability to accommodate missing data is important, as participants are not expected to respond to every EMA signal.

1.9.2 *Previous Eating Research Using EMA*

EMA has successfully been used to study eating attitudes and behaviors. One of the first eating-related EMA studies was designed to test the role of situational and mood variables in promoting temptation and lapse in dieting (Carels et al., 2001). Thirty overweight college students who were reportedly on a diet to lose weight were given pencil-and-paper questionnaires, as well as a cellular pager. The participants were instructed to complete an assessment whenever they (a) felt a temptation to go off their diets, (b) experienced a lapse in dieting, or (c) were prompted by the cellular pager to complete an assessment. Generalized estimating equations were used to analyze the data because of their ability to account for random sampling of persons and repeated measures within persons. Recent consumptive activities, presence of others, activities, mood state, and location were reliably associated with reports of temptation and lapse using EMA. So-called negative emotions including stress, sadness, and nervousness were more common during lapses. In contrast, participants reported fewer lapses when feeling "in control," and relaxed. Socializing, eating with others, and watching TV were all more common during episodes of lapse, than during the random signals via the cellular pager. The authors' findings are congruent with other research that found negative affect and situational variables predict lapses in dieting.

Notably, the outcome variable, a lapse in dieting, was defined simply as "An incident where you felt that you broke your diet (e.g., overeat, eat a forbidden food, etc.)." The use of such a definition, which is inherently ambiguous, allows for variation between participants in what they consider to be a lapse. This approach allows the maximum number of lapses to be recorded, as any incident of eating could "count" as a lapse as long as the participant considered it to be one. However, there is a tradeoff in that variation between persons in the definition of a lapse may introduce an added element of error and uncertainty into the findings. Carles et al. (2001) apparently decided that it was most important to capture the range of lapses in dieting, and their detection of multiple relationships may have been related to their employment of user-based definition of lapses.

Carels et al. (2004) conducted a similar follow-up study that included in active intervention based on the LEARN program (Brownell, 2000). EMA was conducted during the last week of the weight loss program. The results suggested that elevations in positive *or* negative affect were contemporaneous with lapses in dieting. However, situational variables such as the location during the lapse (e.g., home versus office) were not significantly associated with lapses in dieting.

Le Grange, Gorin, Catley, and Stone (2001) set out to use EMA to study the correlates of binge eating in overweight (i.e., BMI > 27.3) women. Participants were recruited through a newspaper advertisement seeking overweight women *or* women with a binge eating problem. Potential participants were screened using a battery of self-report questionnaires and a diagnostic interview that included the eating disorders section of the Structured Clinical Interview for the DSM-IV (SCID). Two groups of women were

identified, including those with a diagnosis of Binge Eating Disorder (BED; n = 18), and a control group (n = 17) who denied any episodes of binge eating within the past 6 months, no subjective sense of loss of control during overeating episodes, no purging, and no other behavior that might meet criteria for a diagnosis of Eating Disorder Not Otherwise Specified. Participants were given pencil-and-paper questionnaires to complete at the beginning and end of every episode of eating, as well as when they were randomly signaled by a wristwatch alarm.

As predicted, the BED group was more likely to binge in response to negative emotions such as stress. There was also a general main effect for group such that the BED group was characterized by higher negative affect scores and restraint ratings. However, the most remarkable result of the study was that the two groups did not differ in frequency of binge eating. The mean number of binges per person during the two week monitoring period was 4.5 (SD = 5.45), regardless of group. This is an especially striking finding, given that participants received detailed training regarding the DSM-IV criteria for a binge, and did not simply rely on their own colloquial definition of a binge when completing the EMA assessments. The authors concluded that it is unlikely that the control group suddenly experienced a sudden surge in binge eating due to the EMA procedure. Rather, it is more likely that the control participants, whether knowingly or unknowingly, gave a highly inaccurate account of their eating behavior during the rigorous screening process. Of course, this conclusion assumes that data collected during EMA was more valid than the pre-experimental screening procedures If one were to accept that assumption, then this provides strong evidence of the advantage of using EMA over retrospective self-report.

In a follow-up study, le Grange, Gorin, Dymek, and Stone (2002) compared CBT treatment for BED to CBT plus monitoring of binge episodes via EMA. It was thought that the addition of EMA to traditional cognitive therapy might help the client and therapist identify antecedents to binge episodes, which might prove useful in breaking the binge cycle. However, at post-treatment assessment, there were no statistically significant difference between the two groups on symptom severity.

Similarly to le Grange et al. (2001), Wegner et al. (2002) used an EMA design to study antecedents and consequences of binge eating. This study is the first in the eating literature to describe the use of a PDA for EMA. Twenty-eight female college students who reported binge eating at least once per week, and significant distress over these binge eating episodes, were invited to participate. In addition to baseline measures, EMA data were collected with seven prompts daily for two nonconsecutive weeks. The results revealed significantly higher ratings on scales measuring negative emotions (i.e., Depression, Anger, and Guild/Self-Blame) on days when a binge took place, compared to non-binge days. However, negative emotions did not spike just before or after a binge, which contradicts the theory that binge eating is used as a strategy to cope with negative emotions. Interestingly, in retrospective reports hours later, participants recalled experiencing a significant worsening of mood directly following the binge. The high quality of these temporal data, which have been instrumental in fostering a greater understanding of the antecedents and consequences of binge eating, could not have been collected without EMA.

Stein et al. (2007), using PDAs for EMA, confirmed many of the findings of le Grange et al. (2001). Negative affect was associated with binge eating in a sample of

overweight women with binge eating disorder. Nevertheless, negative affect did not decrease significantly following a binge. Of particular interest was the finding that the "breaking of a food rule" was not associated with binge eating. The authors conclude that this finding contradicts restraint theory, as disinhibition would be expected to occur following an episode of abstinence violation (i.e., a failure of restraint). However, laboratory studies have repeatedly documented that overweight individuals do not show the same patterns of disinhibited eating as normal weight restrained eaters.

Boseck et al. (2007) used EMA to study nocturnal ingestions in a sample of 14 self-described night eaters. Participants completed a morning, evening, and end of day EMA assessment via PDA for about 16 days. Also, participants were instructed to complete an assessment if any night-eating behaviors that occurred. The results confirmed some aspects of night eating that had been reported in non-EMA studies, such as a link between low mood and night eating, and a high level of awareness during night eating episodes. This study also revealed new information about night eating, including the surprising finding that hunger was higher in the morning than during night eating episodes.

Engel et al. (2007) studied the relationships among mood, impulsivity, and bulimic behavior by providing 133 patients who met criteria for bulimia nervosa with PDAs that were used to assess mood, eating, and purging behaviors. Participants were signaled to complete an assessment at six semi-random times during the day. Additionally, participants were instructed to complete an assessment at bedtime and after binge eating or compensatory behavior. The data were analyzed via HLM. Anger was associated with binge eating and vomiting. However, neither the degree of lability in

anger throughout the day, nor the rate of increase in anger over time, was significantly associated with binging or purging. While impulsivity alone did not predict binging or purging, impulsivity interacted with anger in the prediction of binge eating, such that participants with a moderate or high degree of impulsivity were likely to respond to increases in anger by binge eating, whereas participants low on impulsivity were less likely to binge eat in response to increasing anger. Of particular interest was the finding that an episode of vomiting was rarely followed closely by an episode of binge eating. Engel et al. hypothesize that an episode of binging and purging may be followed by a period of increased dietary restraint, which would make binge eating less likely. This study is similar to the current investigation in that it measured antecedents and consequences of episodes of overeating, albeit in an eating disordered sample.

A very similar study by Smyth et al. (2007) yielded results consistent with Engel et al. (2007). In addition to finding that increased anger was associated with bulimic behavior, Smyth et al. reported that anger, negative affect and stress increased prior to a BN event, whereas positive affect decreased leading up to a BN event. Following the event, there were immediate reductions in anger and negative affect, and an increase in positive affect. These findings contrast with those of Wegner et al. (2002) who reported that within-day variability in affect did not predict binge eating. However, Smyth et al. (2007) recruited more than four times as many participants as Wegner et al (2002), and all of the participants in the Smyth et al. study met DSM-IV criteria for BN, where as the participants collected by Wegner et al. were college undergraduates who did not necessarily meet criteria for an eating disorder. The results of Smyth et al. (2007)

highlight the ability EMA to reveal patterns of within-subject variability that may help explain the cause and consequence of undesirable eating behavior.

1.9.3 *EMA in the Present Study*

In the current investigation, participants completed a series of self-report measures at baseline, including a measure of restraint (i.e., the TFEQ-R) and appetitive responsiveness to the food environment (i.e., the PFS), among others. EMA was used to measure eating events, the food environment associated with eating events, and fluctuations in dietary restraint. Scores from the baseline questionnaires and EMA observations were combined in an HLM model to examine the hypotheses described in the previous section. The predictors measured at baseline assessed characteristics of the individual that are thought to be trait-like (e.g., dietary restraint and appetitive responsiveness to the food environment), whereas the measures collected during EMA reflected situational variables that were expected to change across the day (e.g., eating situations). The primary analyses examined how the trait-like variables interacted with the environment to produce behavior.

As is typical, restraint was treated as a trait-like variable and was measured by the TFEQ-R at baseline. However, EMA also measured fluctuations in restraint to explore the temporal relationship with episodes of greater-than-usual intake. While restraint is not typically thought of as a state-like variable, it has been suggested that restraint may fluctuate in response to internal and external events (Fairburn, 2002).

EMA was crucial for the current study, as the hypotheses to be tested pertain to events and associations in temporal sequence that a participant would be hard pressed to recall accurately. Furthermore, the average individual may not even be aware of the

relationships among events that this investigation sought to study. For example, it would be futile to ask a participant if the presence of highly palatable food increased the likelihood of an episode of overeating unless the participant was herself aware of the relationship. On the other hand, EMA captured the sequence of events without a requirement for the participant to ever be aware the relationship between the food environment and overeating. Finally, each EMA assessment was "stamped" with the time and date, which ensures that participants did not, intentionally or unintentionally, misrepresent the time when each assessment was completed.

CHAPTER 2: STUDY DESIGN AND TEST PROCEDURES

2.1 Participants

A sample of 43 undergraduate women were enrolled to participate in the current study. Although overweight affects both men and women, the majority of research on restrained eating has been conducted with women. Furthermore, the constructs under investigation in the current study may perform differently by gender. Thus, only women were included for this initial study, with the anticipation that these constructs will also be studied in men in another investigation at a later date.

Potential participants were recruited via the web-based recruitment system of the Drexel University Psychology Department. During the initial web-based screening, participants were excluded from the study if they report a height and weight outside of the normal weight range (i.e., BMI > 24 or BMI < 18), current dieting, or symptoms of an eating disorder. Overweight participants were excluded from the current study, as the goal was to investigate the development of overweight and not its maintenance. The acceptable upper value for BMI was held slightly lower than the WHO guideline for overweight because of overweight individuals' tendency to underreport their weight (Palta, Prineas, Berman, & Hannan, 1982). Current dieters were allowed to participate. However, secondary analyses were completed in which their data, and the data of non-dieters, was analyzed separately. As described previously, current dieters exhibit patterns of eating behavior that are significantly different than restrained non-dieters. Participants who reported any experience with an eating disorder were given the opportunity to receive referral information for treatment.

2.2 Measures

See Appendix A for the complete text of questionnaire measures used in this study.

Restrained Eating: Dietary restraint is typically assumed to be a trait-like characteristic that does not vary substantially over time. In support of this conceptualization of restraint are studies that find a high test-retest correlation for measures of restraint across weeks and years (Lowe & Thomas, in press). In the current investigation, restraint was assessed before EMA via the Cognitive Restraint scale of Stunkard and Messick's Three Factor Eating Questionnaire (TFEQ-R). Herman and Polivy's restraint scale will also be administered because, historically, it is the most well known and widely used measure of restraint. The TFEQ-R was used as the measure of restraint in primary outcome analyses because it is considered a "pure" measure of dietary restraint, which unlike the RS, is not confounded by tendency toward overeating or fluctuations in weight. The psychometric properties of the RS and TFEQ-R have been studied extensively, and are considered acceptable (Lowe & Thomas, in press). However, because the work of Herman and Polivy, and Lowe, make different predictions about the temporal order of fluctuations in restraint and eating behavior, restraint was also treated as a state-like variable during EMA. Fluctuations in restraint during EMA were measured via four questions depicted in Figure 2, which assessed plans to restrict intake, avoid fattening foods, eat smaller portions, and eat less frequently.

Appetitive Responsiveness to the Food Environment: Lowe et al.'s (2007) Power of Food Scale (PFS) was administered before EMA to measure hedonic appetitive drive. This measure has only recently been developed. However, early reports indicate that it exhibits

acceptable reliability and validity. Lowe et al. (2007) report that the PFS was found to be unifactorial in a variety of samples varying in weight and age. The corrected item-total correlations (i.e., the correlation of the item with the sum of all other items) ranged from .50 to .76. Chronbach's alpha for all 15 items was .94. Four-month test-retest reliability in a sample of female college students (n = 70) was adequate (r =.79, p = .001).

Body Size: Height and weight was measured by the experimenter during the initial meeting with each participant.

EMA: The EMA program was developed and executed on PalmPilot PDAs running the Satellite Forms software application. See Figure 2 for the full EMA path including all questions asked during EMA. The primary outcome variable for most analyses was "eating episodes." Eating episodes was a dichotomous variable, with one category for normal intake, and another category for greater-than-usual intake (GTUI). Greater-than-usual intake was defined in the current investigation as an episode of eating that occurred outside of the normal pattern of eating, or an episode that occurred during the normal pattern of eating, during which a greater-than-usual amount of food was consumed. Eating episodes that did not meet these criteria were considered normal intake. As mentioned above, EMA studies on binge eating frequently allow participants to define for themselves what constitutes an episode of overeating in the absence of objective criteria. This method is used because the amount of food constituting an episode of overeating varies from person to person. Furthermore, leaving the term "overeating" somewhat ambiguous allows participants to capture the greatest number of episodes of overeating.

Similarly, during EMA training the definition of "greater-than-usual intake" was left somewhat ambiguous in order to allow for variation between participants and the largest possible number of observations. However, participants completed a 24-hour food recall with the experimenter before beginning the EMA, so as help the participant establish their "usual" pattern of intake.

