The Relationships Among Momentary Emotion Experiences, Personality Descriptions, and Retrospective Ratings of Emotion.

by LISA FELDMAN BARRETT

In psychological research, respondents often make retrospective ratings of their emotional experiences after an extended period of time. The present study sought to determine whether such memory-based ratings are influenced by respondents' descriptions of their own emotionality, over and above a summary of their momentary emotion ratings. Participants completed self-report measures of neuroticism and extraversion and made momentary ratings of their emotions across 90 days. At the end of the study, participants recalled what their emotions had been during the course of the study. Findings indicated that retrospective ratings of emotion contained accurate information about momentary emotion reports. Also, the retrospective ratings were influenced in the direction of respondents' personality descriptions. Individuals who described themselves as neurotic remembered experiencing more negative emotion than they reported on a momentary basis; individuals who described themselves as extraverted displayed a trend to remember more positive emotion than they reported on a momentary basis.

Much of psychological knowledge is based on what respondents tell researchers about themselves. When psychologists measure emotion, they often ask respondents to judge their emotional experiences over the past day, the past week, or in general. Self-reports of this type are assumed to be based on participants' accurate recall of their past experiences. But are retrospective ratings accurate?

Some researchers believe that when completing questionnaires requiring global responses, respondents produce biased reports of their actual experiences (for a discussion, see Reis & Wheeler, 1991; Stone & Shiffman, 1994). To describe themselves using such questionnaires, respondents must remember, summarize, and integrate their past experiences into a consistent set of responses to the questionnaire items. Recalling information is a reconstructive process, however. The research on social cognition is full of evidence that heuristics, cognitive structures, implicit theories, and motivations lead to recollections that can be inaccurate (for reviews, see Fiske & Taylor, 1991; Greenwald & Banaji, 1995; Nisbett & Ross, 1980; Ross, 1989). Although researchers vary in the degree to which they believe that recall-based information is reconstructed versus retrieved from memory, most researchers would not disagree with a moderate reconstructive view, in which respondents rely on both processes to produce recall-based ratings (see Schwarz & Sudman, 1994).

The purpose of this study was to explore the meaning of recall-based ratings of emotional states. Do people accurately retrieve their moment-to-moment emotion experiences across some period of time and integrate this information into a set of responses, or do they reconstruct this information based on beliefs about their own emotionality?

Retrospective Ratings of Affective States

A small amount of research has examined the relationship between momentary and recall-based ratings of emotion. Much of this research has tried to answer the question of how respondents use multiple, momentary experiences to arrive at a single, global set of emotion ratings. Some findings suggest that the most intense (peak) momentary emotion experience of the day determines recall-based ratings of emotion (Hedges, Jandorf, & Stone, 1985), whereas other findings suggest that the average of momentary emotion experiences are more influential (Parkinson, Briner, Reynolds, & Totterdell, 1995). Some findings suggest that final emotions have a small but significant relationship to retrospective emotion ratings (Holmberg & Holmes, 1994; Parkinson et al., 1995), whereas others suggest that respondents use a combination of the most intense and the final moments of extended affective episodes as proxies for making retrospective judgments of their affective states (Fredrickson & Kahneman, 1993; Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993). When peak and end effects are present, the average of momentary affective experiences offers no new information over and above the peak and end effects, suggesting that both account for the retrospective judgments equally well (Fredrickson & Kahneman, 1993). The duration of affective states across time plays a small role in respondents' evaluations of their prior affective experiences (Fredrickson & Kahneman, 1993; Kahneman et al., 1993), suggesting that stability of emotions across time (e.g., Izard, Libero, Putnam, & Hayes, 1993) should not be related to increased accuracy of retrospective emotion ratings.

The Relationships Among Momentary Emotion Experiences, Personality Descriptions, and Retrospective Ratings of Emotion.

Although momentary ratings are related to retrospective ratings, bias in recall-based emotion ratings has been found (Parkinson et al., 1995; Thomas & Diener, 1990). When making daily and weekly reports, some respondents tend to overestimate the amount of positive emotion they experienced, and some tend to overestimate the amount of negative emotion they experienced, relative to that reported on a momentary basis. One possible explanation for these biases is that, in part, respondents constructed their recall-based rating in line with their beliefs about their own emotionality.

Research suggests that people's beliefs guide their reconstruction of their past (Ross, 1989). For example, women's recall-based reports of their menstrual symptoms are biased so as to be consistent with their theories of menstrual distress (Boyle & Grant, 1992; MacFarlane, Martin, & Williams, 1988; McFarland, Ross, & DeCourville, 1989). Spouses' recall-based reports of their daily assessments of feelings for their partners are biased by their beliefs about how much they trust their partners (Holmberg & Homes, 1994). People's descriptions of their personality also influence their recall-based ratings. People who describe themselves as high in neuroticism remember experiencing more physical symptoms than they actually experienced (e.g., Larsen, 1992; Watson & Pennebaker, 1989).

A person's responses to personality questionnaires provide a picture of that person's self-perceptions (i.e., what the person thinks he or she is like). When respondents describe themselves using items such as "I rarely feel fearful or anxious" (appearing on the anxiety facet of neuroticism), "Too often, when things go wrong, I get discouraged and feel like giving up" (appearing on the depression facet of neuroticism), "Even minor annoyances can be frustrating to me" (appearing on the hostility facet of neuroticism), and "I am a cheerful,

high-spirited person" (appearing on the positive emotions facet of extraversion), they have few immediate behavioral cues with which to base their ratings. Therefore, they generate their ratings in their head. In part, they may be retrieving, summarizing, and integrating their past experiences when responding to the questionnaire items, respondents' ratings are accurate when they represent retrievals of their actual behavior. Note that even if their personality ratings are significantly related to their behavior, the ratings themselves reflect respondents' descriptions of themselves that are abstracted from their actual behaviors.

