Published in final edited form as:
Body Image. 2013 January ; 10(1): 35–44. doi:10.1016/j.bodyim.2012.08.006.

Is Intensive Measurement of Body Image Reactive? A Two-Study Evaluation Using Ecological Momentary Assessment Suggests Not

Kristin E. Heron, Ph.D.[1] and **Joshua M. Smyth, Ph.D.**[1,2]
[1]Pennsylvania State University, University Park, PA
[2]Syracuse University, Syracuse, NY

Abstract

Intensive assessment methods (e.g., Ecological Momentary Assessment [EMA]) are increasingly used to capture body image experiences in daily life. One concern with EMA is multiple assessments may increase reactivity to internal or external cues, potentially biasing measurement. Reactivity to EMA was evaluated in two studies (Study 1: $N = 63$ female undergraduates, Study 2: $N = 131$ women with high body dissatisfaction/disordered eating). Participants completed five daily surveys on handheld computers for 1–2 weeks and body image-related questionnaires at the start and end of each study. Results showed no systematic changes in pre- and post-EMA measures or momentary EMA reports, suggesting women were not reactive to the EMA protocols. Completing 1–2 weeks of EMA does not appear to affect body dissatisfaction, mood, or attitudes in non-clinical or at-risk samples of women. These studies provide evidence that EMA methods can be used to assess real-world body image experiences without undue concern about measurement reactivity.

Keywords

Ecological Momentary Assessment (EMA); measurement reactivity; body image; body dissatisfaction; disordered eating; college women

Body image is a multifaceted construct that consists of a person's perceptions of, and attitudes towards, his or her body and appearance (Cash, Fleming, Alindogan, Steadman, & Whitehead, 2002). One cognitive-affective component of body image is body dissatisfaction, which has been defined as displeasure with some aspect of one's appearance (Rosen, Crowther, Tennenbaum, Hobfoll, & Stephens, 1992), and can occur when inconsistencies exist between perceptions of one's actual physical attributes and those one would like to, or think one should possess. The development of body image dissatisfaction can be influenced by many factors, including the extent to which women internalize the thin beauty ideal portrayed and reinforced in Western societies (Stice, 2002). Known as thin-ideal internalization, this process occurs when women assimilate the societal thin-body ideal

Corresponding Author: Kristin E. Heron, Ph.D., Pennsylvania State University, The 330 Building, Suite 105, University Park, PA 16802 Phone: 814-863-1069, Fax: 814-863-7678, keheron@psu.edu.

into their personal view of the world and these beliefs become guiding principles in the women's own lives (Stice, Ziemba, Margolis, & Flick, 1996; J. K. Thompson, van den Berg, Roehrig, Guarda, & Heinberg, 2004).

Research also suggests that the extent to which women objectify their bodies may be related to both thin-ideal internalization and body dissatisfaction (Moradi & Huang, 2008; Myers & Crowther, 2007). The self-objectification framework was born out of objectification theory, which argues that in Western societies women's bodies are treated as objects to be viewed and evaluated based on appearance. Girls and women learn to internalize an appearance-based evaluative perspective and begin to self-objectify their bodies even in the absence of others (Fredrickson & Roberts, 1997; Moradi, Dirks, & Matteson, 2005; Noll & Fredrickson, 1998). It is argued that women who have internalized the thin ideal may develop body dissatisfaction in part because their appearance plays a central role in their self-evaluation. As such, both thin-ideal internalization and self-objectification are important factors that are associated with the development of body image dissatisfaction (Myers & Crowther, 2007; Stice et al., 1996).

There is growing interest in understanding how state body image and body dissatisfaction may fluctuate in everyday life using a variety of intensive data collection methods (Cash et al., 2002; Melnyk, Cash, & Janda, 2004; Rudiger, Cash, Roehrig, & Thompson, 2007). In an effort to understand the dynamic, everyday changes in body image experiences, Ecological Momentary Assessment [EMA] methods are increasingly being used to study these processes (Colautti et al., 2011; Leahey, Crowther, & Ciesla, 2011; Leahey, Crowther, & Mickelson, 2007; LePage & Crowther, 2010; Ridolfi, Myers, Crowther, & Ciesla, 2011). Broadly, EMA refers to assessment techniques that involve collecting data in people's natural environments during their daily lives. These methods typically use mobile electronic technology (e.g., palmtop computers, cellular phones) and involve asking participants to report on internal states, behaviors, cognitions, activities, and/or events multiple times daily.

Although the advantages of EMA methods have been discussed at length elsewhere (Smyth et al., 2001; Smyth & Heron, 2012), in brief, this methodology offers three primary benefits over traditional self-report measures. First, self-report data are typically collected via retrospective assessments, requiring participants to summarize their experiences over some time period. This recall and summarization process is prone to systematic biases due to cognitive heuristics used in memory search and reconstruction (Smyth & Stone, 2003; Stone & Shiffman, 1994). Because EMA requests participants to report on current or very recent experiences, retrospective recall (and the associated biases) is greatly reduced. Second, EMA occurs as people are going about their daily lives, thus increasing generalizability and ecological validity. Research and clinical settings (e.g., laboratory, hospital, etc.) are artificial environments in the sense that they cannot reflect all aspects of individuals' lives. It is not always clear if or how behavior evaluated in these settings is related to participants' behavior outside of the laboratory. Third, because multiple assessments occur over a relatively short timeframe (i.e., minutes to hours), temporal relationships among variables can be explored. For instance, data regarding the associations between cognitive, affective, and behavioral aspects of body image or body dissatisfaction in everyday life can be collected, and how these relationships may differ between people or change over relatively short periods of time (e.g., minutes, hours, days) can be examined. Such data allow more complex and nuanced research and clinical questions about dynamic associations and processes that occur over time to be addressed (e.g., Smyth et al., 2007).

One concern with EMA methods is that frequent, real-time assessment may change people's experience of their natural environment. It is possible that repeated assessments could introduce cues that would alter social, psychological, and behavioral aspects of their lives,

and thus the data would not necessarily accurately reflect real-world processes (Hufford, 2007). Known as measurement reactivity, this process has been described as “the systematic biasing effects of instrumentation and procedures on the validity of one's data” (Barta, Tennen, & Litt, 2012, p. 108). Concerns regarding the effects of measurement are not unique to EMA methods, and in fact, the effects of observing or monitoring people on their subsequent behavior, attitudes and beliefs has been studied in various fields (e.g., mere measurement effects in consumer behavior research, guinea pig effect in performance research). The direction of expected reactive effects in response to intensive measurement protocol is often unclear. In most cases, reasonable explanations could be made for either the increase (i.e., sensitization) or decrease (i.e., habituation) of reports of cognitions, attitudes, and behaviors.

Studies designed to evaluate for reactivity to EMA have been conducted in a variety of health behaviors, including alcohol use (Hufford, Shields, Shiffman, Paty, & Balabanis, 2002), smoking (Rowan et al., 2007), and pain (Stone et al., 2003). In the most systematic study of EMA measurement reactivity, Stone and colleagues (2003) had pain patients complete either no EMA, or EMA diaries with different sampling frequencies: 3, 6, or 12 times daily. Pain ratings did not systematically vary as a function of assessment frequency, thus demonstrating no evidence for measurement reactivity. In research areas more closely related to the study of body image, findings have similarly shown completing an EMA protocol does not appear to affect reports of eating disorder behaviors in general (Stein & Corte, 2003) or binge eating more specifically (Le Grange, Gorin, Dymek, & Stone, 2002; Munsch et al., 2009). Leahey and colleagues (2007) have also tested for reactivity to EMA in a study of appearance-related social comparisons. In this study, women were assigned to either complete EMA or no EMA. Results showed no group differences in measures of body dissatisfaction, mood, or weight-related cognitions; although all participants self-reported increased awareness of social comparisons, no group differences emerged. Overall, across various topical areas and using different research designs, there is very little evidence that participants' self-reports are reactive to EMA protocols. As researchers have pointed out, however, merely citing this lack of reactivity in other research domains is insufficient when using intensive assessment methods in new research areas (Barta et al., 2012). Given the increased interest in using EMA methods to study dynamic fluctuations in body image constructs, it seems prudent to demonstrate intensive assessment methods can be used without undue concern for reactive processes on dynamic assessments of body image.

