



Published in final edited form as:

Int J Eat Disord. 2017 October ; 50(10): 1222–1230. doi:10.1002/eat.22770.

Negative affect and binge eating: Reconciling differences between two analytic approaches in ecological momentary assessment research

Kelly C. Berg, PhD¹, Li Cao, MS², Ross D. Crosby, PhD^{2,3}, Scott G. Engel, PhD^{2,3}, Carol B. Peterson, PhD¹, Scott J. Crow, MD^{1,4}, Daniel Le Grange, PhD^{5,6}, James E. Mitchell, MD^{2,3}, Jason M. Lavender, PhD^{2,3}, Nora Durkin, MA¹, Stephen A. Wonderlich, PhD^{2,3}

¹Department of Psychiatry, University of Minnesota, Minneapolis, Minnesota

²Neuropsychiatric Research Institute, Fargo, North Dakota

³Department of Psychiatry and Behavioral Science, University of North Dakota, Fargo, North Dakota

⁴The Emily Program, St. Paul, Minnesota

⁵Department of Psychiatry and Behavioral Neuroscience, The University of Chicago, Chicago, Illinois

⁶Department of Psychiatry, University of California San Francisco, San Francisco, California

Abstract

Objective: Ecological momentary assessment (EMA) research has produced contradictory findings regarding the trajectory of negative affect after binge-eating episodes. Given the clinical implications, the objective of the current study was to reconcile these inconsistencies by comparing the two most commonly employed statistical approaches used to analyze these data.

Method: Data from two EMA studies were analyzed separately. Study 1 included 118 adult females with full- or subthreshold DSM-IV anorexia nervosa. Study 2 included 131 adult females with full-threshold DSM-IV bulimia nervosa. For each dataset, the single most proximal negative affect ratings preceding and following a binge-eating episode were compared. The times at which these ratings were made, relative to binge-eating episodes, were also compared.

Results: The results indicate that the average proximal pre-binge ratings of negative affect were significantly higher than the average proximal post-binge ratings of negative affect. However, results also indicate that the average proximal post-binge ratings of negative affect were made significantly closer in time to the binge-eating episodes (~20 min post-binge) than the average proximal pre-binge ratings of negative affect (~2.5 hr pre-binge). A graphical representation of the results demonstrates that the average proximal pre-binge and post-binge ratings map closely onto the results of previous studies.

Correspondence Kelly C. Berg, Department of Psychiatry, University of Minnesota, Minneapolis, MN. bergx143@umn.edu.

CONFLICT OF INTEREST

The authors indicate no conflict of interest.

Discussion: These data provide one possible explanation for the inconsistent findings regarding the trajectory of negative affect after binge eating. Moreover, they suggest that the findings from previous studies are not necessarily contradictory, but may be complementary, and appear to bolster support for the affect regulation model of binge eating.

Keywords

affect regulation; binge eating; ecological momentary assessment; negative affect

1 | INTRODUCTION

Robust empirical support exists for an association between negative affect and binge eating, regardless of eating disorder diagnostic status (Berg, Crosby, & Cao, 2015; Berg, Peterson, & Crosby, 2013, 2014; Crosby et al., 2009; Engel, Wonderlich, & Crosby, 2013; Haedt-Matt & Keel, 2015; Kenardy, Arnow, & Agras, 1996; Smyth et al., 2007; Telch & Agras, 1996). The affect regulation model, a theory that hypothesizes that binge eating is cued by negative affect and also functions to mitigate negative affect, provides one explanation for this association. An important clinical implication of this model is that binge eating is maintained through negative reinforcement (Polivy & Herman, 1993). Validating this model requires data demonstrating that negative affect is relatively high or increasing before binge eating occurs and relatively low or decreasing afterwards (Haedt-Matt & Keel, 2015). The evidence in favor of the first part of the hypothesis—that negative affect is relatively high or increasing before binge eating—is extensive and largely unequivocal (Berg et al., 2013; Crosby et al., 2009; Engel et al., 2013; Smyth et al., 2007). In clinical terms, the data consistently suggest that high or increasing negative affect serves as an antecedent or trigger for binge eating. However, research on the second part of the hypothesis is more mixed, with some evidence suggesting that negative affect decreases after binge eating (Engel et al., 2013; Smyth et al., 2007) and other evidence suggesting that negative affect actually increases after binge eating (Haedt-Matt & Keel, 2015). In other words, some studies suggest that binge eating effectively reduces negative affect, whereas other studies suggest that binge eating further intensifies it. Importantly, these inconsistent findings are not simply an academic conundrum; the two findings (whether negative affect increases or decreases after binge eating) suggest entirely different mechanisms by which binge eating is maintained, and as a result, have different implications for treatment. Thus, reconciling these mixed findings is essential.