People may do more than merely summarize their past experiences, however; they may magnify or enhance them. Given that memory is a selective and reconstructive process (Ross, 1989), respondents' questionnaire ratings may contain variance that does not accurately represent their actual thoughts, feelings, and behaviors over time. This variance represents a component of their self-perceptions that are biased relative to the experiences that occur on a moment-to-moment basis. Most likely, personality ratings reflect both accurate variance (i.e., a summary of respondents' actual behavior) and inaccurate variance (i.e., respondents' beliefs about themselves that are not representative of their actual behavior).

Recall-based emotion ratings, too, may reflect both respondents' retrieval of their actual emotional experiences across time and their less accurate beliefs about their emotional lives. The latter may lead respondents to overestimate or underestimate the amount of positive and negative emotions that they actually experienced during some previous time span.

Personality Descriptions and Emotion Ratings

In the past decade, systematic links have been made between self-reported personality measurements and affective experiences. Much of this research has focused on neuroticism and extraversion. Measures of neuroticism contain items that lead respondents to describe their degree of negative emotional experiences and have been associated with reports of negative emotions; in contrast, measures of extraversion contain items that lead respondents to describe their degree of chronic positive emotional experiences and have been associated with reports of positive emotions (e.g., Costa & McCrae, 1992; Izard et al., 1993; Larsen & Ketellar, 1991; McCrae & Costa, 1991; Watson & Clark, 1992). These findings have been interpreted to mean that neuroticism predisposes or causes individuals to experience negative emotions, whereas extraversion predisposes or causes individuals to experience positive emotions. In fact, neuroticism has been referred to as "negative emotionality" or "negative temperament," whereas extraversion has been referred to as "positive emotionality" or "positive temperament" (e.g., Tellegen, 1985; Watson & Clark, 1992, in press).

An alternative view is that neuroticism and extraversion ratings represent respondents' beliefs about their previous

The Relationships Among Momentary Emotion Experiences, Personality Descriptions, and Retrospective Ratings of Emotion.

negative and positive emotional experiences. From this perspective, ratings of neuroticism and extraversion represent respondents' beliefs about their tendency to have negative and positive emotional experiences, rather than the causal factors that produce future experiences. Respondents' ratings likely represent the emotional experiences that they summarize and reconstruct from memory. For example, individuals who have experienced a lot of positive emotion in the past will believe themselves to be the type of person who experiences a lot of positive emotion and will subsequently rate themselves as high on extraversion. Part of their belief may be accurate: These individuals likely experience more positive emotion than those who describe themselves as low in extraversion. In addition, however, the ratings may represent respondents' inaccurate beliefs about their own emotionality: Individuals high in extraversion may conclude that they felt more positive emotion in the past than they actually experienced. Similarly, individuals who experience a lot of negative emotion will believe themselves to be the type of person who experiences a lot of negative emotion (i.e., they will have accurate self-beliefs) and will subsequently rate themselves as high on neuroticism. These individuals may also overestimate the amount of negative emotion that they experienced in the past (i.e., some portion of their self-beliefs are inaccurate).

Figure 1 portrays one way in which momentary emotional experiences and personality descriptions are related to recall-based ratings of emotion. Respondents must recall and summarize their momentary emotion experiences when describing themselves using both self-report measures of personality and recall-based ratings of emotion. Thus, momentary emotion ratings contribute directly to recall-based ratings of emotion ([p.sub.31]), as well as to the personality descriptions ([p.sub.21]). Momentary emotion ratings may also contribute indirectly to the recall-based ratings through personality descriptions; respondents describe their personality based in part on their momentary experiences, and this description leads them toward an accurate recall of their emotions. In addition, personality descriptions may also contribute directly to the recall-based ratings of emotion ([p.sub.32]); this path represents the relationship between respondents' personality descriptions and their recall-based ratings of emotion that is not due to momentary emotion ratings. That is, this path represents the degree to which ratings of recall-based emotion are related to the inaccurate portion of respondents' beliefs about their personality.

[ILLUSTRATION OMITTED]

OVERVIEW OF THE PRESENT STUDY

The major goal of this study was to determine whether retrospective ratings of emotion were influenced by the inaccurate portion of respondents' beliefs about their personality. Toward this end, self-report measures of neuroticism and extraversion were collected at the beginning of the study. Participants then rated their momentary emotion three times a day for 90 days. At the end of the study, participants were asked to make general ratings of how they had felt across the 90 days of the study. To take a disconfirming approach to testing the inaccuracy hypothesis, I first identified the strongest link possible between the momentary and the retrospective ratings of emotion so as to maximize the accurate variance estimated in the retrospective ratings. I then predicted that respondents' personality ratings would be associated with their retrospective ratings after all the accurate variance in the retrospective ratings had been accounted for.

Method

SUBJECTS

Participants were 56 undergraduates (16 male and 40 female) from the Department of Psychology at the Pennsylvania State University. Participants both received course credit and earned lottery tickets for their participation. The data for these participants were obtained from a larger longitudinal study on affective differentiation (Barrett, 1995). The hypotheses tested and analyses covered in that report does not directly overlap with those presented here. Only the materials relevant to this report are presented here.

The study began with 77 potential participants. Given the time-consuming nature of this study, 16 participants (21% of the original sample) dropped out in the first few weeks of the study. The partial data sets from these subjects were not included in the analyses. Five more participants were deleted for making a large number of retrospective mood ratings. All of the remaining 56 participants used memory less than 20% of the time and missed completing fewer than 13% of their mood measurements (the average percentage of missed observations was 1.6%). The number of usable observations for

The Relationships Among Momentary Emotion Experiences, Personality Descriptions, and Retrospective Ratings of Emotion.

each participant ranged from 235 to 270, with a mean of 266 and a standard deviation of 6.8. No participant reported awareness of the hypotheses under investigation. Missing data points were treated as missing and were not replaced with estimated values.

