

ULRICH SCHIMMACK

AFFECT MEASUREMENT IN EXPERIENCE SAMPLING RESEARCH

ABSTRACT. Experience-sampling studies are used to study the emotional component of subjective well-being (hedonic balance). This manuscript examines conceptual and methodological issues in the measurement of hedonic balance, and it relates aspects of affective experiences (frequency, intensity, and duration) to affective dispositions (extraversion, neuroticism) and life-satisfaction. Aggregates of experience-sampling data are influenced by response styles, but the effect is negligible. Pleasant affects and unpleasant affects show high discriminant validity. Extraversion is more highly related to aspects of pleasant affects than unpleasant affects, and neuroticism is more highly related to aspects of unpleasant affects than pleasant affects. Mean levels (i.e., frequency * intensity) of affects are the aspects that best predict life-satisfaction. The specific item happiness is a better predictor of life-satisfaction than the average of all pleasant affects.

The scientific study of happiness has a long history (Andrews and Withey, 1976; Beebe-Center, 1932; Diener, 1984). Over the past three decades, psychologists have made immense progress in this endeavor (Diener et al., 1999). One major contribution has been the scientific conceptualization of happiness in terms of subjective well-being (Diener, 1984). Subjective well-being has a cognitive and an affective component. The cognitive component is assessed by means of life-satisfaction judgments (e.g., "I am satisfied with my life;" Diener et al., 1985). The affective component is assessed as the *hedonic balance* of pleasant and unpleasant experiences.

A series of studies have examined the relation between life-satisfaction and hedonic balance (Schimmack et al., 2002; Schimmack et al., 2002; Suh et al., 1998). These studies demonstrate that hedonic balance determines life-satisfaction because people rely on the amount of pleasant and unpleasant experiences to evaluate their lives. This relationship is universal, although hedonic balance is a stronger predictor of life-satisfaction in individualistic cultures (e.g., United States) than in collectivistic cultures (e.g., Ghana). In sum, the determinants of hedonic balance are at the core of scientific investigations of happiness both because hedonic balance is an indicator of happiness and because it is a strong determinant of life-satisfaction (Kahneman et al., 1999).



Typically, well-being researchers assess hedonic balance by means of retrospective ratings with time frames ranging from a day to several years (e.g., Withey and Andrews, 1976; Oishi et al., 2001; Schimmack et al., 2002; Suh et al., 1998; Watson et al., 1988; Zelenski and Larsen, 2000). Numerous studies have demonstrated that these ratings are partially valid reflections of people's actual affective experiences, and partially influenced by biases (cf. Barrett, 1997; Parkinson et al., 1995; Schimmack, 2002; Thomas and Diener, 1990). Experience-sampling studies promise a better assessment of actual hedonic balance than retrospective judgments (Diener and Emmons, 1984; Kahneman et al., 1999). However, several methodological issues need to be addressed before experience-sampling studies can fulfill this promise.

I first review some of these issues, such as the choice of appropriate response formats, the influence of response styles, different sampling procedures of affect, the selection of affects, and the different aggregation procedures. Then I present data from an electronic experience-sampling study to demonstrate how experience-sampling studies can further our understanding of happiness.

EXPERIENCE-SAMPLING STUDIES OF HEDONIC BALANCE

Although experience-sampling studies have a long history (Flügel, 1925), they only became popular at the dawn of the affective revolution in psychology (Brandstätter, 1983; Csikszentmihalyi and Larson, 1987; Epstein, 1982; Diener and Emmons, 1984; Hedges et al., 1985; Hormuth, 1986; Larson and Csikszentmihalyi, 1983; McAdams and Constantian, 1983; Stone and Neale, 1984; Watson et al., 1984; Zevon and Tellegen, 1982). One reason for the increasing popularity of experience-sampling studies may have been the waning influence of behaviorism. Behaviorism either denied the existence of internal feelings or assumed that they cannot be studied scientifically. Experience-sampling studies also faced concerns about the validity of self-report measures because people may have limited access to their feelings. Extensive validation research has demonstrated that these fundamental concerns are unfounded (cf. Csikszentmihalyi and Larson, 1987; Diener et al., 1991; Emmons and Diener, 1984; Schimmack and Diener, 1997). For example, the average level of affect reported in experience-sampling studies shows convergent validity with peer reports (e.g., Diener et al., 1995).

Time Frame

Experience-sampling studies typically ask respondents to report how they feel at one moment in time (e.g., right now, right before the signal). However, this advantage also creates a methodological problem because participants are unable or unwilling to report their feelings at any random moment. For example, participants are unable to report current feelings while they are engaged in a task that cannot be interrupted (e.g., giving a talk, operating a bus, conducting a concert, etc.). Participants are also unlikely to report their feelings in extremely emotional situations (e.g., during a fight or while having sex). To obtain data on these experiences in an experience-sampling study, it may be necessary to use a longer time frame (e.g., during the past 10 min, half hour). As a result, reports at convenient times still reflect emotional experiences that occurred recently but could not be captured in momentary reports. Alternatively, it is possible to ask for memory-based reports as soon as possible after an emotional episode ended (e.g., Flügel, 1925; Schimmack and Diener, 1997). There exists no evidence, which time frame provides a better way of assessing hedonic balance. The present article examines this issue by comparing two time frames within the same study. When signaled, participants made momentary judgments of their affects right before the signal and extended judgments of their affects over the past 30 min before the signal.

Response Formats

Since the first studies of self-reported affect, affective experiences have been assessed with adjective lists (Nowlis, 1965). To assess affective experiences, adjectives are paired with one of numerous response formats, ranging from simple checklists to analog scales. However, the most common response formats are rating scales with four to seven response options (e.g., Diener and Emmons, 1984; Schimmack and Diener, 1997; Watson et al., 1988). Researchers have to be careful about their choices of a response format because response formats influence participants' interpretation of items and therewith their responses (Meddis, 1972; Schwarz et al., 1991; Schimmack, et al., 2002; Svensson, 1977). The most important distinction is between symmetrical response formats and asymmetrical response formats. An example of a symmetrical response format is the typical 5-point Likert scale ranging from $-2 =$ strongly disagree to $+2 =$ strongly agree. Symmetrical scales are interpreted by most participants as bipolar scales and produce strong negative correlations between ratings of

pleasant affects and unpleasant affects (Schimmack et al., 2002). The reason is that positive items are interpreted as bipolar negative–positive items, whereas negative items are interpreted as bipolar positive–negative items. As both ratings are the mirror images of each other, they provide essentially redundant information. Researchers who want to use bipolar scales should use scales that clearly communicate bipolarity to the respondents. For example, they should label both ends of the scale with opposing labels (e.g., pleasant–unpleasant) and use positive and negative numbers to indicate that the scale assesses both positive and negative experiences.

However, many researchers treat pleasant experiences and unpleasant experiences as conceptually distinct constructs (cf. Diener, 1984; Diener and Emmons, 1984; Diener et al., 1995; Watson et al., 1988; Zevon and Tellegen, 1982; Schimmack, 1997, 2001; Schimmack et al., 2000). In this case, it is important to assess affective experiences with asymmetrical or unipolar scales (cf. Schimmack et al., 2002). The labels of asymmetrical scales differ depending on the time frame. For studies of extended time frames, it is common to assess the amount of time that an affect was experienced ranging from zero (e.g., 0 = no, not at all) to all the time (e.g., 6 = yes, all the time). For studies of momentary affect, asymmetrical scales are intensity scales that range from the absence of an affect (e.g., 0 = no, not at all) to the maximum intensity of an affect (e.g., 6 = yes, with maximum intensity). To increase the validity of respondents' reports on these scales, it is advantageous to provide respondents with instructions on the use of unipolar response formats. For example, Schimmack (1997) asked participants to first consider whether a feeling was absent or present and to respond with zero if the feelings was absent. If the feelings were present, participants were instructed to use one of the remaining response categories ranging from one to six to indicate the intensity of the feeling.