Research on the association between negative affect and binge eating has been heavily influenced by ecological momentary assessment (EMA). EMA is a method used to assess variables of interest in the natural environment and in real-time, thus enhancing ecological validity and minimizing retrospective recall bias (Shiffman, Stone, & Hufford, 2008; Stone & Shiffman, 1994). Importantly, contemporary EMA protocols typically utilize computerized assessments that have the capability to time- and date-stamp each assessment. Although EMA data appear particularly well suited for determining the trajectory of negative affect after binge eating, these data have been fraught with inconsistent findings. A summary of these investigations is described below. As illustrated in the summary, two different statistical approaches have been used. The first approach uses multilevel modeling

to examine the trajectory of all available negative affect ratings made during the 4 hr before and the 4 hr after the occurrence of binge-eating episodes. For simplicity's sake, this will be referred to as "the multilevel modeling approach" (for more information on the multilevel modeling approach, please refer to the statistical methods section below). The aim of the second approach is to determine whether negative affect is higher right before or right after a binge-eating episode. This approach identifies the single negative affect ratings made closest in proximity to the binge-eating episode—both before the binge (prebinge) and after the binge (post-binge). The average proximal pre-binge negative affect rating is then compared to the average proximal post-binge negative affect rating using various statistical analyses (e.g., *t* tests, Cohen's *d*, etc.). For simplicity's sake, this will be referred to as "the single points approach".

In 2007, Smyth et al. published a seminal EMA study on the relationship between binge eating and negative affect in adult females with bulimia nervosa (BN) using the multilevel modeling approach. The results demonstrated that negative affect decreased significantly during the 4 hr after binge eating, providing support for the affect regulation model of negative affect. The authors concluded that, "...mood improves rapidly following a BN-event, suggesting the event (binge or vomit) has negatively reinforcing properties. ... The present findings offer empirical support for clinical techniques that emphasize emotion regulation skills and alternative action choices in the face of stress and/or NA." (p. 637).

In 2011, Haedt-Matt and Keel published a comprehensive meta-analysis of EMA data on the relationship between binge eating and negative affect. For each study included in the meta-analysis, the single-points approach was used. Overall, they found a moderate, positive effect size ($d = .50$, $SE = .07$), indicating that the post-binge rating of negative affect was higher than the pre-binge rating of negative affect. Thus, in contrast to the Smyth et al. (2007) study, this meta-analysis suggested that negative affect increased after binge eating. Perhaps most surprisingly, the meta-analysis reported that the individual effect size for the Smyth et al. (2007) study was also moderate and positive ($d = .56$, $SE = .08$), indicating that even in the Smyth et al. (2007) dataset, the single points approach suggested that negative affect increased after binge eating. Haedt-Matt and Keel (2015) concluded that their findings "challenge the widely accepted affect regulation model of binge eating" (p. 22). Clinically, they suggested that the "current results may be used to change how information regarding the function of binge eating is conveyed in treatment. Individuals with eating disorders may benefit from psychoeducation regarding predictable increases in negative affect as a consequence of binge eating. Given that the most likely outcomes are either no changes in negative affect or a worsening of negative affect, patients should be informed that binge eating is not an effective means of affect regulation in order to modify their beliefs about the effects of binge eating on mood" (p. 20).

Unfortunately, despite subsequent studies in additional samples, confusion over the trajectory of negative affect following binge eating has not abated. An additional EMA study was published that used similar methods and statistical analyses as those employed by Smyth et al. (2007). The study included adult women with anorexia nervosa (AN; Engel et al., 2013), and similar to Smyth et al. (2007), utilized the multilevel modeling approach to examine the trajectory of negative affect relative to binge eating. The results demonstrated a

significant decrease in negative affect after binge eating, replicating the Smyth et al. (2007) findings and adding support to the hypothesis that negative affect decreases after binge eating. However, in an interesting twist, Engel et al. (2013) also analyzed their data using the single points approach and found that the post-binge ratings were higher than the pre-binge ratings of negative affect. Results from the single point analysis approach thus replicated the Haedt-Matt and Keel (2015) findings, suggesting that negative affect increased after binge eating.

Given that the logical conclusions and clinical implications of these discrepant findings are in direct conflict with one another, reconciling these contradictory findings is critical. To this end, it is notable that the results, although discrepant, do not seem to be random. Instead, whether results demonstrate that negative affect increases or decreases after binge eating appears to be dependent on the type of statistical analyses used and which data points were included in these analyses. When the multilevel modeling approach has been used, the data consistently indicate that negative affect decreases after binge eating. In contrast, when the single points approach has been used, the data consistently indicate that negative affect increases after binge eating. Thus, it appears that the inconsistent findings may be the result of the different approaches to data analysis. However, it is unclear why results from these different analytic approaches would suggest such divergent conclusions about the same data. And importantly, it remains uncertain which conclusion better reflects the true, real world experience of emotions surrounding binge eating among individuals with eating disorders.