PROCEDURE

Participants began the study by completing a battery of personality questionnaires including a measure of neuroticism and extraversion. Upon completing the personality measures, participants completed an emotion measure every morning (7 a.m. to 12 p.m.), afternoon (12 p.m. to 5 p.m.), and evening (5 p.m. to 12 a.m.) of each day for 90 consecutive days. Participants chose when they completed each measure during the specified time periods, and they returned completed forms on Monday, Wednesday, and Friday of each week. Participants who failed to return their ratings on time were contacted by an experimenter within 48 hrs. In addition, participants were interviewed three times during the study to ensure their compliance with the research procedures. One day after respondents completed the longitudinal reporting phase of the study, they met in several groups and completed a retrospective emotion assessment. Only 52 subjects completed the recall phase of the study.

MEASURES

Personality assessment. The battery of personality measures included the revised NEO Personality Inventory (NEO PI-R; Costa & McCrae, 1992), along with several other personality questionnaires. Because the hypotheses concern neuroticism and extraversion, only these variables are considered in relation to the emotion ratings. The Neuroticism and Extraversion Scales of the NEO PI-R are widely used in research on personality and emotional experience. Both domain scales consist of six facets made up of eight items each (facets of neuroticism are anxiety, hostility, depression, self-consciousness, impulsiveness, and vulnerability; facets of extraversion are warmth, gregariousness, assertiveness, activity, excitement-seeking, and positive emotions). Composites of the neuroticism and extraversion facets that most directly related to experiences of negative emotion (anxiety, hostility, and depression) and to experiences of positive emotion (warmth, activity, and positive emotions) were used in the main analyses of the study because they reflected most specifically respondents' self-descriptions of their own emotionality. Some analyses were rerun using the domain scores (scores involving all six facets). Participants responded to each NEO PI-R item using a 5-point Likert-type ratings scale (1 = strongly disagree, 3 = neutral, 5 = strongly agree).

Momentary emotion questionnaire. The emotion measure, containing 88 affect terms, was designed to sample all octants of the affective circumplex (Feldman, 1995; Barrett, 1995). A total of 60 terms belonged to the Positive Affect Negative Affect Schedule-Expanded Form (PANAS-X; Watson & Clark, 1994), and others were sampled from the remaining octants of the circumplex (Larsen & Diener, 1992). Participants indicated on a 7-point Likert-type scale the extent to which each adjective described their emotion at the moment of rating (0 = not at all, 3 = a moderate amount, 6 = a great deal).

Several sets of scales were constructed from the PANAS-X items (Watson & Clark, 1994): A Negative Emotion Scale was computed by averaging the scores from several subscales measuring negatively valenced emotional states: Sadness (sad, blue, downhearted, alone, and lonely), Hostility (angry, hostile, irritable, scornful, disgusted, loathing), Guilt (guilty, ashamed, blameworthy, angry with self, disgusted with self, dissatisfied with self), and Fear (afraid, scared, frightened, nervous, jittery, and shaky). The Positive Emotion Scale was computed by using the Joviality subscale of the PANAS-X (enthusiastic, excited, lively, energetic, happy, delighted, joyful, and cheerful).

Retrospective emotion assessment. Participants completed a version of the emotion measure but were asked to rate the 88 affect items according to the following instructions: "Please rate your feelings over the past three months in general." Participants used the same 7-point Likert-type scale as described above. Each of the PANAS-X scales was constructed using the recall-based ratings.

Results

This section begins with analyses that identify the strongest momentary predictors of the retrospective emotion ratings. Once indexes were identified that would maximize the amount of accurate variance accounted for in the retrospective ratings, I then assessed whether respondents' inaccurate beliefs about their own emotionality were associated with their

The Relationships Among Momentary Emotion Experiences, Personality Descriptions, and Retrospective Ratings of Emotion.

retrospective emotion ratings, over and above the summary of their momentary emotional experiences.

SELECTING THE MOST PREDICTIVE SUMMARY INDEX OF MOMENTARY EMOTION RATINGS

Previous research has suggested that either averaged momentary emotion, peak emotion, or some combination of peak- and end-state emotions will best predict retrospective emotion ratings (Fredrickson & Kahneman, 1993; Hedges et al., 1985; Parkinson et al., 1995). It was not possible to investigate a true recency effect in the present data set because participants made their retrospective ratings approximately 24 hrs after their last momentary assessment. It was possible, however, to compare several summary indexes to determine which accounted for the most variance in the retrospective ratings, to then allow for the subsequent analyses.

I computed three types of summary indexes to determine which best predicted the retrospective emotion judgments: The mean (average), the peak (most intense), and the modal (most frequent) scores were calculated for both the Positive and the Negative Emotion Scales. The mean scores were computed for the final day of the study, the final week of the study, the final month of the study, and the entire period of the study. Peak emotion was defined as the score at the 90th percentile for each participant, and modal scores were calculated across the entire period of the study.

The Pearson product-moment correlations between the various summary indexes and the recall-based emotion ratings are presented in Table 1. Averaging across more observations increased the relationship between the average of the momentary ratings of negative emotion and the retrospective ratings of negative emotion. Peak momentary emotion was also strongly related to the retrospective ratings of negative emotion. The correlations between the average of the momentary ratings of positive emotion and the retrospective ratings of positive emotion were all substantial, suggesting that averaging positive ratings across the entire period of the study did not improve the prediction of the recall-based ratings over and above averaging across the last week of the study.

TABLE 1: Correlations Between Summary Scores of the Momentary Emotion Ratings and the Recall-Based Ratings of Emotion

	Recall-Based Negative Emotion	Recall-Based Positive Emotion
Averaged emotion		
Final day	.36(***)	.57(***)
Final week	.45(***)	.74(***)
Final month	.58(***)	.72(***)
Entire study	.63(***)	.76(***)
Peak and modal emotion		
Peak	.64(***)	.68(***)
Mode	.36(***)	.65(***)

NOTE: N = 52. Average = the mean of the momentary emotion ratings; Peak = the most intense 10% of the momentary emotion ratings; Mode = the most frequent momentary emotion rating.

(***) p < .01, two-tailed.