Most unipolar response formats clearly distinguish the absence of an affect from experiences of mild affect. However, the response format of the widely used positive affect negative affect schedule (PANAS; Watson et al., 1988), combines absence and very mild feelings of an affect in the lowest response category (1 = very slightly or not at all). This approach is unproblematic in studies with long time frames (in general), but it may be problematic in experience-sampling studies. In experience-sampling studies, small differences in a single rating can become large in aggregates of multiple ratings. Hence, it may

be important to know whether repeated endorsement of the lowest category reflect the absence of an affect or chronically low levels of an affect. Furthermore, combining the absence of an affect and mild experiences of an affect in one response category make it impossible to separate frequency and intensity of affective experiences (Schimmack and Diener, 1997). In the present study, I used asymmetrical response formats that used the lowest response category to report the absence of an affect only.

Measurement Error

All measures are influenced by measurement error, some more than others. Two types of measurement error can be distinguished, namely random measurement error and systematic measurement error. Typically, researchers use multiple items and/or mathematical measurement models to control or eliminate random measurement error. However, this approach is less feasible in experience-sampling studies, which require multiple assessments from each participant. Each assessment is intrusive and adding redundant items to reduce measurement error may be annoying and reduce the quality of the data. Hence, researchers often try to use a minimum number of items for each construct. The use of a few items is not a problem when experience-sampling data are aggregated (e.g., Diener and Larsen, 1984; Epstein, 1982). In this case, the repeated assessments serve as multiple indicators that reduce random measurement error. However, the results of within-subject analysis across an individual's samples of experiences will be attenuated by random measurement error. The amount of random error can be estimated from measurement models, which relate an observed measure with random error to a latent factor without random error. For homogeneous scales (e.g., positive, pleasant, good) these correlations can be very high (0.72 to 0.92), indicating that often two or three items are sufficient (cf. Schimmack and Grob, 2000). However, more items are needed for heterogeneous scales that sample a broad range of distinct affects. For example, the PANAS positive affect scale uses 10 items with item-factor correlations ranging from 0.75 to 0.52 (cf. Watson et al., 1988).

Systematic measurement error poses the exact opposite problems. That is, it does not pose a problem for within-subject analysis across sampled experiences because systematic measurement error is by definition constant across repeated assessments (cf. Diener and Emmons, 1984; Watson and Clark, 1997). If errors were changing

from assessment to assessment, they would be no longer systematic. As constants do not influence the relation of variables, systematic measurement error cannot bias within-subject correlations between measures. However, systematic measurement error can cause problems when experience-sampling data are aggregated to yield trait-like measures. In this case, systematic measurement variance is part of the variance in the aggregates and can bias correlations with other measures. Furthermore, small effects of systematic measurement error on a single rating can accumulate across repeated assessments and have a substantial effect (Fisher et al., 1985). Hence, it is important to determine the amount of systematic measurement error in affect ratings.

The opinions about the importance of systematic measurement error have shifted over the past decades. After initial concerns about dramatic effects of response styles on self-ratings, Rorer's (1965) review alleviated most of these concerns. Subsequent work by Diener and colleagues also suggested that researchers had little to worry about systematic measurement error (Diener and Emmons, 1984). However, Green et al. (1993) suggested that systematic measurement error has a substantial influence on affect ratings. The authors assessed pleasant affect and unpleasant affect with four different response formats. They compared the correlations between pleasant and unpleasant affect for individual formats and for latent factors based on multiple formats. Whereas the single-format correlations were weak to moderate, the latent factors were strongly negatively correlated. Based on this pattern of results, the authors concluded that systematic measurement error could have dramatic influences on affect ratings. However, subsequent multi-method studies revealed much weaker negative correlations between latent factors of pleasant affect and unpleasant affect (Diener et al., 1995). Schimmack et al. (2002) carefully examined the evidence and found that the nature of the response format explained the discrepant results. Green et al. (1993) used symmetric formats that respondents interpreted as bipolar scales (see above), whereas Diener et al. (1995) used asymmetrical formats that most respondents interpreted as unipolar scales. None of the reviewed multi-method studies revealed noticeable systematic measurement variance in self-ratings of affect (cf. Schimmack et al., 2002). However, aggregates of repeated daily diary data revealed a significant method effect that can be attributed to stronger effects of response styles on aggregated affect ratings. The present article introduces a new measure to examine the influence of response styles on aggregated experience-sampling data.

Affect Selection

Researchers have used a diverse range of adjectives to assess hedonic balance. Some lists are based on emotion theories that postulate a set of four to ten basic emotions (Diener et al., 1995; Izard et al., 1993; Oatley and Duncan, 1994; Shaver et al., 1987). Other lists are based on factor analyses (Nowlis, 1965; Watson et al., 1988). The most widely used affect measure is the PANAS (Watson et al., 1988). The PANAS was developed to assess the two major factors in factor analyses of affect ratings after VARIMAX rotation (Watson et al., 1988, 1984; Zevon and Tellegen, 1982). The positive affect scale includes the items enthusiastic, interested, determined, excited, inspired, alert, active, strong, proud, and attentive. The negative affect scale includes the items scared, afraid, upset, distressed, jittery, nervous, ashamed, guilty, irritable, and hostile.

The PANAS scales show high convergent validity with other measures of pleasant and unpleasant affects (Watson, 1988). However, several studies have demonstrated that different affect measures can produce divergent results. Kennedy Moore, et al. (1992) observed that day of the week had different effects on the PA scale of the PANAS and a different measure of pleasant affect. Whereas pleasant affect increased in accordance with the higher number of positive events over the weekend, levels of PA on the PANAS were lower on the weekend than on weekdays. This pattern of results was replicated and extended in a study by Egloff et al. (1995). Once more, pleasant affect was higher on weekends than on weekdays, whereas PANAS's PA was higher on weekdays than on weekends. In addition, the two measures showed different effects of time of day. Whereas PA peaked around noon and decreased in the evening, pleasant affect increased only slightly from morning to noon and further increased in the evening.

The choice of pleasant and unpleasant adjectives also influences the amount of discriminant validity (Egloff, 1998; Watson, 1988). Whereas unipolar ratings of pleasant affects and unpleasant affects typically yield weak negative correlations (e.g., Schimmack, 2001), the PANAS PA and NA scales are nearly perfectly orthogonal. The reason for this discrepancy is that the PANAS items were based on VARIMAX rotation of factor analyses which maximizes discriminant validity. For that reason, the PANAS does not include items of happiness and sadness, which produce negative correlations between scales of pleasant and unpleasant affects.

In sum, the choice of affects can influence the results, and researchers should carefully choose adjectives based on theoretical considerations of the affective component of subjective well-being. The present study explores this issue further, by examining the predictive validity of specific affects for life-satisfaction.

