Journal of Personality and Social Psychology
2002, Vol. 83, No. 1, 198–215

Copyright 2002 by the American Psychological Association, Inc.
0022-3514/02/$5.00 DOI: 10.1037//0022-3514.83.1.198

Episodic and Semantic Knowledge in Emotional Self-Report: Evidence for Two Judgment Processes

Michael D. Robinson
North Dakota State University

Gerald L. Clore
University of Virginia

Three studies involving 3 participant samples (*N*s = 39, 55, and 53) tested the hypothesis that people retrieve episodic emotion knowledge when reporting on their emotions over short (e.g., last few hours) time frames, but that they retrieve semantic emotion knowledge when reporting on their emotions over long (e.g., last few months) time frames. Support for 2 distinct judgment strategies was based on judgment latencies (Studies 1 and 2) and priming paradigms (Studies 2 and 3). The authors suggest that self-reports of emotion over short versus long time frames assess qualitatively different sources of self-knowledge.

People are commonly asked by psychologists to report on emotions that they are not currently feeling. Retrospective reporting formats ask people to characterize their emotions in the past, aggregated reporting formats ask them to characterize their emotions over a specific time interval, and generalized reporting formats ask them to characterize their emotions in general. Regardless of the time frame used, participants seem to have little trouble giving reliable and (to a certain extent) valid emotion ratings. On the basis of such considerations, it is easy to assume that participants are drawing on similar knowledge regardless of the specified time frame.

Emotional experiences, however, fluctuate considerably over time as well as from situation to situation (Brandstätter, 1983; Csikszentmihalyi & LeFevre, 1989; Diener & Larsen, 1984; Fredrickson & Kahneman, 1993). Because of this, it is reasonable to question people's capacity to (a) remember and (b) integrate the subtle nuances of their experiences when they are reporting on emotions that occurred in the past (Kahneman, 1999; Loewenstein, 1996; Stone, Shiffman, & DeVries, 1999). Indeed, at least four types of designs have yielded evidence for retrospective biases in emotion reporting.[1]

Retrospective Biases

In the first type of design, on-line and retrospective reports cover the exact same time period, but there is a delay between the two sets of ratings. This type of design has been particularly useful in showing that erroneous beliefs about particular situations appear to bias retrospective emotion ratings relative to on-line ones (Arntz, van Eck, & Heijmans, 1990; Levine, 1997; McFarland, Ross, & DeCourville, 1989). For example, McFarland et al. (1989) compared emotion reports on a day of menstruation with retrospective reports concerning that same day. As expected, retrospective emotion reports were biased in a negative direction (relative to on-line reports), consistent with the belief that menstruation causes negative affect.

In the second type of design, multiple on-line ratings are averaged and compared with a single set of aggregated ratings covering the same time period. In one such study, Thomas and Diener (1990) obtained daily emotion reports for 6 weeks and then asked people to characterize their emotions over this entire 6-week period. They found that, compared with an average of on-line ratings, people retrospectively underestimated the frequency of their positive affect and overestimated the intensity of both their positive and their negative daily experiences. The tendency for people to retrospectively overestimate emotional intensity appears to be robust (Hedges, Jandorf, & Stone, 1985; Kahneman, 1999; Parkinson, Briner, Reynolds, & Totterdell, 1995), although different accounts of this phenomenon have been offered (Kahneman, 1999; Thomas & Diener, 1990; Winkielman, Knäuper, & Schwarz, 1998).

In the third type of design, multiple on-line reports are averaged and compared with trait reports. The most consistent focus of such

Michael D. Robinson, Department of Psychology, North Dakota State University; Gerald L. Clore, Department of Psychology, University of Virginia.

This research was supported by National Science Foundation Grant 9817649 awarded to Michael D. Robinson and Gerald L. Clore. We thank Joel Johnson for help in drafting the manipulation used in Study 3.

Correspondence concerning this article should be addressed to Michael D. Robinson, Department of Psychology, North Dakota State University, Fargo, North Dakota 58105-5075.

[1] On-line reports concern current emotional experiences, whereas retrospective reports concern past emotional experiences. The distinction is not absolute but relative, as many investigators treat daily reports of emotion (e.g., "How have you felt today?") as on-line ones. We also follow this convention in reviewing the literature.

Additionally, it is useful to define aggregated reports relative to retrospective ones. The necessary feature of a retrospective report is that people are being asked to report on their emotions in the past. The necessary feature of an aggregated report is that people are being asked to report their emotions over a period of time (e.g., "How have you felt over the last few weeks?"). All aggregated reports are retrospective because the relevant time frame extends into the past. However, not all retrospective reports are aggregated (e.g., "How did you feel last Monday at 12?").

research has been on sex differences in emotion. Three literature reviews (Eisenberg & Lennon, 1983; LaFrance & Banaji, 1992; Shields, 1991) have concluded that on-line ratings of emotionality tend to show few consistent sex differences, whereas trait ratings of emotionality are often more consistent with sex stereotypes. In a recent study, for example, Feldman Barrett, Robin, Pietromonaco, and Eyssell (1998) asked participants to complete trait scales of emotionality as well as report on their emotions following daily social interactions. In this study, women scored higher on trait scales measuring affective intensity, openness to feelings, anxiety, sadness, positive emotionality, and interpersonal warmth. However, sex differences following daily social interactions were minimal (for similar findings in the area of emotional coping, see Porter et al., 2000).

The fourth type of design is a combination of the previous two. Specifically, on-line reports are averaged over time and compared with retrospective and trait ratings. For example, Feldman Barrett (1997) asked participants to characterize their emotions three times a day for 90 days. After this period, participants were asked to characterize their emotions during this entire 3-month period and then complete trait scales of neuroticism and extraversion. Feldman Barrett found that, when online reports were controlled, neurotics overestimated their negative affect in retrospection. A similar, though less pronounced, tendency was found for extraverts to overestimate their positive affect in retrospection. Similar results have been found in other studies (Cutler, Larsen, & Bunce, 1996; Diener, Larsen, & Emmons, 1984; Schimmack, 1996). In related research, several studies have shown that neurotics overestimate their experience of physiological symptoms in retrospection (K. W. Brown & Moskowitz, 1997; Larsen, 1992; Watson & Pennebaker, 1989).

There have been several theoretical attempts to account for the types of retrospective biases reported above. Three theories have been proposed for the normative tendency to retrospectively exaggerate the intensity of one's emotions (Kahneman, 1999; Thomas & Diener, 1990; Winkielman et al., 1998). Although these theories can account for retrospective exaggeration (Design 2), they are relatively mute with respect to the belief-consistent biases obtained with Designs 1, 3, and 4. By contrast, Ross (1989; see also Levine, 1997) proposed a theory that emphasizes the role of belief-based reconstruction in retrospective emotion reports. To date, Ross's framework has been most successful in accounting for the type of retrospective biases obtained with Design 1.

Of particular importance, it appears that a comprehensive theory of retrospective biases would need to account for both (a) biases due to the recall of specific moments of an experience (e.g., recent and intense moments; Kahneman, 1999) and (b) biases due to belief-based reconstruction (e.g., erroneous beliefs about stability; McFarland & Ross, 1987). Because no extant theory covers both of these sources of bias, Robinson and Clore (in press) recently proposed a new, more comprehensive theory of emotion reporting. The theory can explain both (a) the memory-based distortions that frequently result from Design 2 and (b) the belief-consistent biases that frequently result from Designs 1, 3, and 4. The theory accounts for both phenomena by proposing that not all emotional self-reports are made in the same manner. Whereas some self-reports (e.g., on-line reports) appear to be made on the basis of episodic emotion knowledge, others (e.g., trait reports) appear to be made on the basis of semantic emotion knowledge.

Episodic and Semantic Emotion Knowledge

Robinson and Clore (in press) suggested that, in understanding the factors that contribute to any type of emotional self-report, it is important to distinguish *episodic* and *semantic* emotion knowledge. Episodic knowledge is experiential in nature and inextricably bound with details of time and place (Tulving, 1984). Semantic knowledge, by contrast, is conceptual in nature and decontextualized—that is, divorced from the details of time and place (Tulving, 1984). Episodic knowledge, because it is event specific and only loosely organized, is subject to a great deal of interference and forgetting. Semantic knowledge, because it is event independent and very tightly organized, is thought to be relatively immune to interference and forgetting (Tulving, 1993).

Episodic emotion knowledge is knowledge about one's emotions in a particular place at a particular time. The prototypical use of episodic emotion knowledge likely occurs when one reports on momentary emotional states. Within a moment, emotions are (a) experienced (Loewenstein, 1996), (b) heavily dependent on time and place (Diener & Larsen, 1984), and (c) subject to the ebb and flow of conscious awareness (Csikszentmihalyi, 1990). These are all characteristics of episodic knowledge (Tulving, 1984). Semantic emotion knowledge, by contrast, consists of beliefs that one has concerning one's emotions. The prototypical use of semantic emotion knowledge likely occurs when one reports on emotional traits. In support of this claim, the ability to characterize oneself "in general" does not appear to depend on the ability to retrieve particular events (Klein, Babey, & Sherman, 1997; Klein, Loftus, & Kihlstrom, 1996; Klein, Loftus, Trafton, & Fuhrman, 1992). This suggests that people have formed beliefs about themselves that become dissociated from their behavior and/or experiences in everyday life (see also Marsh & Yeung, 1998).

In accounting for retrospective biases, it is important to note that one's ability to retrieve episodic information declines quickly with the passage of time (Tulving, 1984). Therefore, any delay between a particular emotional episode and its reporting will result in the loss of certain details of the episode (Rubin & Wetzel, 1996). When the delay is short enough, participants are likely to still use an episodic retrieval strategy, but their ratings are likely to be biased by peak and recency effects (Kahneman, 1999). However, when the delay is particularly long, episodic details will become sufficiently inaccessible that participants will shift to a semantic retrieval strategy (Robinson & Clore, in press). That is, they will access their beliefs about their emotions rather than episodic emotion knowledge.