Again, similar to studies of binge eating that collect information regarding the environment in which a binge takes place, the current study assessed the availability of palatable foods during episodes of eating. During EMA training, participants were given examples of "good tasting, high-calorie foods" (i.e., pizza, ice cream, potato chips) and then quizzed to ensure their understanding of the term.

2.3 Procedures

The study protocol was submitted for review by the Drexel University Office of Research Regulatory Compliance before any participants were approached for recruitment. Participants were invited to participate via the web-based recruitment system of the Drexel University Psychology Department. The study was described to participants as an investigation of the health behavior of female college undergraduates. The web-based system was used to exclude participants who reported a weight outside of the normal weight range, and who reported a current or historical eating disorder. As mentioned above, participants who reported an eating disorder were given the opportunity to contact the experimenter for referral information for treatment. After answering the initial screening questions, participants were invited to schedule a time to meet with the experimenter to begin the enrollment process.

When meeting with the experimenter, participants first underwent informed consent procedures. Next, they were asked to complete a 24-hour food recall with the experimenter, to establish their "usual" pattern of eating, so as to help them define greater-than-usual intake. This was followed by completion of baseline measures, including the TFEQ-R, the RS, and the PFS. Finally, participants were given instructions regarding the use of the PDA device. Each participant practiced by responding to two practice EMA assessments. Participants were informed that the EMA protocol began immediately after leaving the laboratory. Participants were instructed to not complete entries at any times when they feel unable to reply or if safety is a concern. Instead, participants will be instructed to delay ("snooze") the device and complete the entry as soon as possible. Participants were told that any rating completed within 45 minutes of the beep would be considered "eligible," and that they would be paid $1 for any "eligible" rating made during the EMA. Lastly, before beginning EMA, participants were questioned to ensure that the planned EMA would not occur during a holiday, and that no other unusual events were planned that could cause a systematic disruption in their usual pattern of eating

During the EMA protocol, the PDA emitted a tone to indicate that the participant should complete an assessment. This method is known as signal contingent recording (Wheeler & Reis, 1991). As described in Engel et al. (2007), participants were signaled six times per day at semi-random intervals near the following anchor points: 8:30 a.m., 11:10 a.m., 1:50 p.m., 4:30 p.m., 7:10 p.m., and 9:50 p.m. Signal times were randomly dispersed around these anchor points in a normal distribution with a mean of zero and a standard deviation of 30 minutes. Random signaling prevents participants from altering

their behavior in anticipation of the signal. The EMA protocol lasted for 7-10 days for each participant, depending on when they were able to return the PDA. Given that the first question during the EMA (see Figure 2) asked participants if they had eaten since their last rating, the device was programmed to show the time of their previous rating.

2.4 Data Analysis

In the current investigation, EMA signals were nested within participants, and participants were characterized by their scores on baseline measures of restraint, appetitive responsiveness to the food environment, and body size (i.e., BMI). Such multilevel data are not easily accommodated by traditional applications of the general linear model such as ANOVA and regression. Multilevel techniques including hierarchical linear modeling (HLM) are better able to accommodate the data, and have the additional advantages that (a) participants with partially missing data may be included in the analysis, (b) participants need not be measured at the same time points, and (c) variance is accurately accounted for both within and between individuals. Thus, HLM was used to evaluate the hypotheses in the current investigation. Each of the following analyses was repeated an additional three times to determine whether (a) the addition of BMI as a covariate had any effect on the analysis, and (b) the results differ for dieters and non-dieters, and (c) the addition of caloric intake reported in the 24-hour food recall had an effect on the analysis. None of the analyses were affected by the addition of caloric intake during 24-hour food recall, so these models are not reported below.

2.4.1 *Hypothesis 1 a*

The model for hypothesis 1 a included type of eating episode as a dichotomous outcome (0 = normal intake, 1 = greater-than-usual intake) and PFS score as the

predictor. An episode of greater-than-usual intake was defined by intake outside of the normal pattern of eating ("I usually eat at this time" = No;) or intake that was within the normal pattern of eating, but was distinguished because of the amount of food consumed ("I usually eat at this time" = Yes; "I ate" = "more than usual"). Normal intake was characterized by eating that occurred during the normal pattern of eating and was of a usual or small size ("I usually eat at this time" = Yes; "I ate" = "same as usual" or "less than usual") or an instance when food was available, but was not consumed ("I have eaten since the last time I completed ratings" = No; "I have had the opportunity to eat since the last time I competed a rating" = Yes). The test of hypothesis 1 rested on the statistical significance of the coefficient for PFS on the prediction of the likelihood that an eating episode will be characterized by greater-than-usual intake. PFS was mean-centered prior to the analysis.

2.4.2 *Hypothesis 1 b*

The model for hypothesis 1 b included type of eating episode as a dichotomous outcome (0 = normal intake, 1 = greater-than-usual intake) and the number of highly palatable foods available as the predictor. The test of hypothesis 1 b rested on the statistical significance of the coefficient for number of highly palatable foods available on the prediction of the likelihood that an eating episode will be characterized by greater-than-usual intake.

2.4.3 *Hypothesis 1 c*

The model for hypothesis 1 c included type of eating episode as a dichotomous outcome (0 = normal intake, 1 = greater-than-usual intake), and PFS score, the number of highly palatable foods available, and the interaction between PFS and the number of

highly palatable foods, as the predictors. PFS was mean-centered prior to the analysis. The test of hypothesis 1 b will rest on the statistical significance of the coefficient for the interaction.

2.4.4 *Hypothesis 2*

Two analyses compared the EMA restraint ratings before and after eating episodes which were characterized by greater-than-usual consumption. The first analysis tested for a mean difference in aggregated restraint ratings before the eating episode and aggregated restraint ratings after the episode. However, this analysis would not capture a pattern in which EMA restraint ratings increased preceding the eating event, but decreased at an equivalent rate following the eating event. Thus, the second analysis tested for a difference in the direction and rate of change of EMA restraint ratings before and after the eating episode. This type of analysis was modeled by Engel et al. (2007). Because trajectories in restraint that precede greater-than-normal intake might differ from trajectories following the eating event, pre- and post-event, trends will be modeled separately using piecewise linear and quadratic foundations centered on the eating event. The predictor variables included hours prior to event, (hours prior to event)^2, hours following event, and (hours following the event)^2. Hours prior to event and hours following event will capture the linear trends, and the squares of these terms will capture the quadratic trends. If multiple episodes of greater-than-usual intake are reported in a single day, only the first event will be used to avoid confounding the effects of restraint as a consequence of the first event with restraint as an antecedent to the second event.

2.4.5 *Hypothesis 3*

The model for hypothesis 3 will include type of eating episode as a dichotomous outcome (0 = normal intake, 1 = greater-than-usual intake), and PFS score, TFEQ-R score, and the interaction between PFS and TFEQ-R, as the predictors. PFS and TFEQ-R scores will be mean-centered prior to the analysis. The test of hypothesis 3 will rest on the statistical significance of the coefficient for the interaction.

2.5 Power Analysis

The statistical power of a test of significance (represented as $1 - \beta$) is the likelihood that a relationship will be detected, given that such a relationship exists in the population. Before undertaking a study, an *a priori* power analysis is often conducted to determine the number of participants that should be recruited to ensure that the anticipated relationships observed during the study will be detected as statistically significant. A power analysis for simple applications of the general linear model, such as ANOVA and regression, are relatively straight-forward, and given three components (the tolerance for Type I error - α, the anticipated effect size, and the desired power) it is simple to find the corresponding sample size (n) that should be recruited (Cohen, 1998).

The calculation of an *a priori* power analysis for the multilevel models to be used in the current investigation is substantially more involved. The experimenter is required to specify additional parameters, such as the correlation among longitudinal observations within participants, the number of measurements per person, and the frequency of the event of interest (i.e., in most cases, episodes of greater-than usual intake), which are difficult to estimate without access to a data set that approximates the one to be conducted during the study. Furthermore, the science of power analysis for multilevel

models is in its infancy compared to the techniques that have been developed for simpler applications of the general linear model. This is especially true for applications with a dichotomous outcome (e.g., the prediction of whether an eating episode will be characterized by greater-than-usual intake, or not).

Given that most of the parameters of the current study are unknown, and cannot be reliably estimated because of the lack of a similar dataset from which to approximate the unknown parameters, a projected sample size of 60 was selected for the current study, as this was presumed to be the largest sample could be recruited within the practical limits of the available resources including EMA devices and experimenter time. Furthermore, a sample size of 60 was presumed to provide statistical power of at least .60 for simple applications of the general linear model that approximate the most strenuous test in the current investigation (i.e., the prediction of a dichotomous outcome from the interaction of two continuous predictors, with an medium effect size, $\alpha = .05$, base rate of positive outcome = .25, and a two-tailed test; Hsieh, Bloch, & Larsen, 1998). While a power of .60 may seem low, it was assumed that if each participant responded to 80% of the signals (a conservative estimate obtained from the EMA studies reviewed above) for an average of 8 days of EMA, each participant would provide 40 individual measurements of the outcome variable and the state-like predictors. Thus, given 60 participants, the total number of observations of the outcome would be 2,400. The large number of observations of the outcome increases the likelihood of detecting a significant effect, even if the observations within participants are moderately correlated. Finally, other studies of eating behavior using EMA have obtained significant effects with sample sizes of 30, 42, 28, and 33 (Carels et al., 2001; le Grange et al., 2001; Wegner et al.,

2002; Stein et al., 2007, respectively). Thus, with 60 participants, the current investigation will likely have sufficient power to detect any relationships of theoretical or clinical significance.

Due to restrictions in time, resources, and participant interest, the study ultimately recruited 43 participants. The effect of this reduced sample size on statistical power is addressed in the following sections.

CHAPTER 3: RESULTS

3.1 Descriptives

Participant enrollment began in April, 2008, and ended in December, 2008. Over that 9-month period, 43 participants were enrolled. Of those, 39 completed the EMA protocol without incident, and produced sufficient data for analysis. A further four participants failed to complete the EMA protocol, and consequently their data could not be analyzed. Three of these participants experienced technical failures (one accidentally smashed the EMA device, one removed the batteries from the device, and one experienced an unidentified hardware failure), and one completed only six ratings. The following results are based on the data contributed by the 39 participants who successfully completed the EMA protocol.

The sample was predominately Caucasian (61%), and their mean age was 20.1 (*SD* = 2.0) years-old. Their mean height and weight was 126.5 (*SD* = 14.0) lbs and 64.1 (*SD* = 2.1) inches, respectively. The mean BMI was 21.6 (*SD* = 1.8) kg/m^2. Their scores on the TFEQ cognitive restraint (*M* = 8.5, *SD* = 5.8) and disinhibition (*M* = 6.1, *SD* = 3.1) scales were similar to the scores obtained from normal normal-weight non-dieters (*M* = 6.0, *SD* = 5.5 and *M* = 5.6, *SD* = 4.3 for cognitive restraint and disinhibition, respectively) during the initial TFEQ validation study. The mean total score on the PFS was 2.4 (*SD* = 0.8). The mean subscale scores for PFS Factors 1, 2, and 3 were 2.2 (*SD* = 1.0), 2.7 (*SD* = 1.0), and 2.4 (*SD* = 0.8), respectively. BMI was not correlated with the TFEQ-R (*r* = .182, *p* = .269) or PFS total score (*r* = .177, *p* = .281). The 24-hour food recall conducted during EMA training yielded an average caloric intake of 2071.5 k/cal (*SD* = 773.5), with

an average fat consumption of 78.6 g (SD = 44.1), which represents 33% of all calories consumed.

Eight participants stated that they were currently dieting. Due to the possibility that current dieters may behave differently than non-dieters, all outcome analyses were conducted once with the full sample, and again with current dieters segregated from non-dieters. In an effort towards brevity, the results from the separate dieter/non-dieter analyses will only be reported in instances where they differ from the results obtained when using the complete sample.

Each participant carried the EMA device for an average of 7.8 (*SD* = 1.7) days, which gave them the opportunity to respond to an average of 47 (*SD* = 10.2) prompts during the EMA protocol. Compliance with the EMA protocol was fair, with participants responding to an average of 67.0 % (*SD* = 15.0%) of prompted beeps within 45 minutes, for an average of 31.3 (*SD* = 9.4) "eligible" ratings per participant. The average latency between the beep and the participant's response, for eligible ratings, was 8.6 (*SD* = 11.4) minutes. Participants also completed an average of 8.7 (*SD* = 5.7) "ineligible" ratings. Ratings were categorized as ineligible if they, (a) were completed more than 45 minutes after a beep, or (b) constituted a response to a beep for which a rating had already been completed (e.g., a participant completed a rating 10 minutes *and* 30 minutes after a singe beep). In the case that two or more ratings were completed after a single beep, only the first rating after the beep was considered eligible for analysis. A total of 1221 eligible and 338 ineligible ratings were completed by the 39 participants. Compliance did not appear to be affected by time of day, as an equal number of eligible ratings (*n* = 611) were completed before and after the average beep time (i.e., 3:10 PM).

During each rating, participants were asked whether they had eaten since the previous rating, and if not, whether they had had an opportunity to eat. The mean number of eating episodes and eating opportunities per participant was 18.2 (*SD* = 6.2) and 5.7 (*SD* = 5.3), respectively. The average number of eating episodes per participant that were characterized by GTUI was 5.4 (*SD* = 3.6). Likewise, of the 1221 total eligible ratings, 709 (58.0%) described an eating episode, 221 (18.1%) described an opportunity to eat, and 292 (23.9%) described no eating episode or opportunity to eat. Of the 709 episodes involving an eating episode, 211 (29.8%) were characterized by greater than usual intake.

The four EMA restraint items displayed good reliability (*Cronbach's Alpha* = .927). These items were averaged within each rating to create an EMA restraint total score. The average EMA restraint total score was 5.60 (*SD* = 1.21). Given that a score of 5 represents a level of restraint that is "same as usual," the average EMA restraint total score observed in this study suggests that participants were somewhat more restraint than usual during the EMA protocol. Participants' average EMA restraint total score was significantly correlated with their TFEQ-R scores ($r = .605$, $p < .001$).