Facets of Affective Experiences

Schimmack et al. (2000) proposed a facet framework for the conceptualization of affective experiences. The authors defined a facet as a combination of a quality (pleasant vs. unpleasant), a type (mood vs. emotion), and an aspect (frequency, intensity, and duration) of affective experiences. The frequency of an affect is defined as the number of experience samples, in which an affect was present (i.e., the intensity was rated as greater than zero; the duration was rated as greater than zero minutes). Intensity is defined as the average intensity during samples when an affect was present (cf. Schimmack and Diener, 1997). Duration is defined as the average duration of an affect when it is reported as present.

It is difficult to determine which facets of affective experiences should be used in measures of subjective well-being (see Flügel, 1925, for an early discussion of this problem). Diener et al. (1991) proposed using the time an affect is experienced (i.e., frequency and duration) as indicator of subjective well-being and to ignore intensity. They argued that intensity is more difficult to compare across individuals, and that intensity can have undesirable properties. For example, people who experience pleasant emotions intensely also experience unpleasant emotions more intensely (Larsen and Diener, 1987; Schimmack and Diener, 1997). Hence, the benefits of intense pleasant experiences may be offset by intense negative experiences.

The present article examines this issue by relating frequency, intensity, and duration of affects to life-satisfaction. Furthermore, I examine whether interactions between aspects may be better predictors of life-satisfaction. It is possible that intensity interacts with frequency in the prediction of life-satisfaction. That is, frequency may be more important than intensity, but for two individuals with the same frequency of pleasant experiences, the one with more intense pleasant experiences may have higher life-satisfaction.

A RANDOM EXPERIENCE-SAMPLING STUDY OF HEDONIC BALANCE

Subsequently, I am going to present the results of a random experience-sampling study of hedonic balance. Participants received hand-held computers for one week to record their affective experiences. In addition, participants completed questionnaires on affective dispositions and life-satisfaction throughout the semester. The data are used to explore some of the questions raised in the Introduction.

Method

Participants

The analyses are based on a sample of 127 students (38 male, 89 female), who took part in an experience-sampling study as part of a course on personality and subjective well-being. The average age was 20 years.

Experience-Sampling Materials

The experience-sampling questions asked for participants' feelings at the moment right before the alarm and the time an affect was experienced during the past 30 min. Momentary affect was assessed with a 7-point unipolar intensity scale ranging from 0 = not at all to 6 = maximum intensity (Schimmack, 1997; Schimmack and Diener, 1997). Extended affect was assessed with an 8-point scale ranging from 0 = zero minutes to 8 = 30 min. Affect adjectives were sampled to cover a range of pleasant affects (happy, affectionate, proud, excited) and unpleasant affects (sad, worried, guilty, irritated).

Experience-Sampling Procedure

Participants received hand-held computers for eight days. Participants were instructed on how to use the experience-sampling software and started with some practice trials on that day. The experience-sampling study started with the morning report the next day and ended with the evening report of the last day. Hence, the experience-sampling study covered one full week. Each day participants had to report at least seven experiences. The first report was self-initiated at the beginning of a day as soon as possible after waking up. Then five random alarms were scheduled during the next ten hours after the morning report. When participants were unable to respond to an alarm right away, they were

allowed to enter a report within 30 min after an alarm. Participants could always check when the last alarm was sounded. If participants did not respond within 30 min, a new random alarm was scheduled later during the day. At the end of the day, participants entered a self-initiated report before they went to sleep. Finally, participants were also allowed to self-initiate additional reports. Additional reports did not reduce the number of random alarms. The computer recorded the type of report (morning, evening, direct response to alarm, delayed response to alarm, voluntary). In the present article, all reports are included and I do not distinguish between the different types of reports.

Affect ratings were made in a fixed sequence. The sequence started with the intensity ratings, which were followed by the duration ratings. Morning reports did not include duration ratings because they were supposed to be made right after waking up.

Affective Dispositions

Individual differences in affective dispositions were assessed at the beginning of the semester with the 300-item version of Goldberg's IPIP (Goldberg, 1997). For the present study, I used the 60-item neuroticism and extraversion scales, which have good psychometric properties ($\alpha > 0.80$).

Life-Satisfaction

Life-satisfaction was assessed with monthly ratings of life-satisfaction on a five-item scale adapted from the satisfaction-with-life-scale ("I was satisfied with my life in the past month;" cf. Schimmack et al., 2002). The monthly satisfaction judgments were made three times during the semester. We aggregated these judgments to obtain a measure of life-satisfaction during the semester. We choose this measure over general life-satisfaction judgments because general life-satisfaction judgments cover a longer time period and are likely to be more weakly correlated with the experience-sampling data. The life-satisfaction scale also had good internal consistency at each measurement moment ($\alpha > 0.80$).

Response Style Measure

Participants also completed a new response style measure. The measure is based on affect ratings of prototypical facial expressions of emotions. Ratings are made on the same intensity scale as the ones used in the experience-sampling study. Hence, any individual differences in the use

of the intensity scale produce shared variance between ratings of participants' own experiences and ratings of facial expressions. In addition, the facial expression ratings can control for individual differences in the interpretation of affect words. The list of affect words overlapped with the affect words in the experience-sampling study. Pleasant affects were happy and affectionate, whereas proud and excited were not included in the facial expression ratings. Unpleasant affects were sad, guilty, and irritated, whereas worried was not included in the facial expression ratings. Participants rated a set of nine different facial expressions covering a range of basic emotions and emotion blends. The pleasant affect ratings and the unpleasant affect ratings were averaged ($\alpha = 0.60, 0.84$, respectively).

RESULTS

Descriptive Statistics

I begin with descriptive statistics of affects in students' everyday lives. Table I shows the mean intensity ratings and duration ratings of affects. It also shows a decomposition of these ratings into frequency and intensity and frequency and duration, respectively. Finally, Table I shows the percentages of each affect for the 7-point intensity scale and the 8-point duration scale. The results are generally consistent with previous experience-sampling studies (Carstensen et al., 2000; Shapiro et al., 2001). Frequencies derived from the intensity ratings and the duration ratings are very similar. The results for happiness confirm that most people are happy most of the time (Diener and Diener, 1996). Furthermore, the distribution for intensity of happiness shows an approximately normal distribution with a mean of 2.94. All other affects show low frequencies; the modal response for these affects is zero on both the intensity and the duration scale. The frequencies of distinct affects are important for happiness researchers because rare affects have a minor impact on overall hedonic balance.

Response Styles and Correlations Between Affects

There are several ways of examining systematic measurement error in affect ratings (Schimmack et al., 2002). Experience-sampling data provide the opportunity to do so by comparing within-subject and between-subject correlations (Watson and Clark, 1997). Whereas