Several key assumptions of the Robinson and Clore (in press) model are consistent with prior cognitive research. First, there is considerable evidence for the independence of episodic and semantic knowledge in memory (Begg & Nicholson, 1994; Dosher, 1984; Herrmann & Harwood, 1980; Neely & Durgunoglu, 1985; Shoben, Wescourt, & Smith, 1978; for reviews, see Tulving, 1993, and Wheeler, Stuss, & Tulving, 1997). In the area of autobiographical memory, particularly compelling evidence for this independence comes from studies of amnesiacs, who can recall abstract facts about their lives (e.g., "I got my undergraduate degree at the University of California, Santa Cruz") despite their inability to recall any specific events from that time period (see Conway & Pleydell-Pearce, 2000, for a review). Second, it is clear that episodic knowledge declines quickly with the passage of time (Rubin

& Wetzel, 1996). In the area of autobiographical memory, Burt, Watt, Mitchell, and Conway (1998) have estimated that event-specific knowledge becomes inaccessible within a week or 2 unless rehearsed. And third, there is also evidence for the key assumption that semantic intrusions become more common with the passage of time. Particularly useful in this regard are studies by Dooling and colleagues (Dooling & Christiaansen, 1977; Sulin & Dooling, 1974). For example, Sulin and Dooling presented participants with a text about a dictator named Gerald Martin (control condition) or Adolph Hitler (experimental condition). Memory for the text was assessed either 5 min later or after a 1-week delay. In the short delay condition, there were few intrusions related to general knowledge about Hitler. However, after a week, such false intrusions increased dramatically. Such results are highly compatible with those of Robinson, Johnson, and Shields (1998), who asked participants to report on their emotions immediately following a competitive game or 1 week later. As expected, emotion ratings were not consistent with sex role stereotypes in the on-line condition but were consistent with these stereotypes in the delayed condition.

It is important to point out one final aspect of our model. We do not propose that retrospective ratings are more consistent with semantic emotion knowledge because semantic knowledge biases the retrieval of episodic details. Such a proposal is consistent with Bartlett's (1932) schema model, a model that we know to be incorrect (Alba & Hasher, 1983; Stangor & McMillan, 1992). Instead, we propose that semantic emotion knowledge fills in when episodic emotion knowledge becomes inaccessible. In this respect, our model is more consistent with script (Schank & Abelson, 1977) or frame (Minsky, 1975) theories, which emphasize this fill-in function, than with schema (Bartlett, 1932) theories, which emphasize biases in episodic retrieval. In the context of text comprehension, Dooling and Christiaansen (1977) stated our view succinctly: "With the passage of time, subjects have increased difficulty in retrieving passage-specific episodes. They compensate by using related information from semantic memory (p. 428)."

Testing the Robinson and Clore (in press) Model

Although dissociations between on-line and retrospective ratings are informative concerning the Robinson and Clore (in press) model, ratings by themselves are relatively mute on questions of process. Any given rating may reflect episodic emotion knowledge, semantic emotion knowledge, or both, with no obvious signs of how to differentiate these sources of accessed knowledge. As a result, the qualitative shift from the use of episodic to the use of semantic emotion knowledge proposed by Robinson and Clore (in press) is impossible to discern solely on the basis of ratings. For these reasons, we turned to judgment times and priming paradigms, both of which should be more sensitive to dissociations in knowledge use.

As indicated above, previous investigators have used at least four retrospective frames in their designs. The first type of design that we reviewed uses retrospective frames that are relatively narrow and discontinuous with the present (e.g., participants are asked to rate the emotions they had on a day 2 weeks earlier; McFarland et al., 1989). In Designs 2 and 4, by contrast, the retrospective frames are aggregated in nature (e.g., participants are asked to rate their emotions over the last 3 months; Feldman Barrett, 1997). Finally, in Design 3, the retrospective frames concern emotions in general.

In this initial test of our model, we decided to use aggregated (Designs 2 and 4) and trait (Design 3) frames but not narrow frames from the past (Design 1). We asked participants to rate their emotions over seven distinct frames—at this moment, in the last few hours, in the last few days, in the last few weeks, in the last few months, in the last few years, and in general—that differed both in their width and in their proximity to the present. There were two primary reasons for using aggregated rather than narrow retrospective frames. First, such frames are quite common in the personality and emotion literature, and we wanted to make contact with this literature (see Watson, 2000, for a review). Second, we thought participants would have an extremely difficult time rating their emotions on a nondistinct day in the distant past (e.g., "Rate your emotions on a day exactly 2 years ago"). Indeed, they might find such a task ridiculous. By contrast, we expected them to find aggregated frames meaningful and reasonable. By including seven time frames in the current investigation, we expected to find dramatic evidence for two distinct retrieval strategies for narrow (episodic) versus wide (semantic) time frames. The multiple time frames should allow us to more precisely document this qualitative shift in retrieval strategies.

In Study 1, we drew on the literature on event frequency estimation. In this literature, people are asked to respond to survey questions by estimating such things as how often they go to movies, make phone calls, or visit the bank. A clear conclusion from this literature is that not all answers are made on the same basis: People sometimes use episodic strategies (i.e., recall instances and count them) and sometimes use nonepisodic strategies (e.g., estimate rate of occurrence and multiply; Menon, 1993). Investigators have been interested in determining the conditions under which people use episodic versus nonepisodic estimation strategies. Particularly useful in the present context are studies by Blair and Burton (1987) and N. R. Brown (1995). Blair and Burton (1987) obtained verbal protocols while participants were answering the questions. Strategies were coded from the protocols. When there were fewer than five relevant instances, a majority of the participants (86%) appeared to use a retrieve-and-count strategy. However, when there were more than nine relevant instances, no participant (0%) appeared to use this strategy. Blair and Burton also examined the influence of time frame width. A majority of the participants (56%) used a retrieve-and-count strategy for a 2-week frame, compared with 25% for a 2-month frame and 4% for a 6-month frame. On the basis of these findings, we might expect our participants to use an episodic retrieval strategy for time frames shorter than the last few weeks and a semantic retrieval strategy for time frames longer than this.[2]

N. R. Brown (1995) used a laboratory paradigm to gain a more precise understanding of frequency estimation. Although his hypotheses are not directly relevant to the present study, his method

[2] Our episodic retrieval strategy is obviously similar to the retrieve-and-count strategy investigated in this literature. However, there is really no semantic retrieval strategy in this literature that is comparable to ours. Probably this is because people do not have well-established beliefs about how frequently they engage in mundane behaviors (e.g., eat peanut butter). Emotions, however, are presumably more central to the self-concept.

is. Brown presented participants with category names (e.g., *city*) either 2, 4, 8, 12, or 16 times. His predictions concerned the use of episodic versus nonepisodic estimation strategies as a function of a between-subjects manipulation. In addition to relying on verbal protocols to assess his hypotheses, Brown also measured the length of time that it took participants to make their estimations. He reasoned that if participants attempted to retrieve and count individual instances, then reaction time should increase linearly with presentation frequency (i.e., it would take more time to retrieve and count events that were more frequent). By contrast, Brown reasoned that if participants instead used a nonepisodic estimation strategy, there would be no increase in reaction time with presentation frequency. Brown's results supported this key prediction. Specifically, in the condition in which he expected an episodic estimation strategy, there was a linear rise in judgment time with presentation frequency; in the other condition, there was not.[3]

In Study 1, we used N. R. Brown's (1995) reasoning to make predictions concerning emotion judgment latencies. Because we expected participants to rely on episodic emotion knowledge for short time frames, we expected to see a linear rise in judgment latencies from the time frame "at this moment" to the time frame "last few hours" to the time frame "last few days." This is because a longer time frame necessitates that more instances of emotion be retrieved and aggregated. By contrast, we expected participants to rely on semantic emotion knowledge for long time frames. This would be seen in a flat latency curve for time frames longer than the last few weeks. Together, these distinct patterns for short versus long time frames should create a curvilinear pattern of latencies that is highly consistent with the Robinson and Clore (in press) model.

In Study 2, we built on research within the semantic priming literature (see Neely, 1991, for a review). A robust finding in this literature is that participants are faster to categorize or pronounce words if the words are preceded by semantically related (vs. unrelated) words. For example, a lexical decision regarding the word *doctor* is faster if the preceding trial involved the word *nurse* (related) versus an unrelated word (e.g., *couch*). Such a pattern of reaction time facilitation is evidence that the same source of knowledge (i.e., knowledge concerning medical professionals) had been used in making the consecutive judgments.

In the present context, we examined emotion judgment latencies as a function of the width of prime and target frames. If our model is correct in proposing that people access semantic emotion knowledge when reporting on their emotions over long time frames, then judgments for long time frames should be faster if the preceding trial involved another long (vs. short) time frame judgment. This would suggest that the same source of knowledge (i.e., semantic emotion knowledge) was used in making the consecutive judgments. This prediction was tested in Study 2.

In Study 3, we took a different approach to priming. According to our model, people use semantic emotion knowledge to report on their emotions over long but not short time frames. If this is true, it should be possible to influence emotion ratings by covertly priming a source of emotion-related beliefs. Furthermore, this priming manipulation should have dramatically different effects on ratings for short versus long time frames.

To manipulate a source of emotion-related beliefs, we built on the literature concerning sex and emotion. Women, in comparison with men, are widely regarded as the more emotional sex (Shields, 1987; Widiger & Settle, 1987; Williams & Bennett, 1975). Furthermore, such beliefs are incorporated into the self-concepts of women and men to some extent (Cross & Madson, 1997; Eagly & Wood, 1991; Josephs, Markus, & Tafarodi, 1992; Spence, Helmreich, & Stapp, 1975). If so, and if emotion reports for longer time frames do draw on such beliefs, then a manipulation of sex-linked beliefs should lead to more stereotypic ratings for long time frames. However, the same manipulation should produce a different pattern of results for short time frames. In sum, we expected a manipulation of the accessibility of sex role stereotypes to have distinctly different effects on emotion ratings for short versus long time frames.

Altogether, the current studies constitute efforts to examine how people make emotion judgments as directly as possible. Central predictions were that (a) judgment latencies would increase linearly with time frame width for short but not long time frames (Study 1), (b) emotion judgments for long time frames would be faster when the preceding trial also involved a long (vs. short) time frame (Study 2), and (c) a manipulation of beliefs concerning sex and emotion would produce an assimilation effect on ratings of emotion for long but not short time frames (Study 3).

A Methodological Note

The present studies are organized on the basis of major findings rather than on the basis of participant sample. We feel that this manner of presentation produces the clearest and most parsimonious exposition. The data reported in the article come from three independent samples, but results from all three samples are reported together in Studies 1 and 2. Study 3, however, is based exclusively on additional results from Sample 3. The effects reported in the three studies are statistically orthogonal.

Study 1

In Study 1, we tested several rather basic assumptions of our model that have not been previously examined. Three independent samples were asked to report on their emotions over different time frames. We systematically varied the width of these time frames from at this moment to the last few months to in general. Our particular interest was in the length of time required to judge one's emotions over the different frames. To measure judgment time, we presented all of the items by computer.