3.2 Reactive Effects

Reactive effects occur when participants change their behavior in response to the assessment procedures. A series of simple tests were conducted to evaluate for reactive effects during EMA. A dummy coded variable was created with a score of 0 for ratings that occurred during the first four days of a participants EMA, and a score of 1 for ratings that occurred on the fifth day or later. This variable was used in three separate HLM models to predict GTUI, the number of delicious, high-calorie foods in the environment, and EMA restraint. Day (≤ 4 vs. > 4) did not predict the probability of GTUI ($p =$

.343),or the number of delicious, high-calorie foods in the environment (p = .908). These findings suggest that participants' report of GTUI and the food environment was not influenced by the EMA protocol. However, day (≤ 4 vs. > 4) was a borderline significant predictor of EMA restraint (*coefficient* = 0.124857, *SE* = 0.062305, *t-ratio* = 2.004, p = 0.052), with a trend towards decreasing restraint during the latter phase of the EMA.

3.3 Missing Data

Hierarchical linear models (HLM) were the primary method of analysis used in this study. One advantage of HLM over simpler methods of analysis is the ability to include participants with missing data in the analysis (Hedeker & Gibbons, 1997). This is a critical advantage when analyzing an EMA, as it is unlikely that any participant will respond to 100% of the prompts. However, the ability to include participants with missing data does not come without cost. There is a possibility that the missing observations are systematically different in some way than the observations that are not missing. In other words, missingness may be associated with the value of the outcome variable. For example, participants may have been less likely to complete a rating following an eating episode that was characterized by GTUI.

When the probability of missingness is associated with the value of the outcome variable, the data are said to be "missing not at random" (MNAR) or "non-ignorable missing." MNAR may bias estimators, resulting in invalid conclusions (Hedeker & Gibbons, 1997). HLM assumes that missing data are "missing completely at random" (MCAR; the outcome is not related to any variable in the analysis), or "missing at random" (MAR; the outcome is associated with one or more predictor variables). Under conditions of MCAR and MAR, estimators remain unbiased.

Pattern-mixture modeling is an approach that is sometimes used to evaluate the effect of missing data within an HLM analysis (Hedeker & Gibbons, 1997). This technique involves two general steps. First, participants are categorized by their pattern of missing data, which is represented in the dataset by one or more dummy coded variables. Second, these variable(s) are added to the outcome analyses, both as main effects and interactions with the predictors. If these missing data terms are significant in any analysis, there is cause for concern that missingness is MNAR, and that the results may be biased by the pattern of missingness. However, it should be noted that pattern-mixture modeling will not detect all cases of MNAR. Therefore, an absence of a significant result is not a guarantee that MNAR is not present.

Each of the following analyses was tested for MNAR using pattern-mixture modeling. A dummy coded variable was created with a value of 0 for participants with ≤ 25% missing data (n = 15), and a value of 1 for participants with > 25% missing data (n = 24). Each of the following analyses was repeated to include this variable as a main effect and an interaction term (with each predictor). In none of the analyses were any of these terms statistically significant (smallest p = .109), and so the results of these analyses are not described in any further detail. These findings can be interpreted as support for the conclusion that the following results are not biased by the pattern of missing data. However, they do not represent conclusive proof that MNAR was not present.

3.4 Hypothesis 1

It was hypothesized that the likelihood of GTUI during an eating episode could be predicted from a participant's PFS score, the number of highly palatable foods available during the eating episode, and the interaction between these two variables. A series of

HLM analyses was used to evaluate these hypotheses. For each analysis, the type of eating episode (GTUI = 1, non-GTUI = 0) was used as the dependent variable. A Bernoulli distribution was used to model the probability of this dichotomous outcome. Eating episodes (represented at level-1) were nested within participant (represented at level-2).

As a first step, the unconditional means model was generated to describe the overall probability of GTUI, and to determine whether the probability of GTUI was characterized by significant inter-participant variability which might be predicted by level-2 covariates. Participants were treated as a random effect. The results are illustrated in Table 1. Across participants, and accounting for non-independence within participants, the average probability that an eating episode would be characterized by GTUI was estimated at 29%. This estimate was reliable (reliability = 0.554), and there was significant inter-participant variability in the probability of GTUI.

In the second step, the PFS 15-item total score was added to the unconditional model at level-2. This covariate was mean centered prior to analysis. The results are illustrated in Table 2. The PFS was not found to predict the probability of GTUI. The intercept remained reliable (reliability = .551), and significant inter-participant variability remained in the probability of GTUI. A similar pattern of results was obtained when the analysis was repeated using only non-dieters. However, when this analyses was repeated for the subsample of eight current dieters, PFS was found to predict the probability of GTUI (Table 2). Interestingly, higher PFS scores predicted a lower likelihood of GTUI. In this analysis, significant inter-participant variability was not found in the probability of GTUI, and the intercept was less reliable (reliability = .551). The lack of significant inter-

participant variability in GTUI indicates that, after accounting for the effect of predictors, there are no significant individual differences that contribute to the probability of GTUI. The lower reliability of the intercept indicates lower confidence in the accuracy of the estimated intercept. The lack of inter-participant variability in GTUI, and the less reliable intercept, are both likely the result of the small sample size of current dieters.

In order to confirm the finding that the PFS is a predictor of GTUI for current dieters, but not non-dieters, a model was created that included all 39 participants, in which GTUI was predicted from PFS score, dieter status (dummy coded; 0 = non-dieter, 1 = current dieter), and the interaction of these two variables. The results are illustrated in Table 4, and confirm that PFS has an inverse relationship with the probability of GTUI for dieters. Conversely, the PFS has no reliable relationship with the probability of GTUI for non-dieters.

In the third step, the number of highly palatable foods available during the eating episode was added to the unconditional model at level-1. This variable was not treated as a random effect, as preliminary analyses indicated the absence of significant inter-participant variability. The PFS 15-item total score was not included as a predictor in this model. The results of this analysis are illustrated in Table 3. The number of palatable foods available during the eating episode did not predict the probability of GTUI. The intercept remained reliable (reliability = .565), and significant inter-participant variability remained in the probability of GTUI. A similar pattern of results was obtained when the analysis was repeated using only non-dieters. However, when this analyses was repeated for the subsample of eight current dieters, PFS was found to predict the probability of GTUI (Table 3). As the number of highly palatable foods available increased, the

likelihood of GTUI also increased. Significant inter-participant variability was not found in the probability of GTUI, and the intercept was reliable (reliability = .591).

In the fourth step, the PFS 15-item total score and the number of highly palatable foods available during the eating episode were added simultaneously to the unconditional model at levels 2 and 1, respectively. The PFS was mean centered prior to analysis. The results are illustrated in Table 4. Neither PFS nor the number of highly palatable foods available predicted GTUI. Furthermore, the interaction between these two variables was non-signficant. The intercept remained reliable (reliability = .559), and significant inter-participant variability remained in the probability of GTUI.

As a fifth step, the above analyses were repeated with BMI as a covariate. BMI was mean centered prior to analysis. The pattern of findings was completely consistent when BMI was added, with one exception. When BMI was added to the model containing the number of highly palatable foods available, it became evident that there was a significant interaction between these two variables (see Table 5). The specific nature of the interaction is depicted in Figure 3. It appears that participants become more likely to experience GTUI in response to the presence of highly palatable foods, as BMI increases. Lean participants appear to have a relatively low probability of GTUI regardless of the number of highly palatable foods available. In contrast, heavier participants appears to have an even lower probability of GTUI when palatable foods are *not* available, but they experience a much higher probability of GTUI when highly palatable foods *are* available. This specific finding did not reach statistical significance in the subsamples of current dieters and non-dieters, but the trend was in the same direction

in both cases. This lack of statistical significance is likely the product of the reduced sample sizes for the subsample analyses.

3.5 Hypothesis 2

It was hypothesized that the mean EMA restraint total score would be significantly greater after an episode of GTUI than before GTUI. A two-tailed independent samples t-test supported this hypothesis. The mean restraint score after GTUI (M = 5.742, SE = 0.189) was significantly higher than the mean restraint score before GTUI (M = 5.618, SE = 0.191; t (686) = 3.008, p = .003). This relationship was not due to a general trend towards increased restraint scores later in the day, as the same pattern of results was found when the time of day of the rating was included in the analysis. Furthermore, there was no relationship between restraint and time of day.

An additional analysis was conducted via HLM to determine whether a significant trend in the EMA restraint total score could be observed across days in which GTUI occurred. For these analyses, the EMA restraint total score was used as the dependent variable. If multiple episodes of greater-than-usual intake occurred within in a single day, only the first event was used to avoid confounding the effects of restraint as a consequence of the first event with restraint as an antecedent to the second event. The time of GTUI was set at 0, and the time of all ratings during the same day as the GTUI were coded based on the number of hours before or after GTUI. Linear and quadratic trends in restraint over time were evaluated. Intercepts were treated as random. Time slopes were treated as fixed, as preliminary analyses revealed non-significant variability between participants. A dummy coded variable representing pre- and post- GTUI was added to these models to determine whether the trajectory of restraint before and after

GTUI differed. Three participants were excluded from this analysis because of insufficient episodes of GTUI.

The dummy coded variable representing pre- and post- GTUI was non-significant in all of the models that were evaluated. The best-fitting model was one in which a quadratic trend was used to describe the pattern of restraint across days in which GTUI occurred (see Table 6). It appears that restraint begins to increase about 8 hours prior to an episode of GTUI, and continues to accelerate for an additional 8 hours afterward (see Figure 4).

As a final step, the above analyses were repeated with BMI added as a covariate. This addition produced no change in the pattern of results. Therefore, these additional analyses are not described in further detail.

3.6 Hypothesis 3

It was hypothesized that the likelihood of GTUI during an eating episode could be predicted from the interaction between the PFS and the TFEQ-R. The analytic approach and unconditional model described in the testing of Hypothesis 1 was used as the foundation for the following analyses.

As a first step, the TFEQ-R was added to the unconditional model at level-2. This covariate was mean centered prior to analysis. The results are illustrated in Table 7. The TFEQ-R was not found to predict the probability of GTUI. The intercept remained reliable (reliability = .535), and significant inter-participant variability remained in the probability of GTUI.

As a second step, the PFS 15-item total score, the TFEQ-R, and the interaction between these two variables was added to the unconditional model at level-2. Both the

PFS 15-item total score and the TFEQ-R were mean centered prior to analysis. The results are illustrated in Table 8. There was no significant interaction between the PFS and the TFEQ-R in the prediction of the probability of GTUI. The intercept remained reliable (reliability = .511), and significant inter-participant variability remained in the probability of GTUI.

As a final step, the above analyses were repeated with BMI added as a covariate. This addition produced no change in the pattern of results. Therefore, these additional analyses are not described in further detail.

CHAPTER 4: DISCUSSION

To our knowledge, this is the first study to use data collected via palmtop computer to explore associations between participant characteristics and momentary environmental variables in the prediction of eating behavior that may cause weigh gain leading to eventual overweight. Questionnaires were used to assess participants' sensitivity to the food environment (via the PFS) and their tendency to restrict their food intake via cognitive rules (via the TFEQ-R). EMA was used to collect information about participants' eating behavior, the food environment, and dietary restraint, six times daily. This information was used to evaluate the theory of hedonic hunger (Lowe & Butryn, 2007; Lowe & Levine, 2005), which suggests that, (a) individuals with a high degree of appetitive responsiveness are the most susceptible to excessive energy intake leading to weight gain, (b) that restraint is a reaction toward this predisposition to excessive intake, and (c) that restraint may interact with appetitive responsiveness to determine food intake

4.1 Compliance with EMA

All of our primary outcome analyses used data collected via EMA. Therefore, to be confident in our conclusions, it is important that the EMA protocol was completed with high fidelity. However, participants were only moderately compliant with the EMA protocol. The response rate of 67% observed in this study is less than reported in other studies using EMA, which often achieve rates in excess of 80%. There are several potential explanations for our low response rate. First, most prior EMA studies have used adults and/or a medically or psychologically disordered population. Compared to these groups, undergraduates may simply be less motivated to comply with an EMA-type protocol that requires regular responding. Second, it is possible that the EMA program

used in this study was not sufficiently flexible to accommodate the irregular lifestyle of undergraduates. Specifically, it is conceivable that undergraduates who sleep late into the morning may have missed a majority of the morning prompts. However, our analyses were able to eliminate time of day effects as an explanation for incomplete responding. Participants were equally likely to respond to an EMA prompt in the first and second halves of the day. Nevertheless, the possibility remains that undergraduates may experience more difficulty completing an EMA because of the unstructured nature of their lifestyles (Rowland & Wesselhoft, 1998). The third and most likely explanation for the suboptimal compliance seen in this study is participant fatigue. It is possible that adherence to the protocol would have been better if the total duration of the EMA protocol had been shorter. One other EMA which used college students for participants reported an 83-87% response rate with a protocol that lasted only 5 days (Leahey & Crowther, 2008). No statistically significant trend in compliance was observed across the 7-10 days of participation in the current study. However, anecdotal participant reports suggest that adherence was best during the first and last days of the EMA. Thus, EMAs with college students may obtain optimal response rates by shortening the total duration.

4.2 Excess Dietary Intake in College Undergraduates

Our primary outcome measure for this study was self-reported episodes of Greater Than Usual Intake (GTUI). GTUI was defined subjectively as greater than usual food consumption during a regularly scheduled meal, or any food intake that occurred out of a participant's usual pattern of eating.

As was described in greater detail in previous sections, a decision was made to leave the definition of GTUI somewhat ambiguous in order to capture the largest

percentage of overeating episodes. This approach also avoided incurring excessive participant burden inherent to more accurate measures of food intake (e.g., food diaries) that may have negatively affected adherence. Despite a training session to orient participants to the concept of GTUI in relation to their pattern of "usual" intake, there are several potential concerns associated with this study's measure of GTUI. Given these concerns, caution should be used when drawing conclusions from analyses incorporating the GTUI variable.

While this study assumed that GTUI represented excessive dietary intake that could lead to eventual weight gain, this assumption was not empirically tested. The reliability and validity of participants' estimates of their intake, and whether they are able to accurately determine when they have consumed "more than usual," is unknown. Furthermore, the assumption that GTUI causes a positive energy balance only holds true if participants did not compensate for GTUI by restricting their intake before or after GTUI, and did not engage in any other compensatory behavior such as additional physical activity.

Another potential problem with our measure of GTUI is that "usual" intake likely varies from person to person. Likewise, what is considered "greater than usual" will also vary. Therefore, an amount of food (or number of calories) that "usual" for one participant may be "greater than usual" for another participant. Of further concern is the possibility that participants' perception of what is "usual" and "greater than usual" may be related to their scores on the PFS and TFEQ-R. Such an association would introduce an unmeasured confound into the study, which we currently have no way of assessing.