TABLE I
Descriptive statistics of intensity and duration ratings

Affect	Mean	Freq.	Int/Dur	% 0	% 1	% 2	% 3	% 4	% 5	% 6	% 7
<i>Intensity</i>											
Happy	2.94	0.93	3.16	0.07	0.09	0.18	0.33	0.19	0.09	0.05	—
Affectionate	1.49	0.59	2.55	0.42	0.17	0.14	0.14	0.07	0.03	0.03	—
Proud	1.20	0.51	2.35	0.49	0.17	0.14	0.12	0.05	0.02	0.02	—
Excited	1.37	0.60	2.29	0.40	0.20	0.17	0.14	0.07	0.02	0.01	—
Sad	0.49	0.28	1.76	0.72	0.16	0.07	0.03	0.01	0.00	0.00	—
Worried	1.00	0.51	1.97	0.49	0.22	0.15	0.08	0.03	0.01	0.01	—
Guilty	0.45	0.26	1.75	0.75	0.15	0.06	0.03	0.01	0.00	0.01	—
Irritated	0.85	0.44	1.93	0.56	0.21	0.13	0.06	0.03	0.01	0.01	—
<i>Duration</i>											
Happy	4.27	0.93	4.60	0.07	0.09	0.10	0.11	0.14	0.13	0.08	0.28
Affectionate	1.91	0.60	3.18	0.40	0.17	0.14	0.08	0.06	0.04	0.02	0.10
Proud	1.41	0.51	2.80	0.49	0.19	0.11	0.06	0.04	0.03	0.01	0.07
Excited	1.71	0.64	2.69	0.36	0.22	0.15	0.09	0.07	0.03	0.02	0.05
Sad	0.66	0.30	2.20	0.70	0.16	0.06	0.03	0.02	0.01	0.00	0.02
Worried	1.27	0.52	2.44	0.48	0.22	0.13	0.06	0.04	0.02	0.01	0.05
Guilty	0.51	0.26	1.99	0.74	0.15	0.05	0.02	0.01	0.00	0.00	0.02
Irritated	0.99	0.47	2.12	0.53	0.24	0.10	0.05	0.03	0.01	0.00	0.02

between-subject correlations can be contaminated by response styles, within-subject correlations remove response styles along with the between-subject variance. Hence, it is possible to examine the influence of response styles by comparing unstandardized correlations across all observations that include between-subject and within-subject variance with correlations based on all observations after within-subject standardization. These correlations are displayed in Table II. Table II also shows the between subject correlations of the aggregated ratings. For these analyses, intensity ratings as well as duration ratings were simply averaged across individuals' reports.

Correlations between affects of opposing valence are all very similar for standardized (WS) and unstandardized (BS-WS) data, indicating that systematic measurement error has a negligible influence on these correlations. This finding is consistent with previous studies (Schimmack et al., 2002; Watson and Clark, 1997). However, correlations among pleasant affects as well as those among unpleasant affects are stronger when between-subject variance is included than when it is statistically removed. This finding can be interpreted in two ways. Perhaps response styles are valence specific. That is, stylistic tendencies in reports of pleasant affects are independent of stylistic tendencies in reports of unpleasant affects. As a result, response styles would inflate same-valence correlations, but they would not attenuate correlations between affects of opposite valence. Alternatively, the results may reflect valid individual differences in the between-subject variance. Accordingly, individuals who have higher levels of one pleasant affect also tend to have higher levels of other pleasant affects, and individuals who have higher levels of one unpleasant affect also tend to have higher levels of other unpleasant affects. The influences of substance and style on between-subject variance are explored in the following analyses of aggregated experience-sampling data.

First, however, it is also important to note that the pattern of correlations conforms to the well-known fact that correlations between affects of the same-valence (pleasant-pleasant and unpleasant-unpleasant) are much stronger than correlations between affects of opposite valence (pleasant-unpleasant). This finding is consistent with previous studies, although some studies obtained moderate negative correlations in within-subject analyses (Emmons and Diener, 1984; Moskowitz and Coté, 1995). Furthermore, there is no evidence that happiness and sadness differ dramatically from other affects in this regard. Whereas the

TABLE II
Correlations between affects within and between subjects

Affect pair	Moment			Past 30 min		
	WS-BS	WS	BS-A	WS-BS	WS	BS-A
<i>Across-valence</i>						
Happy-sad	-0.18	-0.22	-0.02	-0.22	-0.25	-0.11
Happy-worried	-0.20	-0.23	-0.12	-0.23	-0.24	-0.17
Happy-guilty	-0.11	-0.11	-0.10	-0.14	-0.12	-0.15
Happy-irritated	-0.27	-0.32	-0.11	-0.29	-0.33	-0.17
Affectionate-sad	0.05	-0.04	0.24	0.05	-0.03	0.26
Affectionate-worried	-0.03	-0.10	0.11	-0.03	-0.08	0.12
Affectionate-guilty	0.06	0.01	0.16	0.06	0.03	0.14
Affectionate-irritated	-0.06	-0.16	0.20	-0.05	-0.13	0.22
Proud-sad	0.02	-0.05	0.16	0.01	-0.03	0.15
Proud-worried	0.02	-0.05	0.12	0.00	-0.03	0.09
Proud-guilty	0.08	0.00	0.17	0.09	0.03	0.19
Proud-irritated	-0.02	-0.11	0.19	-0.01	-0.06	0.21
Excited-sad	-0.01	-0.07	0.16	-0.02	-0.07	0.22
Excited-worried	-0.03	-0.10	0.12	-0.05	-0.07	0.09
Excited-guilty	0.04	-0.03	0.19	0.04	-0.01	0.18
Excited-irritated	-0.03	-0.11	0.24	-0.03	-0.08	0.26
<i>Among pleasant</i>						
Happy-affectionate	0.50	0.41	0.58	0.42	0.37	0.45
Happy-proud	0.42	0.33	0.47	0.31	0.28	0.33
Happy-excited	0.48	0.44	0.53	0.35	0.35	0.35
Affectionate-proud	0.52	0.33	0.69	0.48	0.27	0.67
Affectionate-excited	0.45	0.35	0.58	0.45	0.37	0.56
Proud-excited	0.49	0.33	0.71	0.47	0.31	0.68
<i>Among unpleasant</i>						
Sad-worried	0.46	0.33	0.64	0.48	0.38	0.66
Sad-guilty	0.36	0.25	0.52	0.30	0.24	0.44
Sad-irritated	0.38	0.28	0.60	0.36	0.30	0.58
Worried-guilty	0.54	0.33	0.77	0.50	0.31	0.72
Worried-irritated	0.43	0.32	0.60	0.48	0.40	0.62
Guilty-irritated	0.34	0.22	0.53	0.35	0.25	0.53
<i>Average r</i>						
Cross-valence	-0.04	-0.11	0.11	-0.05	-0.09	0.10
Among pleasant	0.48	0.37	0.60	0.42	0.33	0.52
Among unpleasant	0.42	0.29	0.62	0.41	0.31	0.60
<i>Scales</i>						
<i>r</i> (PAS-UAS)	-0.08	-0.21	0.13	-0.11	-0.19	0.13
α -PAS	0.78	0.69	0.80	0.73	0.65	0.80
α -UAS	0.74	0.62	0.84	0.74	0.65	0.85

Note. WS-BS within and between subject, unstandardized; WS = within-subject standardized, BS-A = between-subject aggregated data.

average within-subject correlation between pleasant and unpleasant affects was -0.11 , the one between happy and sad was -0.22 .

Facets of Affective Experiences

I created variables that reflect six different facets of affective experiences, namely the frequency, (typical) intensity, and (typical) duration of pleasant affects and unpleasant affects. I also kept more traditional measures of individual differences in affective experience. Namely, I computed the mean level of intensity ratings (i.e., the product of frequency and intensity; Schimmack and Diener, 1997), and the time participants experienced affects (Diener et al., 1995). The time was the mean level of the duration ratings, which represents the product of frequency and duration (Schimmack et al., 2000).

Table III shows the correlations between these indicators of individual differences in affective experiences. The pattern of correlations reveals positive associations between facets of the same valence. Some of these associations are due to conceptual overlap. For example, the mean level incorporates frequency and intensity in a multiplicative manner. Similarly, the time facets combine frequency and duration in a multiplicative manner. However, other correlations are more substantive. Namely, intensity and duration are based on different ratings, but still correlate highly with each other. There are several possible explanations for this correlation. First, intense experiences could last longer (Frijda et al., 1992). Second, respondents may partially rely on immediate feelings of intensity to judge duration. Third, the two measures may be influenced by a common response style.