If episodic emotion knowledge contributes to self-reports of emotion, we would expect to find evidence for a linear trend in judgment latencies across the different time frames (N. R. Brown, 1995). That is, on-line judgments of emotion, because they concern a particular moment in time, should be relatively fast when compared with judgments of emotion over the period of several hours or several days. As the time frames become progressively longer, more and more instances of felt emotion become relevant, and, thus, it should take more and more time to retrieve and

[3] Researchers in the frequency estimation literature have also been interested in the question of whether episodic or nonepisodic estimation strategies are more accurate. Relevant results are inconclusive. Whereas some have found estimation strategies to be more accurate (N. R. Brown, 1995; Burton & Blair, 1991), others have found the opposite pattern of results (Menon, Raghubir, & Schwarz, 1995).

aggregate the relevant episodic knowledge. It is often assumed that people do use an episodic retrieval strategy in reporting on their emotions over longer time frames, but no one has previously shown that judgment latencies exhibit the predicted linear increase with time frame width.

More important, however, we expected to find evidence for the premise that people abandon an episodic retrieval strategy for time frames that are relatively long or, in the case of trait reports, unspecified. If this is the case, we would expect to see a curvilinear trend in judgment latencies across the different time frames. Such a pattern of latencies would offer simple but, we hope, convincing evidence that people abandon an episodic retrieval strategy when the time frame is too general or too long to support effective episodic retrieval.

Method

Participants. Participants were undergraduates from the University of Illinois who reported on their emotions in return for credit for their introductory psychology classes. Study 1 includes three independent samples (Ns = 39, 55, and 53) from this population.[4]

Manipulation, procedures, and dependent measures. Participants reported on the extent to which they had experienced eight (Samples 1 and 3) or nine (Sample 2) emotions over seven distinct time frames (at this moment; last few hours; last few days; last few weeks; last few months; last few years; in general). Ratings were made on a 5-point intensity scale (1 = *none*; 2 = *a small amount*; 3 = *a moderate amount*; 4 = *a large amount*; 5 = *an extreme amount*). Each emotion was crossed with each time frame, such that participants in Samples 1 and 3 made 56 judgments (8 emotions × 7 time frames), whereas participants in Sample 2 made 63 judgments (9 × 7). Trials were presented on a computer and randomized for each participant. Participants were instructed to make their judgments quickly but accurately.[5]

We were particularly interested in the length of time required to judge emotions over the different time frames. To unconfound ratings and judgment latencies (see Fazio, 1990), we used the following procedures. A given trial began with a time frame (e.g., last few weeks). Participants were given 2 s to read and comprehend the specified time frame, at which point the specific emotion appeared (e.g., *anger*). Once the emotion word appeared, we began timing the judgment latency. Participants were instructed to determine the extent to which they had experienced the given emotion over the given time frame, at which point they were to press the space bar when they were ready to give their answer. Time to hit the space bar therefore represented the time necessary to make the judgment.

To ensure that participants had actually determined their answer prior to hitting the space bar, we removed all information about the trial (i.e., time frame and emotion) once the space bar was pressed. Analyses revealed that these procedures were successful, as rating time—that is, time to make a rating after pressing the space bar—did not vary by emotion, time frame, or their interaction in any of the samples (all ps $> .15$). In addition, these ratings times were quite short (.5–.7 s), consistent with the amount of time that would be required to find and hit the intended rating key (1–5). We therefore felt confident that we could treat each space bar press as the time required to make a particular judgment.

Across samples and trials, the mean correlation between latencies and ratings was only .038. Thus, under the procedures, latencies and ratings were functionally independent.

Results

Time frame effects on judgment latencies. Prior to analyzing space bar times, we log transformed them to normalize the distributions. Additionally, because it is reasonable to think that people judge their positive and negative emotions differently, we collapsed across individual emotion terms to create positive versus negative emotion scales. The resulting latency means were analyzed, separately for each sample, in a 2 (emotion scale: positive vs. negative) × 7 (time frame) analysis of variance (ANOVA). Although participants in Sample 3 judged their negative emotions more quickly than their positive ones, $F(1, 52) = 9.47$, $p = .003$, this main effect for emotion scale was not significant in the other two samples (Fs < 1). Of more importance, the main effect for time frame was significant in each of the three samples, $F(6, 234) = 4.19$, $p = .001$ in Sample 1, $F(6, 324) = 2.88$, $p = .010$ in Sample 2, and $F(6, 312) = 8.08$, $p = .000$ in Sample 3. As expected, participants took longer to judge their emotions in certain time frames than in others.

In the same analyses, there were no Emotion Scale × Time Frame interactions (ps $> .05$). That is, the effects of time frame were the same regardless of emotion valence. Time frame main effects are graphically displayed in Figure 1. For the sake of clarity, means are reported in terms of seconds rather than as log-transformed values.

Linear and curvilinear effects on judgment latencies. From Figure 1, it appears that participants displayed two distinct tendencies in reporting on their emotions across time frames. First, consistent with an episodic retrieval strategy, there appears to have been a linear increase in judgment latencies as the time frames became increasingly longer. Such a pattern is consistent with the adoption of an episodic retrieval strategy, as longer time frames necessitate that more instances of felt emotion be retrieved and aggregated. Second, however, it appears that this linear rise in judgment latency is true only for relatively recent time frames. Beyond a certain time frame (months in Sample 1, days in Sample 2, and weeks in Sample 3), judgment latencies both dropped and flattened out. This second pattern is inconsistent with an episodic retrieval strategy and instead suggests the adoption of a different, arguably semantic, retrieval strategy. Because semantic emotion knowledge (e.g., about one's emotions in general) does not vary with the length of the time frame, use of it would not produce the linear change in judgment latency that we observed for the shorter time frames.

If our account is correct, we would expect to see not only a linear relation between time frame and judgment latencies, reflecting the use of an episodic retrieval strategy, but also a curvilinear relation, reflecting the abandonment of this strategy as time frames became increasingly wide. To determine whether both of these components were significant, we used the following procedures. First, for each participant separately, we performed a multiple regression predicting judgment latencies. With emotion ($n = 8$ or

[4] We did not collect participant sex data for Samples 1 and 2. In Sample 3, however, participant sex was a variable of key interest (see Study 3).

[5] Participants in Sample 1 judged *happiness*, *joy*, *affection*, and *pride* (positive emotions) as well as *sadness*, *anxiety*, *distress*, and *anger* (negative emotions). Participants in Sample 2 judged *enthusiasm*, *happiness*, *confidence*, *affection*, and *empathy* (positive) as well as *distress*, *insecurity*, *worry*, and *anger* (negative). Participants in Sample 3 judged *calmness*, *excitement*, *caring*, and *pride* (positive) as well as *anxiety*, *insecurity*, *sadness*, and *anger* (negative).

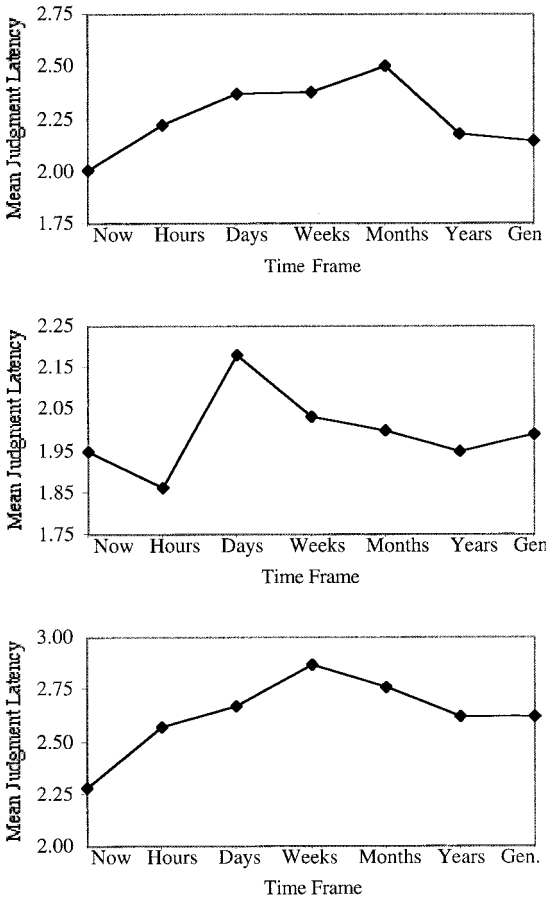

Figure 1. Effects of time frame on judgment latencies, Study 1; Samples 1 (top), 2 (middle), and 3 (bottom). Gen. = general.

9) as the unit of analysis, we entered a linear time frame predictor (coded from 1 = *at this moment* to 7 = *in general*) as well as a curvilinear time frame predictor (time frame squared). The linear and curvilinear predictors were simultaneously controlled. Second, participant-level betas from the regressions were entered into a new data set and tested for significance against the null hypothesis that the true population beta was equal to zero (see Lorch & Myers, 1990, for details concerning this statistical procedure).

There was a positive relation between the width of the time frame and judgment latencies (mean βs = .51, .25, and .56 in Samples 1, 2, and 3, respectively). In each sample, the relevant coefficient was significantly different than zero (all ps $< .05$). There was also a curvilinear relation between the width of the time frame and judgment latencies (mean βs = $-.50$, $-.21$, and $-.50$ in Samples 1, 2, and 3, respectively). In each sample, the relevant coefficient was significantly different than zero (all ps $< .05$). Thus, as the time frames progressively widened, there was evidence for both a linear rise in judgment latency, consistent with an episodic retrieval strategy, and a curvilinear pattern, consistent with the abandonment of an episodic retrieval strategy for relatively long time frames.

To more precisely understand the linear and curvilinear trends, we performed two further sets of regressions. In one set of regressions, we obtained linear within-subject coefficients for the short time frames (1 = moment; 2 = hours; 3 = days). In the other set, we performed parallel regressions for the long time frames (5 = months; 6 = years; 7 = in general). We then conducted three ANOVAs, one for each sample, to confirm that linear trends were distinctly different for short versus long time frames. They were (ps $< .001$). For the short time frames, latencies increased with time frame width (mean $\beta = .11$ across participants and samples; ps $< .05$). For the long time frames, on the other hand, latencies exhibited no systematic pattern in two of the three samples (ps $> .05$) and significantly decreased in one ($p < .05$; mean $\beta = -.06$ across participants and samples).

In summary, time frame width had distinctly different effects on judgment latencies for short versus long time frames. For short time frames, wider time frames took longer to judge, consistent with the adoption of an episodic retrieval strategy. For long time frames, by contrast, judgments were equally fast regardless of the time frame under consideration. This flattening is consistent with our proposal that participants are retrieving the same information regardless of the width of the time frame. As we show next, parallel effects emerged with rating means.

Time frame effects on rating means. Our account of two retrieval strategies possesses implications for rating means. If a person sampled semantic knowledge, one would expect similar answers regardless of the time frame under consideration. This is because semantic knowledge is general and decontextualized in nature. By contrast, if a person sampled episodic knowledge, one would expect different answers depending on the time frame under consideration. This pattern would reflect the fact that episodic knowledge is contextual and should vary by the particular frame in question.