Allowing for the caveats described above, GTUI during EMA constituted about 29% of eating episodes, making it quite common. This suggests that even normal-weight college women are quite prone to patterns of eating that may be associated with weight gain. It is unknown whether this finding can be generalized to other populations. However, this finding is particularly notable, as no other study using real-time data collection in the natural environment has reported on rates of excess intake in undergraduate women.

4.3 Predictors of GTUI

4.3.1 *Sensitivity to the Food Environment and the Number of Highly Palatable Foods Present*

The theory of hedonic hunger (Lowe & Butryn, 2007; Lowe & Levine, 2005) suggests that individuals with a high degree of appetitive responsiveness are the most susceptible to excessive energy intake leading to weight gain, and that this appetitive responsiveness would be activated in the presence of delicious, high calorie, foods. Therefore, our first set of hypotheses concerned participants' sensitivity to the food environment (measured via the PFS) and the number of high calorie palatable foods in the environment during an eating episode, as predictors of GTUI. It was predicted that participants who exhibited increased sensitivity to the food environment would be at greatest risk for GTUI when confronted with an eating episode involving an abundance of high calorie, highly palatable, foods. This hypothesis was not supported in the full sample, or the subsample of non-dieters. Neither participants' sensitivity to the food environment, nor the number of highly palatable foods available, nor the interaction between the two, predicted the probability of GTUI during an eating episode. However,

the number of highly palatable foods in the environment was a borderline-significant predictor, such that the more highly palatable foods available, the greater probability of GTUI.

A different pattern of results emerged when these analysis were repeated for the subsample of eight current dieters. Despite the very small sample size, sensitivity to the food environment was found to be a significant predictor of the probability of GTUI. Surprisingly, greater sensitivity to the food environment predicted a *lower* probability of GTUI. Additionally, the number of highly palatable foods present during the eating episode was a statistically significant predictor of GTUI in this subsample. In this case, the effect was in the expected direction; the probability of GTUI increased as the number of highly palatable foods present increased. Again, however, there was no interaction between sensitivity to the food environment and the number of highly palatable foods available in the prediction of the probability of GTUI.

Insufficient statistical power is a likely explanation for the lack of a relationship between the number of highly palatable foods present and GTUI in the full sample. Assuming an unchanged effect size, the addition of 50 participants would likely have produced a significant effect for this variable. If this allowance is made, our hypothesis that GTUI would be more likely in the presence of highly palatable foods was supported. This finding coincides with lab studies of external cues over eating, which have found that the more highly palatable foods present, the more food is consumed (Rolls et al., 1981).

These findings appear to support one aspect of the theory of hedonic hunger. That is, if it can be assumed that participants were not energy deprived when they encountered

highly palatable foods, then their tendency to eat more when more palatable food was available would presumably be based on the reward value, rather than the energy value, of the palatable foods. Our support for the theory of hedonic hunger would have been stronger if we were able to show that such hedonically motivated eating occurred even when participants were already stated. The current study did not collect information about feelings of hunger or satiety during eating episodes, but this would be an excellent avenue for future research.

It is especially interesting that the effect of highly palatable foods on the probability of GTUI was detected in dieters, but not non-dieters, particularly in light of the small sample size of dieters. The two explanations for this phenomenon that seem most likely are that, (a) dieting increases sensitivity to the effect of palatable foods, or (b) the individuals who are most susceptible to the effects of palatable food in the environment are also the individuals who are most likely to diet. The former seems unlikely because previous studies have found that being on a diet protects against overeating (Lowe et al., 1991; Lowe 1993). There are several ways to approach the latter possibility, including looking for evidence of a stronger hedonic appetitive system in dieters (i.e., higher PFS scores). However, the most promising possibility is that individuals whose eating is highly motivated by palatable food in the environment may be especially prone to weight gain, and consequently need to diet to control their weight. Our findings on the relationship between the PFS and GTUI lend credence to this argument.

At first, it seems counterintuitive that higher PFS scores would predict a lower probability of GTUI in current dieters. However, if the purpose of the diet is, at least in

part, to resist a sensitivity to the food environment, *and* those dieters with the highest PFS scores are working hardest to avoid eating in response to the sensory qualities of food, *then* we would expect that dieters with the *highest* PFS scores would show the *least* GTUI. Again, this is consistent with previous findings that dieting protects against overeating (Lowe et al., 1991; Lowe, 1993).

The question still remains as to why the PFS predicts GTUI in current dieters but not in non-dieters. As mentioned previously, statistical power is always a concern. However, even if 1,000 additional participants were added to the study, the relationship between PFS and GTUI would not likely reach statistical significance in non-dieters. Thus, insufficient power is not likely responsible for our null findings. There are several plausible alternatives. The following paragraphs address those that seem most likely.

First, it's conceivable that PFS scores predicted GTUI only in dieters because dieters tend to have higher PFS scores, and PFS only affects eating behavior once some high threshold has been met. In that case, one would expect current dieters to exhibit significantly higher PFS scores than non-dieters. This was not the case ($p = .864$).

Second, sensitivity to the food environment may fluctuate throughout the day. Perhaps the *variability* in sensitivity to the food environment predicts GTUI, but the absolute value of sensitivity to food environment at a one-time measurement does not. In this case, for the PFS to predict GTUI, it would have to be administered concurrently with eating episodes. This is certainly an avenue for future research.

Third, the PFS may be an imperfect measure of sensitivity to the food environment. This is possible despite some evidence for an association between the PFS and BMI (Cappelleri, Bushmakin, Gerber, Leidy, Sexton, et al., (in press). Laboratory

studies incorporating empirical assessment of actual eating behaviors are needed to determine whether the PFS predicts food intake, as all previously published studies using the PFS have been restricted to correlations between the PFS and other questionnaire measures of eating constructs. In the event that the PFS is not an adequate predictor of sensitivity to the food environment, additional measure development research may be needed.

Fourth, the PFS may not capture the unknown mechanism which causes overweight individuals – or those prone to overweight - to overeat. The third and fourth possibilities are both supported by our finding that BMI interacted with the number of highly palatable foods present to predict the probability of GTUI in the full sample. Evidently, *some* type of trait-like characteristic is causing heavier individuals to reliably eat more than usual in the presence of delicious, high calorie food. It was assumed that this characteristic is a sensitivity to the food environment, as measured by the PFS. However, if this were true, then we would expect that PFS would interact with the number of delicious, high calorie foods available to predict GTUI. Also, we might expect PFS to correlate with BMI. Again, neither of these was the case, which leads to the conclusion that either (a) the PFS is not an adequate measure of a sensitivity to the food environment, or (b) a sensitivity to the food environment is not the personal characteristic that accounts for the relationship between BMI and the number of delicious, high-calorie food available in the prediction of GTUI. Unfortunately, the results of the current study are not sufficient to differentiate between these two possibilities. Future research will be needed to refine our measures of hedonic hunger, and to investigate other constructs that

may help to explain why the eating behavior of heavier individuals is more susceptible to food in the environment.

The interaction between BMI and highly palatable foods in the prediction of GTUI is one of this study's most interesting findings. Heavier individuals were *less* likely to experience GTUI than leaner individuals in the absence of delicious, high-calorie foods, but were *more* likely to experience GTUI when such foods were present. This is very reminiscent of Schachter's internal-external theory of obesity, which posits that the eating behavior of obese individuals is driven more by external factors, such as the availability of delicious food, than internal factors, such as feelings of hunger and satiety. While none of the participants in the current study were obese, or even overweight, it has been shown that a higher BMI during adolescence is a significant risk factor for the development of overweight and obesity (Deckelbaum & Williams, 2001). In fact, the detection of the relationship between BMI and delicious, high-calorie foods in a normal-weight sample leads credence to the idea that externally-motivated eating may play a causal role in the eventual development of overweight. As noted above, it will be important to determine which physiological and/or psychological mechanisms are responsible for externally driven eating, as our current measures are less than perfect.

In summary, the number of highly palatable foods present appears to encourage food intake that is greater than usual among normal weight women, and especially those at the higher end of normal weight. This constitutes partial support for the theory of dual appetitive systems. A sensitivity to the food environment, as measured by the PFS, predicted GTUI only in current dieters, and in the inverse direction. This seems plausible only if dieters are working particularly hard to resist their own propensity to overeat

when delicious, high calorie, foods are present. Additional research is needed to determine whether: (a) sensitivity to the food environment fluctuates throughout the day, (b) our measures can be refined to better assess a sensitivity to the food environment, and (c) we can identify other constructs which explain why heavier individuals eat more in the presence of highly palatable food.

4.3.2 *Restraint as a Cause or Consequence of Greater Than Usual Intake*

The theory of hedonic hunger described by Lowe and Levine (2005) suggests that restraint is an appropriate response to excessive food intake, and is not itself a cause of excess intake. This distinction is important, as some critics have suggested that cognitive control over eating should be discouraged because it can help cause binge eating. The current study measured dietary restraint throughout the day in an effort to determine whether fluctuations in restraint are associated with GTUI.

As predicted, on days in which GTUI occurred, restraint was significantly higher after an episode of GTUI than before. While this finding supports the role of GTUI as a causal factor, rather than a consequence, of restraint, it does not constitute proof that GTUI causes restraint to increase. It also fails to describe a specific trajectory of restraint on days in which GTUI occurred.

A second and more detailed analysis was conducted to examine trajectories in restraint before and after GTUI. We observed a pattern such that restraint began to slowly increase about six hours before GTUI. As an episode of GTUI neared, the increase in restraint began to accelerate. The peak in restraint was reached between 6 and 8 hours after GTUI, at which time the level of restraint began to return to normal.

This pattern of restraint is open to two types of interpretations. The increase in restraint prior to GTUI could be seen as a causative factor in the onset of GTUI. In that case, we would expect restraint to reach, or at least approach, a peak *before* GTUI occurred. However, the peak in restraint clearly occurs several hours after the GTUI has concluded. This suggests that the modest increase in restraint that is seen in anticipation of GTUI represents an awareness that GTUI is coming, *and* an effort to prevent it. In fact, it is possible that instances in which restraint accelerates more quickly *before* a possible episode of GTUI may in fact *prevent* the GTUI from occurring. This study was not able to test this hypothesis, as we did not assess for "near misses" of GTUI.

Regardless of whether the pattern of restraint observed in this study is better interpreted as a cause or consequence of GTUI, it is not consistent with Herman & Polivy's (1984) Boundary Model, which drives their recommendation that restraint should be discouraged because of its association with disinhibited eating and binge eating. GTUI did not reliably occur after levels of markedly increased restraint. Interestingly, there is another pattern of restraint that could be consistent with the Boundary Model. Namely, one may conclude that the Boundary Model would predict a sharp *decrease* in restraint just prior to GTUI, as it is the violation of the "cognitive diet boundary," rather than restraint itself, which is ultimately responsible for overeating. However, no such trend was observed. Restraint reliability increased from shortly before to well after GTUI.

In summary, restraint was shown to build to modest levels prior to GTUI, but the peak in restraint was not reached until several hours after GTUI. Although this pattern could be interpreted to mean that increases in restraint cause GTUI, it could also be

interpreted to mean that GTUI could be avoided by facilitating faster acceleration of restraint *prior* to GTUI. If the increase in restraint prior to GTUI that was observed in this study represents an anticipation of overeating, it seems plausible that novel interventions could be developed to enhance this "early response system" to make it more effective in preventing GTUI, and thereby reducing overall caloric intake in an effort to prevent weight gain.

4.3.3 *Dietary Restraint and Sensitivity to the Food Environment*

The theory of hedonic hunger described by Lowe and Levine (2005) suggests that dietary restraint may interact with appetitive responsiveness to determine food intake. Therefore, participants' sensitivity to the food environment was allowed to interact with their self-reported dietary restraint (both were measured via questionnaire at the start of the study) to predict the likelihood of GTUI. Neither participants' sensitivity to the food environment, nor their dietary restraint, nor the interaction between the two, predicted the probability of GTUI during an eating episode. However, the non-significant trend was such that restraint did not impact the probability of GTUI at low levels of PFS, but at higher levels of PFS, there was an inverse relationship between restraint and the probability of GTUI. Specifically, as PFS increased, TFEQ-R became more protective against GTUI.

Though the relationship between sensitivity to the food environment and dietary restraint was non-significant in the prediction of GTUI, the trend was in the hypothesized direction. Namely, high dietary restraint may protect against overeating in individuals who are particularly susceptible to eating in response to highly palatable foods. The trend was similar regardless of whether current dieters were included or excluded in the

analysis, but it is worth mentioning that this conclusion might be influenced by the tendency of those who are most highly restrained (i.e., current dieters) to curtail their GTUI if they are also high on the PFS. Regardless, this finding runs contrary to the Boundary Model, which suggests that high restraint is a liability because it is associated with disinhibited eating and binge eating. If this finding were to be reach statistical significance in a larger study, it would constitute additional support for Lowe & Levine's (2002) premise that restraint protects against, rather than causes, obesogenic eating behavior, especially for the individuals who are most susceptible overeating in response to high-calorie foods. It also suggests that high levels are restraint are associated with overeating because restrained eaters employ high levels of restraint in attempts to *avoid or compensate for* overeating. In fact, without high levels of restraint, restrained eaters might overeat more often than they already do. The Boundary Model does not seem to account for this possibility, and therefore any recommendation against restrained eating which comes from it must be viewed skeptically.

4.4 Limitations

The suboptimal response rate observed in this study is a significant limitation. As proposed by Hufford (2007), compliance with EMA may be increased by providing feedback on rates of participation in real-time, thus fostering a sense of accountability. Also, it may help to shorten the duration of EMA when used with undergraduate populations, to reduce participant burden and fatigue.

Episodes of GTUI were rated subjectively by participants. Even though participants were trained to identify episodes of GTUI within their specific pattern of eating, we cannot determine whether episodes of GTUI truly represented an increase in

caloric intake. Future research should include more objective measures of food consumption.

The sample is small, and is limited to normal-weight undergraduate women. This sample is advantageous in that overweight individuals were excluded, so there is no concern that participant behavior was driven by the presence of overweight. However, it is not known to what degree this study's findings can be generalized to other populations including men, older individuals, and the overweight. Also, the small sample may have resulted in insufficient statistical power for some tests.