Revisiting the assumptions underlying the two types of analyses suggests one possible explanation for the discrepant findings. The multilevel modeling approach uses all available data points and locates each data point in time. In other words, each data point represents a given participant's negative affect of X intensity level at Y min before (or after) the binge-eating episode. In contrast, the single points approach takes the between-participants average of the single most proximal rating of negative affect made before each binge episode and compares it to the between-participants average of the single most proximal post-binge rating of negative affect. Notably, there is an implicit assumption in the single points approach that the most proximal pre-binge negative affect ratings are all made immediately prior to the binge and that the most proximal post-binge negative affect ratings are all made immediately after the binge. Or, at the very least, there is an implicit expectation that pre-binge and post-binge ratings are made at time points that are equidistant relative to the binge-eating episode (e.g., 15 min before and 15 min after). However, this critically important assumption has not been evaluated empirically.

1.1 | Objectives

The overarching objective of the current study was to examine the timing of pre-binge and post-binge negative affect ratings relative to binge-eating episodes, with the ultimate goal of reconciling the conflicting EMA findings related to negative affect after binge eating. The specific aims of the current study were as follows: (a) to compare the magnitude of the average proximal pre-binge and post-binge negative affect ratings, (b) to compare the timing, relative to the occurrence of binge-eating episodes, of the average proximal pre-binge and post-binge negative affect ratings, and (c) to graphically compare the average

proximal pre-binge and post-binge ratings of negative affect to the trajectory of negative affect before and after binge eating found using multilevel modeling. Accomplishing these aims will help to reconcile the conflicting findings between negative affect and binge eating consistently demonstrated between the two different types of analyses most commonly employed with EMA data.

2 | METHOD

Data from two EMA datasets (Engel et al., 2013; Smyth et al., 2007) were analyzed separately. The methods for these studies are described below in brief given that they have been described in detail elsewhere. The two studies will be referenced as follows: Study 1-AN (Engel et al., 2013) and Study 2-BN (Smyth et al., 2007). It should be noted that the datasets used for study 1-AN and study 2-BN have been utilized in several other studies. For example, the AN dataset has been used in a variety of papers examining specific aspects of the emotion-behavior relationship in AN (Engel et al., 2013; Haynos, Berg, & Cao, 2017; Lavender, Utzinger, and Crosby, 2016). Similarly, the BN data set has been utilized in several studies, which again have examined emotion-behavior patterns (e.g., Berg et al., 2013; Crosby et al., 2009; Smyth et al., 2007). We wish to emphasize that we believe that this current article is a unique contribution to the literature and that neither of these datasets have previously been used to examine the fundamental research question regarding the temporal characteristics of pre-binge and post-binge ratings of negative affect in an EMA study or to reconcile the inconsistent results of the single points and multilevel modeling approaches.

Importantly, the single points approach and the multilevel modeling approach have been conducted in both datasets previously (Engel et al., 2013; Haedt-Matt & Keel, 2015; Smyth et al., 2007). However, there were substantial inconsistencies between the statistical methods used in these studies (e.g., whether binge-only and binge-purge episodes were distinguished, whether significance testing or effect sizes were used in the single points approach). Given that reconciling the results of the single points and multilevel modeling approach require parallel analyses within each study as well as between studies, the analyses had to be rerun in both datasets.

2.1 | Participants

Study 1 included 118 adult females meeting criteria for full- or subthreshold DSM-IV AN (Engel et al., 2013) and Study 2 included 131 adult females meeting criteria for full-threshold DSM-IV BN (Smyth et al., 2007). Participants in the two studies were similar in terms of mean age (Study 1-AN: 25.3 ± 8.4 ; Study 2-BN: 25.3 ± 7.6), but differed with regard to average BMI (Study 1-AN: 17.2 ± 1.0 ; Study 2-BN: 23.2 ± 4.9). The majority of participants in both samples identified themselves as Caucasian (Study 1-AN: 96.0%; Study 2-BN: 96.9%).

2.2 | Measures

Both studies used the measures described below and all measures were utilized in the same way, except where explicitly noted.

2.2.1 | Structured clinical interview for DSM-IV axis I disorders, patient

edition—The structured clinical interview for DSM-IV axis I disorders, patient edition (SCID-I/P) (First, Spitzer, Gibbon, & Williams, 1995) is a semi-structured interview that assesses current and lifetime history of Axis I psychopathology as defined by DSM-IV. The eating disorder module was administered at baseline by trained assessors and was used to establish eating disorder diagnostic status. The kappa coefficient for current AN diagnoses (full vs. subthreshold) in Study 1-AN was .93 and the kappa coefficient for current BN diagnoses in Study 2-BN was 1.00.

2.2.2 | Positive and negative affect schedule

—The positive and negative affect schedule (PANAS) (Watson, Clark, & Tellegen, 1988) is a dimensional, self-report measure of affect. Participants were asked to rate the extent to which they currently felt specific emotions on a 5-point Likert scale, ranging from 1 (“not at all”) to 5 (“extremely”). Both studies utilized an abbreviated version of the PANAS to measure negative affect; however, the two studies used different subsets of variables. In Study 1-AN, the following eight emotions were assessed: nervous, disgusted, distressed, ashamed, angry at self, afraid, sad, and dissatisfied with self. In Study 2-BN, the following 11 emotions were assessed: afraid, lonely, irritable, ashamed, disgusted, nervous, dissatisfied with self, jittery, sad, angry at self, and distressed. The internal consistency of these abbreviated negative affect subscales (Study 1-AN: $\alpha = .94$; Study 2-BN: $\alpha = .92$) were consistent with the internal consistency of the full global negative affect scale when assessed at the momentary level (range of α s = .85–.91) (Watson & Clark, 1994).