I chose to focus future analyses on the summary indexes that had the highest zero-order correlations with the retrospective ratings. It did not seem necessary to statistically compare the summary indexes because this first set of analyses was designed solely to identify which summary indexes of momentary emotion accounted for the most variance in recall ratings so as to then allow for a test of the inaccuracy hypothesis (i.e., the goal was not to determine the best predictor of the retrospective emotion ratings in some absolute sense). Moreover, it was difficult to statistically compare the predictive power of the various summary indexes because they were highly correlated with one another and were often summarizing overlapping data. On the basis of the correlations presented in Table 1, both the total study average of the momentary ratings and the peak momentary emotion were selected for further analyses of negative emotion, whereas both the final week average and the total study average of the momentary ratings were selected for further analyses of positive emotion.

The Relationships Among Momentary Emotion Experiences, Personality Descriptions, and Retrospective Ratings of Emotion.

RELATIONSHIPS AMONG PERSONALITY AND MOMENTARY AND RECALL-BASED RATINGS OF EMOTION

Descriptive statistics for the summary indexes of momentary emotion, the retrospective ratings of emotion, their discrepancies, and the personality ratings are presented in Table 2. The discrepancies between the recall-based ratings and the momentary ratings of emotion were calculated by subtracting the latter from the former for both positive and negative emotion scales. A negative number indicates that the participant recalled less emotion than he or she reported on a momentary basis, and a positive number indicates that the participant recalled more emotion than he or she reported on a momentary basis. None of the mean discrepancies scores differed significantly from zero, but the range of the scores indicated that some people overestimated, whereas others underestimated, both the amount of negative emotion and the amount of positive emotion that they reported on a momentary basis. The discrepancy scores for positive and negative emotion were significantly correlated (rs ranged from .30, p [is less than] .05, to .37, p [is less than] .01), indicating that when participants underestimated their reports of negative emotion, they also underestimated their reports of positive emotion; similarly, when participants overestimated their reports of negative emotion, they also overestimated their reports of positive emotion.

TABLE 2: Descriptive Statistics for Emotion and Personality Ratings

	M	SD	Min.	Max.
Negative emotion				
Retrospective	0.6	0.7	0	2.5
Momentary--Peak	1.2	0.7	0	3.0
Discrepancy--Peak	-0.6	0.6	-2.2	0.9
Momentary--Mean entire study	0.4	0.3	0	2.0
Discrepancy--Mean entire study	0.2	0.6	-0.7	2.1
Positive emotion				
Retrospective	2.5	1.3	0.3	5.6
Momentary--Mean final week	2.0	0.9	1.0	4.0
Discrepancy--Mean final week	0.5	0.9	-1.7	3.0
Momentary--Mean entire study	2.0	0.9	0	4.0
Discrepancy--Mean entire study	0.5	0.8	-1.2	3.5
Personality ratings				
Neuroticism	17.6	4.5	7.3	28.0
Extraversion	23.1	3.0	15.0	29.3

NOTE: N = 52. A negative discrepancy value indicates that the momentary index of emotion is larger than the recall-based rating; a positive discrepancy value indicates that the recall-based rating is larger than the momentary index.

Recall of negative emotions. Pearson product-moment correlations were first computed between the neuroticism composite, the summary indexes of momentary negative emotion, and retrospective ratings of negative emotion. These are presented in the top half of the second and third data columns in Table 3. Although the neuroticism composite appeared to be more strongly related to the recall-based ratings (.38) than to either of the summary indexes (.24 for the averaged ratings and .29 for the peak ratings), these differences were not statistically significant (zs = 1.2 and 0.8, respectively; both ns;, Meng, Rosenthal, & Rubin, 1992).

TABLE 3: Correlational Analyses of Personality, Momentary Emotion Indexes, and Recall-Based Emotion Ratings

	Negative Emotion	
	Mean-E	Peak
Zero-order correlation		
M-P	.24 (.21)	.29(**) (.28(**))
P-R	.38(***)	.38(***)

The Relationships Among Momentary Emotion Experiences, Personality Descriptions, and Retrospective Ratings of Emotion.

M-R	.63(***)	.64(***)
Path		
[p.sub.21]: Direct effect of M on P	.24 (.21)	.29(**) (.28(**))
[p.sub.31]: Direct effect of M on R	.57(***) (.58(***))	.58(***)
[p.sub.21] X [p.sub.32]: Indirect effect (M--P--R)	.06 (.05)	.06
[p.sub.32]: Unique influence of P on R	.25(**)	.21(*)
	Positive Emotion	
	Mean-W	Mean-E
Zero-order correlation		
M-P	.18 (.19)	.29(***)
P-R	.34(**)	.34(**)
M-R	.74(***)	.76(***)
Path		
[p.sub.21]: Direct effect of M on P	.18 (.19)	.29(***)
[p.sub.31]: Direct effect of M on R	.70(***)	.72(***)
[p.sub.21] X [p.sub.32]: Indirect effect (M--P--R)	.04	.04
[p.sub.32]: Unique influence of P on R	.21(**) (.20(**))	.13

NOTE: N = 52. Mean-E = momentary emotion averaged over the entire study; Mean-W = momentary emotion averaged over the final week; Peak = the most intense 10% of the momentary emotion ratings; M = summary index for momentary emotion ratings; P = personality scores (neuroticism facets for analyses of negative emotion; extraversion facets for analyses of positive emotion); R = recall-based emotion ratings. Coefficients in parentheses indicate changes in results when using neuroticism or extraversion domain scores rather than facet composites.

(*) p < .10. (**) p < .05, two-tailed. (***) p < .01, two-tailed.

