Psychological Assessment
2007, Vol. 19, No. 4, 449–458

1040-3590/07/$12.00 DOI: 10.1037/1040-3590.19.4.449

Are Dietary Restraint Scales Valid Measures of Moderate- to Long-Term Dietary Restriction? Objective Biological and Behavioral Data Suggest Not

Eric Stice
Oregon Research Institute

Jamie A. Cooper and Dale A. Schoeller
University of Wisconsin—Madison

Karyn Tappe and Michael R. Lowe
Drexel University

Prospective studies indicate that elevated scores on dietary restraint scales predict bulimic symptom onset, but experiments indicate that assignment to dietary restriction interventions reduces bulimic symptoms. One possible explanation for the inconsistent findings is that the dietary restraint scales used in the former studies are not valid measures of dietary restriction. The authors previously found that dietary restraint scales were not inversely correlated with objective measures of short-term caloric intake (E. Stice, M. Fisher, & M. R. Lowe, 2004). In this follow-up report, 3 studies indicated that the Three-Factor Eating Questionnaire dietary restraint scale was not correlated with doubly labeled water estimated energy intake over 2-week periods or with observationally measured caloric intake over 3 months. Results from this study and others suggest that dietary restraint scales may not be valid measures of moderate- to long-term dietary restriction and imply the need to reinterpret findings from studies that have used dietary restraint scales.

Keywords: dietary restraint, dieting, validity

Dieting theoretically increases risk for onset and maintenance of binge eating and bulimia nervosa (Fairburn, 1997; Polivy & Herman, 1985). Dieting, or dietary restraint,[1] has been defined as intentional and sustained restriction of caloric intake for the purposes of weight loss or maintenance (Herman & Mack, 1975; Wadden, Brownell, & Foster, 2002; Wilson, 2002). Dietary restriction must result in a negative energy balance for weight loss or a balance between intake and output for weight maintenance. Polivy and Herman (1985) suggested that the chronic hunger experienced by dieters increases the risk that they will binge eat. They also argued that a reliance on cognitive controls over eating leaves dieters vulnerable to uncontrolled eating when these cognitive processes are disrupted. These binge-eating episodes theoretically precipitate redoubled dietary efforts and use of radical weight control techniques (e.g., vomiting), which may cascade into the binge–purge cycle (Fairburn, 1997).

Prospective studies have indicated that females with elevated scores on dietary restraint scales are at greater risk for future onset of binge eating, bulimic symptoms, and bulimic pathology (Field et al., 1999; Killen et al., 1996; Stice, Killen, Hayward, & Taylor, 1998; Stice, Presnell, & Spangler, 2002) and future increases in bulimic symptoms (Johnson & Wardle, 2005; Stice, 2001) than their lower-scoring counterparts. Thus, elevated dietary restraint scores show a positive main effect on risk for future onset of bulimic pathology. These studies primarily used the Restraint Scale (Polivy, Herman, & Warsh, 1978) and the Dutch Restrained Eating Scale (van Strien, Frijters, van Staveren, Defares, & Deurenberg, 1986). Although prospective studies are useful because they can establish temporal precedence, they cannot rule out the possibility that some unmeasured third variable confound explains any prospective relations.

Experiments have found that rats randomized to extreme caloric deprivation conditions (in which they lost 7%–20% of their weight) consumed significantly more calories during ad lib feeding immediately after deprivation relative to nondeprived control rats (Hagan, Chandler, Wauford, Rybak, & Oswald, 2003; Ogawa et al., 2005; Sterritt, 1962). However, other experiments found that rats assigned to cycles of caloric restriction and refeeding did not show significantly different ad lib caloric intake after refeeding relative to control rats (Boggiano et al., 2005; Hagan et al., 2003). Thus, there is evidence that caloric deprivation results in elevated caloric intake immediately after deprivation, which is logical given that the animals did not elect to restrict their caloric intake and were therefore probably compensating for the missed calories.

Eric Stice, Oregon Research Institute; Jamie A. Cooper and Dale A. Schoeller, Department of Nutritional Sciences, University of Wisconsin—Madison; Karyn Tappe and Michael R. Lowe, Department of Psychology, Drexel University.

This study was supported by Research Grants DK30031, MH064560, and HL073775 from the National Institutes of Health. We are indebted to Amy F. Subar from the National Cancer Institute for allowing us to conduct secondary analyses of data from the Observing Protein and Energy Nutrition Study, which are reported in Study 3 of this article.

Correspondence concerning this article should be addressed to Eric Stice, Oregon Research Institute, 1715 Franklin Boulevard, Eugene, OR 97403. E-mail: estice@ori.org

[1] We use the terms *dieting* and *dietary restraint* interchangeably on the basis of the practice of other investigators in the literature (e.g., Polivy & Herman, 1985).

Although experiments with animals allow greater experimental control over caloric restriction manipulations and are immune to demand characteristics, they have questionable generalizability to humans, given that the animal studies involve involuntary food restriction, which may be very different than the voluntary dietary restriction practiced by humans in Western culture. In addition, these studies did not assess bulimic symptoms, so it is unclear whether these findings provide information about whether dietary restriction is related to bulimic pathology.

Because of the consistency of the prospective findings, it is widely accepted that dieting plays a causal role in the onset of bulimic pathology (Fairburn, 1997; Huon, 1996; Levine & Smolak, 2006; Neumark-Sztainer, 2005). Some have even called for a moratorium on dieting and have evaluated interventions that decrease dietary restriction (Bacon et al., 2002; Polivy & Herman, 1992). In addition, many eating disorder prevention programs advise against dieting (Smolak, Levine, & Schermer, 1998; Stewart, Carter, Drinkwater, Hainsworth, & Fairburn, 2001).

In contrast to the findings from prospective studies, experiments with humans that have increased dietary restriction over periods ranging from 6 weeks to 3 years have found significant reductions in binge eating and bulimic symptoms. Obesity treatment trials have indicated that assignment to 6-month dietary restriction interventions resulted in significantly greater decreases in binge eating relative to waitlist control conditions for obese women (Goodrick, Poston, Kimball, Reeves, & Foreyt, 1998; Reeves et al., 2001). Assignment to a 20-week energy-restriction diet similarly resulted in significantly greater decreases in binge eating relative to a waitlist control condition for overweight women (Klem, Wing, Simkin-Silverman, & Kuller, 1997). Experimental psychopathology trials have also indicated that assignment to 6-week energy-restriction diets, versus waitlist control conditions, reduced bulimic symptoms in normal-weight adolescent girls and young women (Groesz & Stice, 2007; Presnell & Stice, 2003) and in women with bulimia nervosa (Burton & Stice, 2006). Assignment to a lower intensity weight-maintenance prevention program, which significantly reduced risk for weight gain and obesity onset over a 3-year period, reduced bulimic symptoms in normal-weight adolescent girls with body image concerns relative to an assessment-only control condition and a placebo control condition (Stice, Shaw, Burton, & Wade, 2006; Stice, Spoor, Presnell, & Shaw, 2007). Participants in the latter five trials showed decreases in weight, verifying that they achieved a negative energy balance. Thus, dietary restriction interventions result in a negative main effect on bulimic symptoms. Although experiments offer greater inferential confidence because they more effectively rule out confounding variables (unlike prospective studies), experiments can have limited ecological validity and are subject to demand characteristics.