2.2.3 | Eating episode recordings

—Participants were asked to record all binge-eating episodes on palmtop computers. Participants were provided with a standard definition of binge eating and were asked to indicate whether they considered each eating episode a binge; thus, binge-eating episodes were self-identified by participants.

2.3 | Procedures

Participants for Study 1-AN were recruited at the Neuropsychiatric Research Institute (Fargo, ND), the University of Minnesota (Minneapolis, MN), and The University of Chicago (Chicago, IL). Participants for Study 2-BN were recruited at the Neuropsychiatric Research Institute (Fargo, ND). Both studies were approved by the relevant Institutional Review Boards and all participants provided written, informed consent.

Participants who were interested and eligible based on the phone screen were invited to a baseline assessment visit during which they received information about the study, provided written informed consent, completed baseline assessments, and received instructions for using the palmtop computers. Participants were given 2 days to practice making EMA ratings after which they received feedback before completing the 2-week study protocol. Participants received monetary compensation of \$200 and could earn an extra \$50 for 80% compliance in signal response. Both studies implemented an EMA protocol that included three types of daily self-report methods: (a) *signal-contingent recordings* (i.e., after six semi-random prompts), (b) *interval-contingent recordings* (i.e., at the end of the day), and (c) *event-contingent recordings* (i.e., immediately after binge eating) (Wheeler & Reis, 1991).

2.4 | Statistical analyses

The statistical analyses described below were used in both Study 1-AN and Study 2-BN. Only binge-eating episodes that were not accompanied by purging were included in the analyses. Given that the objective of this study is to reconcile the conflicting EMA findings related to negative affect after binge eating, the decision was made to examine binge-eating-only episodes so that the impact of purging on negative affect would not confound the results.

2.4.1 | Specific aim 1: Comparison of the average proximal pre-binge and post-binge negative affect ratings—The single most proximal pre-binge and post-binge negative affect ratings were identified separately for each individual and each binge-eating episode. A mixed effects model with momentary negative affect ratings nested within participants was used to estimate the average proximal pre-binge negative affect rating, which will be referred to as the “pre-binge rating”. The same procedure was used to estimate the average proximal post-binge negative affect rating, which will be referred to as the “post-binge rating”. The intensity of negative affect was then compared between the pre-binge rating and the post-binge rating using estimated marginal means. The size of the effect was determined by converting the standard errors to standard deviations [$SD = SE * \sqrt{N}$] and then calculating Cohen’s *d*.

2.4.2 | Specific aim 2: Comparison of the timing of the average proximal pre-binge and post-binge negative affect ratings—The times, relative to the binge-eating episode, at which the pre-binge ratings were made were identified separately for each momentary affect rating used in Specific Aim 1. A mixed effects model with time nested within participants was used to estimate the average timing of the pre-binge ratings, relative to the binge-eating episode. In other words, the results represent the average amount of time prior to the binge-eating episode that the pre-binge rating was made. The same procedure was conducted to estimate the timing of the post-binge ratings, relative to the binge-eating episode. The timing of the pre-binge ratings was then compared to the timing of the post-binge ratings using estimated marginal means. The size of the effect was determined by converting the standard errors to standard deviations ($SD = SE * \sqrt{N}$) and then calculating Cohen’s *d*.

2.4.3 | Specific aim 3: Visual comparison of results using the multilevel modeling approach and the single points approach—To compare the results of the multilevel modeling approach and the single points approach, the multilevel modeling approach was first applied to these data. The trajectory of negative affect before and after binge eating was modeled separately using piecewise linear, quadratic, and cubic functions centered on the time at which the binge-eating episode occurred. The linear function represents the rate at which negative affect changes relative to the binge-eating episode, the quadratic function represents the acceleration/deceleration in the rate at which negative affect changes, and the cubic function represents further acceleration/deceleration in the rate at which negative affect changes. To account for intra-participant correlations between binge-eating episodes, episodes were nested within participants. If multiple binge-eating episodes occurred in the same day for a participant, only the first binge-eating episode was

used in the analyses and only ratings made after the first binge-eating episode and prior to the second binge-eating episode were used in the analyses.

Once the multilevel modeling approach was applied to the data, the results were depicted graphically. The results of the single points approach were then superimposed onto the same figure. The timings of the pre-binge and post-binge ratings were located on the *x*-axis and the intensities of negative affect made at the pre-binge and post-binge ratings were located on the *y*-axis.

3 | RESULTS

3.1 | EMA measurements

3.1.1 | Study 1-AN—On average, participants in Study 1-AN made 126.7 ± 31.2 recordings (range = 63 to 224 recordings) over 15.1 ± 1.3 days (range = 11 to 20 days). Thus, the average numbers of recordings made per day was 8.5 ± 3.2 recordings (range = 1 to 22 recordings). The average frequency of binge-eating-only episodes per person during the two-week EMA procedure was 1.5 ± 3.5 episodes (range = 0 to 26 episodes) and analyses were based on observations of 173 binge-eating-only episodes.