The zero-order correlations were decomposed using path analysis (Pedhazer, 1982). One set of analyses was conducted for the averaged momentary ratings, and another set for the peak momentary ratings. According to the model presented in Figure 1, the correlation between momentary emotion ratings and the neuroticism composite was equal to their zero-order correlation. The correlations between momentary and recall-based ratings of negative emotion were decomposed into a direct effect ([p.sub.21]) and an indirect effect ([p.sub.31] x [p.sub.32]). The direct effect is the standardized regression coefficient between the summary score and the recall-based ratings, controlling for neuroticism. The decompositions appear in the bottom half of the second and third data columns of Table 3. The direct effect of both summary indexes on the retrospective ratings of negative emotion was substantial and significant; this is not surprising given that the summary indexes were selected to maximize the accuracy of the retrospective ratings. The direct effects were larger than the indirect effects, suggesting that respondents' momentary emotional experiences did not strongly influence their recall-based ratings indirectly through their personality descriptions.

The correlation between the neuroticism composite and the recall-based ratings of negative emotions was decomposed into direct ([p.sub.32]) and spurious effects.(1) The direct effects of neuroticism on recall-based negative emotions (final data row of Table 3) were significant (.25) or marginally significant (.21). As predicted, the retrospective ratings of emotion

The Relationships Among Momentary Emotion Experiences, Personality Descriptions, and Retrospective Ratings of Emotion.

contained some inaccurate variance. Respondents' descriptions of their level of neuroticism influenced their recall-based ratings of negative emotion independently of their momentary ratings. Individuals who believe that they chronically experience negative emotions tended to overestimate the degree of negative emotion that they experienced relative to that which they reported on a moment-to-moment basis.

Thus, retrospective ratings of negative emotion contained both accurate and inaccurate variance. The total effects (direct plus indirect effects) of the momentary ratings of negative emotion on the recall-based ratings were larger than the effect of the neuroticism composite on the recall-based ratings, indicating that recall-based ratings of negative emotions contain largely accurate information. Nevertheless, the direct effect of neuroticism on the recall-based ratings was statistically significant. This effect represents inaccuracy in the recall-based ratings, because it is the component of self-reported neuroticism that predicted the recall-based ratings after the accurate variance (due to momentary ratings) was accounted for.

All path analyses were conducted again using the entire neuroticism domain scores, rather than the composite of the three negative emotion facets. These analyses are presented in Table 3 and appear in parentheses when they differ from the findings already presented. Overall, the results using the domain scores were virtually identical to those using the facet composite. In addition, all analyses were rerun separately for ratings of discrete emotions (e.g., sadness, fear, guilt, and hostility). These analyses are available from the author on request. Once again, the findings were highly similar, in that all retrospective ratings of emotion were strongly related to the momentary ratings of emotion. Respondents' descriptions of themselves on neuroticism made an independent contribution to the recall-based ratings of guilt and fear but not to ratings of sadness and hostility.

Some personality researchers claim that neuroticism predisposes or causes an individual to experience negative emotional states, rather than merely representing a summary of those experiences. The inaccuracy hypothesis is supported even if this theory is used to decompose the zero-order correlations. From the perspective of this theory, neuroticism would have direct effects on both momentary and recall-based negative emotion, and an indirect effect on recall-based negative emotion through the momentary ratings. The direction of the [p.sub.21] path is reversed to represent the total effect of neuroticism on the momentary negative ratings, but the value of the path remains the same. The total effect of neuroticism on the recall-based ratings would increase to equal the zero-order correlation between the two (.38, p [is less than] .01). And the total effect of momentary negative emotion on the recall-based ratings would decrease slightly (equaling only [p.sub.31] rather than [r.sub.31]); the remainder of the zero-order correlation between the momentary ratings and the recall-based ratings would be considered a spurious effect of neuroticism. Therefore, even when the directionality in the original theoretical model is challenged, the results are essentially identical. Recall-based ratings of negative emotion are better predicted by the momentary ratings of negative emotion than by neuroticism, but neuroticism ratings remain uniquely related to the recall-based ratings.

Recall of positive emotions. Pearson product-moment correlations were computed between the extraversion composite, the summary indexes for the momentary ratings of positive emotion, and the recall-based ratings of positive emotion. These are presented in the upper portion of the fourth and fifth data columns in Table 3. Although the extraversion composite appeared to be more strongly related to the recall-based ratings (.34) than to either of the summary indexes (.18 for the final week average and .29 for the total study average), these differences were not statistically significant (zs = 0.5 and 1.6, respectively; both ns;, Meng et al., 1992).

As with the negative emotion ratings, the correlations associated with positive ratings were decomposed using path analysis and appear in the lower part of the fourth and fifth data columns in Table 3. The direct effects between momentary and recall-based ratings of positive emotion were large and significant. Furthermore, the indirect effects were small, suggesting that respondents' momentary ratings of positive emotion did not influence their recall-based ratings indirectly through their personality description. The direct effect of the extraversion composite on the recall-based ratings of positive emotion was significant for analyses that included the final week average of positive emotion but was not significant for the analyses that included the total study average. The results from analyses using the entire Extraversion Scale were highly similar to those presented for the composite of three extraversion facet scores; those findings appear in parentheses in Table 3. Furthermore, the findings are virtually identical even if the directionality of the effects is challenged. Taken together, these analyses provide partial support for the hypothesis that respondents who described themselves as warm, active, and positive remembered more positive emotion than they reported on a momentary basis.

The Relationships Among Momentary Emotion Experiences, Personality Descriptions, and Retrospective Ratings of Emotion.

The recall-based ratings of positive emotions contained largely accurate information, but participants' descriptions of their level of extraversion influenced their recall-based ratings of positive emotion over and above their momentary ratings.

Discussion

SUMMARY OF RESULTS

The findings of this study suggest that recall-based ratings of emotion contain accurate information about momentary emotion experiences. In the present sample, averaged ratings of positive emotion for the final week and for the entire study period strongly predicted the retrospective ratings of positive emotion, whereas both the peak and the total study averages of the negative emotion ratings strongly predicted the retrospective ratings of negative emotion. Thus, the more participants reported emotions like sadness, hostility, guilt, fear, and joviality on a momentary basis, the more they remembered experiencing those emotions. The data analytic strategy employed in the present study does not allow the claim that one of the summary indexes best represents the information that people rely on when making retrospective judgments of emotion because the indexes were not statistically compared with one another. If anything, the findings do suggest that the momentary ratings of emotion were strongly related to the retrospective ratings, regardless of the summary index chosen.