The contradictory findings are troubling because they have opposing public health implications. If dieting increases risk for bulimic symptoms, interventionists should attempt to decrease dieting, but if dieting reduces bulimic symptoms and facilitates weight control, interventionists should help individuals diet more effectively. The fact that 45% of adolescent girls report dieting underscores the importance of determining whether dieting has negative effects (Neumark-Sztainer, 2005).

One possible explanation for the inconsistent findings is that the dietary restraint scales used in the prospective studies may not be valid measures of dietary restriction. If the scales used in those studies do not identify individuals who are actually achieving the energy-deficit diet necessary for weight loss, it might explain why those studies produced findings that are discrepant from the findings from experimental trials that confirmed that participants achieved energy-deficit diets. Two lines of evidence suggest this is a reasonable hypothesis. First, individuals with elevated dietary restraint scores often gain more weight over time than those with lower scores (French, Jeffery, Forster, et al., 1994; Klesges, Isbell, & Klesges, 1992; Stice, Presnell, Shaw, & Rohde, 2005), suggesting that these scales may not be valid measures of energy-deficit dieting. Second, people often underreport caloric intake, particularly those with high dietary restraint scores (Bandini, Schoeller, Dyr, & Dietz, 1990; Lichtman et al., 1992; Prentice et al., 1986).

We conducted four studies that tested whether five dietary restraint scales showed inverse correlations with directly observed caloric intake (Stice, Fisher, & Lowe, 2004). Because the original validity studies used self-reported caloric intake as the criterion (French, Jeffery, & Wing, 1994; Kirkley, Burge, & Ammerman, 1988; Neumark-Sztainer, Jeffery, & French, 1997; van Strien et al., 1986; Wardle & Beales, 1987), we also used caloric intake as the criterion for our validation studies. We examined five dietary restraint scales that were developed to assess intentional dietary restriction for the purposes of weight control: the Three-Factor Eating Questionnaire (TFEQ) restraint scale (Stunkard & Messick, 1985), Dutch Restrained Eating Scale (van Strien et al., 1986), Restraint Scale (Polivy et al., 1978), Eating Disorder Examination Questionnaire—Restraint subscale (Fairburn & Beglin, 1994), and Dietary Intent Scale (Stice et al., 2004). In contrast to the original validity studies, these scales showed weak and generally nonsignificant correlations with objectively measured caloric intake during unobtrusively observed eating episodes across the four studies ($M\ r = -.07$, range = –.34 to .20; Stice et al., 2004). For example, the average correlation between three dietary restraint scales and observed caloric intake of students consuming meals in a cafeteria was –.09. Results did not change when we controlled for body mass—a potential confound. Other studies have found that dietary restraint scales did not show inverse correlations with objectively measured caloric intake during single eating episodes (e.g., Hetherington, Bell, & Rolls, 2000; Jansen, 1996; Ouwens, van Strien, & van der Staak, 2003; van Strien, Cleven, & Schippers, 2000; Wardle & Beales, 1987). Based on the evidence that individuals with elevated dietary restraint scores do not consume fewer calories than those with low scores, we concluded that these scales do not appear to be valid measures of short-term caloric restriction and that it might be prudent to reinterpret findings from studies that use these scales.

In response to our validity findings, van Strien, Engels, van Staveren, and Herman (2006) noted that short-term caloric intake may not be representative of long-term caloric intake and suggested that researchers test whether dietary restraint scales show inverse correlations with objective measures of longer term caloric intake. Four studies have investigated the concordance between dietary restraint scales and unobtrusively observed caloric intake during multiple eating episodes, which should provide a more representative index of caloric intake. Rolls and colleagues (1997) found that lean and overweight adults with high versus low scores on the TFEQ–Restraint scale did not show significant differences in caloric intake during three meals and a snack consumed during

a 20-hr monitoring period in the lab. Jansen and associates (2003) found that Eating Disorder Examination—Restraint scale was not significantly correlated with observed caloric intake during three separate taste tests involving snack foods for normal-weight preadolescents. Martin and associates (2005) found that the TFEQ–Restraint scale was not correlated with observed caloric intake during four separate healthy meals consumed by normal-weight young women. Sysko, Walsh, Schebendach, and Wilson (2005) found that the TFEQ–Restraint scale, Dietary Intent Scale, Eating Disorder Examination Questionnaire—Restraint subscale, and Eating Disorder Examination—Restraint scale (Fairburn & Cooper, 1993) were not significantly correlated with observed caloric intake of a yogurt shake eaten by women with anorexia nervosa or normal-weight women during two sessions.

Two studies used doubly labeled water (DLW) to assess caloric intake over a 2-week period. DLW uses isotopic tracers to assess total carbon dioxide production, which can be used to generate accurate estimates of habitual caloric intake over this period (Schoeller et al., 1986). Bathalon et al. (2000) found that normal-weight women with high TFEQ–Restraint scores did not consume significantly fewer calories over a 2-week period than weight-matched women with low scores. Tuschl, Laessle, Platte, and Pirke (1990) also found a nonsignificant relation between the TFEQ–Restraint scale and DLW estimates of caloric intake over a 2-week period.

Although these six validity studies suggest that dietary restraint scales are not valid measures of moderate-term dietary restriction, we felt this question warranted further examination because many researchers use these scales. Thus, we conducted secondary analyses of three data sets to test whether one widely used dietary restraint scale—the TFEQ–Restraint scale—showed inverse correlations with objective measures of moderate- to long-term caloric intake. We examined this scale because it should constitute a more demanding validity test, given that it has been argued that it is a measure of successful dietary restriction (Heatherton, Herman, Polivy, King, & McGree, 1988). Study 1 examined the correlation between the TFEQ–Restraint scale and observationally measured caloric intake during lunch meals consumed at work cafeterias over a 3-month period. Because Study 1 focused solely on caloric intake during lunch, Study 2 tested whether the TFEQ–Restraint scale correlated with total caloric intake assessed over a 2-week period with DLW. Because Study 2 involved a small, atypical sample of obese women who had recently lost weight, Study 3 examined the correlation between the TFEQ–Restraint scale and DLW-assessed total caloric intake over a 2-week period in a larger, more representative sample of healthy women.

Study 1

Method

Participants and Procedures

Female employees ($N = 87$) of two hospitals in Philadelphia who regularly ate lunch at the hospital cafeterias constituted the sample (*M* age = 44.9, *SD* = 10.3; 43% Black, 52% White, 1% Hispanic, 2% Asian, and 2% other). The average body mass index (BMI; kg/m^2) was 30.1 (*SD* = 6.1); 26% were normal weight (BMI $<$ 25), 30% were overweight (BMI = 25–30), and 44% were obese (BMI $>$ 30). These data were from a trial testing whether an intervention that provided education about reduced energy-density eating and nutrition labels for cafeteria foods affected eating behavior. However, the present report concerns only the baseline preintervention measures for this study, so the intervention is not discussed further here. Immediately before the 3-month period of lunchtime dietary intake monitoring, participants completed the TFEQ–Restraint scale, a 24-hr food recall, and height and weight measurements. Lunchtime food intake at cafeterias was then monitored by means of computerized cash registers and employee identification cards, which participants swiped when they purchased their lunch at the hospital cafeterias. Participants were only included in the analysis if they swiped their identification card at lunchtime once a week or more (*M* lunches consumed in the cafeteria = 6.8 per month, *SD* = 3.9). Lunch purchases were automatically read to a database that was linked to the programmed cash registers. Three months of lunch purchases were collected and analyzed for caloric and nutritional content. The study was reviewed and approved by the local institutional review board, and informed written consent was obtained from participants, as was the case for the other two studies reported here.