The present studies were designed to systematically evaluate potential measurement reactivity in body image-related assessments in response to intensive EMA protocols in two samples of young women. In both studies, participants completed traditional individual difference measures of body image-related constructs (body discrepancies, thin-ideal internalization, self-objectification, body image quality of life) before and after EMA. Palmtop computers were used to collect EMA, allowing compliance to be electronically verified. In Study 1, a general sample of college women was recruited, and in Study 2, a sample of college women who were screened for high body dissatisfaction and/or disordered eating behavior was recruited. These two samples were used because it is possible that EMA is non-reactive among a general sample of young women, but that individuals who already experience high levels of body- and eating-pathology may be more or less sensitive to intensive measurement procedures. The first goal of these studies was to examine whether individual difference measures of body image and related constructs were reactive to completing 1–2 weeks of an EMA protocol. We expected that there would not be a consistent pattern of changes in any individual difference measures before and after EMA.

The second aim of this study was to evaluate for evidence of reactivity within the momentary EMA data. This was done in two ways. First, we examined women's responses

to EMA body image items across the course of the studies to identify whether systematic responses patterns emerged, suggesting measurement reactivity. Although these analyses were exploratory, we did not expect to see changes in average daily body discrepancies measured via EMA. Second, Study 1 included an experimental manipulation of EMA content to test for reactivity in momentary reports. Specifically, participants completed two weeks of EMA, with mood measured during both weeks, and body discrepancy measures included only during the second week. This design allowed us to test whether multiple daily assessments of body dissatisfaction would alter participants' EMA reports of negative mood. If reports of negative mood differed across weeks this could suggest mood is reactive to body image EMA survey content. Women are consistently surrounded by messages about body image (e.g., from media, peers) and, therefore, we did not expect that five daily EMA surveys would have a significant additional influence on these college women's negative affect.

These studies expand previous research in several important ways. First, these are the first studies to objectively track EMA compliance (using palmtop computers) in an evaluation of body image reactivity. The findings reported by Leahey and colleagues (2011, 2007) demonstrating limited reactivity in appearance-related social comparisons are promising, but a limitation noted by the authors was the use of a paper diary system, which precluded objective tracking of EMA compliance. In the absence of objective compliance measures (i.e., in studies requiring participant to self-report the date/time when completing each assessment), research shows participants can grossly overestimate their actual compliance with the protocol (Stone, Shiffman, Schwartz, Broderick, & Hufford, 2002). Compliance concerns are particularly critical in studies of measurement reactivity because if participants do not complete assessments when prompted or instructed to do so, it would not be surprising to find minimal reactive effects. Second, these studies were designed to assess for potential reactive effects in both retrospective and momentary (i.e., EMA) assessments. To our knowledge, this is the first study to experimentally manipulate EMA content in order to evaluate for EMA reactivity in momentary reporting of body image constructs. Such an evaluation is important for establishing that intensive assessment methods can be used to capture body image in everyday life without significantly altering women's real-world experiences.

Study 1

Method

Participants—Female undergraduates were recruited from introductory level psychology classes. Any interested female student could volunteer to participate in this study as part of her course requirements. Sixty-three college women between the ages of 18 and 22 (M = 19.04, SD = 0.79) completed the study. The majority of women were Caucasian (78%, n = 49), five (8%) were Black or African American, four (6%) were Latina or Hispanic, four (6%) were Asian American, and one (2%) was of mixed race.

Measures

Body-Image Ideals Questionnaire **(BIQ; Cash, 2000; Cash & Szymanski, 1995; Szymanski & Cash, 1995)—**The BIQ is a self-report measure of body image discrepancy or dissatisfaction that asks participants to rate how much 11 actual body attributes (e.g., weight, facial features, body proportions, etc.) resemble their ideal (discrepancy score) and separately rate the importance of their personal ideal for each attribute (importance score). Each discrepancy item is rated on a 4-point scale (0 = *exactly as I am*, 3 = *very unlike me*; sample item: “My ideal weight is:”). The importance of each attribute is also rated on a 4-point scale (0 = *not important*, 3 = *very important*; sample item: “How important to you is

your ideal weight?"). A composite score is calculated as the mean of the item-by-item product of discrepancy and importance ratings. Prior to calculating the composite scores, discrepancy scores of 0 are recoded to a value of −1 to allow for self-congruent items (*exactly as I am*) to be weighted by importance (Cash, 2000). Composite scores range from −3 to +9 with higher scores indicating a larger actual:ideal discrepancy with a greater emphasis on physical ideals. Concurrent validity of the BIQ has been demonstrated, as it is significantly correlated with body dissatisfaction (Cash, 1989) and body image dysphoria (Cash, 1994) among samples of young women. In the present sample, the Cronbach's alpha of the composite index was .73 and the alphas for the discrepancy and importance subscales were .78 and .76, respectively.

Contour Drawing Rating Scale **(CDRS; M. A. Thompson & Gray, 1995)—**This scale includes a series of nine female figures of increasing size. Participants selected the figure that most closely resembled the one they perceived to represent their actual figure (*actual*), ideal figure (*ideal*), and the figure they believe they should or ought to look like (*ought*). Women's understanding of these constructs was assessed prior to completing the measure. Discrepancy scores were calculated as the difference between the *actual* figure rating and the *ideal/ought* figure ratings, resulting in two discrepancy scores (actual:ideal [A:I] and actual:ought [A:O] discrepancies). Larger positive discrepancies reflect a desire to (A:I) or belief that one should (A:O) be thinner. Actual:ideal discrepancies from this measure are significantly correlated ($r = .74$) with discrepancies reported using a digital image manipulation technique, suggesting concurrent validity for its scores among a sample of young women (Rowe, McDonald, Mahar, & Raedeke, 2005).

Ideal-Body Stereotype Scale **(IBSS; Stice et al., 1996)—**The purpose of this measure is to assess the extent to which individuals have internalized a thin ideal body image portrayed in Western societies. Participants indicate on a 5-point scale the extent to which they agree with 10 items (1 = *strongly disagree*, 5 = *strongly agree*; sample item: "Thin women are more attractive."). The total scale mean was calculated, with higher scores reflecting greater agreement with the thin ideal body image. The reliability and validity of this measure's scores have been previously established in female adolescent samples (Stice & Agras, 1998; Stice et al., 1996). Cronbach's alpha in Study 1 for the IBSS was .87.

Self-Objectification Questionnaire **(SOQ; Noll & Fredrickson, 1998)—**This measure assesses the extent to which individuals view their bodies in appearance-based terms. Respondents rank order 12 body attributes based on how relevant each is to their personal self-concept. Six attributes are appearance based (e.g., physical attractiveness) and six are competence-based (e.g., physical fitness). The rank orders are summed for the appearance and competence based items separately and a difference score is computed (ranging from −36 to +36). Higher positive scores indicate a greater personal emphasis on appearance, which is interpreted as greater self-objectification (Noll & Fredrickson, 1998). Adequate concurrent validity for this measure's scores has been documented in samples of undergraduate women, as it is significantly correlated with body shame ($r = .29$), bulimia symptoms ($r = .37$) and anorexia symptoms ($r = .31$; Noll & Fredrickson, 1998).