3.1.2 | Study 2-BN—On average, participants in Study 2-BN made 98.2 ± 17.8 recordings (range = 28 to 152 recordings) over 14.9 ± 1.5 days (range = 5 to 19 days). Thus, the average numbers of recordings made per day was 6.6 ± 1.8 recordings (range = 1 to 12 recordings). The average frequency of binge-eating-only episodes per person during the two-week EMA procedure was 2.2 ± 3.5 episodes (range = 0 to 22 episodes) and analyses were based on observations of 254 binge-eating-only episodes.

3.2 | Specific aim 1: Comparison of the average proximal pre-binge and post-binge negative affect ratings

In Study 1-AN, the pre-binge rating of negative affect was 22.81 ($SE = 1.10$) and the post-binge rating of negative affect was 27.17 ($SE = 1.10$). In Study 2-BN, the pre-binge rating of negative affect was 22.63 ($SE = 0.54$) and the post-binge rating of negative affect was 26.72 ($SE = 0.54$). These differences were statistically significant in both samples (Study 1-AN: $p < .001$, $ES = .31$; Study 2-BN: $p < .001$, $ES = .46$). These data are presented in Table 1.

3.3 | Specific aim 2: Comparison of the timing of the average proximal pre-binge and post-binge negative affect ratings

In Study 1-AN, the pre-binge rating was made 145.56 min ($SE = 7.74$ min) before a binge-eating episode and the post-binge rating was made 13.48 min ($SD = 7.53$ min) after a binge-eating episode. In Study 2-BN, the pre-binge rating was made 146.96 min ($SD = 5.56$ min) before a binge-eating episode and the post-binge rating was made 19.56 min ($SD = 4.90$ min) after a binge-eating episode. These differences were statistically significant in both samples (Study 1-AN: $p < .001$; $ES = 1.35$; Study 2-BN: $p < .001$, $ES = 1.49$). These data are presented in Table 2.

3.4 | Specific aim 3: Visual comparison of results using the multilevel modeling approach and the single points approach

To visually compare the results of the multilevel modeling approach to the results of the single points approach, the results of the multilevel modeling and single points approach are depicted graphically in Figure 1. The figure illustrates that although the post-binge ratings of negative affect were significantly higher than the pre-binge ratings of negative affect, the post-binge ratings were made significantly more proximally to the binge-eating episodes than the pre-binge ratings. The figure also illustrates that the pre-binge and post-binge ratings of negative affect found using the single points approach are relatively consistent with the magnitude of negative affect at that particular moment in time found using the multilevel modeling approach.

4 | DISCUSSION

The results indicate that post-binge ratings of negative affect were significantly higher than the pre-binge ratings of negative affect for women in both Study 1-AN and Study 2-BN. These data are consistent with the findings from the Haedt-Matt and Keel (2015) meta-analysis and imply that, for women with AN and BN, negative affect is higher after binge eating than before. However, the results also indicate that the post-binge ratings were made significantly closer in time to the binge-eating episode than the pre-binge ratings. Thus, these data do not support the implicit assumption that the pre-binge ratings are made immediately prior to the binge episode. Furthermore, these data do not support the implicit assumption that the pre-binge and post-binge ratings are made at time points that are equidistant relative to the binge-eating episode. Finally, the average proximal pre-binge and post-binge ratings of negative affect mapped very closely onto the trajectories of negative affect before and after binge-eating episodes found using the multilevel modeling approach (see Figure 1).

A recent EMA study describing the relationship between negative affect and purging behavior reports a pattern of results similar to studies of binge eating. Similar to the Engel et al. (2013) study, Haedt-Matt and Keel (2015) found that negative affect appeared to decrease after purging when using multilevel modeling to examine the trajectory of negative affect after purging, but negative affect appeared to increase after purging when comparing the most proximal pre-purge and post-purge ratings of negative affect. Consistent with the current study, the authors found that the post-purge ratings of negative affect were made significantly closer in time to the purge episodes than the pre-purge ratings. When interpreting these findings, Haedt-Matt and Keel (2015) stated, “Given that negative affect increases over time prior to purging, it is likely that negative affect continued to increase during the 2 hr prior to purging behavior and may have been higher immediately before purging compared with immediately after purging.” (pp. 406–407). Haedt-Matt and Keel’s (2015) interpretation is also applicable to the findings in the current study. In Study 1-AN and Study 2-BN, the average proximal pre-binge ratings of negative affect were made approximately 2.5 hr before the binge episode occurred. An examination of the trajectories of negative affect suggests that negative affect continued to increase in the 2.5 hr prior to binge episodes in both of these samples. Given that the average proximal post-binge ratings

of negative affect were made < 20 min after the binge episode, negative affect could have been at least as high, if not higher, before the binge episodes than afterwards. Thus, the results of the multilevel modeling approach and the single points approach may not be contradictory after all.