Measures

TFEQ–Restraint scale. The TFEQ–Restraint scale (Stunkard & Messick, 1985) assesses dietary behaviors designed to produce weight loss or maintenance, monitoring of body shape, and importance of thinness (sample item: "I count calories as a conscious means of controlling my weight"). This scale has been found to be internally consistent (alphas range from .85 to .93) and temporally reliable (1-month test–retest = .98) in past studies (French, Jeffery, & Wing, 1994; Stunkard & Messick, 1985).

Self-reported caloric intake. Self-reported caloric intake was assessed with a 24-hr dietary recall that was completed over the phone by the Diet Assessment Center at The Pennsylvania State University. Participants reported what foods they had eaten that day, how the foods had been prepared, and how much of each food had been consumed. They were given a poster that provided pictures of sizes and amounts of food to help them estimate how much food they had consumed. Although the 24-hr dietary recall is widely used, self-reported caloric intake on these recalls does not correlate significantly with total caloric intake over several-week periods as assessed by biological procedures (*M r* = .12; Bathalon et al., 2000; Bingham et al., 1995), suggesting that this measure is not a particularly valid measure of caloric intake.

Objectively assessed caloric intake. Meal purchase data were collected by the electronic cash registers developed with Sodhexo Corporation (Gaithersburg, MD), which manages the cafeterias investigated in this study. These meal purchase data were used in combination with the nutritional database maintained by Sodexho Corporation, which reflects the nutritional content of all prepared food served in the cafeterias, to obtain caloric values for each item purchased, thereby permitting estimation of the caloric intake during the meal. To accommodate custom food purchases, such as salads and sandwiches, participants completed an initial survey about how they prepare such foods, including the type and amounts of ingredients that they typically use. These data were stored with the participant identifier and linked to the cafeteria

food database, which allowed accurate data collection regarding the caloric content of both standard and custom food items. Less than 20% of all of the foods consumed by participants involved customized preparation. A subset of participants was asked to take pictures of their food trays after they had completed their meals to determine whether they had consumed all of their food. These photographs confirmed that virtually all participants consumed all of the purchased food, which suggested that it was reasonable to assume that participants usually consumed all of the food they purchased for their lunches. Participants consumed an average of 613 kcal per meal ($SD = 142$).

Body mass. BMI served as a proxy measure of adiposity. After participants removed shoes and coats, their height was measured to the nearest millimeter using a stadiometer and weight was assessed to the nearest 0.1 kg using digital scales. BMI correlates with direct measures of total body fat such as dual energy x-ray absorptiometry (rs = .80–.90; Goran, Driscoll, Johnson, Nagy, & Hunter, 1996; Pietrobelli et al., 1998) and with health measures including blood pressure, adverse lipoprotein profiles, atherosclerotic lesions, serum insulin levels, and diabetes mellitus (Pietrobelli et al., 1998).

Results and Discussion

Table 1 reports the correlations between TFEQ–Restraint scores, observationally measured caloric intake during lunch meals, and self-reported caloric intake from the 24-hr dietary recall. TFEQ–Restraint scores did not show a statistically significant inverse correlation with observationally measured caloric intake during lunches over the 3-month period ($r = .11$). This relation did not become significant when BMI, weight, or ethnicity were included as covariates, which was also the case for the other two studies reported here. TFEQ–Restraint scores did not show a statistically significant or substantively meaningful correlation with self-reported caloric intake from the 24-hr dietary intake measures ($r = –.16$).

Thus, Study 1 suggests that the TFEQ–Restraint scale did not show a statistically significant or substantively meaningful relation with objectively measured caloric intake over a several-month period. Although Study 1 did not find any relation between the TFEQ–Restraint scale and number of calories consumed during cafeteria meals over the 3-month period, one limitation was the sole focus on caloric intake during lunch meals consumed at work cafeterias, which may not be representative of total caloric intake. Accordingly, Study 2 examined the correlation between the TFEQ–Restraint scale and total caloric intake over a 2-week period using DLW.

Table 1
Correlations Between Study Variables, Study 1

Variable	1	2	3	4
1. Three-Factor Eating Questionnaire-Restraint	—	.11	−.16	.02
2. Observationally measured caloric intake		—	.02	.03
3. Self-reported caloric intake[a]			—	−.09
4. Body mass index				—

Note. No effects were statistically significant.
[a] 24-hr dietary recall.

Study 2

Method

Participants and Procedures

In Study 2, previously obese women who had recently lost weight completed the TFEQ–Restraint scale and a self-report measure of caloric intake and underwent a DLW measure of 2-week total caloric intake. Women ages 20 to 50 years (M age = 38, $SD = 7$) were recruited from the Chicago area ($N = 33$) to participate in a weight-maintenance study (see Schoeller, Shay, & Kushner, 1997). The sample was composed of 12% Blacks and 88% Whites. Entry criteria were a weight loss ≥12 kg, maintenance of weight stability within 1 kg for >1 month but not >3 months, and a current BMI of 20 to 30. Weight loss had to have been documented in writing by a physician or weight-loss program director. Exclusion criteria included cigarette smoking, history of a metabolic disease such as diabetes mellitus or a thyroid disorder, hypertension, history of a psychiatric disorder, and a physical handicap that would interfere with exercise. Average BMI was 23.8 ($SD = 2.8$) at baseline. During the 12-month study period, participants made six visits to the clinical research center: two at baseline (to allow urine collection) and one each at 3, 6, 9, and 12 months. Data for the present report were drawn from the baseline assessment. Participants were told to follow whatever weight-maintenance strategy they preferred. To minimize alterations in behavior during the periods of data collection, we told participants that the DLW assessment measured body composition.

Measures

TFEQ–Restraint scale. Participants completed the TFEQ–Restraint scale (Stunkard & Messick, 1985) at baseline (see *TFEQ–Restraint scale* section of Study 1 for psychometric details).

Self-reported dietary intake. Participants recorded their dietary intake by using a pocket diary for 7 days at baseline. The diary contained separate lined pages for each day of the week, with subsections for each meal and snacks. Participants were taught how to estimate portion sizes in common household units. They recorded a description and amount for all foods and energy-containing beverages consumed over the subsequent 7 days and returned the diary by mail. Research staff and participants reviewed the diaries to ensure complete information was obtained. A dietitian reviewed the records and calculated the energy content of the diets using the Nutripractor III program (Practorcare, San Diego).

Objective measure of caloric intake. Energy intake at baseline was calculated from the sum of energy expenditure from DLW and the estimated change in body energy stores from serial body weight measurements performed at 3-month intervals. For the latter, change in body energy stores was calculated assuming 7,800 kcal/kg for any change in weight. This was divided by 180 days to calculate the daily source of energy substrates from weight loss or storage of excess energy intake as weight gain (Forbes, 2000).