EMA surveys—Customized surveys were developed using Satellite Forms MobileApp Designer (Intellisync Corporation, version 6.0.0, http://www.satelliteforms.net), a software development environment that can be used to create applications (i.e., survey programs) for Palm OS devices. The EMA surveys contained items developed for this study and items adapted from validated measures. The items developed for this study included questions regarding participants' current and recent location, activity, exercise, eating, media use, and social interactions; these items were not used for the purpose of the present study. Three

well-validated questionnaires were also adapted to be administered using EMA. These measures are described below:

***EMA-adapted Depression Anxiety and Stress Scale* (DASS; Lovibond & Lovibond, 1995)—**The DASS is a 21-item questionnaire used to assess negative mood (shortened version of the original 42-item scale). It contains three 7-item subscales measuring depression, anxiety, and stress. All items are rated on a 4-point scale (0 = *not at all*, 3 = *very much*), and scale scores are calculated as the sum of all scale items. The DASS was originally tested on undergraduate college students and has since been used with clinical and community samples. This measure was selected because it can be administered over short time periods, as is required for an EMA protocol (Lovibond & Lovibond, 1995), and the instructions for each item were adjusted to inquire about current emotional state. The psychometric properties for the DASS are good, with the depression subscale correlated at r = .77 with the Beck Depression Inventory, and the anxiety subscale correlated with the Beck Anxiety Inventory at r = .84 (Lovibond & Lovibond, 1995). Previous research has demonstrated that this measure yields internally consistent scores among a sample of community volunteers, with Cronbach's alphas for the depression, anxiety, and stress subscales of .94, .87, and .91, respectively (Antony, Bieling, Cox, Enns, & Swinson, 1998). The internal consistency for the EMA-administered DASS subscales in the present sample was .97, .94, and .97 for the depression, anxiety, and stress subscales, respectively.

***EMA-adapted Contour Drawing Rating Scale* (CDRS; M. A. Thompson & Gray, 1995)—**The CDRS (described above) was adapted to be administered on palmtop computers to assess momentary body image discrepancy. The three questions from the CDRS were viewed on the palmtop computer screen. A laminated copy of the nine numbered figures was attached to the inside cover of the computer screen (i.e., above the screen) and participants were directed to look at the figures and record their response to each question on the device. When administered via EMA, the instructions were adapted to inquire about current *actual*, *ideal*, and *ought* bodies.

***EMA-adapted Body-Image Ideals Questionnaire* (BIQ; Cash, 2000)—**Each of the BIQ items (described above) appeared in their original format on the palmtop computer screen where respondents recorded their responses. The instructions for the EMA-administered version were adapted such that participants reported on current body image discrepancies. Cronbach's alphas for the EMA-adapted BIQ composite score, discrepancy subscale, and importance subscale were .88, .90, and .93, respectively.

Equipment—PalmOne m105 handheld computers were used to administer the EMA surveys. These devices have 160×160 pixel LCD display screens and use the Palm OS® operating system. The customized surveys were downloaded onto these palmtop computers and participants completed surveys directly on the display screen using an included stylus. The palmtop computer prompted participants to complete five surveys daily; participants were unaware as to exactly when the palmtop computer would signal. Prompts were scheduled between 9:00 am and 11:00 pm. Alarm times were selected by dividing the 14 hours of possible assessment time each day into five equal time segments (one for each signal) with a 20-minute gap between each epoch to ensure the alarm times did not overlap; one alarm occurred randomly within each epoch. The alarms occurred on average every 2–3 hours. This stratified random sampling schedule ensures assessments will sufficiently sample times across the day, yet not be easily anticipated by participants. After being signaled, participants had the option to “snooze” the alarm for up to 15 minutes before completing the survey and they could take as long as needed to complete the survey. The palmtop computers automatically recorded a time and date stamp for each EMA survey

participants completed, thus allowing for an objective measure of compliance to be calculated.

Procedures—All study procedures were approved by the university's Institutional Review Board. College women were invited to take part in a study about body image and everyday experiences of young women. During the initial meeting (conducted in groups of 4–12 women), study procedures were described and all women provided informed consent to participate. All participants then completed the demographic information and the questionnaires described above. Special care was taken to ensure that participants understood the instructions for all questionnaires prior to completion. They were then provided with palmtop computers programmed with the customized survey and received a tutorial on care, general operation of the palmtop computers, and instructions for completing the daily assessments. Participants began the study on different days of the week to ensure that day of week and day of EMA were not confounded.

In order to evaluate whether completing multiple daily assessments of body image discrepancy influenced momentary negative mood, portions of the EMA survey content were systematically manipulated. During both weeks of the study, the surveys included general questions regarding participants' daily activities, location, and the depression and anxiety subscales of the DASS. In the first week, the EMA survey also contained items measuring stress (stressor occurrence, type, severity, and DASS stress subscale) but none assessing body image. During the second week of EMA, the stress items were removed and replaced with body image discrepancy questions (EMA-adapted CDRS and BIQ items described above). A similar number of stress and body image questions were used, and items had similar response formats to maintain a consistent assessment length across the study and ensure only the content (not content and survey length) were manipulated. EMA content was not counterbalanced (i.e., body discrepancies were always completed the second week), because the duration of any mood changes in response to the EMA discrepancy measures was unknown. We were concerned that if mood changes in response to EMA discrepancy items were present, these effects could carry over into the second week of EMA. This design provided measures of depressed and anxious mood that were completed both with (week 2) and without (week 1) intensive EMA body discrepancy measurement.

Participants completed the EMA for two weeks. At the end of the first week, participants returned to the research office for an individual appointment, during which time data were uploaded, compliance information was reviewed with participants, and the palmtop computers were returned to participants to complete a second week of EMA. After completing two weeks of EMA, participants returned to the research office to complete the questionnaires administered at the start of the study. Participants were debriefed and compensated (research credit toward their psychology course), thus concluding their participation.

Results

Compliance with the EMA protocol—In order to adequately address the aim of these studies, it was important to document that participants were compliant with the EMA protocol so that potential EMA reactivity could be assessed. Using the time and date stamps recorded by the palmtop computers, compliance rates were calculated as the percent of assessments completed by each participant. During the two weeks of EMA, participants completed 78% of all the scheduled assessments. The majority of participants (76%) completed at least 70% of the assessments during the two weeks of EMA. Furthermore, of the surveys completed by participants, 92% were finished within 20 minutes of the initial alarm, suggesting that not only were participants filling out the assessments, but they were

doing so in a timely manner. These compliance rates are consistent with previous studies using similarly demanding EMA protocols (Stone & Shiffman, 1994) and in EMA studies of body image (e.g., Leahey et al., 2007; LePage & Crowther, 2010).

The effect of day of study on compliance was also examined to determine whether there was systematic variation in EMA compliance over time. Compliance rates for day of study were calculated for each participant. A within-person repeated measures ANOVA was conducted using the day of study as the repeated factor and compliance rate was the dependent variable. Results showed that there was no significant effect of day of study on compliance, $F(13,868) = 1.19$, $p = .28$, $R^2 = .01$. Similar analyses were conducted using study week to predict compliance and found no significant difference between the compliance rates for week 1 (during with the EMA contained no body image discrepancy measures) and week 2 (EMA content included body discrepancy measures), $F(1, 124) = 0.47$, $p = .50$, $R^2 = .004$. Figure 1a provides a graphical illustration of compliance rates by day of EMA.