Given that the implicit assumptions of the single points approach are not supported in the current datasets, caution is necessary when interpreting the results of the single points approach in these datasets. With that said, it is worth considering whether the single points approach would be appropriate in other datasets. To answer this question, it is important to have a clear understanding of what the single points approach can and cannot tell us about the data. The single points approach compares the mean affect score at one time point to the mean affect score at a second time point. Thus, the single points approach can tell us whether affect differed between the two time points (and if it did, by how much). However, the single points approach cannot tell us about the trajectory of affect between those two time points. Extrapolating the trajectory of affect over a period of time using only two data points implicitly suggests that affect follows a linear slope between those two time points and ignores the possibility that a quadratic (i.e., the trajectory includes an inflection point at which acceleration or deceleration occurs) or cubic (i.e., a second inflection point at which further acceleration or deceleration occurs) slope could be equally as likely. Thus, if the analytic goal is to describe the *difference* in affect between two time points, then the single points approach *would* be appropriate (as long as the implicit assumptions of the approach were met). However, if the analytic goal is to describe the *trajectory* of affect over a period of time, the single points approach *would not* be an appropriate analytic technique.

In contrast, the multilevel modeling approach is an appropriate strategy for describing the trajectory of affect because it uses all available data within a given timeframe, locates each data point in time, and takes into account that the trajectory could be linear, quadratic, or cubic. However, one limitation of the multilevel modeling approach, as it was used in the current datasets, is that it does not take into account the fact that binge-eating episodes take time and that that time can be variable, both within and across individuals. In other words, the multilevel modeling approach does not describe the trajectory of negative affect *during* a binge-eating episode. And because affect was not measured during binge eating in these studies, the multilevel modeling could not describe the trajectory of negative affect during a binge using the current datasets. It has been suggested that the single points approach could address this limitation by comparing negative affect at the start of the binge to negative affect at the end of the binge. However, as described above, this analysis would only describe whether negative affect was different between the start and end of the binge. The single points approach could not describe the trajectory of negative affect during the binge eating. For example, if the single points approach found that negative affect was higher after the binge, it would still be unclear how this unfolded (e.g., negative affect could have increased at a constant rate throughout the binge, negative affect could have plateaued at the beginning of the binge and then spiked at the end, negative affect could have spiked at the beginning of the binge and then plateaued or even declined at the end of the binge). However, even if the goal was to only describe the difference in affect between the start and end of a binge, the single points approach would only be appropriate if affect was measured immediately before and immediately after the binge (or at least at equidistant times from

the binge). In the current datasets, these implicit assumptions of the single points approach were violated. Thus, the single points approach could not be used to describe the difference between negative affect at the start and end of the binge.

This violation of the implicit assumptions of the single points approach could be addressed by limiting the data points included in the analyses to those that occur immediately before and immediately after a binge-eating episode (e.g., within 5 or 10 min of a binge-eating episode); however, there are two problems with this strategy. First, few ratings were made immediately before or immediately after a binge-eating episode (even when participants were explicitly instructed to do so), which substantially limits power. Second, the participants who made ratings immediately before a binge-eating episode were not necessarily the same participants who made ratings immediately after a binge-eating episode, which means that the analysis may provide a between-person rather than a within-person comparison. As such, it is possible, but unlikely that a modified version of the single points approach could accurately describe the difference between negative affect at the start and end of a binge in the current datasets. Future studies could address this problem by modifying the assessment methodology; however, such changes could impact the external validity of the data. For example, one could increase the frequency of interval-contingent recordings (e.g., every half hour), but as the frequency of measurement points increases, so does the burden of the assessment and the likelihood that participants' responses may be impacted by the assessments themselves. Increasing the frequency of assessments and simultaneously decreasing the length of the assessment (e.g., one week instead of two) could limit the assessment burden, but it would potentially result in fewer identified binge-eating episodes, which would also limit power. Other solutions may include using passive, physiological assessments, which can track certain variables (e.g., heart rate, breathing) continuously and/or taking advantage of advances in computer learning to identify specific times at which a specific participant is most likely to binge eat and then signal the participant to respond at those times.

For the purposes of determining whether binge eating functions to regulate negative affect, knowing the trajectory of negative affect before and after a binge is imperative; however, knowing the trajectory of negative affect during a binge may be of limited utility. For example, it is well-accepted that the rituals often observed in OCD (e.g., excessive hand washing) function to regulate negative affect and that these rituals can vary in terms of how long they take to complete. However, even though some rituals can take hours to complete and negative affect may not begin to decrease until after the ritual has been completed, it is unquestioned in the literature that they function to regulate negative affect. Thus, it is unclear whether the duration of a binge or the trajectory of negative affect during a binge are necessary considerations in describing binge eating as a behavior that functions to regulate affect.