Also noteworthy are the positive cross-valence correlations for intensity, frequency, and duration. These correlations support earlier findings that people with frequent pleasant experiences also have more frequent unpleasant experiences (Schimmack et al., 2000), and people with more intense pleasant experiences also tend to have more intense unpleasant experiences (Schimmack and Diener, 1997). The positive correlation for durations of pleasant and unpleasant affects is a new finding. Interestingly, the mean level and the time facets of pleasant and unpleasant affects were uncorrelated. This finding seems to be at odds with the positive correlations for their components. However, it has to be noted that the components are combined multiplicatively. As a result, mean levels can be unrelated although both frequency and intensity are positively correlated.

TABLE III

Correlations between facets of affective experiences

	PF	PI	PD	PM	PT	UF	UI	UD	UM	UT
<i>Pleasant affects</i>										
Frequency (PF)	—									
Intensity (PI)	0.48*	—								
Duration (PD)	0.26*	0.78*	—							
Mean intensity (PM)	0.84*	0.85*	0.59*	—						
Time (PT)	0.71*	0.81*	0.81*	0.90*	—					
<i>Unpleasant affects</i>										
Frequency (UF)	0.43*	-0.03	-0.06	0.31*	0.21*	—				
Intensity (UI)	-0.09	0.33*	0.31*	0.04	0.07	-0.05	—			
Duration (UD)	-0.19*	0.36*	0.53*	-0.02	0.13	0.00	0.66*	—		
Mean intensity (UM)	0.16	0.14	0.12	0.15	0.11	0.88*	0.30*	0.32*	—	
Time (UT)	0.03	0.18*	0.29*	0.07	0.13	0.76*	0.34*	0.52*	0.94*	—

* $p < 0.05$.

TABLE IV
Correlations between facets and external correlates

	RS-P	RS-U	E	N	SWLS
<i>Pleasant affects</i>					
Frequency (PF)	0.34*	0.39*	0.10	0.02	0.21*
Intensity (PI)	0.25*	0.26*	0.30*	-0.15	0.28*
Duration (PD)	0.12	0.07	0.18	-0.04	0.19*
Mean intensity (PM)	0.32*	0.36*	0.26*	-0.11	0.32*
Time (PT)	0.26*	0.25*	0.23*	-0.06	0.30*
<i>Unpleasant affects</i>					
Frequency (UF)	0.36*	0.31*	-0.24*	0.35*	-0.25*
Intensity (UI)	0.20*	0.20*	-0.11	0.21*	-0.14
Duration (UD)	0.11	0.12	-0.19*	0.30*	-0.28*
Mean intensity (UM)	0.40*	0.34*	-0.29*	0.39*	-0.37*
Time (UT)	0.32*	0.27*	-0.32*	0.44*	-0.40*

Note. RS-P = response styles pleasant affects, RS-U = response styles unpleasant affects, E = extraversion, N = neuroticism, SWLS = satisfaction with life scale. * $P < 0.05$.

Table IV shows the correlations of the experience-sampling aggregates with response styles, extraversion, neuroticism, and life-satisfaction. The correlations with the response style scales are significant, except for the duration measure, which is conceptually different from the intensity ratings of facial expressions. This finding shows that participants' intensity ratings of their own experiences are contaminated by response styles, which become more prominent in aggregated measures (Fisher et al., 1985). However, duration ratings do not share the same response styles. Hence, correlations between intensity ratings and duration ratings cannot be attributed to response styles.

Even the response styles in intensity ratings have small practical significance. For example, after controlling for the two response style measures, the correlation between frequency of pleasant affect and unpleasant affect changed from 0.43 to 0.34, the correlation between intensity of pleasant affects and intensity of unpleasant affects changed from 0.32 to 0.29, and the correlation between mean level of pleasant affect and unpleasant affect changed from 0.15 to -0.02. Response styles also have no influence on the correlations with external variables such as extraversion, neuroticism, and life-satisfaction because these measures were uncorrelated with the response style measure.

The substantive correlations in Table IV replicate and extend previous findings. First, extraversion correlates more strongly with the pleasant facets than with the unpleasant facets, whereas neuroticism correlates more strongly with the unpleasant facets than with the pleasant facets. This finding is consistent with numerous studies that relate extraversion to pleasant affects and neuroticism to unpleasant affects (Coté and Moskowitz, 1998; Costa and McCrae, 1980; Gross et al., 1997; Izard et al., 1993; Larsen and Ketelaar, 1991; Watson and Clark, 1992). However, the results are not always consistent across studies. Carstensen et al. (2000) found no significant correlation between neuroticism and intensity of unpleasant affect, whereas this correlation was significant in the present study. In contrast, Carstensen et al. (2000) found a significant correlation between extraversion and frequency of pleasant affects, whereas this correlation was not significant in the present study.

All facets except intensity of unpleasant affects predict life-satisfaction during the semester. The lack of an effect for intensity of unpleasant affects is probably due to the fact that many participants reported unpleasant affects rarely. A few intense unpleasant affects do not really undermine subjective well-being. However, intensity seemed to be important when it was considered in combination with frequency because the mean level of unpleasant affect was more strongly correlated with life-satisfaction than frequency of unpleasant affect. This possibility was examined more rigorously in a set of hierarchical regression analyses.

I first entered frequency and intensity of unpleasant affect as predictors of life-satisfaction, and I entered the mean level of unpleasant affect in a second step. Whereas frequency and intensity additively explained 9% of variance, adding mean level significantly increased predictive validity by another 8% of variance, $F(1, 116) = 11.00, p < 0.01$. This finding demonstrates that intensity of unpleasant affect contributes to life-satisfaction but only in combination with frequency. An individual with rare and intense unpleasant affects has higher life-satisfaction than an individual with frequently and mildly unpleasant experiences. However, an individual who experiences unpleasant affect frequently and mildly has higher life-satisfaction than an individual with frequent and strong unpleasant affects.

The same regression analysis was performed for intensity, frequency, and mean level of pleasant affects. Whereas frequency and intensity explained 9% of the variance, adding mean level of pleasant affects

produced a significant increase by 4%, $F(1, 116) = 5.64, p = 0.02$. This finding shows once more that intensity matters when it is seen in the context of frequency. Just a few intense pleasant experiences are not going to increase life-satisfaction, but of two individuals with similar frequencies of pleasant affects, the one with more intense experiences is likely to have higher life-satisfaction. The same analyses were conducted for frequency, duration and time of pleasant and unpleasant affects. Time of pleasant affects increased predictive validity of life-satisfaction over frequency and duration by 4%, $F(1, 116) = 4.22, p < 0.05$. However, the incremental effect for time of unpleasant affects (2% increment) was only marginally significant, $F(1, 116) = 3.82, p = 0.05$. In general, these findings support the dominant role of frequency in subjective well-being. However, once frequency is taken into account, duration or intensity also play a role. It is virtually impossible in the present data to determine whether intensity or duration were the key factors because mean level (frequency \times intensity) and time (frequency \times duration) were highly correlated with each other. Future research needs to examine this question with better measures of intensity and duration.