Although the rating scale that we used (1 = *none*; 5 = *an extreme amount*) was primarily designed to measure intensity rather than frequency, consider what would happen if frequency information also influenced ratings. If an episodic retrieval strategy was used, a longer time frame should produce higher rating means because people would be able to retrieve more instances on which they experienced the particular emotion. If a semantic retrieval strategy was used, by contrast, a longer time frame should not produce higher rating means because such a strategy does not involve an aggregation of instances. We therefore predicted that longer time frames would lead to higher ratings, but only if participants were using an episodic retrieval strategy.

To assess whether time frame influenced rating means, we performed three 2 (emotion scale) $\times$ 7 (time frame) ANOVAs, one for each sample. In all samples, participants gave higher intensity ratings to their positive emotions than to their negative ones (ps $< .001$). Of more importance, time frame had a main effect in each of the samples, $F(6, 234) = 54.18$, $p = .000$ in Sample 1, $F(6, 324) = 78.76$, $p = .000$ in Sample 2, and $F(6, 312) = 105.74$, $p = .000$ in Sample 3. As anticipated, people reported more intense emotions for longer time frames than for shorter ones.

In all three samples there were also significant Emotion Scale $\times$ Time Frame interactions, $F(6, 234) = 8.24$, $p = .000$ in Sample 1, $F(6, 324) = 6.23$, $p = .000$ in Sample 2, and $F(6, 312) = 4.37$, $p = .000$ in Sample 3. The nature of the interaction was parallel in each sample: The extent to which positive emotions were rated higher than negative ones systematically increased by time frame width. Figure 2 displays these interactions.

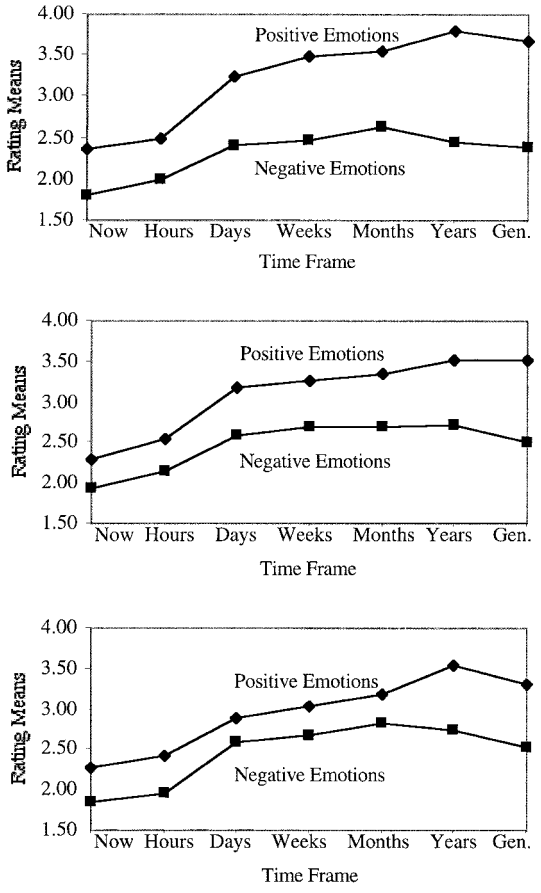

Figure 2. Effects of time frame and emotion scale on rating means, Study 1; Samples 1 (top), 2 (middle), and 3 (bottom). Gen. = general.

To determine the locus of the Emotion Scale × Time Frame interactions, we performed separate 2 (emotion scale) × 3 (time frame) ANOVAs for the short time frames (now, hours, days) on the one hand versus the long time frames (months, years, in general) on the other. In all three of the analyses involving the long time frames, there was a significant Emotion Scale × Time Frame interaction, $F(2, 78) = 5.35$, $p = .007$ in Sample 1, $F(2, 108) = 5.88$, $p = .004$ in Sample 2, and $F(2, 104) = 12.57$, $p = .000$ in Sample 3. In the analyses involving short time frames, by contrast, only one of the Emotion Scale × Time Frame interactions was significant, $F(2, 78) = 3.25$, $p = .044$ in Sample 1, $F(2, 108) = 1.88$, $p = .158$ in Sample 2, and $F < 1$ in Sample 3. In sum, these results are consistent with the idea that the tendency to report more positive than negative emotions increases when participants switch from an episodic retrieval strategy to a semantic retrieval strategy.

Linear and curvilinear effects on rating means. Although positive and negative emotions displayed slightly different patterns, both conformed to predictions. For short time frames, both seemed to increase linearly. This is consistent with an episodic retrieval strategy. However, for long time frames, it appears that participants gave similar answers regardless of the time frame under consideration. The latter pattern is consistent with the premise that participants were using a semantic rather than an episodic retrieval strategy.

If our interpretation of the rating means is correct, we should see both linear and curvilinear effects on rating means, just as we did with judgment latencies. To assess this prediction, we performed within-subject multiple regressions, predicting a person's rating means as a function of time frame width (coded from 1 = *at this moment* to 7 = *in general*). As in predicting judgment latencies, both linear (time frame) and curvilinear (time frame squared) predictors were entered simultaneously. Participant-level betas were then entered into a new data set and tested for significance. Mean betas testing the linear contrast were highly significant in each of the samples (mean βs = 1.02, 1.02, and 1.00 in Samples 1, 2, and 3; ps $< .001$). That is, the wider the time frame was, the more intense the emotion ratings were. In addition, however, mean betas testing the curvilinear contrast were also highly significant in each of the samples (mean βs = −.72, −.73, and −.70 in Samples 1, 2, and 3; ps $< .001$). The latter results indicate that, as time frames became particularly long, the tendency to give higher ratings disappeared.

To further probe the linear and curvilinear trends, we obtained two new sets of regression coefficients. For each participant separately, we entered a linear contrast predictor for short time frames (1 = moment; 2 = hours; 3 = days) and then did the same for long time frames (5 = months; 6 = years; 7 = in general). As expected, the linear contrast was significant for short time frames (mean beta across participants and samples = .293; ps $< .05$). However, the linear contrast was not significant for long time frames (mean beta across participants and samples = −.02; ps $> .05$). Thus, it appears that participants gave different answers for each of the short time frames but gave similar answers for each of the long time frames.

Discussion

It is often assumed that people retrieve and aggregate instances when reporting on their emotions. Study 1 provides a method, previously unused, for assessing this assumption. If we lengthen the period of time involved in the self-report, an episodic retrieval strategy should be associated with longer and longer judgment latencies because, with progressively wider time frames, more and more instances of felt emotion should become relevant. The linear rise in judgment latencies with time frame width, observed in each of our three samples, thus provides positive evidence for the use of this effortful strategy.

Study 1, however, demonstrates that there are important limitations on the use of an episodic retrieval strategy. When the time frame becomes particularly wide or when the time frame is not specified, we might expect people to abandon such a strategy. Study 1 offers novel evidence for this contention. Specifically, the linear rise in judgment latencies with time frame width was qualified by a second curvilinear trend. For relatively long time frames, increasing the width of the time frame did not result in longer latencies but rather resulted in latencies that were insensitive to time frame width. Beyond a certain peak aggregation point (normatively the last few weeks), judgment latencies flattened out. Such results provide some initial latency evidence for our contention that certain reporting formats (retrospective, long aggregated, prospective, and trait) elicit semantic rather than episodic retrieval strategies.

Ratings provided converging evidence for two processes in self-report. For time frames shorter than the last few weeks, ratings increased linearly with time frame width. That is, the longer the time frame was, the more extreme the emotion ratings were. We interpret this effect by proposing that emotion frequency, which increases with the width of the time frame, influenced intensity ratings. Support for this idea can be found in Thomas and Diener (1990) as well as Schimmack and Diener (1997). The present results document an interesting consequence of this tendency. If people implicitly consider emotional frequency when rating emotional intensity, longer time frames should result in more extreme intensity ratings. In essence, our participants appeared to be more emotional beings when longer (vs. shorter) time frames were involved, but only when they based their estimates on an episodic retrieval strategy.

It should be noted, however, that researchers have invoked three distinct principles to account for participants' exaggerated ratings of retrospective emotional intensity. Kahneman (1999) invoked the hypothesis of peak effects in memory, Thomas and Diener (1990) invoked the hypothesis that frequency biases intensity, and Winkielman et al. (1998) invoked the hypothesis of conversational norms. According to the Winkielman et al. framework, exaggerated ratings of retrospective emotional intensity are not due to memory biases per se but to the fact that participants implicitly assume that the researcher is interested in more extreme experiences when more retrospective prompts (e.g., in the last year) are used relative to less retrospective prompts (e.g., today).

The present findings cannot distinguish among these hypotheses, but they do suggest limitations to all of them. All three accounts of retrospective exaggeration assume that participants are using an episodic retrieval strategy and that the bias is due to the particular episodic information that is retrieved (particularly intense instances, as in Kahneman, 1999; frequency information, as in Thomas & Diener, 1990; extreme exemplars, as in Winkielman et al., 1998). To the extent that participants abandon an episodic retrieval strategy for relatively long time frames and instead access semantic emotion knowledge, we expect similar answers regardless of the time frame under consideration. Consistent with this expectation, there was no systematic trend for ratings longer than the last few weeks. Thus, ratings, like judgment latencies, suggested the operation of two distinct retrieval strategies for relatively short (episodic retrieval) versus relatively long (semantic knowledge) time frames, a pattern that is not predicted by the other accounts of retrospective exaggeration.

Finally, there was an interesting Emotion Scale $\times$ Time Frame interaction that replicated across each of the three samples. Ratings for the positive emotions were always higher than ratings of negative emotions (Watson, 2000), but this differential pattern increased with time frame width. Subsequent analyses indicated that the divergence was only reliable for the particularly long time frames—months, years, and in general. It is notable that these frames were those that we felt would elicit a semantic rather than an episodic retrieval strategy. This pattern of findings makes sense when we consider the large literature demonstrating that people tend to view their lives in the most positive possible light given the constraints of the situation (Baumeister, 1993; Dunning, Meyerowitz, & Holzberg, 1989; Liberman & Trope, 1998; Robinson & Ryff, 1999; Taylor, 1991; Taylor & Brown, 1988).

Given an episodic retrieval strategy, such constraints should be readily available, leading to a less rosy view of one's emotional experiences. However, given a semantic retrieval strategy, such situational constraints should not be retrieved, leading to a more rosy view of one's emotional experiences. In essence, it appears that people view themselves as happier in the abstract than in particular situations. Such a pattern of findings is consistent with the other episodic–semantic dissociations reported in Study 1.