For the measurement of energy expenditure, women arrived at the research center at 5 p.m. between the 5th and 12th day after the start of their menstrual period. They were fed dinner but no further intake of food or beverage other than water was allowed after 7:30

p.m. DLW for measurement of total body water (TBW) and total energy expenditure (TEE) was administered at 8 p.m.. The doses were 0.12g D_20 (99.9 atom percent)/kg estimated TBW and 2.0 to 2.5g $H_2{}^{18}O$ (10 atom percent)/kg estimated TBW. The lower dose was used when there was a shortage of $H_2{}^{18}O$. Spot urine samples were collected immediately before deuterium was administered, at 8 a.m. the next morning, and at 8 a.m. on Day 15 at the research center. None of the samples were the first void of the day. Energy expenditure was calculated using Equation A6 as published elsewhere (Schoeller et al., 1986) along with dilution space ratios of Racette et al. (1994). DLW gives the rate of carbon dioxide production. Because carbon dioxide production can vary based on which type of macronutrient is being consumed, information from the 7-day food records was used along with the modified Weir equation (Weir, 1949) to calculate energy expenditure as described by Black, Prentice, and Coward (1986). Respiratory Quotient was assumed to equal the food quotient, which was calculated from macronutrient intake using the 7-day food diary and assuming that all food was completely oxidized to carbon dioxide and water.

The mean (± *SD*) ratio of ^2H and ^{18}O dilution space was 1.034 ± 0.015; the ^2H and ^{18}O elimination rates were 0.10 ± 0.028/day and 0.126 ± 0.027/day, respectively. DLW-estimated kilocalorie intake per day (kcal/day) was calculated using TEE data and weight change over the 3 months adjacent to the measured TEE. The equation used for each participant was as follows: DLW kcal/day = TEE + [(weight change for 90 days × 7,800) / 90 days]. The 7,800 kcal/kg is an estimate of the energy density of adipose tissue (Forbes, 2000). Self-reported energy intake from the 7-day food diaries was subtracted from the DLW-calculated energy intake to determine underreporting in kcal/day.

Estimated energy requirement (EER) in kcal/day was calculated based on the 2002 Dietary Reference Intake (DRI). DRI values were chosen because they are based on the most recent scientific knowledge with respect to energy requirements for healthy populations. The DRI prediction requires an estimate of physical activity level (PAL), which was calculated from TEE divided by basal energy expenditure (PAL = TEE / BEE). The measured PAL values ranged from 1.0 to 2.5 and fit into one of four categories: sedentary (PAL ≥ 1.0 < 1.4), low active (PAL ≥ 1.4 < 1.6), active (PAL ≥ 1.6 < 1.9), and very active (PAL ≥ 1.9 < 2.5). Corresponding values from those categories were applied to the equation to calculate individual EER values.

The ratio of DLW-measured energy expenditure to predicted energy expenditure (TEE/EER) was calculated to permit a comparison of measured energy expenditure with predicted average values to assess whether the participants were hypometabolic. The TEE/EER was used to assess whether these participants had a reduced energy requirement as compared to the general population when matched for age, height, weight, and gender.

Results and Discussion

The average self-reported caloric intake on the food records was 1,683 (*SD* = 307), whereas the average DLW estimate of daily caloric intake was 2,406 (*SD* = 337). Thus, the average participant underreported caloric intake by 733 calories (*SD* = 326). Replicating results from previous studies of overweight and obese individuals (Bandini et al., 1990; Lichtman et al., 1992; Prentice et al., 1986), participants underreported their daily caloric intake by an average of 30%.

Table 2 reports the correlations between TFEQ–Restraint scores, self-reported kcal/day, DLW kcal/day, kcal/day underreporting, the ratio of estimated expenditure to predicted expenditure, and BMI in Study 2. TFEQ–Restraint scores did not show a statistically significant or substantially meaningful inverse correlation with objectively measured caloric intake over the 2-week period (r = .11). TFEQ–Restraint scores were not significantly correlated with self-reported caloric intake (r = –.23). As observed previously (Bathalon et al., 2000; Bingham et al., 1995), underreporting of caloric intake correlated with TFEQ–Restraint scores (r = .33), suggesting that those with elevated dieting scores showed a greater underreporting tendency. DLW caloric intake correlated with underreporting (r = .57), indicating that the more participants ate, the more they underreported intake.

Because many researchers have taken a median split on the continuous dietary restraint scales (e.g., Jansen, 1996; Polivy et al., 1978), we compared those above and below the median on the TFEQ–Restraint scale in the present sample. DLW estimates of caloric intake at baseline for those above and below the median on the TFEQ–Restraint scale were 2,401 (±342) and 2,424 (±335) kcal per day, respectively. These data suggest that those above the median on the TFEQ–Restraint score consumed an average of 23 fewer kcal per day—a difference equivalent to three animal crackers.

Table 2 also reports the correlation between the ratio of the DLW measure of estimated expenditure to predicted energy expenditure and the TFEQ–Restraint scores to determine whether those with a high restraint scores had a low energy expenditure for their gender, weight, height, and age due to a successful long-term calorie restriction or a biologically low energy requirement. However, this relation was not statistically significant or substantive in magnitude (r = .12), suggesting that individuals with elevated dietary restraint scores did not exhibit lower energy expenditure on average.

Thus, Study 2 suggests that individuals with elevated dietary restraint scores did not consume significantly or substantively less calories than do their lower scoring counterparts when a biological measure of total caloric intake over a 2-week period served as the

Table 2
Correlations Between Study Variables, Study 2

Variable	1	2	3	4	5	6
1. Three-Factor Eating Questionnaire-Restraint	—	−.23	.11	.33	.12	−.20
2. Self-reported kilocaloric intake per day[a]		—	.49**	−.43**	.35*	.22
3. Doubly labeled water, kcal intake per day			—	.57***	.16	.19
4. Underreporting, kcal intake per day				—	.38*	−.01
5. Ratio of measured EE to predicted EE					—	.22
6. Body mass index						—

Note. EE = energy expenditure.
[a] From 7-day food records.
* $p < .05$. ** $p < .01$. *** $p < .001$.

criterion. This was true whether the dietary restraint scale was modeled as a continuous variable or a median split. However, because the sample was small and atypical by virtue of recent significant weight loss, we sought to replicate these relations with data from a larger, representative sample of healthy women.

Study 3

Method

Participants and Procedures

Participants in this study were part of the Observing Protein and Energy Nutrition study that was conducted from September 1999 to March 2000 in Montgomery County, Maryland.

Subar and associates (2003) reported details of recruitment procedures. A total of 223 healthy women were recruited (*M* BMI = 27.6, *SD* = 6.0; 10% Black, 78% White, 2% Hispanic, and 10% who specified other or mixed racial heritage). Exclusion criteria included diabetes, congestive heart failure, kidney failure, fluid retention, malabsorption, hemophilia, conditions requiring supplemental oxygen, inability to read English, pregnancy, or being on a weight-loss diet. Participants completed three study visits: baseline, 11 to 14 days later, and 3 months following their baseline visit. The data used in the present report were drawn from these three assessments.

Measures

TFEQ–Restraint scale. Participants completed the TFEQ–Restraint scale (Stunkard & Messick, 1985) at baseline (see *TFEQ–Restraint scale* section of Study 1 for psychometric details).