In summary, these data have several important implications. First, they provide one possible explanation for the inconsistent findings regarding what happens to negative affect after binge eating. Moreover, they suggest that previous findings based on the multilevel modeling and single points approach may not be contradictory at all, but instead, they may be complementary. Second, these data suggest that the single points approach may have

limited utility in the current datasets given that the data appear to violate the assumption that proximal pre-binge and post-binge ratings are equidistant from the episode itself. Finally, these data support the validity of the results from multilevel modeling analyses, and consequentially, bolster support for the validity of the affect regulation model of binge eating.

ACKNOWLEDGMENTS

This research was supported by grants from the National Institute of Mental Health (R01 MH059674: Eating Disorders and Impulsivity—A Longitudinal Study, Ecological Momentary Assessment of Anorexia Nervosa) and NIDDK (P30DK050456, Minnesota Obesity Center).

REFERENCES

- Berg KC, Crosby RD, Cao L, Crow SJ, Engel SG, Wonderlich SA, & Peterson CB (2015). Negative affect prior to and following overeating-only, loss of control eating-only, and binge eating episodes in obese adults. *The International Journal of Eating Disorders*, 48(6), 641–653. 10.1002/eat.22401. [PubMed: 25808854]
- Berg KC, Crosby RD, Cao L, Peterson CB, Engel SG, Mitchell JE, & Wonderlich SA (2013). Facets of negative affect prior to and following episodes of binge-only, purge-only, and binge/purge events in women with bulimia nervosa. *Journal of Abnormal Psychology*, 122, 111–118. [PubMed: 22985015]
- Berg KC, Peterson CB, Crosby RD, Cao L, Crow SJ, Engel SG, & Wonderlich SA (2014). Relationship between daily affect and overeating-only, loss of control eating-only, and binge eating episodes in obese adults. *Psychiatry Research*, 215, 185–191. [PubMed: 24200217]
- Crosby RD, Wonderlich SA, Engel SG, Simmonich H, Smyth J, & Mitchell JE (2009). Daily mood patterns and bulimic behaviors in the natural environment. *Behaviour Research and Therapy*, 47, 181–188. [PubMed: 19152874]
- Engel SG, Wonderlich SA, Crosby RD, ... Gordon KH (2013). The role of affect in the maintenance of anorexia nervosa: Evidence from a naturalistic assessment of momentary behaviors and emotion. *Journal of Abnormal Psychology*, 122, 709–719. [PubMed: 24016011]
- First MB, Spitzer RL, Gibbon M, & Williams JB (1995). *Structured clinical interview for the DSM-IV axis I disorders — patient edition (SCID-I/P, version 2)*. New York: New York State Psychiatric Institute, Biometrics Research Department.
- Haedt-Matt AA, & Keel PK (2015). Affect regulation and purging: An ecological momentary assessment study in purging disorder. *Journal of Abnormal Psychology*, 124, 399–411. [PubMed: 25688426]
- Haynos AF, Berg KC, Cao L, ... , Crow SJ (2017). Trajectories of higher- and lower-order dimensions of negative and positive affect relative to restrictive eating in anorexia nervosa. *Journal of Abnormal Psychology*, 126, 495–505. [PubMed: 27893231]
- Kenardy J, Arnow B, & Agras WS (1996). The aversiveness of specific emotional states associated with binge-eating in obese subjects. *Australian and New Zealand Journal of Psychiatry*, 30, 839–844.
- Lavender JM, Utzinger LM, Crosby RD, ... Le Grange D (2016). A naturalistic examination of temporal patterns of affect and eating disorder behaviors in anorexia nervosa. *International Journal of Eating Disorders*, 49, 77–83.
- Polivy J, & Herman CP (1993). Etiology of binge eating: Psychological mechanisms. In Fairburn CG, Wilson GT (Eds), *Binge eating: Nature, assessment, and treatment* (pp. 173–205). New York, NY: Guilford Press.
- Shiffman S, Stone AA, & Hufford MR (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, 4, 1–32.
- Smyth JM, Wonderlich SA, Heron KE, Sliwinski MJ, Crosby RD, Mitchell JE, & Engel SG (2007). Daily and momentary mood and stress are associated with binge eating and vomiting in bulimia

nervosa patients in the natural environment. *Journal of Consulting and Clinical Psychology*, 75, 629–638. [PubMed: 17663616]