Study 2

Study 1 suggests that people use distinctly different strategies when reporting on their emotions over short versus long time frames. For the momentary, hours, and days time frames, the results were consistent with the use of an episodic retrieval strategy. For the months, years, and in general time frames, the results were consistent with the use of a semantic retrieval strategy. Support for the distinct strategies would be stronger, of course, if other sources of data suggested a similar dissociation. Study 2 constitutes such a test.

Our model suggests that people access episodic emotion knowledge for short time frames and semantic emotion knowledge for long time frames. One way to examine knowledge use is to look for priming effects across consecutive trials (Neely, 1991). If people retrieve the same source of information (e.g., semantic emotion knowledge) on two consecutive trials, the second trial should be facilitated because useful information is already accessible. Returning to the literature on lexical decisions, one might expect lexical decisions for animal words (e.g., *bunny*) to be faster if the preceding trial also involved an animal word (e.g., *skunk*) versus a texture word (e.g., *smooth*). This suggests that both trials activated the same source of knowledge, namely knowledge about animals.

In the current studies, trials were randomly presented for each participant. This resulted in four trial types of primary interest. A given target frame was either short or long, preceded by a given prime frame that was also either short or long. This produced four prime–target pairs (short–short; short–long; long–short; long–long). To investigate the idea that there are two separate knowledge sources (episodic and semantic), we reanalyzed the target latencies from Study 1 as a function of the width of prime and target frames.

If we are correct in assuming that there are two distinct knowledge sources accessed when people report on short versus long time frames, then prime frame should interact with target frame in influencing target latencies. This can be contrasted with the assumption that the same source of information, either episodic or semantic emotion knowledge, is accessed on every trial. Such an alternative, one-process model would not produce a Prime Frame $\times$ Target Frame interaction.

In predicting the exact nature of the expected interaction, it is useful to revisit cognitive research on episodic and semantic knowledge. Semantic knowledge has been characterized as invariant and tightly organized (Tulving, 1984). If we are correct in assuming that people access semantic emotion knowledge in reporting on their emotions over long time frames, then long–long trial pairs should result in faster target latencies than should short–long trial pairs.

In contrast to semantic knowledge, episodic knowledge has been characterized as highly variable and loosely organized (Tulving, 1984). If this is true and if we are correct in assuming that people access episodic emotion knowledge in reporting on their emotions over short time frames, then we might expect short–short target latencies to be equivalent to long–short target latencies. In essence, such results would suggest that, when people make an episodic judgment, information relevant to one time frame–emotion combination (e.g., anger in the last few hours) is irrelevant to another time frame–emotion combination (e.g., sadness in the last few days). For each episodic judgment, in other words, a distinct set of episodic details must be retrieved.

In sum, we predicted that long–long target latencies would be faster than short–long target latencies because semantic knowledge accessed for one trial is useful on the next. By contrast, we predicted that short–short target latencies would be equivalent to long–short target latencies because every episodic judgment concerns a different set of episodic details. Such an asymmetrical interaction would provide strong evidence for the idea that emotion judgments made on the basis of episodic versus semantic emotion knowledge are qualitatively distinct.

Obtaining Sufficient Power

We expected priming effects to be robust but small. There are several reasons for this. First, we predicted an asymmetrical Prime Frame × Target Frame interaction rather than a crossover one. Such an asymmetrical interaction requires more power to detect. Second, there were over 2 s between trials. Because priming effects decay rapidly (Neely, 1991), this would result in a smaller priming effect than would be detected with a shorter delay between trials. And third, primes and targets were not explicitly paired but rather consisted of consecutive trials. The major advantage of this method is that it is sensitive to spreading activation but not to several more controlled mechanisms that often contribute to priming effects (McNamara & Altarriba, 1988; Shelton & Martin, 1992). For example, when primes and targets are explicitly paired, participants use the prime category (e.g., an animal word) to generate a conscious expectancy for the target category (e.g., another animal word), which results in a larger priming effect (Neely, 1991). Such conscious expectancies do not play a role when primes and targets represent consecutive trials (McNamara & Altarriba, 1988; Shelton & Martin, 1992). The major disadvantage of this method is that it results in a smaller priming effect, typically in the neighborhood of 15 ms (Neely, 1991). As such, more power is necessary to detect it.

Because we used the same procedures and time frames for each of the three samples (see Study 1), we were able to obtain more power by combining the data sets. We did, however, make sure that the critical findings did not interact with the particular sample involved. It should be noted that the added power works against our hypothesis that short–short target latencies are equivalent to long–short target latencies.

Method

Participants from all three samples were asked to judge their experience of positive versus negative emotions over short versus long time frames (see Study 1 for details). Of interest in Study 2 is how quickly participants could judge a particular type of emotion (target trial) as a function of the type of emotion judged on the previous (prime trial). The full design consisted of a 2 (prime valence: positive vs. negative) × 2 (target valence: positive vs. negative) × 7 (prime frame) × 7 (target frame) analysis. Needless to say, no participant had observations in every cell.

For purposes of aggregating trials, we adopted several necessary simplifications. First, prime and target trials were classified as short (moment, hours, and days) versus long (months, years, and general), with the intermediate time frame of the last few weeks excluded. Second, prime–target pairs were classified as congruent (positive–positive or negative–negative) versus incongruent (positive–negative or negative–positive) in valence. This resulted in a 2 (valence type: congruent vs. incongruent) × 2 (prime frame: short vs. long) × 2 (target frame: short vs. long) design. Because of these simplifications, all participants could be included in the analysis.[6]

As indicated above, priming effects across sequential trials tend to be slight in magnitude. For this reason, we created a new data set composed of participants from all of the three samples. However, we coded the sample that the participants came from so that we could ensure that the effects were parallel across samples. Analyses were performed on log-transformed latencies, but means are reported in terms of milliseconds.

Results

Judgment latencies were initially examined in a 2 (valence congruence) × 2 (prime frame) × 2 (target frame) × 3 (sample) ANOVA. Because sample did not modify any of the effects of interest ($ps > .10$), we reran the analysis without sample as a variable. The analysis revealed a main effect for valence congruence, $F(1, 146) = 6.32$, $p = .013$, that was not modified by prime frame, target frame, or their interaction ($ps > .20$). When succeeding trials involved emotions that were congruent in valence (vs. incongruent), judgments were faster (in seconds, congruent valence = 2.118 vs. incongruent valence = 2.203; priming effect = 8.5 ms).

There was also a main effect for target frame, $F(1, 146) = 4.82$, $p = .030$, due to the fact that shorter target frames were rated more quickly than were longer target frames (short target frame = 2.124; long target frame = 2.197; difference = 7.3 ms). More intriguing, there was a main effect for prime frame, $F(1, 146) = 10.42$, $p = .002$. Target judgments were quicker when the preceding trial involved a long (vs. short) time frame (short prime frame = 2.204; long prime frame = 2.117; priming effect = 8.7 ms).

The main effect for prime frame, however, must be interpreted in light of a significant Prime Frame × Target Frame interaction, $F(1, 146) = 9.03$, $p = .003$. For long target frames, there appeared to be a substantial effect for prime frames (short prime frame = 2.276; long prime frame = 2.118; priming effect = 15.8 ms). For short target primes, on the other hand, there appeared to be no effect for prime frames (short prime frame = 2.131; long prime frame = 2.117; priming effect = 1.4 ms). These means are graphically displayed in Figure 3.

As a follow-up to the significant Prime Frame × Target Frame interaction, we next performed separate 2 (valence congruence) × 2 (prime frame) ANOVAs on short target judgments, on

[6] We also performed a second set of analyses by characterizing prime and target time frames as short (moment and hours) versus medium (days, weeks, and months) versus long (years and general). Results involving these analyses were highly parallel to those reported below.

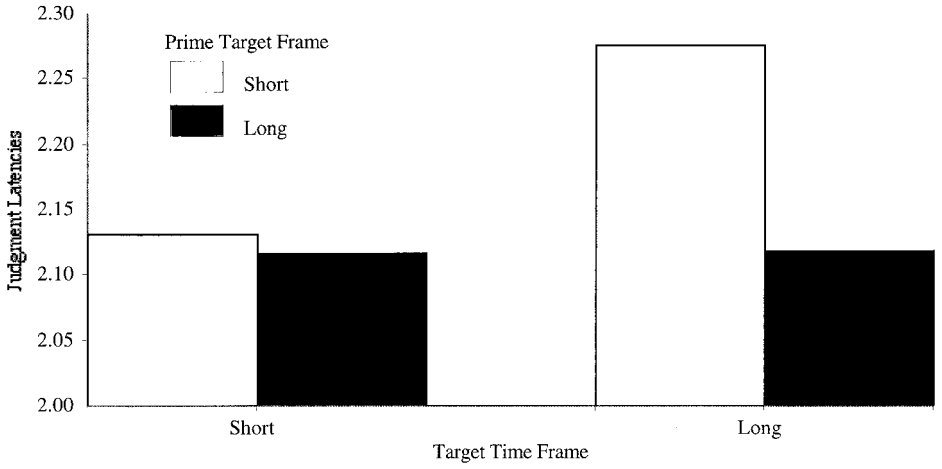

Figure 3. Effect of prime and target frame on judgment latencies, Study 2.

the one hand, and long target judgments, on the other. The analysis performed on short target judgments yielded no significant priming effects. That is, the main effect for prime frame was not significant ($F < 1$), nor was the main effect for valence congruence significant ($p > .10$). Also, the Prime Frame × Valence Congruence interaction was not significant ($F < 1$). Thus, judgments concerning short time frames were uninfluenced by the priming factors.

By contrast, both priming factors influenced the speed with which long target frames were judged. That is, for long target frames, there was a main effect for prime frame, $F(1, 147) = 19.97$, $p = .000$, as well as a main effect for valence congruence, $F(1, 147) = 4.12$, $p = .044$. The interaction between these two priming factors was not significant ($F < 1$).

In summary, these analyses support the claim that people use distinctly different strategies in reporting on short versus long time frame emotions. Of most importance, long time frame judgments were significantly facilitated by preceding trials that were either congruent in valence or congruent in time frame width. By contrast, the same priming factors did not facilitate short time frame judgments.

Discussion

On the basis of our model (Robinson & Clore, in press), we tested the idea that participants access different sources of knowledge when reporting on their emotions over short versus long time frames. When we analyze consecutive trials as a function of prime and target time frame width, such a two-process model should yield a Prime Frame × Time Frame interaction on target latencies. Study 2 confirms this prediction and thus rules out the idea that people access the same source of knowledge—either episodic or semantic—regardless of the width of the time frame.

Study 2 further suggests a fundamental distinction between episodic and semantic retrieval strategies. In this connection, we found that judgments for long time frames differed by the time frame width of the prime trial but that judgments for short time frames did not. Consistent with the idea that semantic knowledge is invariant and tightly organized (Tulving, 1984), long–long target latencies were significantly faster than were short–long target latencies. This suggests that the same source of knowledge—namely, semantic emotion knowledge—was accessed for different judgments of emotion over long time frames. These results provide a further rationale for the flat latency and rating curves found for long time frame reports in Study 1. If the same source of information is accessed regardless of the particular time frame (Study 2), one would expect these flat curves (Study 1).