Pre- and post-EMA analyses—Table 1 presents the pre- and post-EMA means for each of the measures. A within-person MANOVA was performed on the five dependent variables: actual:ideal discrepancy, actual:ought discrepancy, BIQ score, IBSS score, and SOQ score. MANOVA was used because it takes into account the intercorrelations among the dependent variables and protects against inflated Type 1 error caused by conducting multiple repeated measure *t*-tests. All assumptions of linearity, homogeneity of variance-covariance matrices, multivariate normalcy, and the absence of multicollinearity were met, and there were no univariate or multivariate outliers. The independent variable in the model was assessment time (pre-EMA, post-EMA). Using Wilks' lambda criterion ($\lambda = .99$), the multivariate omnibus test was not significant, $F(5,119) = 0.02$, $p = .99$, $\eta^2 = .01$. Given that the overall F test was not significant, we do not report univariate results from the individual DVs (although all ps > .05). These results suggest there are no difference in measures of body image discrepancy, thin-ideal internalization, or self-objectification pre- and post-EMA.

Reactivity in EMA body image discrepancy reports—EMA body image discrepancy reports collected during the second week of the study were also examined for evidence of reactivity to the EMA protocol. We were interested in determining whether there was a systematic pattern of change in the reported mean level of EMA discrepancies during the week of assessment. Hierarchical linear modeling was used to analyze the data, where the multiple observations gathered for each participant are considered nested within assessment days and individuals. Multi-level random intercept models using PROC MIXED in the SAS statistical software package (SAS Institute Inc., Version 9.2) were used to produce maximum likelihood estimates, which control for autocorrelated residuals. The error variance was modeled at three levels, taking into account the autocorrelations at the individual, day, and prompt (beep) level. Results showed there was no effect of day of EMA on actual:ideal, $F(6,348) = 1.22$, $p = .30$, or actual:ought, $F(6,348) = 0.55$, $p = .77$, discrepancies measured using the CDRS, or the BIQ score, $F(6,348) = 0.40$, $p = .88$. These findings are graphical depicted in Figure 2. There is no systematic pattern suggesting women's reports of momentary body discrepancies on EMA-adapted state measures of the CDRS or BIQ change during the course of one week of EMA.

Affect reactivity to EMA—This study contained a design feature that allowed us to test whether completing multiple daily assessments of body discrepancy on palmtop computers would significantly alter women's momentary reports of negative mood (relative to completing EMA items regarding stress). On the measure of negative mood, participants reported no depressed or anxious affect during 61% and 63% of the EMA surveys respectively, resulting in a highly positively skewed distribution (absolute skew > 2.32). In

order to address this skewed distribution, we created dichotomous negative mood variables. Depressed and anxious affect were treated as separate variables, with each coded as present or absent for each EMA assessment point. Depressed affect was coded as present if the average DASS depression subscale score was greater than 0 on the 4-point scale (0 = not at all, 3 = very much), and was coded as absent if the average score equaled zero. Anxious affect was coded in the same manner using the average DASS anxiety subscale items. These two new variables provided an index of whether at a given EMA assessment participants experienced depressed/anxious mood at any level, or reported no negative mood.

Multi-level random intercept logistic regressions using PROC GLIMMIX in the SAS 9.2 statistical package were used to analyze these data. Depressed and anxious affect were reported at higher rates on the first assessment day (~54% of assessments) compared to the remaining 13 days (~36%), thus assessment day was included in models to account for the first day reporting difference. Two models estimated the odds of experiencing momentary depressed and anxious affect based on EMA content (body image discrepancy items vs. no discrepancy items), controlling for assessment day. Results showed there was no effect of EMA content on depressed mood, $OR = 1.21$, 95% CI: 0.83–1.59, $t(2811) = 0.80$, $p = .37$, or on anxious mood, $OR = 1.41$, 95% CI: 0.99–1.81, $t(2811) = 2.68$, $p = .10$. These results suggest that completing an intensive EMA protocol of body image dissatisfaction does not make women any more or less likely to report experiencing negative mood.

Study 2

Method

Participants—As part of a larger study, college women who reported high levels of body dissatisfaction and/or disordered eating behaviors and attitudes were recruited, thus, selecting women who were “at risk” for developing eating disorders. The screening procedures and criteria for inclusion are discussed in the Procedures below. The mean age of participants was 19.6 years old ($SD = 1.18$, range 18–24). The majority of participants were Caucasian (71%, $n = 93$), with other participants self-identifying as Asian (18%, $n = 24$), Hispanic ($n = 5$, 4%), Black ($n = 4$, 3%), Native American ($n = 4$, 3%) and Pacific Islander ($n = 2$, 1%). After beginning the study, two participants (1.5%) dropped out before completing the study, citing time constraints. Participants who dropped out did not differ from those completing the study on any demographic or psychological variables.

Measures

The Eating Disorder Examination-Questionnaire **(EDE-Q; Fairburn & Beglin, 1994)—**The original self-report version of the Eating Disorder Examination interview (Z. Cooper & Fairburn, 1987) was used during the participant screening process to assess for behavioral and attitudinal aspects of disordered eating. A total of twenty-two items measuring disordered eating behaviors and attitudes are rated. Fourteen of the items use a 7-point frequency response rating scale (0 = *no days*, 1 = *1–5 days*, 2 = *6–12 days*, 3 = *13–15 days*, 4 = *16–22 days*, 5 = *23–27 days*, 6 = *everyday*; sample item: “On how many days out of the past 28 days have you been deliberately trying to limit the amount of food you eat to influence your shape or weight?”), and eight items use a 7-point likert scale (0 = *not at all*, 6 = *markedly*; sample item: “Over the past 28 days, has your weight influenced how you think about yourself as a person?”). A total EDE-Q scale score is calculated as the average of all 22 items, and higher scores are interpreted as greater disordered eating behavior/attitudes. Research suggests an EDE-Q score of 2.3 or above (computed as the average of the 22 items) most appropriately identified women with potentially problematic disordered eating behaviors and attitudes (Mond, Hay, Rodgers, Owen, & Beumont, 2004). Correlations between the questionnaire and interview formats of this measure range from .79 to .81 for

the subscales, suggesting concurrent validity in a sample of female undergraduate students (Luce & Crowther, 1999). Cronbach's alpha for the total EDE-Q score in the present sample was .89.

The Body Shape Questionnaire **(BSQ; P. J. Cooper, Taylor, Cooper, & Fairburn, 1987)—**The BSQ is a 34-item self-report questionnaire used to evaluate fear of putting on weight, feelings of low self-esteem because of one's appearance, desire to lose weight, and body dissatisfaction. Higher scores are interpreted as greater body dissatisfaction. A total BSQ score of 110 (computed as the sum of the 34 items rated on a 1 to 6 scale; 1 = *never*, 6 = *always*) has been used to identify women at risk for developing eating disorders (Zabinski et al., 2001); this score was used to screen for potential participants in the present study. Scores on the BSQ have been shown to yield high internal consistency (Cronbach's alpha = .97), 3-week test-retest reliability (r = .88), and concurrent validity (r = .66) with other measures of body dissatisfaction among clinical samples (e.g., obese people seeking weight reduction) and non-clinical college student samples (Rosen, Jones, Ramirez, & Waxman, 1996). Cronbach's alpha for the total BSQ score in the present sample was .95.

The Contour Drawing Rating Scale **(CDRS; M. A. Thompson & Gray, 1995)—** The CDRS was administered and scored in the same way as in Study 1.

Sociocultural Attitudes Towards Appearance Questionnaire-3 (SATAQ-3; J. K. Thompson et al., 2004)—The SATAQ-3 assesses the extent to which women are aware of sociocultural thinness norms, experience pressure related to their appearance, and have internalized appearance standards. Respondents rate the extent to which they agree with 14 items (1 = *completely disagree*, 4 = *neither agree nor disagree*, 7 = *completely agree*) describing sociocultural thinness norms and pressures. Higher scores suggest greater awareness and internalization of appearance standards. In two samples of college women, Cronbach's alphas for the measure were .94–.96 (J. K. Thompson et al., 2004). The SATAQ converges with measures of drive for thinness (r = .54) and body dissatisfaction among female college students (r = .32; J. K. Thompson et al., 2004). Cronbach's alpha in the present sample was .82.