Even when summary indexes of momentary emotional experience were chosen to maximize their correspondence to the retrospective ratings, the correspondence was not perfect. Participants' retrospective ratings of emotion were influenced in the direction of their beliefs about their own emotional lives. Individuals who described themselves as high in neuroticism overestimated, and those who described themselves as low in neuroticism underestimated, the amount of negative emotion that they reported during the course of the study. This finding is consistent with those that have demonstrated neuroticism to bias retrospective physical symptoms reports (Larsen, 1992; Watson & Pennebaker, 1989). Similarly, there was some evidence to suggest respondents who described themselves as extraverted overestimated, and those who described themselves as introverted underestimated, the amount of positive emotion that they reported during the course of the study. We must place less confidence in this latter finding, however, because we did not replicate it across analyses with different summary indexes of momentary emotion ratings.

Although the findings indicate that portions of the retrospective ratings are inaccurate, they do not necessarily indicate that those ratings are invalid. Strongly held beliefs fill in information that is unavailable or inaccessible to individuals (Fiske & Taylor, 1991). This process of filling in missing information will not affect the validity of retrospective judgments if beliefs are related to what occurs in reality. Respondents who described themselves as high in neuroticism may have recalled experiencing more negative emotion than they reported on a moment-to-moment basis because their momentary ratings did not capture all the negative emotional responses that they experienced throughout the day. In this scenario, respondents' beliefs about their own emotionality would not affect the validity of their retrospective ratings. In fact, in this scenario, self-perceptions might have compensated for the information missing in the momentary ratings and thus improved the validity of the retrospective ratings. In contrast, if self-beliefs provided information that is incorrect, then they will have a biasing effect on retrospective judgments. It is only when we regard the momentary ratings as sufficient estimates of everyday emotional experiences that we have grounds to claim that respondents' beliefs about their own emotionality biased the retrospective ratings by having participants remember more emotion than they actually experienced.

In part, then, our ability to determine whether self-descriptions weaken the validity of retrospective ratings depends on our confidence that the momentary emotion ratings are accurate representations of everyday emotional experiences. Participants made only three sets of ratings per day' and did not assess their emotional states continuously throughout each day during the study. Three momentary ratings per day do not completely capture each participant's daily affective experience. Thus, it would be surprising if the recall-based emotion ratings correlated perfectly with the momentary ratings. The lack of perfect agreement between the two may reflect the incomplete sampling of the entirety of each participant's emotional experience. Perhaps the momentary emotion ratings are better considered to be a representative sample of each participant's emotional life. Thus, a more precise interpretation of the present findings is that participants' personality descriptions made an independent contribution to their recall-based ratings of emotion, over and above a representative sample of their daily experience.

The Relationships Among Momentary Emotion Experiences, Personality Descriptions, and Retrospective Ratings of Emotion.

The findings also indicate that people's perceptions of their emotional lives are, in part, derived from their daily experiences. Several summary indexes of momentary ratings were related to self-descriptions of neuroticism or extraversion. The personality ratings were taken before participants began their momentary emotion ratings, and, thus, the relationships can be considered as evidence for the accuracy of participants' self-perceptions about their emotional experiences over time. Interestingly, these effects were obtained both for the facet composites and for the entire domain scores. One could argue, then, that the broad domain scores also communicate people's perceptions of their own emotionality, probably because many of the items across both domains are related to emotion in one form or another. For example, the self-consciousness facet of the Neuroticism Scale includes items such as "At times I have been so ashamed I just want to hide"; the impulsiveness facet includes items such as "I am always able to keep my feelings under control"; and the vulnerability facet includes items such as "I am pretty stable emotionally." Similarly, the gregariousness facet of the Extraversion Scale has items such as "Social gatherings are usually boring for me"; and the excitement-seeking facet has items such as "I often crave excitement." Taken together, these results mirror others that indicate individuals' personality ratings accurately represent the thoughts and feelings that they experience during their daily lives (Barrett & Pietromonaco, in press).

POTENTIAL LIMITATIONS OF THE PRESENT STUDY

The interpretations of the findings are qualified somewhat by the potential limitations of the present study First, the rating procedure used in the present study (three times a day) may not have provided a temporal resolution high enough to best estimate the effect of peak momentary emotions. In a recent study comparing the peak and end effects, participants rated their emotion every 2 hrs across a 2-week period, and a circumscribed role for peak emotion was found (Parkinson et al., 1995); peak emotion made a small but significant contribution to recalled positive emotion ratings after controlling for averaged momentary positive emotion, but the unique effect of peak momentary negative emotion on recalled negative emotion was not significant. Studies finding strong peak and end effects had participants provide continuous ratings of affect while experiencing some circumscribed affective episode (Fredrickson & Kahneman, 1993; Kahneman et al., 1993). Thus, peak effects seem to become more robust as the time interval between momentary assessments decreases.

Second, it was not possible to investigate a true recency effect in the present data set because participants made their retrospective ratings approximately 24 hrs after their last momentary assessment. In studies that show a peak and end effect, participants made retrospective judgments of their overall feelings of pleasure or displeasure either immediately or after a short delay (Fredrickson & Kahneman, 1993; Kahneman et al., 1993). To best assess the magnitude of an end effect, participants would have to make their recall ratings just moments after the experience-sampling procedure ceased.

It is not clear, however, that end effects are even relevant to the design of the present study. The affective episodes being recalled in the Fredrickson and Kahneman (1993) studies were minutes long. In contrast, when people make retrospective judgments of their affect states under "general" or "trait" rating instructions, the emotional episodes to be recalled are measured in days or months, rather than minutes or seconds. It is possible to construct an assessment of end emotion in studies like the present one by aggregating emotion ratings across the final day or the final week of the study. In the present study, combinations of peak and final-day averaged emotion did not predict either of the recall-based ratings better than did combinations of peak and final-week averaged emotion. Furthermore, neither of these combinations predicted the recall-based ratings better than did the peak and first-week aggregates, or the peak and middle-week aggregates. Thus, it remains unclear whether a combination of peak and end states is even appropriate as an indicator of emotion ratings that require a respondent to retrospect over a period of weeks or months.