Specific Affects and Life-Satisfaction

The following analyses examine the contribution of specific affects to life-satisfaction. As the previous analyses showed that mean levels were the best aspect for predicting life-satisfaction, I limited the analyses to this aspect. Table V shows the correlations between mean levels of specific affects with life-satisfaction. The item happy shows a much stronger relationship with life-satisfaction than the other pleasant affects, whereas unpleasant affects are about equally related to life-satisfaction. A regression analysis revealed that happiness was the only specific affect that explained unique variance in life-satisfaction.

A hierarchical regression analysis revealed that happiness alone explained 21% of the variance in life-satisfaction; the average of unpleasant affects added 10%, whereas the average of pleasant affects added a non-significant 1% of variance. When pleasant and unpleasant affects were entered first, happiness still added a significant 4% of variance in the prediction of life-satisfaction, $F(1, 118) = 6.71, p < 0.01$. This finding demonstrates that out of the four pleasant affects, happiness is necessary and sufficient to predict life-satisfaction. No single unpleasant affect had a similar status.

TABLE V
Prediction of life-satisfaction (SWLS) by
specific pleasant and unpleasant affects

	<i>r</i>	β
<i>Pleasant affects</i>		
Happy	0.46*	0.38*
Affectionate	0.17	-0.14
Proud	0.25*	0.21
Excited	0.21*	0.03
<i>Unpleasant affects</i>		
Sad	-0.27*	-0.05
Worried	-0.33*	-0.11
Guilty	-0.29*	-0.06
Irritated	-0.34*	-0.19

* $p < 0.05$; + $p = 0.06$.

DISCUSSION

This article examined affect measurement in experience-sampling studies. Based on a review of the existing literature and new analyses of data from an electronic experience-sampling study several conclusions are well supported by data spanning 20 years of research.

Measurement Error

To use experience-sampling data in happiness research, it is important that affect ratings are valid. That is, affect ratings have to reflect individuals' internal experiences and not individual differences in the interpretation of emotion words or response scales. Fortunately, there is ample evidence that measurement error has a negligible influence on experience-sampling data that does not threaten their validity. This does not mean that experience-sampling data are free of systematic measurement error. In fact, the present study presents the strongest evidence so far for an influence of response styles on experience-sampling data. However, response styles are too weak to produce major biases in the data. In within-subject analyses across samples of experiences stylistic variance is small. It had virtually no influence on correlations between affects of opposing valence. It produced slightly stronger positive correlations between affects of the same valence. For example, the average correlation between unpleasant affects decreased from 0.42 to 0.29

after between-subject variance was removed. However, some of the removed between-subject variance was due to true individual differences in affective traits. Hence, this reduction overestimates the effect of response styles. Overall, it is safe to conclude that response styles do not change correlations between momentary measures of affect by more than 0.10. It is difficult to imagine a situation, in which such a small difference would have practical implications.

Measurement error has a stronger impact on aggregates of experience-sampling data. However, even for aggregated measures systematic measurement error has limited practical implications, especially when aggregates are related to measures that are based on different methods such as personality questionnaires. These correlations are hardly influenced by response styles in experience-sampling data because they are not shared with the other measures. Caution may be necessary when aggregates of an experience-sampling study are correlated with each other. However, even in this case, shared method variance due to response styles has a rather small effect. For example, Schimmack et al. (2002) estimated that response styles biased correlations between aggregates of daily-diary data by about 0.20. In the present study, controlling for the response style measures produced even smaller changes in correlations between aggregates of experience-sampling data. In sum, empirical data strongly support the conclusion that effects of measurement error are negligible, especially when researchers use clear and unambiguous response formats (Diener and Emmons, 1984; Schimmack et al., 2002; Watson and Clark, 1997).

Structure of Affect

Numerous experience-sampling studies have demonstrated weak within-subject and between-subject correlations between pleasant affects and unpleasant affects (Watson, 1988; Zevon and Tellegen, 1982). This finding has generated much confusion and controversy. However, most of this confusion is due to misinterpretations and ambiguous terminology and not due to inconsistent findings. Diener and Emmons (1984) proposed that pleasant affects and unpleasant affects are independent because variation in one affect can occur without reciprocal variation in the opposite affect. For example, negative affects can decrease without a similar increase in positive affects. "But if these two dimensions are indeed independent, then the distress of clients could be alleviated without producing positive affect and vice-versa" (p. 1106).

To avoid misunderstandings, this type of independence should be called discriminant validity (Schimmack et al., 2002) or uncoupled activation (Cacioppo and Berntson, 1994).

Independence of the former type should not be confused with another form of independence that is better called co-activation (Cacioppo and Berntson, 1994) or co-occurrence of pleasant and unpleasant affects (cf. Schimmack, 2001). Indeed, Diener and Emmons (1984) arguing for uncoupled activation and against co-activation. "It is highly unlikely that one can feel both positive and negative affect at the same time, especially at strong levels" (p. 1112).

The co-activation of pleasant affects and unpleasant affects is the topic of several recent investigations (see, e.g., Larsen et al., 2001; Schimmack, 2001), but it is less important for well-being researchers. Well-being researchers have typically focused on independence in the sense of discriminant validity. High discriminant validity of pleasant affects and unpleasant affects means that a complete understanding of happiness requires the assessment of both pleasant and unpleasant affects. Researchers cannot simply assume that low levels of happiness imply high levels of unhappiness and vice versa.

Discriminant validity of pleasant affects and unpleasant affects is also supported by the pattern of correlations with extraversion and neuroticism. Aggregates of pleasant affects were more highly correlated with extraversion than with neuroticism, whereas aggregates of unpleasant affects were more highly correlated with neuroticism than with extraversion. Surprisingly, the correlation between extraversion and frequency of pleasant affects was not significant, although this correlation was significant in an earlier study (Carstensen et al., 2000).

Facets of Affective Experiences and Subjective Well-Being

The present study replicated and extended earlier findings regarding the relations between aspects of affective experiences in experience-sampling data (Carstensen et al., 2000; Schimmack and Diener, 1997). Frequency and intensity of pleasant affects were positively correlated, whereas frequency and intensity of unpleasant affects were uncorrelated. The new finding was that intensity correlated highly with duration. Additional analyses revealed that response styles could not explain this correlation. A more plausible explanation is that duration judgments are influenced by intensity. This explanation is suggested by

a comparison of the frequency and duration of happiness. Frequency suggests that participants nearly always experienced at least a very mild level of happiness. However, duration judgments suggest that participants experienced happiness only for about 20 of the 30 min. The discrepancy suggests that respondents had a higher threshold for happiness when they made duration judgments than when they indicated the presence or absence of happiness. Similarly, participants may have only reported other affects on the duration scale when it exceeded a certain level of intensity on the intensity scale, which would create a positive correlation between intensity and duration ratings. Future research needs to examine this relation.

Another important finding was the relation between aspects and life-satisfaction. The present study demonstrated that life-satisfaction is best predicted by the mean level of intensity ratings of pleasant and unpleasant affects. This finding indicates that intensity ratings of affects include valid information beyond the information about the presence or absence of affects. It also suggests that intensity should be considered in measures of subjective well-being. If two individuals experience pleasant affects with the same frequency, then the one with more intense experiences can be regarded as having a happier life. Previous findings that intensity is irrelevant can be explained by different statistical procedures to separate frequency and intensity (cf. Diener et al., 1991). Furthermore, extended affect judgments may implicitly contain intensity information because participants only incorporate experiences of moderate or high intensity in frequency judgments (cf. Schimmack, 2002; Winkelman et al., 1998).