Self-reported caloric intake. Two methods of self-reported dietary intake were used in this study. At baseline and 3 months later, participants completed a 24-hr dietary recall that used a standardized five-pass method described elsewhere (Subar et al., 2003). Total energy intake (kcal/day) was averaged between those 2 days. Participants also completed the Diet History Questionnaire (DHQ), a food frequency questionnaire developed by the National Cancer Institute (Subar et al., 2001) that is described elsewhere (Subar et al., 2003). Briefly, the DHQ examines intake for 124 food items over the past 12 months and asks about portion sizes by providing a choice of three ranges. Participants completed the DHQ prior to their baseline visit.

Objective measure of caloric intake. Energy intake was calculated from TEE plus 3-month change in body energy stores in the same manner as in Study 2. TEE was measured by DLW as described in detail elsewhere (Trabulsi et al., 2003). The measurement was performed on an outpatient basis. Participants were instructed not to eat or drink for 3 hr prior to their visit. After a baseline spot urine specimen was obtained, participants were dosed with DLW, providing approximately 2 g of 10 atom percent $H_2{}^{18}O$ and 0.14 g 99.9 atom percent ^2H_2O per kg TBW. Spot urine specimens were also collected 2, 3, and 4 hr after dosing, and participants were allowed to leave the clinic. During this clinic visit, participants were not allowed liquids or foods except for up to 600 mL of a liquid replacement meal between 1 and 3 hr after the dose. Participants returned to the clinic after a 3-hr fast approximately 14 days after DLW dosing and provided two additional spot urines.

Results and Discussion

Table 3 reports the correlations between the TFEQ–Restraint scores, self-reported caloric intake from the 24-hr dietary recall measures, self-reported caloric intake from the DHQ, and the DLW estimate of kcal per day. TFEQ–Restraint scores did not show a statistically significant or substantially meaningful inverse correlation with objectively measured caloric intake over a 2-week period ($r = -.10$). TFEQ–Restraint scores were significantly inversely correlated with self-reported caloric intake as assessed by both the 24-hr recall and the food frequency questionnaire, though the relations were modest (rs = –.17 and −.21, respectively).

General Discussion

Results collectively suggest that one widely used dietary restraint scale does not appear to be a valid measure of dietary restriction. In Study 1, TFEQ–Restraint scores did not correlate with caloric intake during lunch meals consumed at work cafeterias over a 3-month period among adult women. In Study 2, TFEQ–Restraint scores did not correlate with DLW estimates of caloric intake over a 2-week period among women who had recently lost weight. In Study 3, TFEQ–Restraint scores did not correlate with DLW estimates of caloric intake over a 2-week period in a larger, representative sample of healthy women. Because we had adequate power to detect a medium effect ($r = .30$) in Study 1 (power = .82) and Study 3 (power > .99), but had less power in Study 2 (power = .40), it seems unlikely that the null findings in all three studies can be attributed to a lack of power, particularly given that the average correlation between the TFEQ–Restraint scale and objectively measured caloric intake was .04. These results converge with prior evidence that dietary restraint scales did not show substantively meaningful correlations with caloric intake assessed over multiple eating episodes (Jansen et al., 2003; Martin et al., 2005; Rolls et al., 1997; Sysko et al., 2005) or over multiple-week periods using DLW (Bathalon et al., 2000; Tuschl et al., 1990) in various populations.

The results from the three studies reported here converge with those of our previous dieting validity report (Stice et al., 2004), which found that five dietary restraint scales did not correlate with short-term caloric intake in four studies. The average correlation between dietary restraint scales and objectively measured short-

Table 3
Correlations Between Study Variables, Study 3

Variable	1	2	3	4	5
1. Three-Factor Eating Questionnaire-Restraint	—	−.17*	−.21**	−.10	−.15*
2. Self-reported caloric intake[a]		—	.16*	.21**	.10
3. Self-reported caloric intake[b]			—	.02	.14
4. Doubly labeled water, kcal intake per day				—	.43
5. Body mass index					—

[a] Repeated 24-hr recall. [b] Food frequency questionnaire.
* $p < .05$. ** $p < .01$.

term caloric intake ($r = -.07$) was similar to the average validity coefficient from the present study ($r = .04$). The fact that numerous studies conducted by independent labs have not provided evidence that dietary restraint scales show substantively meaningful inverse correlations with objective measures of short-, moderate-, and long-term caloric intake suggests that these scales are not valid measures of dietary restriction. It could be argued that dietary restraint scales would show stronger inverse correlations with caloric intake if dietary intake were assessed over even longer periods of time, yet this seems unlikely because the typical energy-restriction diet only lasts 4 to 6 weeks (French, Jeffery, & Murray, 1999; D. F. Williamson, Serdula, Anda, Levy, & Byers, 1992), though patients with bulimia nervosa retrospectively report dieting for 4 to 6 years before the onset of their eating disorder (Brewerton, Dansky, Kilpatrick, & O'Neil, 2000). The fact that individuals with elevated scores on dietary restraint scales show significantly greater weight gain over the subsequent 1 to 4 years than those with low scores on these scales (French, Jeffery, Forster, et al., 1994; Klesges et al., 1992; Stice et al., 2005) also seems to imply that these scales are not valid measures of very long-term caloric intake.

One limitation is that we only examined one dietary restraint scale. Although this should provide a demanding test of the validity question because it has been argued that the TFEQ–Restraint scale assesses successful dietary restraint (Heatherton et al., 1988), it would have been preferable if multiple dietary restraint scales had been included. However, the correlations across dietary restraint measures are typically quite high. For instance, the average correlation across the five dietary restraint scales examined in our first four validity studies was .63 (Stice et al., 2004). In addition, our first validity report examined five dietary restraint scales and other validity reports from independent labs examined additional dietary restraint scales; collectively, these studies found that seven dietary restraint scales used in the literature did not show substantively meaningful correlations with objective measures of caloric intake. Still, it would be reassuring if additional studies examined the concordance between long-term dietary intake and the dietary restraint scales. Although our objective measures of caloric intake (observed food purchases and DLW estimates of caloric intake) were an improvement over the self-report measures used in the original validity studies, they were not error free. It is certain that some participants did not consume all of the food that was purchased at work cafeterias. Likewise, error was introduced into DLW estimates of caloric intake by week-to-week variation in physical activity, random error associated with the measurement of DLW and body composition, and random as well as systematic errors associated with variation in energy balance in Study 3, in which body composition change was not assessed.

Another potential limitation is that because each of these three studies involved monitoring of body weight and eating, participants might have been likely to underreport caloric intake because of social desirability bias. However, the fact that Study 2, which blinded participants to the fact that DLW can be used to estimate caloric intake, produced similar validity findings to Studies 1 and 3, which did not blind participants, suggests that this issue did not dramatically influence the results.

The findings from our seven validity studies and from other validity studies that examined objectively measured caloric intake (Hetherington et al., 2000; Jansen, 1996; Jansen et al., 2003; Martin et al., 2005; Ouwens et al., 2003; Rolls et al., 1997; Sysko et al., 2005; van Strien et al., 2000) stand in contrast to the findings from past validation studies that documented significant negative correlations between dietary restraint scales and self-reported caloric intake (e.g., French, Jeffery, & Wing, 1994; Kirkley et al., 1988; Wardle & Beales, 1987). The most likely explanation for the discrepant findings is that self-reported caloric intake is inaccurate. Studies that used biological measures of actual caloric intake have revealed that people often underreport caloric intake (Bandini et al., 1990; Lichtman et al., 1992; Prentice et al., 1986). Such reporting inaccuracies are probably partially rooted in social desirability biases, as this underreporting is greatest for overweight individuals (Prentice et al., 1986). Indeed, underreporting of dietary intake correlates positively with social desirability scales (Maurer et al., 2006; Tooze et al., 2004).