Body Image Quality of Life Inventory **(BIQOL; Cash & Fleming, 2002)—**The BIQOL quantifies how body image experiences affect various life domains, including sense of self, social functioning, emotional well-being, eating, exercise, and grooming. Respondents rated the effect of body image on 19 different aspects of their lives using a 7-point scale (−3 = *very negative effect*, −2 = *moderate negative effect*, −1 = *slight negative effect*, 0 = *no effect*, 1 = *slight positive effect*, 2 = *moderate positive effect*, 3 = *very positive effect*). Larger positive scores are interpreted as body image having a more positive impact on quality of life, and larger negative scores indicated a more negative effect of body image on quality of life. A score of zero suggests a neutral effect of body image on quality of life. Scores on the BIQOL are internally consistent (Cronbach's alpha = .95), have adequate 2-week test-retest reliability (r = .79), converge with body satisfaction (r = .66) and diverge with body image dissatisfaction (r = −.20) and BMI in a sample of female college students (r = −.21; Cash & Fleming, 2002). In the present sample, Cronbach's alpha was .93.

EMA survey—As with Study 1, the EMA survey was developed using Satellite Forms MobileApp Designer® and was downloaded to the palmtop computers. At each assessment, items regarding current location, activity, social interaction, stress, exercise, eating behaviors, media exposure, and mood were measured. The responses to the EMA measures were not used in analyses for the present study.

Equipment—Palm m105 and Z22 handheld computers were used for the EMA data collection. Both devices are of similar size, have 160×160 pixel LCD display screens and use the Palm OS® operating system. The most apparent difference between these two devices for participants was the color. As with Study 1, the customized surveys were downloaded onto the palmtop computers and participants completed surveys directly on the display screen using an included stylus. The devices time and date stamped all EMA surveys, allowing for compliance rates to be objectively tracked. The EMA sampling method was similar to Study 1, with women completing five daily surveys (at semi-random intervals) based on a signal-contingent assessment protocol. The palmtop computer provided auditory prompts between 9:00am and 10:00pm at semi-random times; a slightly earlier end time was used because some participants in Study 1 commented on the late alarm times. Participants were again unaware of the exact alarm times in advance, and the procedures for signaling participants were the same as Study 1.

Procedures—All study procedures were approved by the university's Institutional Review Board. Undergraduate women were recruited to participate in a larger research study regarding health experiences and life in college. Campus-wide flyers, class announcements, and online postings were used to recruit women to complete an online screening survey. The survey included a variety of questionnaires regarding college life, health, and psychosocial well-being (e.g., college adjustment, health behaviors, stress, mood, social support, etc.), in addition to the target screening measures, the EDE-Q and BSQ. These measures were used to identify women who reported current disordered eating behavior and/or body dissatisfaction. In this study, women completing the screening procedures were eligible to participate in the study if they had an average EDE-Q score ≥ 2.30 and/or a total BSQ score ≥ 110, did not report current diagnosis or treatment for an eating disorder, were between the ages of 18 and 24, and agreed to be contacted about additional studies. Both the EDE-Q and BSQ were used as screening measures to allow for a broader range of women with disordered eating behaviors or body image concerns to enroll in the study.

Seven hundred ninety-five women completed the online screening procedures. Of these, 37% (n = 296) met inclusion criteria. These individuals were contacted via email, provided with a more detailed description of the study (e.g., general study activities, frequency of appointments, duration of study), and instructed to contact the researcher via email or telephone if interested in participating. Of the eligible women, 48% (n = 141) did not respond to two email contact attempts by the researcher and 8% (n = 24) replied indicating they were not interested. Forty-four percent of eligible women (n = 131) enrolled in the study. Among participants in this study, 82 (63%) met both the EDE-Q and BSQ requirements for eligibility, 7 (5%) met only the BSQ criteria, and 42 (32%) met only the EDE-Q criteria. The EDE-Q and BSQ were correlated at r = .70. Participants in the study did not reliably differ from eligible women who did not participate on measures of disordered eating behavior or body dissatisfaction (ps < .26).

Women eligible to participate in the study based on the screening procedures were invited to participate in a study about daily experiences and life in college. Interested women attended an initial appointment in groups of 2−6 people in the research office, the study procedures were reviewed, and they provided informed consent. Participants completed the study questionnaires on a computer. The computers were separated with dividers to ensure privacy. Next, all participants were provided with palmtop computers and took part in a detailed training session regarding how to use and care for the equipment and information about completing the EMA surveys was reviewed. For the following week, participants carried the palmtop computer and completed the survey five times daily. As with Study 1, participants began the study on different days of the week to ensure that day of week and day of EMA were not confounded. Participants returned to the research office after one

week with the palmtop computers and again completed questionnaires (CDRS, SATAQ, BIQOL). Participants received monetary compensation for completing the research office visits and week of EMA surveys.

Results

Compliance with the EMA protocol—As with Study 1, compliance rates with the EMA protocol were calculated based on the time and date stamps provided by the palmtop computer, and the percent of assessments completed by each participant was calculated. Although participants completed the EMA for one week, they began the study at different times of day and, therefore, data from the first day of EMA were not used in compliance calculations. Participants in Study 2 completed 90% of all assessments during the study. The vast majority of women (91%) responded to at least 70% of the EMA prompts, suggesting very good compliance with the study protocol. A more conservative method for calculating compliance rates was also used in which only assessments completed within 30 minutes of the initial prompt were counted; using this method, 85% of the EMA surveys were completed, suggesting women both completed the EMA and did so promptly after being signaled.

The effect of day of EMA on compliance rates was examined using a within-person repeated measures ANOVA, where day of study was the repeated (within-person) factor and EMA compliance was the dependent variable. The percent of EMA surveys completed was used for these analyses. There was a significant effect of day of study on compliance, $F(5, 744) = 4.73$, $p < .001$, $R^2 = .03$. The direction of this effect suggests that compliance rates declined during the week of EMA as is seen in Figure 1b. The size of this effect is small and compliance remained very high throughout the week (>86%).

Pre- and post-EMA analyses—Total scores on measures of actual:ideal and actual:ought discrepancies (CDRS), thin-ideal internalization (SATAQ), and body image quality of life (BIQOL) were used in these analyses. Missing post-EMA data for dropouts ($n = 2$) were imputed using the mean replacement method. Although more sophisticated techniques are available for imputing missing data (e.g., maximum likelihood methods), the mean replacement method is appropriate and should not significantly bias results given the very low attrition rate (Tabachnick & Fidell, 2007). Analyses were conducted both with and without dropouts and there was no significant pattern of differences between these analyses; the reported results include all participants.

Table 1 presents the pre- and post-EMA means for each of the measures administered in Study 2. As is shown in Table 1, the baseline level of body discrepancies of participants in Study 2 is higher than those in Study 1, suggesting these samples differ in their overall levels of body dissatisfaction as intended. We tested whether there was a significant change in body image-related constructs before and after women completed EMA using a within-person MANOVA with four dependent variables: actual:ideal discrepancy, actual:ought discrepancy, SATAQ score, and BIQOL score. Assumptions of linearity, homogeneity of variance-covariance matrices, multivariate normalcy, and the absence of multicollinearity were checked and all assumptions were met. There were no univariate or multivariate outliers. The independent variable was assessment time (pre-EMA, post-EMA). Using Wilks' lambda criterion ($\lambda = .99$), the multivariate omnibus test was not significant, $F(4,257) = 0.02$, $p = .68$, $\eta^2 = .03$. We do not report the univariate results because the omnibus test was not significant, although all were non-significant ($p < .05$). Findings suggest no difference in measures of discrepancy, thin-ideal internalization, or body image quality of life before and after completing EMA.