Whereas semantic knowledge is tightly organized, episodic knowledge is highly variable and loosely structured (Tulving, 1984). Consistent with the idea that every episodic judgment involves a distinct set of episodic details, we found that short–short target latencies were equivalent to long–short target latencies. Thus, answering a question about anger over a period of several days does not allow one to more quickly characterize one's sadness in the last couple of hours. In short, an episodic retrieval strategy seems to entail the retrieval of different emotional instances for every emotion judgment. There appears to be no such thing as an organized memory structure for the episodic emotional self.

In addition to the interaction showing priming effects for time frame (Prime Frame × Target Frame), we also obtained an interaction showing priming effects for emotion valence (Prime Valence × Target Valence). The latter interaction indicates that people retrieve different sets of knowledge when reporting on their negative versus positive emotions. It would be interesting to use this paradigm to access individual differences. For example, it might be the case that neurotics have more tightly interconnected knowledge concerning their negative emotions, whereas extraverts have more tightly interconnected knowledge concerning their positive emotions. This possibility should be investigated in future studies.

Although the valence interaction did not interact with the time frame interaction, additional analyses provide some support for distinctions between episodic and semantic retrieval strategies. Specifically, there was a significant Prime Valence × Target Valence interaction for long time frames but not for short ones. The differential results again suggest a tighter organization for semantic emotion knowledge than for episodic emotion knowledge. However, this conclusion should be viewed with caution, as

the critical Prime Valence × Target Valence × Target Frame interaction was not significant.

Study 3

In Study 2, we examined priming on a trial-by-trial basis and found support for the idea that people access distinct sources of information in reporting on their emotions over short (episodic) versus long (semantic) time frames. In Study 3, we take a different approach to priming. On the basis of the idea that people use semantic emotion knowledge when reporting on their emotions over long time frames, we decided to manipulate the accessibility of a given source of emotion-related beliefs prior to the study. Because most people agree that women experience more intense emotions than do men (Robinson & Johnson, 1997; Shields, 1987) and because such beliefs are incorporated into the self-concept (see Cross & Madson, 1997, for a review), it should be possible to influence ratings by manipulating the accessibility of sex role stereotypes.

Of critical importance, we expected that the manipulation would interact with time frame. If people use semantic emotion knowledge in reporting on their emotions over long time frames, the manipulation should produce the following pattern of results. Men in the belief priming condition, relative to men in the control condition, should report less emotionality for long time frames, consistent with the stereotype of the stoic man. Women in the belief priming condition, relative to women in the control condition, should report more emotionality for long time frames, consistent with the stereotype of the emotional woman.

We initially entertained the idea that the priming manipulation would not influence ratings of emotion for short time frames. However, a review of the literature suggested that this was implausible. When people are primed with knowledge that they regard as irrelevant to the judgment at hand, the typical result is a contrast effect (Herr, Sherman, & Fazio, 1983; Martin & Achee, 1992; Schwarz & Bless, 1992; Stapel, Koomen, & van der Pligt, 1997; Strack, 1992). Indeed, Wilson and Brekke (1994) have asserted that priming always or nearly always taints judgments, either in a prime-congruent or in a prime-incongruent direction (e.g., see Herr et al., 1983). In their view, it is nearly impossible to avoid this mental contamination.

A number of specific models have been proposed to account for assimilation and contrast effects in social judgment. What nearly all of them share is an emphasis on prime–target similarity. If the prime and the target involve similar knowledge sources, assimilation is likely. However, if they involve dissimilar knowledge sources, contrast is likely (Herr, 1986; Herr et al., 1983; Schwarz & Bless, 1992; Stapel et al., 1997; Strack, 1992).

In the present context, we manipulated a source of semantic emotion knowledge. If people use semantic emotion knowledge in reporting on their emotions over long time frames, the prime and target information involve similar (i.e., semantic) emotion knowledge. On this basis, we expected an assimilation effect; that is, priming should result in more stereotypic ratings for long time frames. If people use episodic rather than semantic emotion knowledge in reporting on their emotions over short time frames, the prime (i.e., semantic) and target (i.e., episodic) information involve dissimilar emotion knowledge. On this basis, we expected a contrast effect; that is, priming should result in less stereotypic ratings for short time frames.

In sum, we predicted a 2 (participant sex) × 2 (priming condition) × 7 (time frame) interaction on emotion ratings. If the priming manipulation does indeed have such dramatically opposite effects for short versus long time frame emotion reports, this would provide strong support for the model.

Method

Participants. Participants were the same 53 individuals (19 men and 34 women) described as Sample 3 in Study 1. By random assignment, 10 men and 18 women were assigned to the control condition, whereas 9 men and 16 women were assigned to the experimental (priming) condition.

Procedures and dependent measures. Participants engaged in what they believed to be two separate studies. The first one concerned "people's ideas about different groups of people," whereas the second one concerned "your experience of emotions over different time frames." In reality, we used the first task to prime sex role stereotypes about emotion and examined the effects on self-reported emotions. Procedures for the computerized task are presented above. All participants in this sample were run individually.

Priming sex role stereotypes. To increase the accessibility of sex role stereotypes of emotion, we used a priming manipulation modeled after Trafimow, Triandis, and Goto (1991). Participants were asked to write about "how you differ from a specific group of people" in terms of "your behaviors, emotions, values, and/or lifestyle." Participants were asked to list at least five important differences in 4 min and were randomly assigned to control ($n = 28$) versus experimental ($n = 25$) conditions. Those in the control condition contrasted themselves with Canadians, a reference group that elicited few systematic beliefs (aside from group differences in saying "eh"). By contrast, those in the experimental condition contrasted themselves with their opposite-sex peers, a reference group that should elicit strong and systematic beliefs. As expected, a majority of participants in the experimental condition mentioned women's greater emotionality in their answers.

Results

Overview. Initial analyses included emotion scale (positive vs. negative) as a variable. However, none of the effects below interacted with this variable ($ps > .20$). For this reason, we collapsed across this variable. Higher rating means represent perceptions of more intense felt emotion, whereas longer latencies represent relative difficulties in making emotion judgments. Mean ratings and judgment latencies were separately examined in 2 (sex) × 2 (priming condition) × 7 (time frame) ANOVAs.

Effects on rating means. The analysis on rating means revealed a main effect for sex, $F(1, 49) = 4.32$, $p = .043$ (Ms = 2.60 vs. 2.76 for men and women) and time frame, $F(1, 49) = 101.05$, $p = .000$, but not priming condition ($F < 1$). Among the possible two-way interactions, the Sex × Priming Condition and the Priming Condition × Time Frame interactions were not significant ($ps > .35$). There was, however, a significant Sex × Time Frame interaction, $F(6, 294) = 3.54$, $p = .002$. The extent to which women reported more intense emotion than men was somewhat more pronounced for longer time frames, particularly relative to those time frames that were quite short (mean differences by sex were .10, −.11, .16, .34, .21, .27, and .15 for now, hours, days, weeks, months, years, and in general time frames). Such a pattern is consistent with prior research indicating that sex differences in

emotional intensity become more exaggerated with more retrospective reporting formats (Robinson et al., 1998; Shields, 1991), likely because sex-related beliefs play a larger role when people retrospect on their emotions (Robinson & Clore, in press).

However, the effects reported above were qualified by the predicted Sex × Priming Condition × Time Frame interaction, $F(6, 294) = 3.64$, $p = .002$. As revealed by the means reported in Figure 4, the tendency for women to report more intense emotions than men was more pronounced in the belief priming condition than in the control condition, but only for relatively long time frames. For shorter time frames, the pattern of means reversed, such that belief-primed men reported more intense emotions (relative to men in the control condition), whereas belief-primed women reported less intense emotions (relative to women in the control condition). In other words, activated beliefs were assimilated when participants reported on emotions over long time frames but served as a standard of contrast when participants reported on emotions over short time frames. These findings strongly point to the presence of two processes, one episodic and one semantic, in emotional self-report.

There is a striking convergence involving the latency data reported for Sample 3 in Figure 1 and the rating means reported in Figure 4. The peak latency for this sample occurred for the time frame of the last few weeks (see Figure 1). This same time frame appeared to serve as the normative shifting point from belief contrast to belief assimilation (see Figure 4). To illustrate this fact, we performed an analysis in which rating means were collapsed across Times 1 (now), 2 (hours), and 3 (days), on the one hand, and across Times 5 (months), 6 (years), and 7 (in general) on the other. A 2 (sex) × 2 (priming condition) × 2 (time frame: left vs. right of peak) ANOVA resulted in a significant three-way interaction, $F(1, 49) = 10.02$, $p = .003$. Means for this interaction are graphically displayed in the bottom panel of Figure 4.

Effects on judgment latencies. Judgment latencies were examined in a 2 (sex) × 2 (priming condition) × 7 (time frame) ANOVA. The analysis revealed a significant main effect for time frame, $F(6, 294) = 9.74$, $p = .000$, as well as a Sex × Time Frame interaction, $F(6, 294) = 2.84$, $p = .011$. None of the other effects were significant ($ps > .40$). The Sex × Time Frame interaction was intriguing, as it suggested that women, in comparison with men, found it relatively easy to make certain judgments but relatively difficult to make others. Consistent with the premise that women possess more semantic knowledge concerning their emotions (Feldman Barrett, Lane, Sechrest, & Schwartz, 2000), women were consistently faster to judge their emotions over longer time frames. Specifically, women were 15.0 ms faster for the weeks time frame, 3.2 ms faster for the months time frame, 13.9 ms faster for the years time frame, and 13.2 ms faster for the general time frame. By contrast, there was no systematic pattern for short time frames, as women were 42.7 ms slower for the moment time frame, 16.7 ms faster for the hours time frame, and 8.4 ms slower for the days time frame.