It is noteworthy that the TFEQ–Restraint scale only showed significant correlations with self-reported caloric intake in two out of the four relations examined in this report ($M\ r = -.19$), because the original validity studies found that most dietary restraint scales showed significant inverse relations with self-reported intake. Of interest is the fact that the average correlation between dietary restraint scales and self-reported caloric intake from the original validity studies (French, Jeffery, & Wing, 1994; Neumark-Sztainer et al., 1997; van Strien et al., 1986; Wardle & Beales, 1987) was only slightly larger ($M\ r = -.22$), but these effects were often significant because the average sample sizes used in the validity studies were larger ($M = 314$). Parenthetically, because most of the original validity studies did not report the effect size, we employed meta-analytic procedures to estimate these effects. Thus, it appears that the original validation studies that used self-report caloric intake as the criterion provided only slightly stronger evidence for the validity of these scales.

One implication of the present results, and those from the other studies that have examined objectively measured caloric intake to validate dietary restraint scales, is that it may be wise to reinterpret the findings from studies that used these dietary restraint scales, including those suggesting that dietary restraint increases risk for future onset of bulimic pathology. If dietary restraint scales do not identify people who are restricting their caloric intake, then this seems to imply that it may not be dietary restriction that increases the risk for bulimic pathology, as has been concluded on the basis of prospective studies indicating that dietary restraint scales predict bulimic pathology onset (Fairburn, 1997; Huon, 1996; Levine & Smolak, 2006; Neumark-Sztainer, 2005). Stated differently, studies that use restraint scales to draw conclusions about the effect of diet-induced negative energy balance on the development of bulimic pathology have to be reconsidered because restraint scales do not appear to assess caloric restriction.

Another implication of the validity findings is that dietary restraint scales may assess relative dietary restriction rather than absolute dietary restriction. Because individuals are eating less than they normally eat or than they desire to eat, they may perceive this relative restriction as dietary restraint (Lowe & Levine, 2005; Timmerman & Gregg, 2003), despite the fact that they are not achieving the negative energy balance necessary for weight loss. This interpretation is consistent with the evidence that (a) intermittent dieters temporarily arrest a weight gain trajectory while they are attempting to engage in a weight-loss diet but do not lose weight (Presnell, Stice, & Tristan, in press), (b) dietary restraint

scores often increase among people placed on energy-restriction diets relative to nondieting controls (e.g., D. A. Williamson et al., 2007), and (c) individuals with elevated dietary restraint scores consume significantly more calories than those with low dietary restraint scores but do not feel that they have overeaten (Jansen, 1996). The fact that individuals with elevated dietary restraint scores have been found to be more likely to gain weight over time relative to individuals with lower dietary restraint scores (e.g., French, Jeffery, Forster, et al., 1994; Klesges et al., 1992; Stice et al., 2005) suggests that the former are not only unsuccessful at reducing their caloric intake below their energy needs on a sustained basis but are often eating beyond their energy needs and therefore gain weight over time (Lowe & Levine, 2005).

Future Directions

Although the present findings suggest that dietary restraint scales are not valid measures of caloric intake, it would be useful for independent labs to replicate these findings. It would be particularly useful if these studies used criterion variables that captured even longer term caloric intake and designs that assessed the fine-grained topography of caloric intake.

Given that dietary restraint scales do not appear to be valid measures of dietary restriction, a priority for future research will be to determine the latent construct that is assessed by these scales. The fact that these scales consistently predict bulimic pathology onset suggests that a resolution of this question would have important implications for etiologic and maintenance theories of this condition and for the design of prevention and treatment interventions.

Future research should also explore the possibility that there are qualitatively different types of dieting and that some forms increase and others decrease the risk for onset or persistence of bulimic symptoms (e.g., strict fasting vs. replacing high-fat foods with fruits and vegetables). For example, it has been theorized that severe diet-induced weight suppression may increase risk for persistence of bulimia nervosa (Butryn, Lowe, Safer, & Agras, 2006) based on the observation that patients with bulimia nervosa retrospectively report that significant weight loss preceded the development of bulimia nervosa (Brewerton et al., 2000). It is also possible that dieting that takes the form of acute fasting between episodes of overeating increases risk for bulimic pathology.

Another priority will be to develop a dieting scale that is a valid measure of dietary restriction. This is necessary before it will be possible to determine whether dietary restraint plays a role in the development or maintenance of eating pathology. Although DLW represents one attractive option, it is expensive. Another alternative is to take an empirical-keying approach that identifies items that prospectively predict weight loss but lack face validity (to circumvent reporter biases).

Future experiments should also test whether the effects of assignment to dietary restriction interventions on binge eating and bulimic symptoms are moderated by other risk factors for binge eating (e.g., emotional eating or impulsivity), as it is possible that for a subset of at-risk individuals, actual dietary restriction has deleterious effects (Fairburn, Cooper, Doll, & Davies, 2005). Although several dietary restriction experiments have studied individuals at risk for escalations in binge eating and bulimic symptoms—including those with elevated body mass, body dissatisfaction, and subthreshold bulimic pathology—other at-risk groups have not been systematically studied.

Finally, we hope future studies will experimentally manipulate "dieting as usual" as it naturally occurs in the real world, rather than through monitored energy-restriction interventions. We believe that a paradigm in which intermittent dieters are randomly assigned to engage in their regular weight-loss dieting behaviors for a period of time or to refrain from these behaviors may be a promising approach because it offers the inferential benefits of an experiment while retaining ecological validity by studying real-world dieting. Given the potent social desirability factors operating, we feel it will be vital to make use of experimental paradigms because they permit stronger inferential confidence.

References

Bacon, L., Keim, N. L., Van Loan, M. D., Derricote, M., Gale, B., Kazaks, A., & Stern, J. S. (2002). Evaluating a non-diet wellness intervention for improvement of metabolic fitness, psychological well-being and eating and activity behaviors. *International Journal of Obesity, 26,* 854–865.

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., et al. (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.

Black, A. E., Prentice, A. M., & Coward, W. A. (1986). Use of food quotients to predict respiratory quotients for the doubly-labelled water method of measuring energy expenditure. *Human Nutrition and Clinical Nutrition, 40,* 381–391.

Boggiano, M. M., Chandler, P. C., Viana, J. B., Oswald, K. D., Maldonado, C. R., & Wauford, P. K. (2005). Combined dieting and stress evoke exaggerated responses to opioids in binge-eating rats. *Behavioral Neuroscience, 119,* 1207–1214.

Brewerton, T. D., Dansky, B. S., Kilpatrick, D. G., & O'Neil, P. M. (2000). Which comes first in the pathogenesis of bulimia nervosa: Dieting or bingeing? *International Journal of Eating Disorders, 28,* 259–264.

Burton, E., & Stice, E. (2006). Evaluation of a healthy-weight treatment program for bulimia nervosa: A preliminary randomized trial. *Behaviour Research & Therapy, 44,* 1727–1738.

Butryn, M., Lowe, M. R., Safer, D., & Agras, W. S. (2006). Weight suppression is a robust predictor of outcome in the cognitive-behavioral treatment of bulimia nervosa. *Journal of Abnormal Psychology, 115,* 62–67.