General Discussion

Reactive effects have been described as the degree to which self-report data could be affected by the very act of assessing them (Nelson, 1977), an issue that is of particular concern in EMA studies because of the frequent and intensive assessment protocols used. The goal of the present studies was to evaluate the extent to which traditional measures of body image constructs and state measure of body image discrepancy and negative affect are reactive to an intensive EMA protocol in two samples of young women. In both studies, individual difference measures of body image-related constructs (body image discrepancy, thin-ideal internalization, self-objectification, and body image quality of life) were evaluated before and after participants completed EMA. Across the two studies, no significant changes in pre- to post-EMA scores were seen, providing evidence that traditional measures used to assess individual differences in body image and related constructs are not altered by completing EMA. These findings are consistent with EMA research in other areas, which similarly show limited evidence that EMA changes participants' self-reports of alcohol use (Hufford et al., 2002), smoking (Rowan et al., 2007), pain (Stone et al., 2003), eating disorder behaviors (Le Grange et al., 2002; Munsch et al., 2009; Stein & Corte, 2003), or social comparisons (Leahey et al., 2011, 2007).

The present studies expanded previous research in several important ways. First, Study 1 evaluated for evidence of reactivity within the EMA data itself. Results showed no systematic change in average EMA body discrepancies reported across study days, suggesting that at least over a 1-week period, state discrepancy ratings do not appear to be reactive to frequent assessment. This finding is consistent with research demonstrating frequent EMA pain ratings do not affect reported mean pain levels (Cruise, Broderick, Porter, & Kaell, 1996; Stone et al., 2003). Furthermore, in Study 1, EMA content was experimentally manipulated to evaluate whether EMA mood ratings are reactive to concurrent assessment of state body image discrepancies. It is not unreasonable that requiring women to complete multiple daily assessments of body dissatisfaction may negatively influence either global or momentary measures of body discrepancy and mood. Study 1 findings demonstrated that regardless of whether the EMA contained questions about body image discrepancy or stress, the likelihood that women reported depressed or anxious affect via EMA was similar. The present findings suggesting no reactivity in momentary reports are particularly novel and timely for researchers, as there is a growing interest in using EMA methods to study the real-world relationships between negative mood and body image experiences (Colautti et al., 2011; Leahey & Crowther, 2008; Leahey et al., 2011, 2007) and eating disorder behaviors (Smyth et al., 2007). Preliminary evidence from Study 1 indicates momentary reports of negative mood are not reactive to EMA assessments of body image discrepancies and suggests these constructs can be assessed together using intensive assessment methods such as EMA.

As was mentioned previously, when designing the EMA content manipulation the order of content was not counterbalanced; discrepancy items were always completed during the second week of the study. Counterbalancing was not used because we were concerned that if completing EMA body image measures did negatively influenced mood, this effect could carry over into the second week of data collection, thus diluting the manipulation and confusing interpretation of study results. It is possible that by the time body image items were introduced during the second week participants had already habituated to the EMA, eliminating potential momentary reactive effects that may have been present if administered during the first week. Although additional research is needed, if participants do habituate to completing an EMA protocol, having participants take part in 1–3 days of EMA during a brief “training” or “run-in” period prior to actual data collection, could be a useful strategy for reducing or eliminating reactive effects.

A second important contribution of the present studies is that, to our knowledge, these are the first to objectively assess EMA compliance in studies evaluating EMA reactivity of body image measurement. Objectively measuring compliance is especially important for reactivity because it provides evidence participants actually completed the “manipulation” (i.e., EMA) being evaluated. Research has shown that in the absence of electronic tracking, objective compliance rates are significantly lower than self-reported compliance (Stone et al., 2002). In the present studies, objectively-measured high compliance rates with the EMA protocol were seen (78–90% of all assessments completed). Women were compensated for participating, which could have improved compliance, although both the practice of providing compensation, and the compliance rates seen in these studies, are consistent with other EMA research studies (Leahey & Crowther, 2008; LePage & Crowther, 2010; Stone & Shiffman, 1994). The present studies complement previous work by providing more well documented evidence that reactivity effects in body image discrepancy and related measures are likely minimal even in the presence of objectively measured high EMA compliance.

Concerns regarding reactivity often stem from the idea that EMA methods can be thought of as a self-monitoring activity, and thus, may result in people altering their behaviors, attitudes, or perceptions. Researchers have attempted to capitalize on the potential beneficial effect of self-monitoring by using EMA. In a treatment study of patients with binge eating disorder (BED), researchers hypothesized that adding EMA to group CBT would result in a greater reduction of binge eating episodes than standard group CBT alone. However, they did not find that EMA improved the efficacy of CBT for BED (Le Grange et al., 2002). There appears to be conflicting evidence regarding the extent to which self-monitoring alone can induce behavior change (see Barta et al., 2012 for a review). In some ways, the present findings (and others) suggesting the lack of evidence of reactivity to EMA could seem counterintuitive. When placed in a larger context of behavior change programs, however, they may be less surprising. Consider the multitude of clinical treatments and therapies available, most of which involve multiple treatment components, regular contact with trained professionals, and occur over extended periods of time. Most of these treatments have modest effects on short-term symptom reduction and behavior change. Given this, it seems more reasonable to think that requiring people to complete a few brief assessments every day during relatively short periods of time, would not significantly alter behaviors, symptoms, thoughts, or attitudes.

Despite the lack of evidence supporting EMA measurement reactivity, researchers should be sensitive to the fact that the emergence of reactivity effects in a study likely depends on a several factors. Barta and colleagues (2012) have described a number of conditions that influence the likelihood of seeing EMA reactivity, including perceived desirability of the monitored behavior, number of behaviors being monitored, participant motivation for change, perceived demand for change, and assessment timing (i.e., whether the EMA assessment is completed before or after the target behavior). The duration of the EMA period and frequency of assessments may also influence likelihood of reactivity, although it is not entirely clear whether reactivity would be expected to increase or decrease over longer assessment periods and plausible predictions could be made for either finding. It seems reasonable that the longer and more frequently people are asked to complete assessments, the likelihood of becoming reactive to the high level of reporting would increase (i.e., sensitization). Alternatively, participants may also adjust to completing the EMA over many weeks as they become less sensitive to the monitoring itself and any reactivity effects present early on may disappear (i.e., habituation). Additional research evaluating reactivity over longer periods of time and using different sampling intensities may be useful. However, many EMA studies use a 1–2 week assessment period with several daily assessments, and thus, the present findings are likely fairly generalizable to much of the existing EMA research. Nonetheless, it is important when using EMA or other intensive assessment

methods that researchers are aware of content and design considerations that may be associated with increased likelihood of measurement reactivity.

Several limitations of the present studies should be acknowledged. First, both studies recruited undergraduate female college students, resulting in a relatively narrow age range (18–24 year olds). Previous studies of body image using EMA have also used college samples (Colautti et al., 2011; Leahey & Crowther, 2008; Leahey et al., 2011, 2007; LePage & Crowther, 2010; Ridolfi et al., 2011), although EMA reactivity in disordered eating behavior has been studied in female samples with larger age ranges (Le Grange et al., 2002; Munsch et al., 2009; Stein & Corte, 2003). Second, this study used several measures of body discrepancy and related constructs, but generalizing to other measures should be done with caution. Given there is evidence for limited EMA reactivity in studies of eating disorder behaviors and appearance-based comparisons, similar results may be seen with other body- and weight-related constructs as well. Third, although Study 1 utilized a within-subject comparison condition, these studies did not include a between-subject comparison group. If reactive effects were seen, the lack of a control group would be particularly problematic, as it would be unclear whether effects were due to EMA or other factors (e.g., time). Although no reactivity was seen in these studies, future research using a no-EMA control group is warranted. Fourth, we recognize it is very difficult to make strong conclusions from null findings. Our confidence in these results is strengthened by the fact that these were seen across two samples, effect sizes for changes were very small, and these null findings strongly converge with EMA reactivity research on related constructs (e.g., appearance comparisons, eating behaviors) and in other fields (e.g., alcohol use, smoking, pain).