The present study also examined the relation between specific affects and life-satisfaction. The most interesting finding was obtained for happiness. Mean levels of happiness were a better predictor of life-satisfaction than the average level of all four pleasant affects. This finding suggests that subjective well-being measures should include happiness. Although this recommendation may seem trivial, the most widely used measure, the PANAS-PA scale, does not assess happiness. I recommend that happiness researchers using the PANAS should complement it with an item that assesses happiness. In the present study, sadness did not add predictive validity over the general scale of unpleasant affects. Future studies should examine in a broader set of affects which pleasant and unpleasant affects are necessary and sufficient for the prediction of life-satisfaction.

Future Research

One purpose of experience-sampling data is the assessment of happiness without the influence of memory biases. However, another important purpose of experience-sampling studies is the investigation of the processes that make people's lives pleasant or unpleasant (Kahneman et al., 1999). The present article demonstrated that aggregates of experience-sampling data are important indicators of subjective well-being. Future research needs to examine the factors that determine individuals' levels of hedonic balance. Experience-sampling studies are ideally suited for this purpose.

REFERENCES

- Andrews, F.M. and S.B. Withey: 1976, *Social Indicators of Well-being: America's Perception of Life Quality* (Plenum, New York).
- Barrett, L.F.: 1997, 'The relationships among momentary emotion experiences, personality descriptions, and retrospective ratings of emotion', *Personality and Social Psychology Bulletin* 23(10), pp. 1100–1110.
- Beebe-Center, J.G.: 1932, *The psychology of pleasantness and unpleasantness* (Van Nostrand, New York).
- Brandstaetter, H.: 1983, 'Emotional responses to other persons in everyday life situations', *Journal of Personality and Social Psychology* 45(4), pp. 871–883.
- Cacioppo, J.T. and G.G. Berntson: 1994, 'Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates', *Psychological Bulletin* 115(3), pp. 401–423.
- Carstensen, L.L., M. Pasupathi, U. Mayr and J.R. Nesselrode: 2000, 'Emotional experience in everyday life across the adult life span', *Journal of Personality and Social Psychology* 79(4), pp. 644–655.
- Costa, P.T. and R.R. McCrae: 1980, 'Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people', *Journal of Personality and Social Psychology* 38(4), pp. 668–678.
- Cote, S. and D.S. Moskowitz: 1998, 'On the dynamic covariation between interpersonal behavior and affect: Prediction from neuroticism, extraversion, and agreeableness', *Journal of Personality and Social Psychology* 75(4), 1032–1046.
- Csikszentmihalyi, M. and R. Larson: 1987, 'Validity and reliability of the experience-sampling method', *Journal of Nervous and Mental Disease* 175(9), pp. 526–536.
- Diener, E.: 1984, 'Subjective well-being', *Psychological Bulletin* 95(3), pp. 542–575.
- Diener, E. and C. Diener: 1996, 'Most people are happy', *Psychological Science* 7(3), pp. 181–185.
- Diener, E. and R.A. Emmons: 1984, 'The independence of positive and negative affect', *Journal of Personality and Social Psychology* 47(5), pp. 1105–1117.
- Diener, E., R.A. Emmons, R.J. Larsen and S. Griffin: 1985, 'The satisfaction with life scale', *Journal of Personality Assessment* 49(1), pp. 71–75.

- Diener, E. and R.J. Larsen: 1984, 'Temporal stability and cross-situational consistency of affective, behavioral, and cognitive responses', *Journal of Personality and Social Psychology* 47(4), pp. 871–883.
- Diener, E., E. Sandvik and W. Pavot: 1991, 'Happiness is the frequency, not the intensity, of positive versus negative affect', in F. Strack and M. Argyle (eds), *Subjective Well-being: An Interdisciplinary Perspective*. International series in experimental social psychology, Vol. 21 (Pergamon Press, Inc, Elmsford, NY, US), pp. 119–139.
- Diener, E., E. Sandvik, W. Pavot and D. Gallagher: 1991, 'Response artifacts in the measurement of subjective well-being', *Social Indicators Research* 24(1), pp. 35–56.
- Diener, E., H. Smith and F. Fujita: 1995, 'The personality structure of affect', *Journal of Personality and Social Psychology* 69(1), pp. 130–141.
- Diener, E., E.M. Suh, R.E. Lucas and H.L. Smith: 1999, 'Subjective well-being: Three decades of progress', *Psychological Bulletin* 125(2), pp. 276–302.
- Egloff, B.: 1998, 'The independence of positive and negative affect depends on the affect measure', *Personality and Individual Differences* 25(6), pp. 1101–1109.
- Egloff, B., A. Tausch, C. Kohlmann and H.W. Krohne: 1995, 'Relationships between time of day, day of the week, and positive mood: Exploring the role of the mood measure', *Motivation and Emotion* 19(2), pp. 99–110.
- Epstein, S.: 1982, 'A research paradigm for the study of personality and emotions', *Nebraska Symposium on Motivation* 1982, pp. 91–154.
- Fisher, G.A., D.R. Heise, G.W. Bohrnstedt and J.F. Lucke: 1985, 'Evidence for extending the circumplex model of personality trait language to self-reported moods', *Journal of Personality and Social Psychology* 49(1), pp. 233–242.
- Flügel, J.C.: 1925, 'A quantitative study of feeling and emotion in everyday life', *British Journal of Psychology* 15, pp. 318–355.
- Frijda, N.H., A. Ortony, J. Sonnemans and G.L. Clore: 1992, 'The complexity of intensity: Issues concerning the structure of emotion intensity', in M.S. Clark (ed.), *Emotion. Review of Personality and Social Psychology*, No. 13 (Sage, Thousand Oaks, CA, US), pp. 60–89.
- Goldberg, L.: 1997, A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. [Online] available: <http://www.ipio.ori.org/ipip/>.
- Green, D.P., S.L. Goldman and P. Salovey: 1993, 'Measurement error masks bipolarity in affect ratings', *Journal of Personality and Social Psychology* 64(6), pp. 1029–1041.
- Gross, J.J., S.K. Sutton and T. Ketelaar: 1998, 'Relations between affect and personality: Support for the affect-level and affective reactivity views', *Personality and Social Psychology Bulletin* 24(3), pp. 279–288.
- Hedges, S.M., L. Jandorf and A.A. Stone: 1985, 'Meaning of daily mood assessments', *Journal of Personality and Social Psychology* 48(2), pp. 428–434.
- Hormuth, S.E.: 1986, 'The sampling of experiences in situ', *Journal of Personality* 54(1), pp. 262–293.
- Izard, C.E., D.Z. Libero, P. Putnam and O.M. Haynes: 1993, 'Stability of emotion experiences and their relations to traits of personality', *Journal of Personality and Social Psychology* 64(5), pp. 847–860.