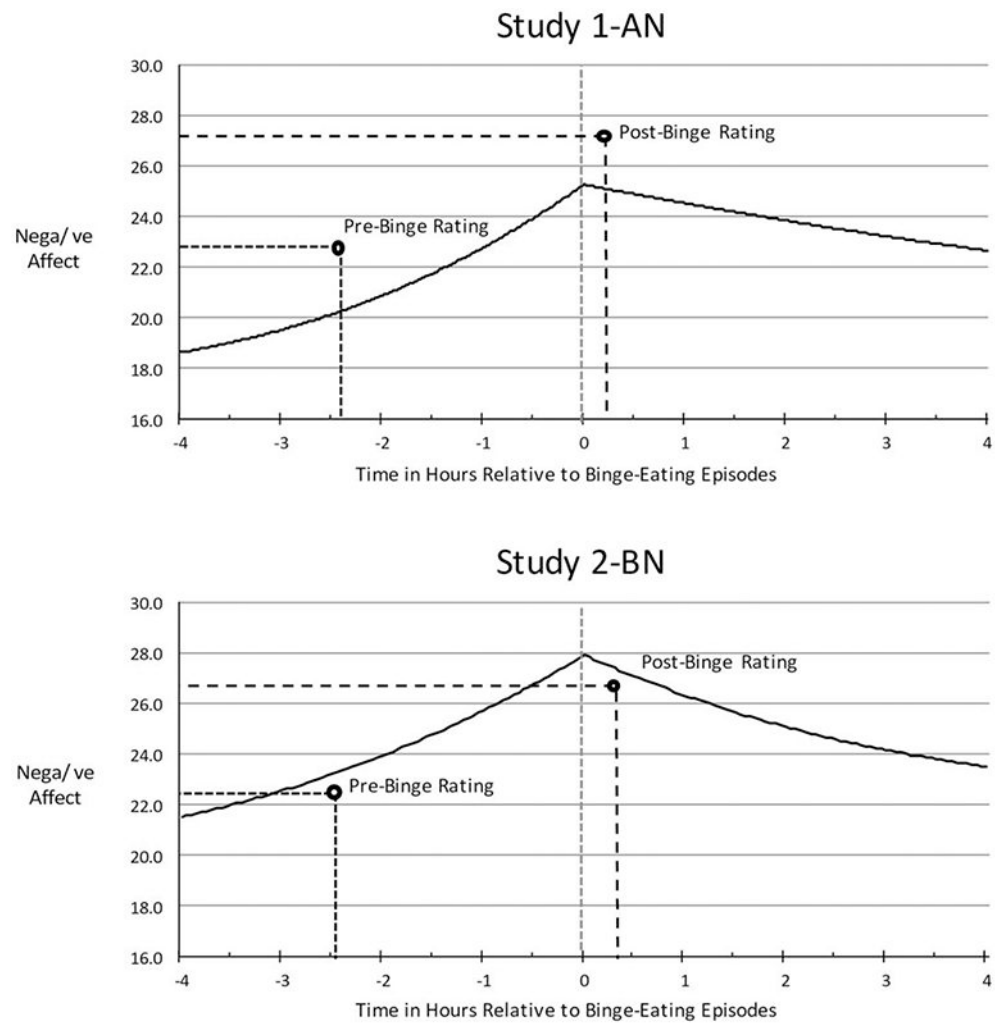
Stone AA, & Shiffman S (1994). Ecological momentary assessment (EMA) in behavioral medicine. *Annals of Behavioral Medicine*, 16, 199–202.

Telch CF, & Agras WS (1996). Do emotional states influence binge eating in the obese? *International Journal of Eating Disorders*, 20, 271–279.

Watson D, & Clark LA (1994). The panas-X: Manual for the positive and negative affect schedule—Expanded form.

Watson D, Clark LA, & Tellegen A (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063–1070. [PubMed: 3397865]

Wheeler L, & Reis HT (1991). Self-recording of everyday life events: Origins, types, and uses. *Journal of Personality*, 59, 339–354.

**FIGURE 1.**

Visual comparison of results using the multilevel modeling approach and the single points approach. The solid curved line depicts results from the multilevel modeling approach. Specifically, it shows the momentary levels and trajectories of global negative affect relative to binge eating in Study 1-AN and Study 2-BN. In these analyses, the pre- and post-event trajectories of global negative affect were modeled separately using piecewise linear, quadratic, and cubic functions centered on the time at which each of the eating episodes occurred. Momentary observations (Level 1) were nested within subjects (Level 2). The black dots depict the average proximal pre-binge and average proximal post-binge ratings of negative affect. The location of these points represents the average time at which the ratings were taken relative to the binge-eating episode and the average intensity of negative affect reported during that rating. For Study 1-AN, the scaling of the y axis ranged from 8.0 to 40.0. For Study 2-BN, the scaling of the y axis ranged from 11.0 to 55.0

TABLE 1
Average proximal ratings of negative affect before and after binge-eating episodes

	Mean	SE	95% CI		t	df	p	ES
			LL	UL				
Study 1-AN								
Pre-binge	22.81	1.10	20.60	25.02	18.88	245.69	>.001	.31
Post-binge	27.17	1.10	24.96	29.37				
Study 2-BN								
Pre-binge	22.63	0.54	21.56	23.69	-5.37	576.00	>.001	.46
Post-binge	26.72	0.54	25.67	27.77				

Timing (in min) of average proximal ratings of negative affect before and after binge-eating episodes relative to the binge-eating episodes

TABLE 2

	Mean	SE	95% CI		t	df	p	ES
			LL	UL				
Study 1-AN								
Pre-binge	145.56	7.74	129.84	161.28	18.88	245.69	<.001	1.35
Post-binge	13.48	7.53	-1.86	28.83				
Study 2-BN								
Pre-binge	146.96	5.56	135.91	158.02	20.41	372.18	<.001	1.49
Post-binge	19.56	4.90	9.79	29.32				