Conclusions

As body image researchers move to understand the dynamic, real-world experience of body image and body dissatisfaction, they will continue to rely on intensive, repeated measure assessments, such as EMA. The goal of the present studies was to evaluate for measurement reactivity to an EMA protocol in various body image constructs in two samples of young women. Results showed no evidence that completing EMA influenced women's response to either traditional measures of body image and related constructs, or momentary EMA responses on palmtop computers. Findings were conceptually replicated in a general sample of undergraduate college women (Study 1) and in a sample of young women with high body dissatisfaction and/or disordered eating behavior (Study 2). These studies are the first to carefully examine EMA reactivity in body image, and suggest such methods can be used to assess the real-world experience of body image without undue concern for reactive processes.

Acknowledgments

A portion of this manuscript was funded by a National Research Service Award from NIHNIMH awarded to the first author (F31MH082564). NIH was not involved in the study design, collection, analysis, or interpretation of data, or preparation of this manuscript.

References

Antony MM, Bieling PJ, Cox BJ, Enns MW, Swinson RP. Psychometric properties of the 42-item and 21-item versions of Depression Anxiety Stress Scales in clinical groups and a community sample. Psychological Assessment. 1998; 10:176–181.

Barta, WD.; Tennen, H.; Litt, MD. Measurement reactivity in diary research. In: Mehl, MR.; Conner, TS., editors. Handbook of research methods for studying daily life. The Guilford Press; New York: 2012. p. 108-123.

Cash TF. Body-image affect: Gestalt versus summing the parts. Perceptual and Motor Skills. 1989; 69:17–18.

Cash TF. The contextual assessment of body-image affect: The Situational Inventory of Body-Image Dysphoria. The Behavior Therapist. 1994; 7:133–134.

Cash, TF. Manual for the Body-Image Ideals Questionnaire. 2000. Retrieved from www.body-images.com

Cash TF, Fleming EC. The impact of body image experiences: Development of the Body Image Quality of Life Inventory. International Journal of Eating Disorders. 2002; 31:455–460. [PubMed: 11948650]

Cash TF, Fleming EC, Alindogan J, Steadman L, Whitehead A. Beyond body image as a trait: The development and validation of the Body Image States Scale. Eating Disorders: The Journal of Treatment & Prevention. 2002; 10:103–113.

Cash TF, Szymanski ML. The development and validation of the Body-Image Ideals Questionnaire. Journal of Personality Assessment. 1995; 64:466–477. [PubMed: 16367722]

Colautti LA, Fuller-Tyszkiewicz M, Skouteris H, McCabe M, Blackburn S, Wyett E. Accounting for fluctuations in body dissatisfaction. Body Image. 2011; 8:315–321. [PubMed: 21840778]

Cooper PJ, Taylor MJ, Cooper Z, Fairburn CG. The development and validation of the Body Shape Questionnaire. International Journal of Eating Disorders. 1987; 6:485–494.

Cooper Z, Fairburn C. The Eating Disorder Examination: A semi-structured interview for the assessment of the specific psychopathology of eating disorders. International Journal of Eating Disorders. 1987; 6:1–8.

Cruise CE, Broderick J, Porter L, Kaell A. Reactive effects of diary self-assessment in chronic pain patients. Pain. 1996; 67:253–258. [PubMed: 8951918]

Fairburn CG, Beglin SJ. Assessment of eating disorders: Interview or self-report questionnaire? International Journal of Eating Disorders. 1994; 16:363–370. [PubMed: 7866415]

Fredrickson BL, Roberts T-A. Objectification theory: Toward understanding women's lived experience and mental health risks. Psychology of Women Quarterly. 1997; 21:173–206.

Hufford, MR. Special methodological challenges and opportunities in ecological momentary assessment. In: Stone, AA.; Shiffman, S.; Atienza, A.; Nebeling, L., editors. The science of real-time data capture: Self-reports in health research. Oxford University Press; New York: 2007. p. 54-75.

Hufford MR, Shields AL, Shiffman S, Paty J, Balabanis M. Reactivity to ecological momentary assessment: An example using undergraduate problem drinkers. Psychology of Addictive Behaviors. 2002; 16:205–211. [PubMed: 12236455]

Le Grange D, Gorin A, Dymek M, Stone AA. Does ecological momentary assessment improve cognitive behavioural therapy for binge eating disorder? A pilot study. European Eating Disorders Review. 2002; 10:316–328.

Leahey TM, Crowther JH. An ecological momentary assessment of comparison target as a moderator of the effects of appearance-focused social comparisons. Body Image. 2008; 5:307–311. [PubMed: 18585108]

Leahey TM, Crowther JH, Ciesla JA. An ecological momentary assessment of the effects of weight and shape social comparisons on women with eating pathology, high body dissatisfaction, and low body dissatisfaction. Behavior Therapy. 2011; 42:197–210. [PubMed: 21496506]

Leahey TM, Crowther JH, Mickelson KD. The frequency, nature, and effects of naturally occurring appearance-focused social comparisons. Behavior Therapy. 2007; 38:132–143. [PubMed: 17499080]

LePage ML, Crowther JH. The effects of exercise on body satisfaction and affect. Body Image. 2010; 7:124–130. [PubMed: 20153709]

Lovibond PF, Lovibond SH. The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behavior Research and Therapy. 1995; 33:335–343.

Luce KH, Crowther JH. The reliability of the Eating Disorder Examination--Self-Report Questionnaire Version (EDE-Q). International Journal of Eating Disorders. 1999; 25:349–351. [PubMed: 10192002]

Melnyk S, Cash TF, Janda L. Body image ups and downs: Prediction of intra-individual level and variability of women's daily body image experiences. Body Image. 2004; 1:225–235. [PubMed: 18089155]

Mond J, Hay P, Rodgers B, Owen C, Beumont P. Validity of the Eating Disorder Examination Questionnaire (EDE-Q) in screening for eating disorders in community samples. Behaviour Research and Therapy. 2004; 42:551–567. [PubMed: 15033501]

Moradi B, Dirks D, Matteson AV. Roles of sexual objectification experiences and internalization of standards of beauty in eating disorder symptomatology: A test and extension of objectification theory. Journal of Counseling Psychology. 2005; 52:420–428.

Moradi B, Huang YP. Objectification Theory and psychology of women: A decade of advances and future directions. Psychology of Women Quarterly. 2008; 32:377–398.

Munsch S, Meyer AH, Milenkovic N, Schlup B, Margraf J, Wilhelm FH. Ecological momentary assessment to evaluate cognitive-behavioral treatment for binge eating disorder. International Journal of Eating Disorders. 2009; 42:648–657. [PubMed: 19197978]

Myers TA, Crowther JH. Sociocultural pressures, thin-ideal internalization, self-objectification, and body dissatisfaction: Could feminist beliefs be a moderating factor. Body Image. 2007; 4:296–308. [PubMed: 18089276]

Nelson, RO. Assessment and therapeutic functions of self-monitoring. In: Hersen, M.; Eisler, RM.; Miller, P., editors. Progress in behavioral modification. Academic Press; New York: 1977. p. 3-41.

Noll SM, Fredrickson BL. A Mediational Model Linking Self-objectification, Body Shame, and Disordered Eating. Psychology of Women Quarterly. 1998; 22:623–636.

Ridolfi DR, Myers TA, Crowther JH, Ciesla JA. Do appearance focused cognitive distortions moderate the relationship between social comparisons to peers and media images and body image disturbance? Sex Roles. 2011; 65:491–505.