Also, this study was designed to test factors that lead to the development of overweight. However, the participants were not selected on characteristics known to increase their chances of excessive weight gain in the future. This may have limited our ability to obtain statistically significant effects for some analysis.

Lastly, although EMA is touted as a method that minimizes reactive effects, it should be mentioned that carrying the device and responding to prompts may have altered the behavior of participants. Any such reactive effects would cast doubt on our conclusions, and limit the generalizability of our findings. However, we have no reason to believe that the current study was unduly influenced by the data collection procedures.

4.5 Implications for Obesity Treatment and Prevention

The results of this study suggest that delicious, high-calorie food in the environment drives excess caloric intake in normal-weight women, and especially among those who are already at risk for eventual overweight or obesity because of their higher BMIs. Additionally, we found some indication that healthy dietary restraint may protect

against obesogenic eating behavior. Only weak evidence was found for restraint as a cause of overeating.

Two clear recommendations for obesity treatment and prevention follow from these findings. First, changing the food environment is a promising avenue for curbing the worsening obesity epidemic. Specifically, making delicious, high-calorie foods *less* available may help prevent excess caloric intake that leads to weight gain. Reducing the availability of such foods could be done within the micro environment (e.g., one's home and workspace) as well as within the macro environment (e.g., within the community, at restaurants and stores). Behavioral weight loss programs routinely recommend that participants remove delicious, high-calorie snacks from the home for this specific reason. Some trials have even made this a focus of treatment, to good effect. Lowe et al. (2008, Obesity), among others, have found that diets which focus on the consumptions of foods that are low in energy density (which are also low-calorie) show promise for facilitating weight control. Gorin et al. (2007) has found that replacing high-calorie foods in the home with low calorie alternatives facilitates weight loss.

There have been fewer attempts to combat the obesogenic effects of high-calorie food within the macro environment, and they have generally focused more on educating consumers about the nature of high-calorie foods, and making healthy foods available, rather than removing high-calorie foods from the environment. Lowe et al. (2009) used a color coding system to alert patrons of two hospital cafeterias of high-calorie foods. This intervention was found to lower the fat content and reduce the energy density of lunches purchased in the cafeterias. Additionally, at least one state has now passed legislation which requires many restaurants to post calorie information on their menus. The impact

of this change on food intake and weight has not been reported. However, recent evidence suggests that merely providing information, and making healthy alternative available, actually *increases* consumption of high-calorie foods. This finding supports the need to actually *remove* obesogenic food from the environment, rather than simply educating consumers or providing healthy alternatives.

A second recommendation which follows from this study is that healthy dietary restraint should be encouraged, as there is only weak evidence that it drives excess caloric intake, at least in non-eating-disordered women. Given our finding that individuals may anticipate episodes of overeating with small increases in restraint, it may be possible to prevent such episodes if we can facilitate *greater* increases in restraint beforehand. To our knowledge, no intervention has been tested that delivers treatment to individuals to increase restraint in the moments before a potential episode of overeating. However, this is an excellent avenue for future research which could incorporate technology to deliver treatment in the moments before eating occurs.

4.6 Directions for Future Research

Our findings suggest several avenues for future research. Future EMAs of eating behavior would benefit from more objective measurements of food consumption, or at least a record of actual intake. Since this study was initiated, advances in technology have been made that facilitate easy recording of food intake via palmtop computers such as the iPhone. More accurate records of intake would allow us to be more specific about which individual differences and environmental conditions are associated with excessive caloric intake.

If this study were repeated, it would be desirable to measure momentary feelings of hunger and satiety, as well as fluctuations in sensitivity to the food environment. This would accomplish two goals. First, it would allow us to better understand whether delicious, high-calorie foods drive overeating even when an individual is sated, as this is a main assumption of the theory of hedonic hunger. Second, it would allow us to know if *variability* in sensitivity to the food environment (rather than the absolute value during a single measurement), is related to overeating.

It is also important to further investigate factors that account for the externally motivated eating observed in individuals with higher BMIs. Our current measure of sensitivity to the food environment helped explain the eating behavior of dieters, but not non-dieters. New types of measures, whether psychological or physiological, may help to further explain the external eating of those individuals at the higher end of the normal BMI range.

Future weight control interventions are encouraged to focus on efforts to remove high-calorie foods from both the micro and macro environment. Our findings suggest that individuals with higher BMIs are better than leaner individuals at avoiding overeating when delicious, high-calorie foods are absent. However, individuals with higher BMIs become much more succeptable to overeating as the number of these foods increases.

Table 1
Hypothesis 1: Prediction of GTUI (Unconditional Means Model)

Parameter	Estimate	*SE*	*df*	*t-ratio*	*OR*	Significance	95% Confidence Interval of the Odds Ratio	
							Lower Bound	Upper Bound
Estimates of fixed effects								
INTERCEPT	-1.237402	0.125549	38	9.856	0.290137	.000	0.225	0.374

	Estimate	*SD*	*df*	χ^2	Significance
Estimates of variance components					
INTERCEPT (subject variance)	0.34041	0.58345	37	90.723	.000

Table 2

Hypothesis 1: Prediction of GTUI from the PFS (Full Sample)

Parameter	Estimate	*SE*	*df*	*t-ratio*	*OR*	Significance	95% Confidence Interval of the Odds Ratio	
							Lower Bound	Upper Bound
Estimates of fixed effects								
INTERCEPT	-1.240408	0.125437	37	9.889	0.289266	.000	0.224	0.373
PFS	0.086391	0.165388	37	0.522	1.090233	.604	0.780	1.523

	Estimate	*SD*	*df*	χ^2	Significance
Estimates of variance components					
INTERCEPT (subject variance)	0.33707	0.58058	37	90.147	.000

Hypothesis 1: Prediction of GTUI from the PFS(Current Dieters Only)

Parameter	Estimate	*SE*	*df*	*t-ratio*	*OR*	Significance	95% Confidence Interval of the Odds Ratio	
							Lower Bound	Upper Bound
Estimates of fixed effects								
INTERCEPT	-1.619403	0.241174	6	6.715	0.198017	.000	0.116	0.337
PFS	-1.330609	0.485063	6	2.743	0.264316	.034	0.091	0.772

	Estimate	*SD*	*df*	χ^2	Significance
Estimates of variance components					
INTERCEPT (subject variance)	0.04020	0.20049	6	7.6193	.266

Table 2 (continued)

Hypothesis 1: Prediction of GTUI from the PFS (Interaction Between Current Dieters and Non-dieters)

							95% Confidence Interval of the Odds Ratio	
Parameter	Estimate	*SE*	*df*	*t-ratio*	*OR*	Significance	Lower Bound	Upper Bound
Estimates of fixed effects								
INTERCEPT	-1.184236	0.135361	35	8.749	0.305980	.000	0.233	0.403
PFS	0.230501	0.171082	35	1.347	1.259231	.187	0.890	1.781
Dieter Status	3.509819	1.461227	35	2.402	33.44222	.022	1.730	646.599
PFS x Dieter Status	-1.678741	0.638691	35	2.628	0.186609	.013	0.051	0.681
	Estimate	*SD*	*df*	χ^2		Significance		
Estimates of variance components								
INTERCEPT (subject variance)	0.30699	0.55407	35	76.916		.000		

Table 3

Hypothesis 1: Prediction of GTUI from the Number of Highly Palatable Foods Available (Full Sample)

							95% Confidence Interval of the Odds Ratio	
Parameter	Estimate	*SE*	*df*	*t-ratio*	*OR*	Significance	Lower Bound	Upper Bound
Estimates of fixed effects								
INTERCEPT	-1.418779	0.172683	38	8.216	0.242009	.000	0.171	0.343
Number of Palatable Foods	0.102595	0.063702	904	1.611	1.108042	.107	0.978	1.255

	Estimate	*SD*	*df*	χ^2	Significance
Estimates of variance components					
INTERCEPT (subject variance)	0.36027	0.60022	38	93.079	.000

Hypothesis 1: Prediction of GTUI from the Number of Highly Palatable Foods Available (Current Dieters Only)

							95% Confidence Interval of the Odds Ratio	
Parameter	Estimate	*SE*	*df*	*t-ratio*	*OR*	Significance	Lower Bound	Upper Bound
Estimates of fixed effects								
INTERCEPT	-2.014241	0.417874	7	4.820	0.133422	.001	0.053	0.336
Number of Palatable Foods	0.315965	0.160599	158	1.982	1.371582	.049	1.010	1.883

	Estimate	*SD*	*df*	χ^2	Significance
Estimates of variance components					
INTERCEPT (subject variance)	0.42572	0.65247	7	17.4087	.015

Table 4

Hypothesis 1: Prediction of GTUI from the PFS and the Number of Highly Palatable Foods Available

							95% Confidence Interval of the Odds Ratio	
Parameter	Estimate	*SE*	*df*	*t-ratio*	*OR*	Significance	Lower Bound	Upper Bound
Estimates of fixed effects								
INTERCEPT	-1.413848	0.171825	37	8.228	0.243206	.000	0.172	0.172
PFS	-0.075173	0.227370	37	0.331	0.387488	.743	0.586	1.469
Number of Palatable Foods	0.094615	0.064145	904	1.475	0.969	.140	0.969	1.246
PFS x Palatable Foods	0.084756	0.084604	904	1.002	1.088452	.371	0.922	1.285
	Estimate	*SD*	*df*	χ^2		Significance		
Estimates of variance components								
INTERCEPT (subject variance)	1.285	0.59285	37	91.725		.000		

Table 5

Hypothesis 1: Prediction of GTUI from the Number of Highly Palatable Foods Available and BMI

							95% Confidence Interval of the Odds Ratio	
Parameter	Estimate	*SE*	*df*	*t-ratio*	*OR*	Significance	Lower Bound	Upper Bound
Estimates of fixed effects								
INTERCEPT	-1.439511	0.178464	37	8.066	0.237044	.000	0.165	0.340
BMI	-0.104994	0.101724	37	1.032	0.900330	.309	0.733	1.106
Number of Palatable Foods	0.110419	0.064351	904	1.716	1.116745	.086	0.984	1.267
BMI x Palatable Foods	0.075804	0.036032	904	2.104	1.078751	.035	1.005	1.158
	Estimate	*SD*	*df*	χ^2		Significance		
Estimates of variance components								
INTERCEPT (subject variance)	0.42676	0.65327	37	101.89		.000		

Table 6

Hypothesis 2: The Trajectory of the EMA Restraint Total Score Relative to Episodes of GTUI

						95% Confidence Interval of the Coefficient	
Parameter	Estimate	*SE*	*df*	*t-ratio*	Significance	Lower Bound	Upper Bound
Estimates of fixed effects							
INTERCEPT	5.692069	0.186974	36	30.443	.000	5.312928	6.071209
Linear	0.029487	0.010287	655	2.866	.004	0.009286	0.049687
Quadratic	-0.001934	0.000819	655	2.360	.019	-0.003544	-0.000324

	Estimate	*SE*	*df*	Wald Z	Significance
Estimates of variance components					
INTERCEPT (subject variance)	1.221474	0.295818	36	4.129	.000

Table 7
Hypothesis 3: Prediction of GTUI from the TFEQ-R

							95% Confidence Interval of the Odds Ratio	
Parameter	Estimate	*SE*	*df*	*t-ratio*	*OR*	Significance	Lower Bound	Upper Bound
Estimates of fixed effects								
INTERCEPT	-1.243424	0.123268	37	10.087	0.28839	0.000	0.225	0.225
TFEQ-R	-0.027657	0.021522	37	-1.285	0.97272	.207	0.931	1.016

	Estimate	*SD*	*df*	χ^2	Significance
Estimates of variance components					
INTERCEPT (subject variance)	0.31560	0.56178	37	87.164	.000

Table 8
Hypothesis 3: Prediction of GTUI from the PFS and the TFEQ-R

Parameter	Estimate	*SE*	*df*	*t-ratio*	*OR*	Significance	95% Confidence Interval of the Odds Ratio	
							Lower Bound	Upper Bound
Estimates of fixed effects								
INTERCEPT	-1.320466	0.131039	35	10.077	0.267011	.000	0.205	0.348
PFS	-0.076734	0.180062	35	0.426	0.926136	.672	0.643	1.334
TFEQ-R	-0.031701	0.022829	35	1.389	0.968796	.174	0.925	1.015
PFS x TFEQ-R	-0.049943	0.031590	35	1.581	0.951284	.123	0.892	1.014
	Estimate	*SD*	*df*	χ^2		Significance		
Estimates of variance components								
INTERCEPT (subject variance)	0.28718	0.53589	35	83.247		.000		

Figure 1: The Boundary Model

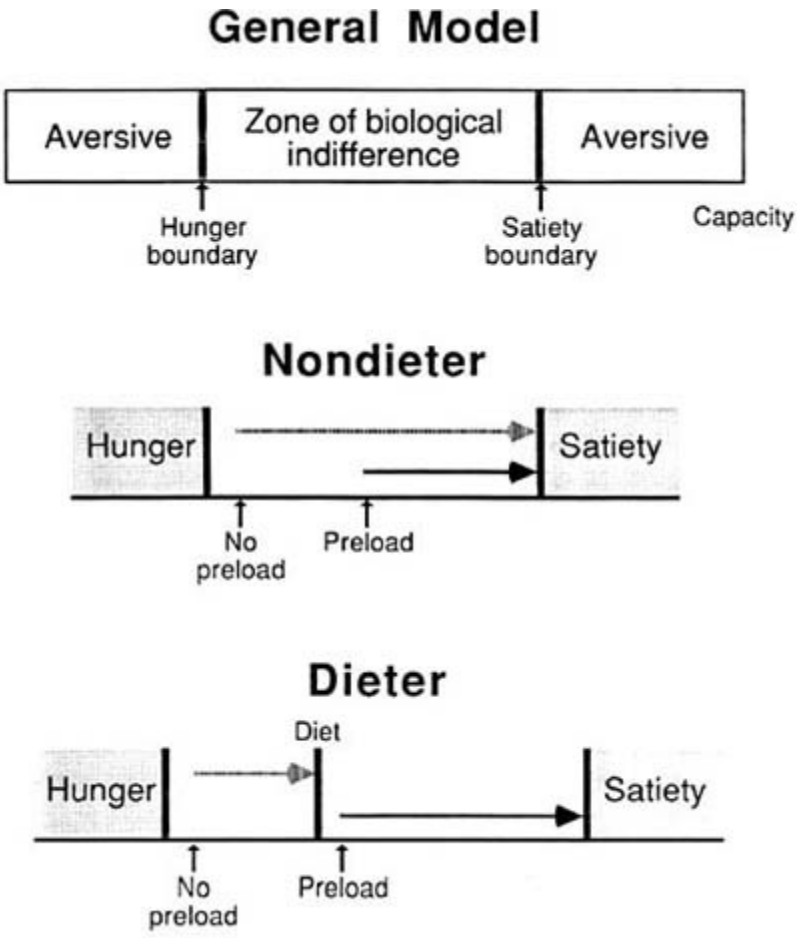


Figure 1. The boundary model of eating regulation. (From "A Boundary Model for the Regulation of Eating" (p. 149) by C. P. Herman and J. Polivy in Eating and Its Disorders by A. J. Stunkard and E. Stellar, [Eds.], 1984, New York: Raven Press. Copyright 1984 by Raven Press.