Fairburn, C. G. (1997). Eating disorders. In D. M. Clark & C. G. Fairburn (Eds.), *Science and practice of cognitive behaviour therapy* (pp. 209–241). Oxford, England: Oxford University Press.

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders, 16,* 363–370.

Fairburn, C. G., & Cooper, Z. (1993). The eating disorder examination (12th ed.). In C. Fairburn & G. Wilson (Eds.). *Binge eating: Nature, assessment, and treatment* (pp. 317–360). New York: Guilford Press.

Fairburn, C. G., Cooper, Z., Doll, H. A., & Davies, B. A. (2005). Identifying dieters who will develop an eating disorder: A prospective, population-based study *American Journal of Psychiatry, 162,* 2249–2255.

Field, A. E., Camargo, C. A., Taylor, C. B., Berkey, C. S., Frazier, A. L., Gillman, M. W., & Colditz, G. A. (1999). Overweight, weight concerns, and bulimic behaviors among girls and boys. *Journal of the American Academy of Child and Adolescent Psychiatry, 38,* 754–760.

Forbes, G. (2000). Body fat content influences the body composition response to nutrition and exercise. *Annals of the New York Academy of Sciences, 904,* 359–365.

French, S. A., Jeffery, R. W., Forster, J. L., McGovern, P. G., Kelder, S. H., & Baxter, J. E. (1994). Predictors of weight change over two years among a population of working adults: The Healthy Worker Project. *International Journal of Obesity, 18,* 145–154.

French, S. A., Jeffery, R. W., & Murray, D. (1999). Is dieting good for you? Prevalence, duration and associated weight and behaviour changes for specific weight loss strategies over four years in US adults. *International Journal of Obesity, 23,* 320–327.

French, S. A., Jeffery, R. W., & Wing, R. R. (1994). Food intake and physical activity: A comparison of three measures of dieting. *Addictive Behaviors, 19,* 401–409.

Goodrick, G. K., Poston, W. S., Kimball, K. T., Reeves, R. S., & Foreyt, J. P. (1998). Nondieting versus dieting treatments for overweight binge-eating women. *Journal of Consulting and Clinical Psychology, 66,* 363–368.

Goran, M. I., Driscoll, P., Johnson, R., Nagy, T. R., & Hunter, G. (1996). Cross validation of body composition techniques against dual-energy X-ray absorptiometry in young children. *American Journal of Clinical Nutrition, 63,* 299–305.

Groesz, L. M., & Stice, E. (2007). An experimental test of the effects of dieting on bulimic symptoms: Impact of eating episode frequency. *Behaviour Research and Therapy, 45,* 49–62.

Hagan, M. M., Chandler, P. C., Wauford, P. K., Rybak, R. J., & Oswald, K. D. (2003). The role of palatable food and hunger as trigger factors in an animal model of stress induced binge eating. *International Journal of Eating Disorders, 34,* 183–197.

Heatherton, T. F., Herman, C. P., Polivy, J., King, G. A., & McGree, S. T. (1988). The (mis)measurement of restraint: An analysis of conceptual and psychometric issues. *Journal of Abnormal Psychology, 97,* 19–28.

Herman, C. P., & Mack, D. (1975). Restrained and unrestrained eating. *Journal of Personality, 43,* 647–660.

Hetherington, M. M., Bell, A., & Rolls, B. J. (2000). Pleasure and monotony: Effects of repeat exposure on pleasantness, preference and intake. *British Food Journal, 102,* 507–521.

Huon, G. F. (1996). Health promotion and the prevention of dieting-induced disorders. *Eating Disorders, 4,* 257–268.

Jansen, A. (1996). How restrained eaters perceive the amount they eat. *British Journal of Clinical Psychology, 35,* 381–392.

Jansen, A., Theunissen, N., Slechten, K., Nederkoorn, C., Boon, B., Mulkens, S., & Roefs, A. (2003). Overweight children overeat after exposure to food cues. *Eating Behaviors, 4,* 197–209.

Johnson, F., & Wardle, J. (2005). Dietary restraint, body dissatisfaction, and psychological distress: A prospective analysis. *Journal of Abnormal Psychology, 114,* 119–124.

Killen, J. D., Taylor, C. B., Hayward, C., Haydel, K. F., Wilson, D. M., Hammer, L., et al. (1996). Weight concerns influence the development of eating disorders: A 4-year prospective study. *Journal of Consulting and Clinical Psychology, 64,* 936–940.

Kirkley, B. G., Burge, J. C., & Ammerman, A. (1988). Dietary restraint, binge eating, and dietary behavior patterns. *International Journal of Eating Disorders, 7,* 771–778.

Klem, M. L., Wing, R. R., Simkin-Silverman, L., & Kuller, L. H. (1997). The psychological consequences of weight gain prevention in healthy, premenopausal women. *International Journal of Eating Disorders, 21,* 167–174.

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

Levine, M. P., & Smolak, L. (2006). *The prevention of eating problems and eating disorders: Theory, research, and practice.* Mahwah, NJ: Erlbaum.

Lichtman, S. W., Pisarska, K., Berman, E. R., 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,* 1893–1898.

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,* 797–806.

Martin, C. K., Williamson, D. A., Geiselman, P. J., Walden, H., Smeets, M., Morales, S., & Redman, S. (2005). Consistency of food intake over four eating sessions in the laboratory. *Eating Behaviors, 6,* 365–372.

Maurer, J., Taren, D. L., Teixeira, P. J., Thomson, C. A., Lohman, T. G., Going, S. B., & Houtkooper, L. B. (2006). The psychosocial and behavioral characteristics related to energy misreporting. *Nutritional Review, 64,* 53–66.

Neumark-Sztainer, D. (2005). *"I'm like so fat": Helping your teen make healthy choices about eating and exercise in a weight-obsessed world.* New York: Guilford Press.

Neumark-Sztainer, D., Jeffery, R. W., & French, S. A. (1997). Self-reported dieting: Who should we ask? What does it mean? Associations between dieting and reported energy intake. *International Journal of Eating Disorders, 22,* 437–449.

Ogawa, R., Strader, A. D., Clegg, D. J., Sakai, R. R., Seeley, R. J., & Woods, S. C. (2005). Chronic food restriction and reduced dietary fat: Risk factors for bouts of overeating. *Physiology and Behavior, 86,* 578–585.

Ouwens, M. A., van Strien, T., & van der Staak, C. P. F. (2003). Tendency toward overeating and restraint as predictors of food consumption. *Appetite, 40,* 291–298.

Pietrobelli, A., Faith, M., Allison, D., Gallagher, D., Chiumello, G., & Heymsfield, S. (1998). Body mass index as a measure of adiposity among children and adolescents: A validation study. *Journal of Pediatrics, 132,* 204–210.

Polivy, J., & Herman, C. P. (1985). Dieting and binge eating: A causal analysis. *American Psychologist, 40,* 193–204.

Polivy, J., & Herman, C. P. (1992). Undieting: A program to help people stop dieting. *International Journal of Eating Disorders, 11,* 261–268.

Polivy, J., Herman, C. P., & Warsh, S. (1978). Internal and external components of emotionality in restrained and unrestrained eaters. *Journal of Abnormal Psychology, 87,* 497–504.