Rosen, JC.; Crowther, JH.; Tennenbaum, DL.; Hobfoll, SE.; Stephens, MAP. The etiology of bulimia nervosa: The individual and familial context. Hemisphere Publishing Corp.; Washington, DC: 1992. Body-image disorder: Definition, development, and contribution to eating disorders; p. 157-177.

Rosen JC, Jones A, Ramirez E, Waxman S. Body Shape Questionnaire: Studies of validity and reliability. International Journal of Eating Disorders. 1996; 20:315–319. [PubMed: 8912044]

Rowan PJ, Cofta-Woerpel L, Mazas CA, Vidrine JI, Reitzel LR, Cinciripini PM, Wetter DW. Evaluating reactivity to ecological momentary assessment during smoking cessation. Experimental and Clinical Psychopharmacology. 2007; 15:382–389. [PubMed: 17696685]

Rowe D, McDonald S, Mahar M, Raedeke T. Multitrait-multimethod investigation of a novel body image measurement technique. Research Quarterly for Exercise and Sports. 2005; 76:407–415.

Rudiger J, Cash TF, Roehrig M, Thompson JK. Day-to-day body-image states: Prospective predictors of intra-individual level and variability. Body Image. 2007; 4:1–9. [PubMed: 18089247]

Smyth, JM.; Heron, KE. Health Psychology. In: Mehl, MR.; Conner, TS., editors. Handbook of research methods for studying daily life. The Guilford Press; New York: 2012. p. 569-584.

Smyth JM, Stone AA. Ecological Momentary Assessment research in behavioral medicine. Journal of Happiness Studies. 2003; 4:35–52.

Smyth JM, Wonderlich SA, Crosby RD, Miltenberger R, Mitchell JE, Rorty M. The use of ecological momentary assessment approaches in eating disorder research. International Journal of Eating Disorders. 2001; 30:83–95. [PubMed: 11439412]

Smyth JM, Wonderlich SA, Heron KE, Sliwinski MJ, Crosby RD, Mitchell JE, Engel SG. 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. 2007; 75:629–638. [PubMed: 17663616]

Stein KF, Corte CM. Ecologic momentary assessment of eating-disordered behaviors. International Journal of Eating Disorders. 2003; 34:349–360. [PubMed: 12949927]

Stice E. Risk and maintenance factors for eating pathology: A meta-analytic review. Psychological Bulletin. 2002; 128:825–848. [PubMed: 12206196]

Stice E, Agras WS. Predicting onset and cessation bulimic behaviors during adolescence: A longitudinal grouping analysis. Behavior Therapy. 1998; 29:257–276.

Stice E, Ziemba C, Margolis J, Flick P. The dual pathway model differentiates bulimics, subclinical bulimics, and controls: Testing the continuity hypothesis. Behavior Therapy. 1996; 27:531–549.

Stone AA, Broderick J, Schwartz J, Shiffman S, Litcher-Kelly L, Calvanese P. Intensive momentary reporting of pain with an electronic diary: Reactivity, compliance, and patient satisfaction. Pain. 2003; 104:343–351. [PubMed: 12855344]

Stone AA, Shiffman S. Ecological momentary assessment (EMA) in behavorial medicine. Annals of Behavioral Medicine. 1994; 16:199–202.

Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient non-compliance with paper diaries. British Medical Journal. 2002; 324:1193–1194. [PubMed: 12016186]

Szymanski ML, Cash TF. Body-image disturbances and self-discrepancy theory: Expansion of the Body-Image Ideals Questionnaire. Journal of Social and Clinical Psychology. 1995; 14:134–146.

Tabachnick, BG.; Fidell, LS. Using multivariate statistics. 5th ed.. Pearson Education, Inc.; New York: 2007.

Thompson JK, van den Berg P, Roehrig M, Guarda AS, Heinberg LJ. The Sociocultural Attitudes Towards Appearance Scale-3 (SATAQ-3): Development and validation. International Journal of Eating Disorders. 2004; 35:293–304. [PubMed: 15048945]

Thompson MA, Gray JJ. Development and validation of a new body-image assessment scale. Journal of Personality Assessment. 1995; 64:258–269. [PubMed: 7722852]

Zabinski MF, Pung MA, Wilfley DE, Eppstein DL, Winzelberg AJ, Celio A, Taylor CB. Reducing risk factors for eating disorders: Targeting at-risk women with a computerized psychoeducational program. International Journal of Eating Disorders. 2001; 29:401–408. [PubMed: 11285577]

Highlights

- Measurement reactivity is a concern with Ecological Momentary Assessment [EMA] methods
- We tested for body image reactivity to an EMA protocol in two samples of young women
- Completing EMA did not affect body image measures in non-clinical or at-risk women
- EMA can be used to assess body image without undue concern for reactive processes

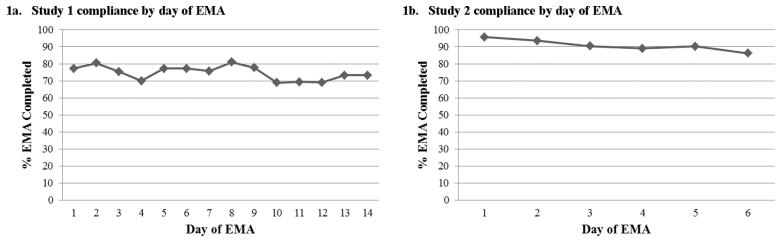

Figure 1.
EMA compliance rates by day of EMA.
Note. EMA = Ecological Momentary Assessment.

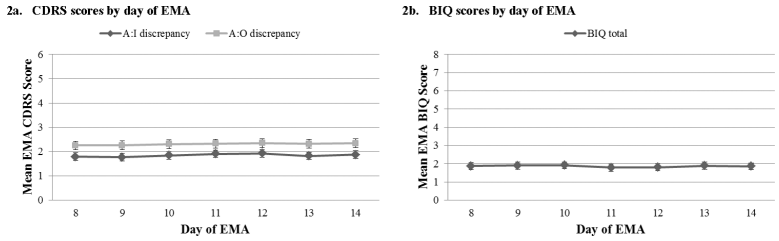

Figure 2.

Mean EMA-adapted body image discrepancy scores by day of EMA.

Note. BIQ = Body-image Ideals Questionnaire, CDRS = Contour Drawing Rating Scale, EMA = Ecological Momentary Assessment.

Table 1

Pre- and Post-EMA Descriptive Statistics for Body Image Measures and Related Constructs

	Study 1: General Undergraduate Sample (*N* = 63)		Study 2: High Body Dissatisfaction/Disordered Eating Sample (*N* = 131)	
	Pre-EMA *M (SD)*	**Post-EMA *M (SD)***	**Pre-EMA *M (SD)***	**Post-EMA *M (SD)***
Actual:ideal discrepancy (CDRS)	1.71 (1.40)	1.86 (1.27)	2.48 (1.03)	2.49 (1.10)
Actual:ought discrepancy (CDRS)	2.31 (1.55)	2.37 (1.48)	2.65 (1.26)	2.66 (1.36)
Ideal body discrepancy (BIQ)	1.91 (1.15)	2.07 (1.44)		
Thin-ideal internalization (IBSS, SATAQ)	3.43 (0.52)	3.41 (0.56)	5.08 (0.83)	5.06 (0.90)
Self-objectification (SOQ)	−0.98 (16.23)	−6.16 (16.40)		
Body image quality of life (BIQOL)			0.19 (1.03)	0.20 (1.03)

Note. EMA = Ecological Momentary Assessment, BIQ = Body-image Ideals Questionnaire, CDRS = Contour Drawing Rating Scale, IBSS = Ideal-Body Stereotype Scale, SATAQ = Sociocultural Attitudes Toward Appearance Questionnaire, SOQ = Self-Objectification Questionnaire, BIQOL = Body Image Quality of Life Questionnaire.