Figure 2. Full Path of Questions Asked During Ecological Momentary Assessment.

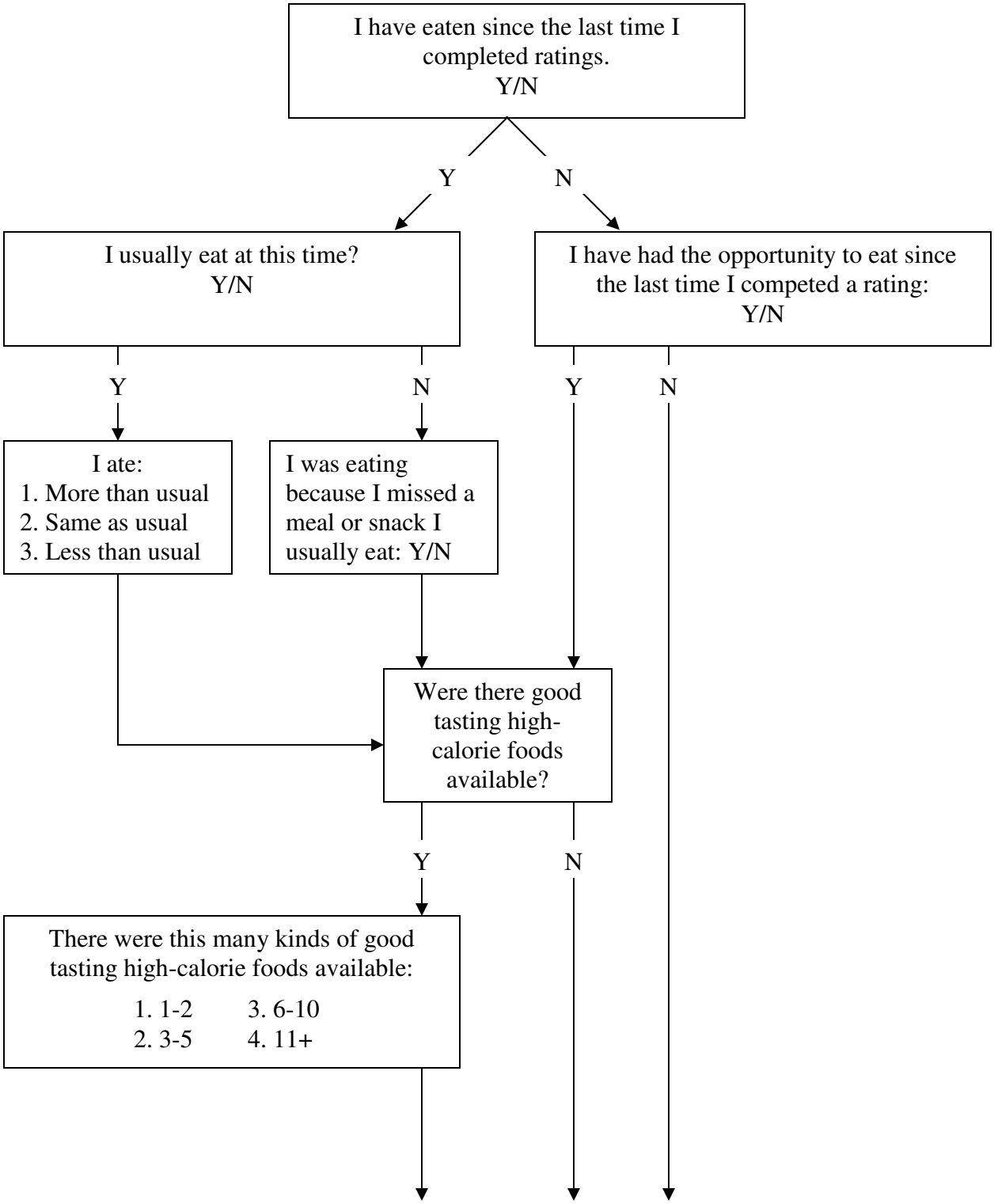
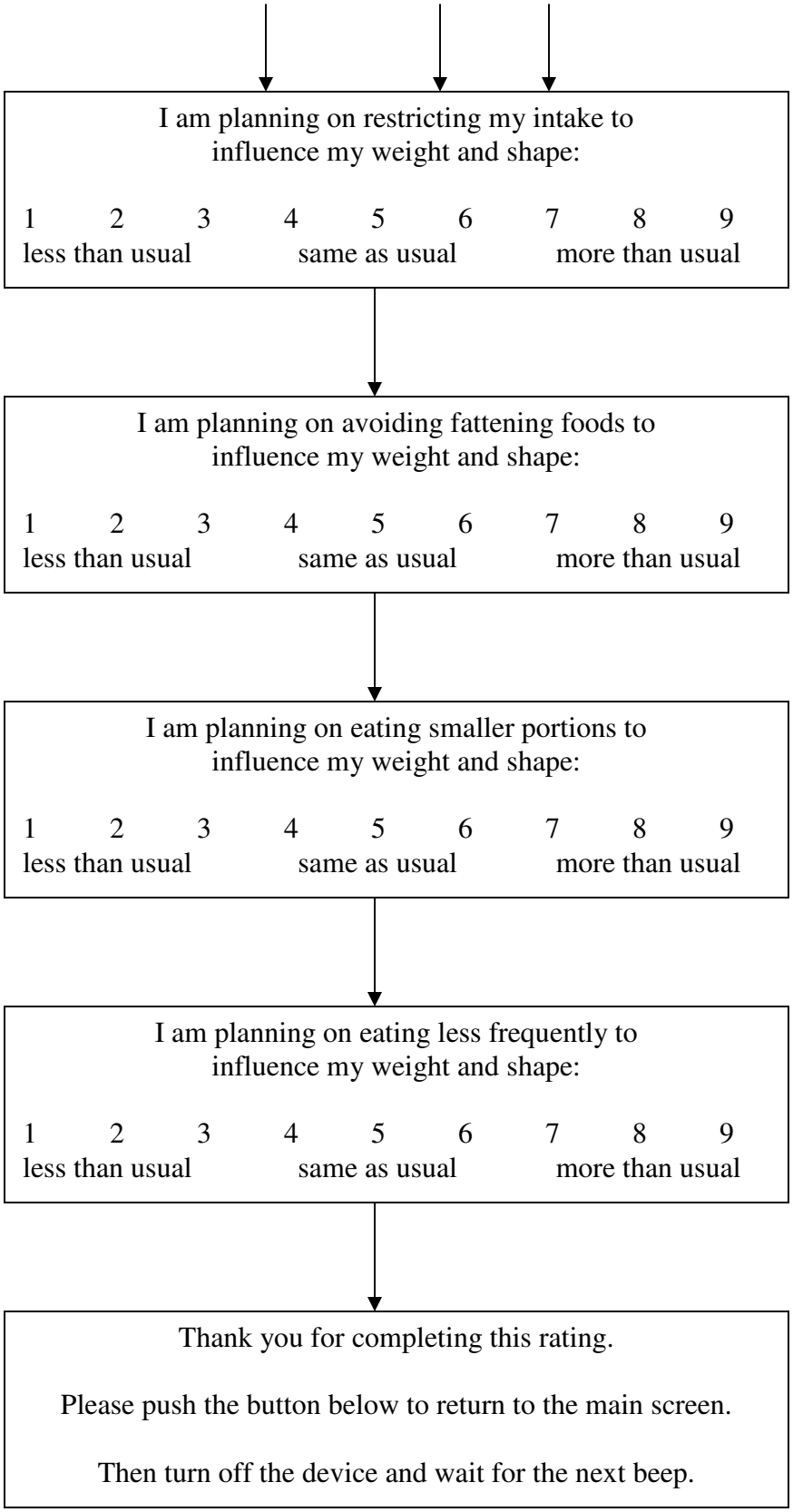


Figure 2. (Continued)



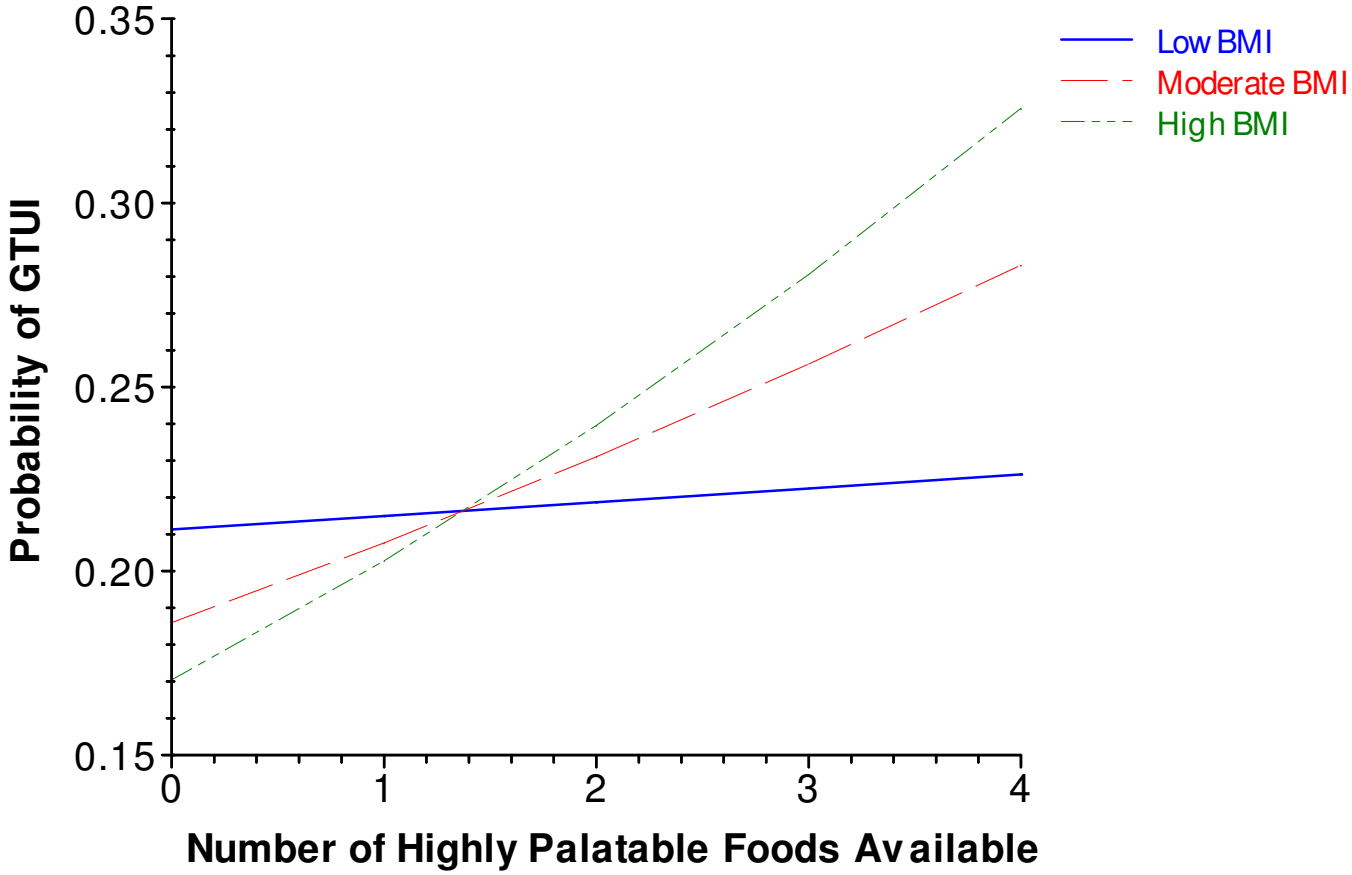


Figure 3. Interaction Between Number of Palatable Foods Available and BMI in the Prediction of the Probability of GTUI.