A third potential limitation is that statistical or methodological artifacts may have, in part, contributed to the differential relationships between averaged momentary emotion, recall-based emotion, and the personality descriptions. The correlation between the averaged momentary and recall-based ratings is likely due, in part, to the fact that both shared the same response scale. Because personality ratings did not make use of the same scale as the recall-based emotion ratings, method variance would differentially inflate the relationship between momentary and recall-based emotion ratings, making it appear stronger.

Finally, the relationship between the momentary and retrospective emotion ratings may be stronger in this study than in situations in which respondents are asked only to make the recall-based ratings or when they recall their emotional experiences spontaneously without prompting from the experimenter (Linton, 1986). Because participants rated their

The Relationships Among Momentary Emotion Experiences, Personality Descriptions, and Retrospective Ratings of Emotion.

emotion experiences for a 90-day period before making their retrospective ratings, their momentary experiences may have been more salient, thus making participants more accurate in their recall-based ratings. Because participants made their personality ratings before they began their momentary emotion ratings, any increases in salience would not affect the relationship between the personality ratings and the recall-based ratings of emotion. Previous research suggests, however, that accuracy does not change appreciably with daily emotion reporting (Thomas & Diener, 1990).

The latter two limitations had a directional effect on the findings: They both act to strengthen the relationship between the momentary and the retrospective ratings. From this perspective, then, the findings of the present study may have produced an upper limit estimate of accuracy in retrospective ratings of emotion. Despite the evidence for accuracy, however, evidence for the inaccuracy hypothesis was also present.

MEMORY-BASED VERSUS ON-LINE STRATEGIES

The present study did not assess the processes through which self-perceptions may influence recall-based emotion ratings. One interesting hypothesis, coming from research on person-perception, is that the retrieval of prior experiences and reconstruction of the past based on beliefs represent two different processing strategies for information about the self. When processing information about a target to make a judgment, individuals may use either "on-line" or "memory-based" processing strategies (Hastie & Park, 1986; McConnell, Sherman, & Hamilton, 1994). If using an on-line strategy, respondents will form an impression about themselves at the time that they experience and encode their affective experiences. This impression is then stored in memory and retrieved when needed. When asked to make a recall-based judgment at a later time, these individuals use a preprocessed impression of their emotionality to make their recall-based judgment, and thus may rely less on trying to reconstruct their previous emotional experiences on the basis of what they can remember in the moment. In contrast, when using a memory-based strategy, respondents would form an impression of their emotionality only when asked. Their impression will be constructed in the moment and will be based on whatever information is available at the time of recall. This information likely includes the most recent emotional experiences but may also include the emotional state at the time of recall (Forgas, 1995).

According to this distinction, the results of the present study might be interpreted as evidence that respondents used both types of processing strategies. Participants may have constructed on-line perceptions of their emotional experiences as they proceeded to make ratings of their experiences across a 90-day period and used these experiences when making the retrospective ratings of their emotional states. Participants might also have relied on beliefs about their own emotionality to construct a memory of their emotional experiences when they made the retrospective ratings. Participants in the present study seemed to use an on-line strategy more than a memory-based strategy, perhaps because the daily rating procedure provided them with a wealth of information about their emotional experience with which to form a strong impression of their own emotionality.

One interesting way to test the effects of on-line and memory-based strategies more systematically would be to vary the length of time between the momentary ratings and the recall-based ratings. With a longer delay between the two types of emotion ratings, respondents might engage in a more memory-based strategy and rely more on their beliefs about their emotionality. In such a scenario, personality descriptions would have a stronger direct effect on the recall-based ratings of emotion than if respondents engaged in a more on-line strategy with a short delay between the two types of ratings. Indeed, Fredrickson and Kahneman (1993) hypothesized that people are likely to reconstruct a judgment from what they can remember about a stimulus when recall is delayed. Note that in the present study, participants made their retrospective emotion ratings only 24 hrs after they made their last momentary emotion rating. This is a relatively short delay considering respondents rated their emotional experiences over a 90-day period; a short delay between momentary and recall-based ratings would constrain the influence that participants' beliefs had on their recall. Once again, the design of the study may have maximized the use of an on-line recall strategy. Despite the fact that an accurate impression of momentary emotional experiences may have been available at the time of recall, however, some variance in the retrospective ratings was left unaccounted for, and this variance was related to participants' personality descriptions.

CONCLUSION

The results of this study suggest two important observations about retrospective measures of emotion experiences. First, recall-based ratings of emotion contain some accurate information about momentary emotion experiences, and this

The Relationships Among Momentary Emotion Experiences, Personality Descriptions, and Retrospective Ratings of Emotion.

seems to be the case regardless of how the momentary ratings of emotion are summarized. Second, respondents' perceptions of their own emotionality often influence their recall-based ratings of emotion, even when assessments of accuracy are maximized.

NOTE

(1.) The spurious effects can be computed as the product of the beta between momentary negative emotion summary scores and neuroticism, and the beta between momentary summary scores and recall-based ratings, controlling for neuroticism. These effects were considered spurious because they are the part of the correlations between neuroticism and recalled emotion that were attributed to the momentary ratings as a common cause.

REFERENCES

Barrett, L. E (in press). Discrete emotions or dimensions: The role of valence focus and arousal focus. Cognition and Emotion.

Barrett, L. F., & Pietromonaco, P. IL (in press). Accuracy of the five factor model in predicting perceptions of daily social interactions. Personality and Social Psychology Bulletin.

Boyle, G.J., & Grant, A. F. (1992). Prospective versus retrospective assessment of menstrual cycle symptoms and moods: Role of attitudes and beliefs. Journal of Psychopathology and Behavioral Assessment, 14, 307-321.