Prentice, A., Black, A., Coward, W., Davies, H., Goldberg, G. R., Murgatroyd, P., et al. (1986). High levels of energy expenditure in obese women. *British Medical Journal, 292,* 983–987.

Presnell, K., & Stice, E. (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,* 166–170.

Presnell, K., Stice, E., & Tristan, J. (in press). An empirical investigation of the effects of naturalistic dieting on bulimic symptoms: Moderating effects of depressive symptoms. *Appetite.*

Racette, S. B., Schoeller, D. A., Luke, A. H., Shay, K., Hnilicka, J., & Kushner, R. F. (1994). Relative dilution spaces of ^{2}H and ^{18}O-labeled water in humans. *American Journal of Physiology, 267,* 585–590.

Reeves, R. S., McPherson, R. S., Nichaman, M. Z., Harrist, R. B., Foreyt, J. P., & Goodrick, G. K. (2001). Nutrient intake of obese female binge eaters. *Journal of the American Dietetic Association, 101,* 209–215.

Rolls, B. J., Castellanos, V. H., Shide, D. J., Miller, D. L., Pelkman, C. L., Thorwart, M. L., & Peters, J. C. (1997). Sensory properties of a nonabsorbable fat substitute did not affect regulation of energy intake. *American Journal of Clinical Nutrition, 65,* 1375–1383.

Schoeller, D. A., Ravussin, E., Schutz, Y., Acheson, K. J., Baertschi, P., &

Jequier, E. (1986). Energy expenditure by doubly labeled water: Validation in humans and proposed calculation. *American Journal of Physiology, 250,* 823–830.

Schoeller, D. A., Shay, K., & Kushner, R. F. (1997). How much physical activity is needed to minimize weight gain in previously obese women? *American Journal of Clinical Nutrition, 66,* 551–556.

Smolak, L., Levine, M., & Schermer, F. (1998). A controlled evaluation of an elementary school primary prevention program for eating problems. *Journal of Psychosomatic Research, 44,* 339–353.

Sterritt, G. M. (1962). Inhibition and facilitation of eating by electric shock. *Journal of Comparative and Physiological Psychology, 55,* 226–229.

Stewart, D. A., Carter, J. C., Drinkwater, J., Hainsworth, J., & Fairburn, C. G. (2001). Modification of eating attitudes and behavior in adolescent girls: A controlled study. *International Journal of Eating Disorders, 29,* 107–118.

Stice, E. (2001). A prospective test of the dual pathway model of bulimic pathology: Mediating effects of dieting and negative affect. *Journal of Abnormal Psychology, 110,* 124–135.

Stice, E., Fisher, M., & Lowe, M. R. (2004). Are dietary restraint scales valid measures of acute dietary restriction? Unobtrusive observational data suggest not. *Psychological Assessment, 16,* 51–59.

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

Stice, E., Presnell, K., Shaw, H., & Rohde, P. (2005). Psychological and behavioral risk factors for onset of obesity in adolescent girls: A prospective study. *Journal of Consulting and Clinical Psychology, 73,* 195–202.

Stice, E., Presnell, K., & Spangler, D. (2002). Risk factors for binge eating onset: A prospective investigation. *Health Psychology, 21,* 131–138.

Stice, E., Shaw, H., Burton, E., & Wade, E. (2006). Dissonance and healthy weight eating disorder prevention programs: A randomized efficacy trial. *Journal of Consulting and Clinical Psychology, 74,* 263–275.

Stice, E., Spoor, S., Presnell, K., & Shaw, H. (2007). *Dissonance and healthy weight eating disorder prevention programs: Long-term effects from a randomized efficacy trial.* Manuscript submitted for publication.

Stunkard, A. J., & Messick, S. (1985). The Three-Factor Eating Questionnaire to measure dietary restraint, disinhibition, and hunger. *Journal of Psychosomatic Research, 29,* 71–83.

Subar, A. F., Kipnis, V., Troiano, R. P., Midthune, D., Schoeller, D. A., Bingham, S., et al. (2003). Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: The OPEN study. *American Journal of Epidemiology, 158,* 1–13.

Subar, A. F., Thompson, F. E., Kipnis, V., Midthune, D., Hurwitz, P., McNutt, S., et al. (2001). Comparative validation of the Block, Willet, and National Cancer Institute food frequency questionnaires: The Eating at America's Table Study. *American Journal of Epidemiology, 154,* 1089–1099.

Sysko, R., Walsh, T. B., Schebendach, J., & Wilson, G. T. (2005). Eating behaviors among women with anorexia nervosa. *American Journal of Clinical Nutrition, 82,* 296–301.

Timmerman, G. M., & Gregg, E. K. (2003). Dieting, perceived deprivation, and preoccupation with food. *Western Journal of Nursing Research, 25,* 405–418.

Tooze, J. A., Subar, A. F., Thompson, F. E., Troiano, R., Schatzkin, A., & Kipnis, V. (2004). Psychosocial predictors of energy underreporting in a large doubly labeled water study. *American Journal of Clinical Nutrition, 79,* 795–804.

Trabulsi, J., Schoeller, D. A., Troiano, R., Subar, A. F., Sharbaugh, C., Kipnis, V., & Schatzkin, A. (2003). Precision of the doubly labeled water method in a large-scale application: Evaluation of a streamlined dosing protocol in the Observing Protein and Energy Nutrition (OPEN) study. *European Journal of Clinical Nutrition, 57,* 1370–1377.

Tuschl, R. J., Laessle, R. G., Platte, P., & Pirke, K. M. (1990). Differences in food-choice frequencies between restrained and unrestrained eaters. *Appetite, 14,* 9–13.

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

van Strien, T., Engels, E., van Staveren, W., & Herman, C. P. (2006). The validity of dietary restraint scales: Comment on Stice et al. 2004. *Psychological Assessment, 18,* 89–94.

van Strien, T., Frijters, J. E., van Staveren, W. A., Defares, P. B., & Deurenberg, P. (1986). The predictive validity of the Dutch Restrained Eating Scale. *International Journal of Eating Disorders, 5,* 747–755.

Wadden, T. A., Brownell, K. D., & Foster, G. D. (2002). Obesity: Responding to the global epidemic. *Journal of Consulting and Clinical Psychology, 70,* 510–525.

Wardle, J., & Beales, S. (1987). Restraint and food intake: An experimental study of eating patterns in the laboratory and in normal life. *Behavior Research and Therapy, 25,* 179–185.

Weir, J. B. (1949). New methods for calculating metabolic rate with special reference to protein metabolism. *Journal of Physiology, 109,* 1–9.

Williamson, D. A., Martin, C. K., York-Crowe, E., Anton, S. D., Redman, L. M., Han, H., & Ravussin, E. (2007). Measurement of dietary restraint: Validity tests of four questionnaires. *Appetite, 48,* 183–192.

Williamson, D. F., Serdula, M. K., Anda, K. F., Levy, A., & Byers, T. (1992). Weight loss attempts in adults: Goals, duration, and rate of weight loss. *American Journal of Public Health, 82,* 1251–1257.

Wilson, G. T. (2002). The controversy over dieting. In C. G. Fairburn & K. D. Brownell (Eds.), *Eating disorders and obesity: A comprehensive handbook* (2nd ed., pp. 93–97). London: Guilford Press.

Received February 19, 2007
Revision received August 8, 2007
Accepted August 9, 2007 ■