Although the sex difference in latencies was not consistent for the short time frames, the overall pattern is suggestive that women, in comparison with men, found it easier to rate their emotions over longer time frames than over shorter ones. This makes sense if one proposes that semantic knowledge about emotion—which women arguably possess more of—facilitates reporting for long time frames but can actually interfere with reporting for short time

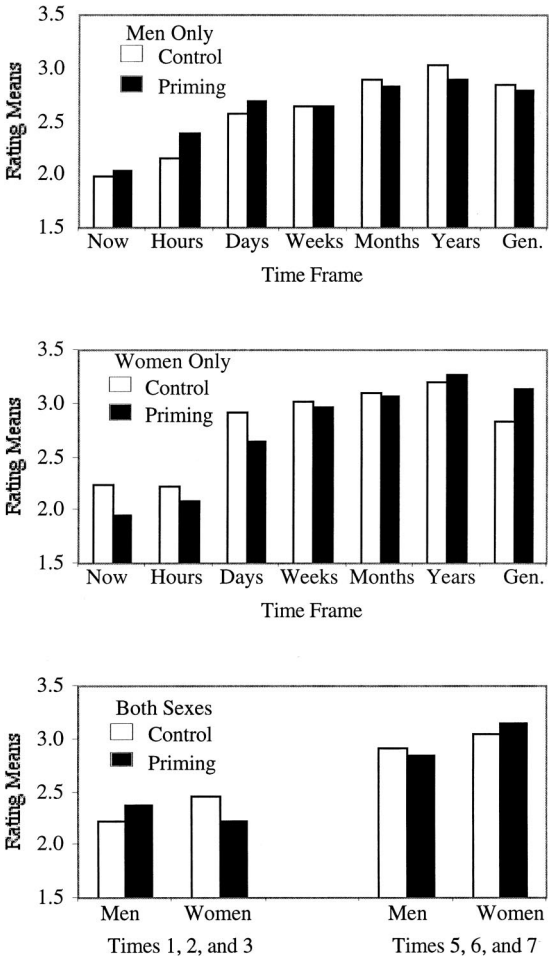

Figure 4. Effects of time frame, participant sex, and priming condition on rating means, Study 3; men only (top), women only (middle), and both sexes, with Time Frames 1, 2, and 3 contrasted with Time Frames 5, 6, and 7 (bottom).

frames. We pursued this sex difference in one additional way. We examined the multiple regression coefficients reported in Study 1 for linear effects of time frame (7) on judgment latencies. If women, in comparison with men, found it easier to report on long time frames than on shorter ones, we would expect to see a lower beta for the linear rise in latency with time frame as a predictor. A 2 (sex) × 2 (priming condition) ANOVA on the linear coefficient betas revealed that this was the case. Although there were no effects due to priming condition ($ps > .25$), the main effect for sex was significant, $F(1, 49) = 6.40$, $p = .015$. As time frames became progressively wider, men found it increasingly difficult to make their judgments (mean $\beta = .91$). The same tendency among women, however, was markedly weaker (mean $\beta = .35$).

Discussion

A number of researchers have documented dissociations between on-line and retrospective reports of emotion. For example, several literature reviews (LaFrance & Banaji, 1992; Shields, 1991) and studies (Feldman Barrett et al., 1998; Robinson et al.,

1998) have shown that self-reports of emotion tend to be more consistent with sex role stereotypes when individuals report on retrospective or generalized emotions than when they report on emotions that are relatively on-line. Such results hint at differences in the sources of information accessed and used under different reporting conditions. That is, it appears that people use sex role stereotypes about emotion, a source of semantic emotion knowledge, particularly when the width of the time frame discourages the use of an episodic retrieval strategy (Robinson & Clore, in press).

Study 3 was designed to provide more direct evidence for our account of dissociations. If we assume that people use semantic emotion knowledge when reporting on their emotions over relatively long but not short time frames, then a manipulation of the accessibility of this source of knowledge should differentially influence long versus short time frame reports. Study 3 provides evidence for this hypothesis. For relatively long time frames and for generalized reports of emotion, participants who had been primed with sex role stereotypes reported emotions that were more stereotypic. For relatively short time frames, by contrast, participants who had been primed with sex role stereotypes reported emotions that were less stereotypic. The differential effects of priming for short versus long time frames thus converges with Studies 1 and 2 in suggesting two fundamentally different processes in emotional self-report.

Why did the priming manipulation produce stereotypic ratings for long time frames? According to the present framework, long time frames as well as generalized reports of emotion discourage the use of episodic emotion knowledge (i.e., knowledge of how one felt at a particular moment in time in a particular setting). In both cases, there are simply too many moments of experience to retrieve and aggregate. People possess various sources of semantic emotion knowledge, and sex role stereotypes are only one such source. Therefore, a manipulation of the relative accessibility of this source of knowledge should influence the degree to which people use sex-related beliefs in making ratings.

Why did the priming manipulation produce less stereotypic ratings for short time frames? One possibility is that activated sex role stereotypes facilitated the retrieval of stereotype-incongruent information. By this account, men who had been primed to think of themselves as relatively unemotional recalled instances in which they felt particularly intense emotions, whereas women who had been primed to think of themselves as relatively emotional recalled instances in which they felt emotions that were not particularly intense. Such a possibility is consistent with the frequent finding that schemas facilitate schema-incongruent recall (Alba & Hasher, 1983; Olson, Roese, & Zanna, 1996; Stangor & McMillan, 1992). However, such a memory-based account encounters difficulties in explaining why the momentary (i.e., "right now") time frame (in which memory likely played a small role) was associated with a similar pattern of findings as were the other short time frames (see Figure 4).

Instead, we assume that participants in the two conditions retrieved similar episodic information. If so, the difference between the two conditions was in the accessibility of sex role stereotypes, a source of semantic emotion knowledge. Given our contention that people prefer to use an episodic retrieval strategy when reporting on recent experiences, then participants in the belief priming condition were faced with removing or correcting for the influence of potentially misleading semantic information. Indeed, the social judgment literature confirms that contrast effects are the norm when people view activated information as irrelevant to the task at hand (Herr, 1986; Herr et al., 1983; Schwarz & Bless, 1992; Stapel et al., 1997; Strack, 1992). Such frameworks clearly predict the obtained finding that belief-primed individuals reported counterstereotypic emotions for recent time frames. In short, participants implicitly viewed the activated semantic knowledge as irrelevant to the task at hand—judging their recent emotional experiences—and discounted it.

A final result from Study 3, although not central to predictions, nonetheless provides suggestive evidence for two processes in emotional self-report. Women were faster than men in rating their emotions over long time frames and in general but slower than men in rating their emotions over short time frames. Thus, there was a smaller effect for time frame on the judgment latencies of women than on the judgment latencies of men. There is considerable support for the idea that women think about and value their emotions to a greater extent (Cross & Madson, 1997; Gasper & Clore, 2000; Robinson & Clore, in press; Shields, 1995). As a result, they possess more elaborated semantic knowledge about their own emotions as well as about the emotions of others (Feldman Barrett et al., 2000). Because our framework assumes that semantic emotion knowledge facilitates self-reports of emotion for long time frames and generalized reporting formats but not for short time frames, the observed sex difference is consistent with such a framework.

General Discussion

Summary of Main Findings

Our aim in these studies was to investigate, as directly as possible, how people make self-reports of emotion. On the basis of previous work (see Robinson & Clore, in press, for a review), we expected to find evidence for two retrieval strategies—episodic and semantic—and also expected that these two strategies would be differentially used for short (episodic) versus long (semantic) time frames. A premise of our model is that people prefer to retrieve and aggregate instances of felt emotion when possible but that certain reporting formats make it very difficult to use such an episodic retrieval strategy. Thus, we expected that when the time frame was particularly long (e.g., the last few years) or when the time frame was not specified (e.g., in general), we would find evidence for a semantic retrieval strategy.

In Study 1, we asked three samples of participants to make emotion judgments over time frames that systematically varied in their width. The use of an episodic retrieval strategy should be associated with a linear rise in judgment latencies with the width of the time frame (N. R. Brown, 1995), whereas the use of a semantic retrieval strategy should be associated with a flat slope. If participants had adopted only an episodic retrieval strategy, we might have expected only a linear pattern. If participants had adopted only a semantic retrieval strategy, we might have expected a flat slope. Instead, there was evidence for both a linear rise for short time frames and a flat slope for long time frames, suggesting the use of two distinct retrieval strategies for short (episodic) versus long (semantic) time frames. Rating means also suggested the use of two different strategies, with a linear rise for short time frames and a flat slope for long time frames.

In Study 2, we provided further support for two separate emotion reporting strategies. Reasoning that if people access the same source of information on two consecutive trials, judgment latencies should be facilitated, we predicted that target latencies for long time frames would be faster when the preceding trial involved another long (vs. a short) time frame judgment. Such results would suggest that the same source of semantic emotion knowledge was used on both trials. Priming results support this prediction. By contrast, we predicted that target latencies for short time frames would not evidence priming effects because each rating required a different source of episodic emotion knowledge (Tulving, 1984). As expected, there was no facilitation for judgments of emotion over short time frames. Together, the findings confirm that people use strikingly different sources of knowledge when reporting on their emotions over short (episodic) versus long (semantic) time frames.

In Study 3, we manipulated the accessibility of sex role stereotypes because there is considerable evidence that both men and women share the belief that women are more emotional than men are (e.g., Shields, 1987; Widiger & Settle, 1987). If participants had adopted only an episodic retrieval strategy, we would expect the manipulation to have little influence or, more likely (Schwarz & Bless, 1992), produce counterstereotypic emotion ratings. If participants had adopted only a semantic retrieval strategy, we would expect the manipulation to produce ratings consistent with sex role stereotypes. Instead, there was evidence for two processes, one occurring for long time frames and one occurring for short ones. For long time frames, the priming manipulation resulted in ratings that were more stereotypic, indicating that the activated semantic emotion knowledge had been used for these time frames. For short time frames, the priming manipulation resulted in ratings that were counterstereotypic, indicating that the activated semantic knowledge had been rejected or discounted for these time frames.

Until the present investigation, evidence for two processes in emotional self-report had come primarily from the literature on dissociations between on-line and retrospective ratings (Levine, 1997; Robinson & Clore, in press). The present results complement this literature but extend it with references to several novel sources of evidence. Judgment latencies, we suggest, can provide us with unique insights into the emotion reporting process. In addition, priming manipulations allow us to examine how people use accessible information in reporting on their emotions. Before outlining some of these future directions for research, however, we discuss several of the central assumptions of our model.

Aggregation in Emotional Self-Report

Compared with reports of emotion for short time frames (e.g., right now), reports of emotion for longer time frames (e.g., in the last month) and generalized reports of emotion both display higher test–retest stabilities (e.g., Watson & Walker, 1996). Similarly, when we aggregate instances of felt emotion, we obtain scores that are more reliable both on a temporal and on a cross-situational basis (Epstein, 1983; Diener & Larsen, 1984). Thus, it is tempting to suggest that we are obtaining similar scores regardless of whether aggregation is done explicitly by the experimenter (i.e., an average of instances) or implicitly by the person making the ratings (i.e., emotion reports for longer time frames or in general).