Costa, P. T., & McCrae, R. R. (1992). The Revised NEO Personality Inventory (NEO PI-R) professional manual Odessa, FL: Psychological Assessment Resources.

Feldman, L. A. (1995). Valence focus and arousal focus: Individual differences in the structure of affective experience. Journal of Personality and Social Psychology, 69, 153-166.

Fiske, S. T., & Taylor, S. E. (1991). Social cognition. Reading, MA: Addison-Wesley.

Forgas, J. P. (1995). The affect infusion model (AIM): Review and an integrative theory of emotion effects on judgments. Psychological Bulletin, 117, 39-66.

Fredrickson, B. L., & Kahneman, D. (1993). Duration neglect in retrospective evaluations of affective episodes. Journal of Personality and Social Psychology, 65, 45-55.

Greenwald, A. G., & Banaji, M. 1L (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. Psychological Review, 102, 4-27.

Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. Psychological Review, 93, 258-268.

Hedges, S. M., Jandorf, L., & Stone, A. A. (1985). Meaning of daily emotion measurements. Journal of Personality and Social Psychology, 48, 428-434.

Holmberg, D., & Holmes, J. G. (1994). Reconstruction of relationship memories: A mental models approach. In N. Schwarz & S. Sudman (Eds.), Autobiographical memory and the validity of retrospective reports (pp. 267-288). New York: Springer-Verlag.

Izard, C. E., Libero, D. Z., Putnam, P., & Hayes, O. M. (1993). Stability of emotional experiences and their relations to traits of personality. Journal of Personality and Social Psychology, 64, 847-860.

Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better end. Psychological Scieno, 6, 401-405.

The Relationships Among Momentary Emotion Experiences, Personality Descriptions, and Retrospective Ratings of Emotion.

Larsen, R. J. (1992). Neuroticism and selective encoding and recall of symptoms: Evidence from a combined concurrent-retrospective study. Journal of Personality and Social Psychology, 62, 480-488.

Larsen, R. J., & Diener, E. (1992). Promises and problems with the circumplex model of emotion. In M. Clark (Ed.), Review of personality and social psychology, 13, 25-59.

Larsen, R. J., & Ketellar, T. (1991). Personality and susceptibility to positive and negative emotional states. Journal of Personality and Social Psychology, 61, 132-141.

Linton, M. (1986). Ways of searching and the contents of memory. In D. C. Rubin (Ed.), Autobiographical memory (pp. 504-57). Cambridge, UK: Cambridge University Press.

MacFarlane, J., Martin, C. L., & Williams, T. M. (1988). Mood fluctuations: Women versus men and menstrual versus other cycles. Psychology of Women Quarterly, 12, 201-223.

McConnell, A. R., Sherman, S.J., & Hamilton, D. L. (1994). On-line and memory-based aspects of individual and group target judgments. Journal of Personality and Social Psychology, 67, 173-185.

McCrae, R. IL, & Costa, P. T. (1991). Adding Liebe und Arbeit: The full five-factor model and well-being. Personality and Social Psychology Bulletin, 17, 227-232.

McFarland, C., Ross, M., & DeCourville, N. (1989). Women's theories of menstruation and biases in recall of menstrual symptoms. Journal of Personality and Social Psychology, 5 7, 522-531.

Meng, X., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. Psychological Bulletin, 111, 172-175.

Nisbett, R., & Ross, L. (1980). Human inference: Strategies and shortcomings of social judgment. Englewood Cliffs, NJ: Prentice Hall.

Parkinson, B., Briner, R. B., Reynolds, S., & Totterdell, P. (1995). Time frames for emotion: Relations between momentary and generalized ratings of affect. Personality and Social Psychology Bulletin, 21, 331-339.

Pedhazer, E.J. (1982). Multiple regression in behavioral research (2nd ed.). New York: Holt, Rinehart, and Winston.

Reis, H. T., & Wheeler, L. (1991). Studying social interaction with the Rochester Interaction Record. In L. Berkowitz, (Ed.), Advances in experimental social psychology (Vol. 24, pp. 269-318). San Diego, CA: Academic Press.

Ross, M. (1989). Relation of implicit theories to the construction of personal histories. Psychological Review, 96, 341-357.

Schwarz, N., & Sudman, S. (1994.), Autobiographical memory and the validity of retrospective reports. New York: Springer-Verlag.

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

Tellegen, A. (1985). Structures of emotion and personality and their relevance to assessing anxiety, with an emphasis on serf-report. In A. H. Tuma & J. D. Maswer (Eds.), Anxiety and the anxiety disorders (pp. 681-706). Hillsdale, NJ: Lawrence Erlbaum.

Thomas, D. L., & Diener, E. (1990). Memory accuracy in the recall of emotions. Journal of Personality and Social Psychology, 59, 291-297.

Watson, D., & Clark, L. A. (1992). On traits and temperament: General and specific factors of emotional experience and

The Relationships Among Momentary Emotion Experiences, Personality Descriptions, and Retrospective Ratings of Emotion.

their relation to the five-factor model. Journal of Personality, 60, 441-475.

Watson, D., & Clark, L.A. (1994). Preliminary manual for the Positive Affect and Negative Affect Schedule (expanded form). Unpublished manuscript, Southern Methodist University, Dallas, Texas.

Watson, D., & Clark, L. A. (in press). Extraversion and its positive emotional core. In S. Briggs, R. Hogan, &W. Jones (Eds.), Handbook of personality psychology. San Diego, CA: Academic.

Watson, D., & Pennebaker, J. W. (1989). Health complaints, stress, & distress: Exploring the central role of negative affectivity. Psychological Review, 96, 234-254.

Received August 31, 1995

Revision accepted November 22, 1996

Author's Note: I thank Michael Ross and Daniel Barrett for their comments on earlier versions of this article. Address correspondence to Lisa Feldman Barrett, Department of Psychology, Boston College, Chestnut Hill, MA 02167, e-mail barretli@bc.edu.

Liza Feldman Barrett Boston College