Note: Low BMI = Lower 25th percentile of the sample BMI distribution, Moderate BMI = 25th to 75th percentile of the sample BMI distribution, High BMI = upper 25th percentile of the sample BMI distribution

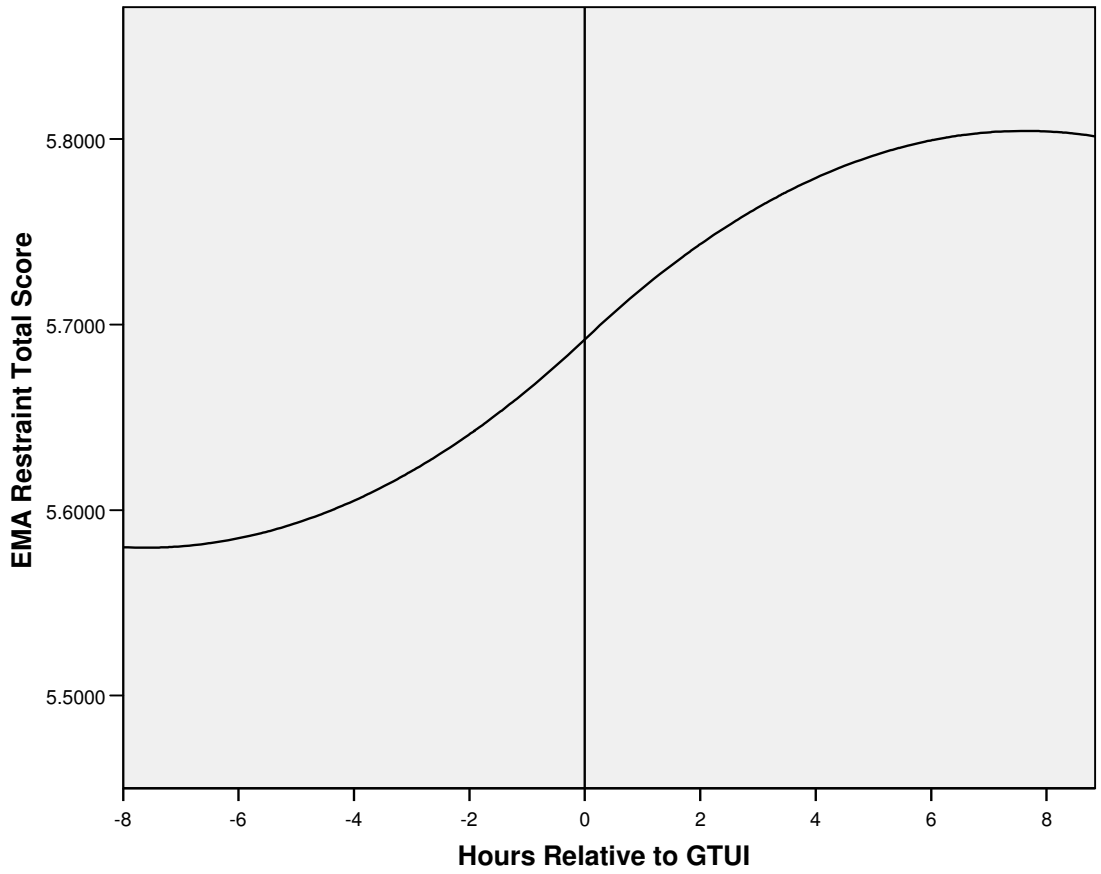


Figure 4. Pattern of Restraint on Days Characterized by Greater-than-usual intake.

References

Bandini, L. G., Schoeller, D. A., Dyr, H. N., & Dietz, W. H. (1990). Validity of reported energy intake in obese and nonobese adolescents. *American Journal of Clinical Nutrition, 52*, 421-425.

Bathalon, G. P., Tucker, K. L., Hays, N. P., Vinken, A. G., Greenberg, A. S., McCrory, M. A., et al. (2000). Psychological measures of eating behavior and the accuracy of 3 common dietary assessment methods in healthy postmenopausal women. *American Journal of Clinical Nutrition, 71*, 739-745.

Bingham, S. A., Cassidy, A., Cole, T. J., Welch, A., Runswick, S. A., & Black, A. E. (1995). Validation of weighted records and other methods of dietary assessment using the 24 h urine nitrogen technique and other biological markers. *British Journal of Nutrition, 71*, 531-550.

Björvell, H., Rössner, S., & Stunkard, A. (1986). Obesity, weight loss, and dietary restraint. *International Journal of Eating Disorders, 5,* 727-734.

Block, G., Hartman, A. M., Dresser, C. M., Carroll, M. D., Gannon, J., & Gardner, L. (1986 Sep). A data-based approach to diet questionnaire design and testing. *American Journal of Epidemiology, 124*(3), 453-469.

Blundell, J. E., & Macdiarmid, J. I. (1997). Passive overconsumption. fat intake and short-term energy balance. *Annals of the New York Academy of Sciences, 827*, 392-407.

Boseck, J. J., Engel, S. G., Allison, K. C., Crosby, R. D., Mitchell, J. E., & de Zwaan, M. (2007). The application of ecological momentary assessment to the study of the night eating. *International Journal of Eating Disorders, 40*(3), 271-276.

Brown, G., & Harris, T. (1978). *Social origins of depression: A study of psychiatric disorder in women*. New York: Wiley.

Brownell, K. D., & Rodin, J. (1994). The dieting maelstrom: Is it possible and advisable to lose weight. *American Psychologist, 49*(9), 781-791.

Cappelleri, J. C., Bushmakin, A. G., Gerber, R. A., Leidy, N. K., Sexton, C. C., Karlsson, J., & Lowe, M. R. (under review). Evaluating the Power of Food Scale in obese subjects and a general sample of individuals: development and measurement properties.

Carels, R. A., Douglass, O. M., Cacciapaglia, H. M., & O'Brien, W. H. (2004). An ecological momentary assessment of relapse crises in dieting. *Journal of consulting and clinical psychology, 72*(2), 341-348.

Carels, R. A., Hoffman, J., Collins, A., Raber, A. C., Cacciapaglia, H., & O'Brien, W. H. (2001). Ecological momentary assessment of temptation and lapse in dieting. *Eating Behaviors, 2*(4), 307-321.

Chaput, J. P., St-Pierre, S., & Tremblay, A. (2007). Currently available drugs for the treatment of obesity: Sibutramine and orlistat. *Mini-Reviews in Medicinal Chemistry, 7*(1), 3-10.

Daeninck, E., & Miller, M. (2006). What can the national weight control registry teach us?. *Current Diabetes Reports, 6*(5), 401-404.

De Castro, J. (1995). The relationship of cognitive restraint to the spontaneous food and fluid intake of free-living humans. *Physiology & Behavior, 57*(2), 287-295.

Deckelbaum, R. J., & Williams, C. L. (2001). Childhood obesity: The health issue. *Obesity, 9,* S239-S243.

Elfhag, K., & Rossner, S. (2005). Who succeeds in maintaining weight loss? A conceptual review of factors associated with weight loss maintenance and weight regain. *Obesity Reviews, 6*(1), 67-85.

Engel, S. G., Boseck, J. J., Crosby, R. D., Wonderlich, S. A., Mitchell, J. E., Smyth, J., et al. (2007). The relationship of momentary anger and impulsivity to bulimic behavior. *Behaviour research and therapy, 45*(3), 437-447.

Evangelista, L. S., & Miller, P. S. (2006). Overweight and obesity in the context of heart failure: Implications for practice and future research. *Journal of Cardiovascular Nursing, 21*(1), 27-33.

Fairburn, C.G. (2002). Cognitive-behavioral therapy for bulimia nervosa. In C.G. Fairburn & K.D. Brownell (Eds.), *Eating disorders and obesity* (pp. 309-307). New York: Guilford Press.

Ferrante, J. M., Chen, P., Crabtree, B. F., & Wartenberg, D. (2007). Cancer screening in women: Body mass index and adherence to physician recommendations. *American Journal of Preventive Medicine, 32*(6), 525-531.

Flegal, K. M., Carroll, M. D., Ogden, C. L., & Johnson, C. L. (2002). Prevalence and trends in obesity among US adults, 1999-2000. *JAMA, 288*(14), 1723-1727.

Foster, G. D., Wadden, T. A., Peterson, F. J., Letizia, K. A., Barlett, S. J., & Conill, A. M. (1992) A controlled comparison of three very-low-calorie-diets: effects on weight, body composition and symptoms. *American Journal of Clinical Nutrition, 55,* 811–817.

Foster, G. D., Wadden, T. A., Swain, R. M., Stunkard, A. J., Platte, P., & Vogt, R. A. (1998). The Eating Inventory in obese women. *International Journal of Obesity, 22,* 778-785.

Gorin, A., Raynor, H., Niemeier, H., & Wing, R. (2007). Home grocery delivery improves the household food environments of behavioral weight loss participants: Results of an 8-week pilot study. *International Journal of Behavioral Nutrition and Physical Activity, 4,* 58.

Greenberg, J. A. (2006). Correcting biases in estimates of mortality attributable to obesity. *Obesity, 14*(11), 2071-2079.

Hedeker, D., & Gibbons, R. D. (1997). Application of random-effects pattern-mixture models missing data in longitudinal studies. *Psychological Methods, 2,* 64-78.

Herman, C. P., & Polivy, J. (1975). Anxiety, restraint, and eating behavior. *Journal of abnormal psychology, 84*(6), 666-672.

Herman, C. P., & Polivy, J. (1980). Restrained eating. *Obesity* (pp. 208-225). Philadelphia: Saunders.

Herman, C. P., & Polivy, J. (1984). A boundary model for the regulation of eating. In A. J. Stunkard, & E. Stellar (Eds.), *Eating and its disorders* (pp. 141-156). New York: Raven Press.

Hetherington, M. M., & Rolls, B. J. (1991). Eating behavior in eating disorders: Response to preloads. *Physiology & Behavior, 50*(1), 101-108.

Hill, J. O., & Peters, J. C. (1998). Environmental contributions to the obesity epidemic. *Science, 280*(5368), 1371-1374.

Hsieh, F. Y., Bloch, D. A., & Larsen, M. D. (1998). A simple method of sample size calculation for linear and logistic regression. *Statistics in medicine, 17*(14), 1623-1634.

Hufford, M. R. (2007). Special methodological challenges and opportunities in Ecological Momentary Assessment. In Stone, A. A., Shiffman, S., Atienza, A. A., & Nebeling, L. (Eds.) *The science of real-time data capture: Self-reports in health research* (1st edition, pp. 54-75). New York, NY: Oxford University Press.

James, W. P. T. (2002). A world view of the obesity problem. In C. G. Fairburn, & K. G. Brownell (Eds.), *Eating disorders and obesity* (2nd ed., pp. 411). New York: Guilford Press.

Jebb, S. A. (2002). Energy intake and body weight. In C. G. Fairburn, & K. D. Brownell (Eds.), *Eating disorders and obesity* (2nd ed., pp. 37). New York: Guilford Press.

Jeffery, R. W., & Utter, J. (2003). The changing environment and population obesity in the united states. *Obesity research, 11*(Suppl), 12S-22S.

Kaplan, R. M., & Atkins, C. J. (1987). Selective attrition causes overestimates of treatment effects in studies of weight loss. *Addictive Behaviors, 12*(3), 297-302.

Killen, J. D., Taylor, C. B., Hayward, C., Haydel, K., Wilson, D., & Hammer, L. (1996). Weight concerns influence the development of eating disorders: A 4-year prospective study. *Journal of consulting and clinical psychology, 64*, 936–940.

Killen, J. D., Taylor, C. B., Hayward, C., Wilson, D., Haydel, K., & Hammer, L. (1994). Pursuit of thinness and onset of eating disorder symptoms in a community sample of adolescent girls: A three-year prospective analysis. *International Journal of Eating Disorders, 16*, 227-238.

Klesges, R. C., Isbell, T. R., & Klesges, L. M. (1992). Relationship between dietary restraint, energy intake, physical activity, and body weight: A prospective analysis. *Journal of abnormal psychology, 101*, 668-674.

Klesges, R., Klem, M., Epkins, C., & Klesges, L. (1991). A longitudinal evaluation of dietary restraint and its relationship to changes in body weight. *Addictive Behaviors, 16*(5), 363-368.

Knight, L. J., & Boland, F. J. (1989). Restrained eating: An experimental disentanglement of the disinhibiting variables of perceived calories and food type. *Journal of abnormal psychology, 98*(4), 412-420.

Kral, T. V., & Rolls, B. J. (2004). Energy density and portion size: Their independent and combined effects on energy intake. *Physiology & Behavior, 82*(1), 131-138.

Laessle, R. G., Tuschl, R. J., Kotthaus, B. C., & Pirke, K. M. (1989). Behavioral and biological correlates of dietary restraint in normal life. *Appetite, 12*(2), 83-94.

Leahey, T. M., & Crowther, J. H. (2008). An ecological momentary assessment of comparison target as a moderator of the effects of appearance-focused social comparisons. *Body Image, 5,* 307-311.

Le Grange, D., Gorin, A., Catley, D., & Stone, A. A. (2001). Does momentary assessment detect binge eating in overweight women that is denied at interview? *European Eating Disorders Review, 9*(5), 309-324.

Le Grange, D., Gorin, A., Dymek, M., & Stone, A. (2002). Does ecological momentary assessment improve cognitive behavioural therapy for binge eating disorder? A pilot study. *European Eating Disorders Review, 10*(5), 316-328.

Lichtman, S. W., Pisarska, K., Berman, E., Pestone, M., Dowling, H., Offenbacher, E., et al. (1992). Discrepancy between self-reported and actual caloric intake and exercise in obese subjects. *New England Journal of Medicine, 327*(27), 1893-1898.

Livingstone, M. B., Prentice, A. M., & Strain, J. J. (1990). Accuracy of weighed dietary records in studies of diet and health. *British Medical Journal, 300*, 708-712.

Lowe, M. R. (1993). The effects of dieting on eating behavior: A three-factor model. *Psychological bulletin, 114*, 100-121.

Lowe, M. R., Annunziato, R. A., Markowitz, J. T., Didie, E., Bellace, D. L., Riddell, L., et al. (2006). Multiple types of dieting prospectively predict weight gain during the freshman year of college. *Appetite, 47*(1), 83-90.

Lowe, M. R., & Butryn, M. L. (2007). Hedonic hunger: A new dimension of appetite?. *Physiology and Behavior, 91,* 432-439.

Lowe, M.R., Butryn, M.L., Diddie, E.R., Annunziato, R.A., Crerand, C.E., Ochner, C.N., et al. (2007). *The Power of Food Scale: A new measure of the psychological influence of the food environment*. Manuscript submitted for publication.

Lowe, M. R., & Caputo, G. C. (1991). Binge eating in obesity: Toward the specification of predictors. *International Journal of Eating Disorders, 10,* 49-55.

Lowe, M. R., & Levine, A. S. (2005). Eating motives and the controversy over dieting: Eating less than needed versus less than wanted. *Obesity Research, 13*, 797-805.

Lowe, M. R., Tappe, K. A., Annunziato, R. A., Riddell, L. J., Coletta, M. C., Crerand, C. E., Didie, E. R., Ochner, C. N., & McKinney, S. (2008). The effect of training in reduced energy density eating and self-monitoring accuracy on weight loss maintenance. *Obesity, 16,* 2016-2023.

Lowe, M. R., & Thomas, J. G.(in press). Measures of Restraint. In D. B. Allison & M. L. Baskin (Eds.), Handbook of Assessment Methods for Obesity and Eating Behaviors

Lowe, M. R., Whitlow, J. W., & Bellwoar, V. (1991). Eating regulation: The role of restraint, dieting, and weight. *International Journal of Eating Disorders, 10*(4), 461-471.

Marcus, M., Wing, R. R., & Lamparski, D. M. (1985). Binge eating and dietary restraint in obese patients. *Addictive Behaviors, 10,* 163-168.

Mela, D. J. (2001). Determinants of food choice: Relationships with obesity and weight control. *Obesity research, 9*(Suppl 4), 249S-255S.