- Kahneman, D., E. Diener and N. Schwarz: (eds) 1999, *Well-being: The Foundations of Hedonic Psychology* (Russell Sage Foundation, New York, NY, US).
- Kennedy-Moore, E., M.A. Greenberg, M.G. Newman and A.A. Stone: 1992, 'The relationship between daily events and mood: The mood measure may matter', *Motivation and Emotion* 16(2), pp. 143–155.
- Larsen, J.T., A.P. McGraw and J.T. Cacioppo: 2001, 'Can people feel happy and sad at the same time?' *Journal of Personality and Social Psychology* 81(4), pp. 684–696.
- Larsen, R.J. and E. Diener: 1987, 'Affect intensity as an individual difference characteristic: A review', *Journal of Research in Personality* 21(1), pp. 1–39.
- Larsen, R.J. and T. Ketelaar: 1991, 'Personality and susceptibility to positive and negative emotional states', *Journal of Personality and Social Psychology* 61(1), pp. 132–140.
- Larson, R. and M. Csikszentmihalyi: 1983, 'The experience sampling method', *New Directions for Methodology of Social and Behavioral Science* 15(Mar.).
- McAdams, D.P. and C.A. Constantian: 1983, 'Intimacy and affiliation motives in daily living: An experience sampling analysis', *Journal of Personality and Social Psychology* 45(4), pp. 851–861.
- Meddis, R: 1972, 'Bipolar factors in mood adjective checklists', *British Journal of Social and Clinical Psychology* 11(2), pp. 178–184.
- Moscowitz, D.S. and S. Cote: 1995, 'Do interpersonal traits predict affect? A comparison of three models', *Journal of Personality and Social Psychology* 69(5), pp. 915–924.
- Nowlis, V: 1965, 'Research with the mood adjective check list', in *Affect, Cognition, and Personality: Empirical Studies* (Springer, New York), pp. 352–389.
- Oatley, K. and E. Duncan: 1994, 'The experience of emotions in everyday life', *Cognition and Emotion* 8(4), pp. 369–381.
- Oishi, S., U. Schimmack and E. Diener: 2001, 'Pleasures and subjective well-being', *European Journal of Personality* 15, pp. 153–167.
- Parkinson, B., R.B. Briner, S. Reynolds and P. Totterdell: 1995, 'Time frames for mood: Relations between monetary and generalized ratings of affect', *Personality and Social Psychology Bulletin* 21(4), pp. 331–339.
- Rorer, L.G: 1965, 'The great response-style myth', *Psychological Bulletin* 63(3), pp. 129–156.
- Schimmack, U: 2001, 'Pleasure, displeasure, and mixed feelings: Are semantic opposites mutually exclusive?' *Cognition and Emotion* 15(1), pp. 81–97.
- Schimmack, U: (in press). 'Frequency judgments of emotions: The cognitive basis of personality assessment', in P. Sedelmeier and T. Betsch (eds), *Frequency Processing and Cognition* (Oxford University Press, Oxford).
- Schimmack, U: 1997, 'The Berlin everyday language mood inventory (BELMI): Toward the content valid assessment of moods', *Diagnostica* 43(2), pp. 150–173.
- Schimmack, U. and E. Diener: 1997, 'Affect intensity: Separating intensity and frequency in repeatedly measured affect', *Journal of Personality and Social Psychology* 73(6), pp. 1313–1329.

- Schimmack, U. and A. Grob: 2000, 'Dimensional models of core affect: A quantitative comparison by means of structural equation modeling', *European Journal of Personality* 14(4), pp. 325–345.
- Schimmack, U., S. Oishi, E. Diener and E. Suh: 2000, 'Facets of affective experiences: A framework for investigations of trait affect', *Personality and Social Psychology Bulletin* 26(6), pp. 655–668.
- Schimmack, U., U. Bockenholt and R. Reisenzein: 2002, 'Response styles in affect ratings: Making a mountain out of a molehill', *Journal of Personality Assessment* 78, pp. 461–483.
- Schimmack, U., E. Diener and S. Oishi: 2002, 'Life-satisfaction is a momentary judgment and a stable personality characteristic: The use of chronically accessible and stable sources', *Journal of Personality* 70, pp. 345–385.
- Schimmack, U., P. Radhakrishnan, S. Oishi, V. Dzokoto and S. Ahadi: 2002, 'Culture, personality, and subjective well-being: Integrating process models of life-satisfaction', *Journal of Personality and Social Psychology* 82, pp. 1313–1329.
- Schwarz, N., B. Knäuper, H. Hippler, E. Noelle-Neumann and F. Clark: 1991, 'Rating scales: Numeric values may change the meaning of scale labels', *Public Opinion Quarterly* 55(4), pp. 570–582.
- Shapiro, D., L.D. Jamner, I.B. Goldstein and R.J. Delfino: 2001, 'Striking a chord: Moods, blood pressure, and heart rate in everyday life', *Psychophysiology* 38(2), pp. 197–204.
- Shaver, P., J. Schwartz, D. Kirson and C. O'Connor: 1987, 'Emotion knowledge: Further exploration of a prototype approach', *Journal of Personality and Social Psychology* 52(6), pp. 1061–1086.
- Stone, A.A. and J.M. Neale: 1984, 'New measure of daily coping: Development and preliminary results', *Journal of Personality and Social Psychology* 46(4), pp. 892–906.
- Suh, E., E. Diener, S. Oishi and H.C. Triandis: 1998, 'The shifting basis of life satisfaction judgments across cultures: Emotions versus norms', *Journal of Personality and Social Psychology* 74(2), pp. 482–493.
- Svensson, E: 1977, 'Response format and factor structure in mood adjective check lists', *Scandinavian Journal of Psychology* 18(1), pp. 71–78.
- Thomas, D.L. and E. Diener: 1990, 'Memory accuracy in the recall of emotions', *Journal of Personality and Social Psychology* 59(2), 291–297.
- Watson, D: 1988, 'The vicissitudes of mood measurement: Effects of varying descriptors, time frames, and response formats on measures of positive and negative affect', *Journal of Personality and Social Psychology* 55(1), pp. 128–141.
- Watson, D. and L.A. Clark: 1992, 'On traits and temperament: General and specific factors of emotional experience and their relation to the five-factor model', *Journal of Personality* 60(2), pp. 441–476.
- Watson, D. and L.A. Clark: 1991, 'Self- versus peer ratings of specific emotional traits: Evidence of convergent and discriminant validity', *Journal of Personality and Social Psychology* 60(6), pp. 927–940.

- Watson, D., L.A. Clark and A. Tellegen: 1988, 'Development and validation of brief measures of positive and negative affect: The PANAS scales', *Journal of Personality and Social Psychology* 54(6), pp. 1063–1070.
- Watson, D., L.A. Clark and A. Tellegen: 1984, 'Cross-cultural convergence in the structure of mood: A Japanese replication and a comparison with U.S. findings', *Journal of Personality and Social Psychology* 47(1), pp. 127–144.
- Watson, D. and L.A. Clark: 1997, 'Measurement and mismeasurement of mood: Recurrent and emergent issues', *Journal of Personality Assessment* 68(2), pp. 267–296.
- Winkielman, P., B. Knäuper and N. Schwarz: 1998, 'Looking back at anger: Reference periods change the interpretation of (emotion) frequency questions', *Journal of Personality and Social Psychology* 75, pp. 719–728.
- Zelenski, J.M. and R.J. Larsen: 2000, 'The distribution of basic emotions in everyday life: A state and trait perspective from experience sampling data', *Journal of Research in Personality* 34(2), pp. 178–197.
- Zevon, M.A. and A. Tellegen: 1982, 'The structure of mood change: An idiographic/nomothetic analysis', *Journal of Personality and Social Psychology* 43(1), pp. 111–122.

Address for correspondence:

ULRICH SCHIMMACK

Department of Psychology

University of Toronto, Mississauga

3359 Mississauga Road North

Mississauga, Ont., Canada, L5L 1C6

E-mail: uli.schimmack@utoronto.ca