Meiselman, H. L., deGraaf, C., & Lesher, L. L. (2000). The effects of variety and monotony on food acceptance and intake at a midday meal. *Physiology & Behavior, 70*(1-2), 119-125.

Nazarian, D., Smyth, J. M., & Sliwinski, M. J. (2006). A naturalistic study of ambulatory asthma severity and reported avoidant coping styles. *Chronic Illness, 2*(1), 51-58.

Nisbett, R. E. (1972). Eating behavior and obesity in men and animals. *Adv Psychosom Med, 7*, 173-193.

Ouwens, M. A., vanStrien, T., & vanderStaak, C. P. (2003). Tendency toward overeating and restraint as predictors of food consumption. *Appetite, 40*(3), 291-298.

Palta, M., Prineas, R. J., Berman, R., Hannan, P. (1982). Comparison of self-reported and measured height and weight. *American Journal of Epidemiology*, *115*, 223-30.

Peters, J. C. (2003). Combating obesity: Challenges and choices. *Obesity research, 11 Suppl*(Suppl), 7S-11S.

Polivy, J., & Herman, C. P. (1983). *Breaking the diet habit: The natural weight alternative*. New York: Basic Books.

Polivy, J., & Herman, C. P. (1985). Dieting and binging: A casual analysis. *American Psychologist, 40*, 193-201.

Powell, L. H., Calvin, J. E.,3rd, & Calvin, J. E.,Jr. (2007). Effective obesity treatments. *American Psychologist, 62*(3), 234-246.

Prentice, A. M., Black, A. E., Coward, W. A., Davies, H. L., Goldberg, G. L., & Murgatroyd, P. (1986). High levels of energy expenditure in obese women. *British Medical Journal, 292*, 983-987.

Presnell, K., & Stice, E. (2003; 2003). An experimental test of the effect of weight-loss dieting on bulimic pathology: Tipping the scales in a different direction. *Journal of abnormal psychology, 112*(1), 166-170.

Pudel, V., Metzdorff, M., & Oetting, M. (1975). Zur personlichkeit adiposer in psychologischen tests unter berucksichtigung latent fettsuchtiger. Zeitschrift fur Psychosomatische Medizin und Psychoanalyse, 21, 345-361.

Raynor, H. A., & Epstein, L. H. (2001). Dietary variety, energy regulation, and obesity. *Psychological bulletin, 127*(3), 325-341.

Rejeski, W. J., Lang, W., Neiberg, R. H., Van Dorsten, B., Foster, G. D., Maciejewski, M. L., et al. (2006). Correlates of health-related quality of life in overweight and obese adults with type 2 diabetes. *Obesity, 14*(5), 870-883.

Rolls, B. J. (2003). The supersizing of america: Portion size and the obesity epidemic.miscellaneous article. *Nutrition Today, 38*(2), 42-53.

Rolls, B. J., Rowe, E. A., Rolls, E. T., Kindston, B., Megson, A., & Gunary, R. (1981). Variety in a meal increases food intake in Man. *Physiology and Behavior, 26,* 215-221.

Rousee, S. V. (2008). Review of implicit measures of attitudes. *Journal of Personality Assessment, 90,* 524-536.

Rowland, D. L., & Wesselhoft, T. (1998). The study of biobehavioral rhythms in a psychology laboratory course. *Teaching of Psychology, 25,* 106-109.

Sarwer, D. B., Allison, K. C., Gibbons, L. M., Markowitz, J. T., & Nelson, D. B. (2006). Pregnancy and obesity: A review and agenda for future research. *Journal of Women's Health, 15*(6), 720-733.

Schachter, S. (1968). Obesity and eating. internal and external cues differentially affect the eating behavior of obese and normal subjects. *Science, 161*(843), 751-756.

Schoeller, D. A., & van Santen, E. (1982). Measurement of energy expenditure in humans by doubly labeled water method. *Journal of Applied Physiology, 53,* 4955-4959

Smiciklas-Wright, H., Mitchell, D. C., Mickle, S. J., Goldman, J. D., & Cook, A. (2003). Foods commonly eaten in the united states, 1989-1991 and 1994-1996: Are portion sizes changing?see comment. *Journal of the American Dietetic Association, 103*(1), 41-47.

Smyth, J., Wonderlich, S., Crosby, R., Miltenberger, R., Mitchell, J., & Rorty, M. (2001). The use of ecological momentary assessment approaches in eating disorder research. *International Journal of Eating Disorders, 30*(1), 83-95.

Smyth, J. M., Wonderlich, S. A., Heron, K. E., Sliwinski, M. J., Crosby, R. D., Mitchell, J. E., et al. (2007). Daily and momentary mood and stress are associated with binge eating and vomiting in bulimia nervosa patients in the natural environment. *Journal of consulting and clinical psychology, 75*(4), 629-638.

Stein, R. I., Kenardy, J., Wiseman, C. V., Dounchis, J. Z., Arnow, B. A., & Wilfley, D. E. (2007). What's driving the binge in binge eating disorder?: A prospective examination of precursors and consequences. *International Journal of Eating Disorders, 40*(3), 195-203.

Stice, E., Cooper, J., Schoeller, D., Tappe, K., & Lowe, M. (in press). Are dietary restraint scales valid measures of longer-term dietary restriction? Objective biological and behavioral data suggest not, *Psychological Assessment*.

Stice, E., Fisher, M., & Lowe, M. R. (2004). Are dietary restraint scales valid measures of dietary restriction? unobtrusive observational data suggest not. *Psychological assessment, 16*(1), 51-59.

Stice, E., Killen, J. D., Hayward, C., & Taylor, C. B. (1998). Age of onset for binge eating and purging during late adolescence: A 4-year survival analysis. *Journal of abnormal psychology, 107*(4), 671-675.

Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavorial medicine. *Annals of Behavioral Medicine, 16*(3), 199-202.

Teasdale, J. D., & Fogarty, S. J. (1979). Differential effects of induced mood on retrieval of pleasant and unpleasant events from episodic memory. *Journal of abnormal psychology, 88*(3), 248-257.

Tiggemann, M. (1994). Dietary restraint as a predictor of reported weight loss and affect. *Psychological reports, 75*(3, Pt 2), 1679-1682.

Tsai, A. G., & Wadden, T. A. (2005). Systematic review: An evaluation of major commercial weight loss programs in the united states. *Annals of Internal Medicine, 142*(1), 56-66.

Van Strien, T., Herman, C. P., Engels, R. C., Larsen, J. K., & van Leeuwe, J. F. (2007). Construct validation of the restraint scale in normal-weight and overweight females. *Appetite, 49*(1), 109-121.

Van Strien, T., Cleven, A., & Schippers, G. (2000). Restraint, tendency toward overeating and ice cream consumption. *International Journal of Eating Disorders, 28*(3), 333-338.

Van Strien, T. (1999). Success and failure in the measurement of restraint: Notes and data. *International Journal of Eating Disorders, 25*(4), 441-449.

Wadden, T. A., Bartlett, S. J., Foster, G. D., Greenstein, R. A., Wingate, B. J., Stunkard, A. J., et al. (1995). Sertraline and relapse prevention training following treatment by very-low-calorie diet: A controlled clinical trial. *Obesity research, 3*(6), 549-557.

Wadden, T. A., Butryn, M. L., & Wilson, C. (2007). Lifestyle modification for the management of obesity. *Gastroenterology, 132*(6), 2226-2238.

Wansink, B. (2004). Environmental factors that increase the food intake and consumption volume of unknowing consumers. *Annual Review of Nutrition, 24*, 455-479.

Ward, A., & Mann, T. (2000). Don't mind if I do: Disinhibited eating under cognitive load. *Journal of personality and social psychology, 78*(4), 753-763.

Wardle, J. (1980). Dietary restraint and binge eating. *Behavior Analysis and Modification, 4,* 201-209.

Wegner, K. E., Smyth, J. M., Crosby, R. D., Wittrock, D., Wonderlich, S. A., & Mitchell, J. E. (2002). An evaluation of the relationship between mood and binge eating in the natural environment using ecological momentary assessment. *International Journal of Eating Disorders, 32*(3), 352-361.

Wheeler, L., & Reis, H. T. (1991). Self-recording of everyday life events: Origins, types, and uses. *Journal of personality, 59*(3), 339-354.

Yao, M., & Roberts, S. B. (2001). Dietary energy density and weight regulation. *Nutrition reviews, 59*, 247-258.

Yeomans, M. R., Blundell, J. E., & Leshem, M. (2004). Palatability: Response to nutritional need or need-free stimulation of appetite. *British Journal of Nutrition, 92*(1S), S3-14.

Zandstra, E. H., De Graaf, C., & Van Trijp, H. C. M. (2000). Effects of variety and repeated in-home consumption on product acceptance. *Appetite, 35*(2), 113-119.

APPENDIX A: QUESTIONNAIRE MEASURES

Power of Food Scale

Instructions: Please indicate the extent to which you agree that the following items describe you.

Response options: (1) I don't agree; (2) I agree a little, (3) I agree somewhat; (4) I agree quite a bit; (5) I strongly agree.

1. I find myself thinking about food even when I'm not physically hungry.
2. I get more pleasure from eating than I do from almost anything else.
3. If I see or smell a food I like, I get a powerful urge to have some.
4. When I'm around a fattening food I love, it's hard to stop myself from at least tasting it.
5. It's scary to think of the power that food has over me.
6. When I know a delicious food is available, I can't help myself from thinking about having some.
7. I love the taste of certain foods so much that I can't avoid eating them even if they're bad for me.
8. Just before I taste a favorite food, I feel intense anticipation.
9. When I eat delicious food I focus a lot on how good it tastes.
10. Sometimes, when I'm doing everyday activities, I get an urge to eat "out of the blue" (for no apparent reason).
11. I think I enjoy eating a lot more than most other people.
12. Hearing someone describe a great meal makes me really want to have something to eat.

13. It seems like I have food on my mind a lot.

14. It’s very important to me that the foods I eat are as delicious as possible.

15. Before I eat a favorite food my mouth tends to flood with saliva.

Three Factor Eating Questionnaire

Cognitive Restraint Subscale

Instructions: Read each of the following statements carefully. If you agree with the statement, or feel that it is true as applied to you, choose "Agree" for the corresponding statement. If you disagree with the statement, or feel that it is false as applied to you, choose "Disagree" next to the corresponding statement. Be certain to answer each question.

Response Options (Questions 1-12): (0) True; (1) False

1. When I have eaten my quota of calories, I am usually good about not eating any more.
2. I deliberately take small helpings as a means of controlling my weight.
3. Life is too short to worry about dieting.
4. I have a pretty good idea of the number of calories in common foods .
5. While on a diet, if I eat a food that is not allowed, I consciously eat less for a period of time to make up for it.
6. I enjoy eating too much to spoil it by counting calories or watching my weight .
7. I often stop eating when I am not really full as a conscious means of limiting the amount that I eat.
8. I consciously hold back at meals in order not to gain weight.
9. I eat anything I want, any time I want.
10. I count calories as a conscious means of controlling my weight.
11. I do not eat some foods because they make me fat.
12. I pay a great deal of attention to changes in my figure.

13. How often are you dieting in a conscious effort to control your weight?

    Response Options: (0) Rarely; (0) Sometimes; (1) Usually; (1) Always

14. Would a weight fluctuation of 5 lbs. affect the way you live your life?

    Response Options: (0) Not at all; (0) Slightly; (1) Moderately; (1) Very much

15. Do your feelings of guilt about overeating help you to control your food intake?

    Response Options: (0) Never; (0) Rarely; (1) Often; (1) Always

16. How conscious are you of what you are eating?

    Response Options: (0) Not at all; (0) Slightly; (1) Moderately; (1) Extremely

17. How frequently do you avoid “stocking up” on tempting foods?

    Response Options: (0) Almost never; (0) Seldom; (1) Usually; (1) Always

18. How likely are you to shop for low calorie foods?

    Response Options: (0) Unlikely; (0) Slightly Likely; (1) Moderately Likely; (1) Very Likely

19. How likely are you to consciously eat slowly in order to cut down on how much you eat?

    Response Options: (0) Unlikely; (0) Slightly Likely; (1) Moderately Likely; (1) Very Likely

20. How likely are you to consciously eat less than you want?

    Response Options: (0) Unlikely; (0) Slightly Likely; (1) Moderately Likely; (1) Very Likely

21. On a scale of 1 to 5, where 1 means no restraint in eating (eat whatever you want, whenever you want it) and 5 means total restraint (usually or constantly limiting food intake and rarely or never "giving in"), what number would you give yourself?

    Response Options: (0) 1-eat whatever you want, whenever you want it;(0) 2-usually eat whatever you want, whenever you want it; (0) 3-often eat whatever you want, whenever you want it; (1) 4-often limit food intake, but often "give in"; (0) 5-usually or constantly limit food intake, rarely or never "give in"

Vita

J. Graham Thomas was born in Florence, Alabama on March 21st, 1980. He received his bachelor's degree from Trinity College in Hartford, CT in 2002, where he conducted research in cognitive psychology under Dr. Karl Haberlandt. Publications from this work include, "Forward and backward recall: Different response time patterns, same retrieval order" in Psychological Science, "Pauses and durations exhibit a serial position effect" in Psychonomic Bulletin and Review, and "Transposition asymmetry in immediate serial recall" in Memory. Following graduation from Trinity College, Mr. Thomas served as Project Coordinator for Dr. Madelyn Gould on a study of adolescent suicide screening and prevention at the New York State Psychiatric Institute. Publications from this work include, "Teenagers' attitudes about coping strategies and help-seeking behavior for suicidality" in the Journal of the American Academy of Child and Adolescent Psychiatry, and "Evaluating the iatrogenic risk of youth suicide screening programs: A randomized controlled trial" in the Journal of the American Medical Association. Mr. Thomas began his doctoral training in clinical psychology at Drexel University in 2004, where he was mentored by Dr. Michael Lowe in the study of obesity and eating disorders. Publications from this work include, "Reductions in Internal Disinhibition during weight loss predict better weight loss maintenance" in Obesity, and "The relationship of weight suppression and dietary restraint to binge eating in bulimia nervosa" in the International Journal of Eating Disorders. Mr. Thomas completed his pre-doctoral internship at Brown University and The Miriam Hospital, where he was mentored by Dr. Rena Wing.