We argue, however, that these two forms of aggregation—by the experimenter versus by the participant himself or herself—are not equivalent. Even when people are asked to aggregate their experiences shortly after an emotional incident and even when the incident is a relatively brief occurrence, the two forms of aggregation do not yield the same answers. In this connection, Kahneman's (1999) research reveals that there are often systematic differences between the two types of reports, such that participant-aggregated reports (relative to experimenter-aggregated ones) tend to be disproportionately influenced by the most intense moments of the experience as well the most recent moments of the experience (Fredrickson & Kahneman, 1993; Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993; Redelmeier & Kahneman, 1996).

Aside from biases in episodic retrieval, there is another, perhaps more fundamental, source of discrepancies between participant-aggregated and experimenter-aggregated reports. When people are asked to report on their emotions over a relatively long time frame or in general, they are likely to shift from an episodic retrieval strategy to a semantic one. By positing such a shift in retrieval strategies, we are able to understand other dissociations that are difficult to explain with reference to episodic retrieval alone. We can understand why self-reports are more consistent with sex role stereotypes when retrospective or generalized reporting formats are used than when on-line reporting formats are used (Feldman Barrett et al., 1998; Robinson et al., 1998), we can understand why the neuroticism–distress relationship is more pronounced when participant-aggregated formats are used relative to experimenter-aggregated ones (K. W. Brown & Moskowitz, 1997; Feldman Barrett, 1997; Larsen, 1992), and we can understand why people believe that they are less happy on Monday than on other days of the week, even though experience-sampling studies fail to establish this relationship (Stone, Hedges, Neale, & Satin, 1985). In each of these cases, participants appear to be drawing on different information when retrospective or participant-aggregated reports are used relative to on-line reports of emotion.

Our results suggest that the time frame of the last few weeks is likely to be the longest one typically associated with an episodic retrieval strategy. By prompting participants with multiple time frames, we found clear evidence for different retrieval processes on either side of this time frame. In Study 1, the time frame of the last few weeks separated the linear rise in judgment latencies for shorter frames from the flat slope for longer frames. In Study 2, the time frame of the last few weeks separated the distinctly different priming effects found for time frames shorter versus longer than this. In Study 3, finally, the time frame of the last few weeks marked the shifting point from contrast effects for time frames shorter than this to assimilation effects for time frames longer than this. In addition, Parkinson et al. (1995) and Schimmack (1997) both reported that week-retrospective reports of emotion were fairly accurate in capturing average daily emotion during the week.

Casting the net a bit wider, we note that there is striking convergence for the idea that episodic information becomes inaccessible after several weeks. In the literature on event frequency judgments, Blair and Burton (1987) found that 56% of their participants appeared to be using an episodic retrieval strategy for a 2-week time frame. For time frames shorter than this, the percentage was much higher; for time frames longer than this, the percentage was much lower. In the literature on autobiographical memory, Burt et al. (1998) have found that specific information

about events from the past becomes inaccessible after a week or 2 unless rehearsed. Finally, there is intriguing evidence for two memory systems, one (centered in the hippocampus) associated with excellent recall of episodic details, and the other (centered in the neocortex) associated with poor recall for episodic details (McClelland, McNaughton, & O'Reilly, 1995). The latter, slow-learning memory system, unlike the fast-learning memory system centered in the hippocampus, maintains long-term memories for regularities in experience and changes only very slowly in response to new events. These two memory systems bear a striking resemblance to our characterization of episodic versus semantic emotion knowledge. In this connection, it is interesting to note that memory traces within the hippocampus have been shown to disappear within a matter of weeks (McClelland et al., 1995).

In sum, these findings suggest that there is something like a 2-week time limit on the episodic self. For time frames longer than this as well as for generalized reports of emotion, we argue that the person must rely on more generalized beliefs about his or her emotions to make emotion judgments.

Belief in Emotional Self-Report

Beliefs about emotion are a source of semantic rather than episodic knowledge. They come from a variety of sources but generally fall into four categories that can be expressed within a 2 (normative vs. individualistic) $\times$ 2 (situation dependent vs. situation independent) design (Robinson & Clore, in press). A consideration of the different possible sources of belief helps us make sense of the wide variety of dissociations that have been reported in the literature.

Consider the following dissociations. Mitchell, Thompson, Peterson, and Cronk (1997) found that people retrospectively reported that they were more happy on their vacations than they actually were. Such a dissociation is likely due to the use of a normative/situation-dependent belief (i.e., vacations are pleasant). Feldman Barrett et al. (1998) found that women scored higher on trait emotionality scales but did not generally feel more intense emotions following daily social interactions. Such a dissociation is likely due to the use of a normative/situation-independent belief (i.e., women are more emotional than men). Arntz et al. (1990) found that trait anxious participants overestimated the degree to which they had experienced pain during dental surgery. Such a dissociation is likely due to the use of an individualistic/situation-dependent belief (i.e., dental surgery is intolerable). Finally, Feldman Barrett (1997) found that neurotics retrospectively overestimated the degree to which they had experienced negative affect over a 90-day period. Such a dissociation is likely due to the use of an individualistic/situation-independent belief (i.e., "I am the kind of person who experiences a lot of negative affect").

Which sources of belief do people access in any particular reporting situation? A review of the literature suggests that certain prompts tend to elicit situation-dependent beliefs, whereas others tend to elicit situation-independent ones. When one asks people to retrospect on their emotions in a particular situation (e.g., while menstruating; McFarland et al., 1989), it seems likely that they will access situation-dependent beliefs (e.g., menstruation causes negative affect). When no particular situation is specified, by contrast, it seems likely that people will access situation-independent beliefs. However, we see the need for more basic research on this issue.

Before suggesting some future directions for research, we should comment on several important issues. By emphasizing dissociations, we do not wish to imply that one retrieval strategy—episodic—always produces veridical ratings, whereas the other retrieval strategy—semantic—always produces biased ratings. As Kahneman's (1999) research demonstrates, people may retrieve a biased sample of emotional moments when using an episodic retrieval strategy. Study 3 of the present article also suggests that, in attempting to correct for irrelevant sources of information, people may sometimes give episodic reports that are biased in a belief-inconsistent direction. The latter effect has been replicated in several unpublished studies within our laboratory (Robinson, 2002), suggesting that even on-line reports of emotion are influenced by accessible but irrelevant information.

Furthermore, semantic emotion knowledge is often valid. People's beliefs about their emotion are no doubt based to a certain extent on their experiences over time. Neurotics do experience more negative affect in their daily lives, and extraverts do experience more positive affect (Feldman Barrett, 1997). However, there may be other beliefs, such as those related to menstruation and emotion, that are exaggerated or largely erroneous (Boyle & Grant, 1992; Olasov & Jackson, 1987; Ruble & Brooks-Gunn, 1979; Slade, 1984). Regardless of whether particular beliefs are valid or erroneous, our major aim is to understand how people judge their emotions. We see the distinction between episodic and semantic knowledge as crucial to this question.

Future Directions

There are many directions for future research, but two tools that we think will play an important role. Judgment latencies as well as priming manipulations probe the reporting process in a relatively direct way. Consider some of the questions that can be investigated by such methods.

What sources of information are cued by different prompts? Such questions are best addressed by manipulation studies. For example, we might phone people on a Monday and ask them about their emotions. One group could be asked "how you feel today"; a second group could be asked "how you feel today, that is, this Monday"; a third group could be asked "how you felt a week ago"; and a fourth group could be asked "how you felt a week ago, that is, last Monday." On the basis of the premise that situation-specific prompts cue situation-specific beliefs, we might predict that the two Monday-salient groups would report more negative emotions than would the two control groups. However, on the basis of the current results, we might predict a reversal similar to that observed in Study 3. If so, the retrospective Monday-salient group would report more negative affect than would the retrospective control group, whereas the on-line Monday-salient group would report less negative affect than would the on-line control group. Such a study would add to our knowledge concerning the distinct sources of information accessed under different reporting conditions.

In Study 3, we manipulated a semantic source of emotion knowledge, but it is also possible to manipulate an episodic source of information. Specifically, we might ask participants to recall event information from a particular day. If such episodic knowledge facilitates reaction times, we might conclude that people used

episodic knowledge in making their ratings. However, on the basis of the premise that people use semantic knowledge when retrospectively reporting on their emotions, we might expect no reaction time facilitation for a time frame such as a week ago. Instead, we might expect a manipulation of semantic knowledge, such as a trait report of typical emotions, to facilitate reaction times for such a time frame. Somewhat counterintuitively, we might expect this manipulation of semantic knowledge (i.e., trait ratings) to slow down ratings for particularly short time frames such as right now. We are currently using similar reaction time paradigms in our lab.

In summary, our understanding of the emotion-reporting process is largely inferential to date. Judgment latencies and manipulation studies can supplement dissociation studies in answering the question of how people judge their emotional experience.

Conclusions

On the basis of a prior literature review (Robinson & Clore, in press), we proposed that people access different types of knowledge when reporting their emotions over short (episodic) versus long (semantic) time frames. The three studies reported in this article provide evidence consistent with this contention. Although one may not realize it, it appears that people possess two distinct emotional selves, one that lives in the moment and one that lives in the abstract. By this account, dissociations between relatively on-line versus relatively retrospective emotion reports arise as a function of the distinct sources of self-knowledge accessed under different reporting conditions.

References

Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin, 93,* 203–231.

Arntz, A., van Eck, M., & Heijmans, M. (1990). Predictions of dental pain: The fear of any expected evil is worse than the evil itself. *Behaviour Research and Therapy, 28,* 29–42.

Bartlett, F. C. (1932). *Remembering.* Cambridge, England: Cambridge University Press.

Baumeister, R. F. (1993). Lying to yourself: The enigma of self-deception. In M. Lewis & C. Saarni (Eds.), *Lying and deception in everyday life* (pp. 166–183). New York: Guilford Press.

Begg, I. M., & Nicholson, S. (1994). Semantic and episodic relations are experimentally dissociable and stochastically independent. *Canadian Journal of Experimental Psychology, 48,* 399–417.

Blair, E., & Burton, S. (1987). Cognitive processes used by survey respondents to answer behavioral frequency questions. *Journal of Consumer Research, 14,* 280–288.

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

Brandstätter, H. (1983). Emotional responses to other persons in everyday life situations. *Journal of Personality and Social Psychology, 45,* 871–883.

Brown, K. W., & Moskowitz, D. S. (1997). Does unhappiness make you sick? The role of affect and neuroticism in the experience of common physical symptoms. *Journal of Personality and Social Psychology, 72,* 907–917.

Brown, N. R. (1995). Estimation strategies and the judgment of event frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1539–1553.

Burt, C. D. B., Watt, S. C., Mitchell, D. A., & Conway, M. A. (1998). Retrieving the sequences of autobiographical event components. *Applied Cognitive Psychology, 12,* 321–338.

Burton, S., & Blair, E. (1991). Task conditions, response formulation processes, and response accuracy for behavioral frequency questions in surveys. *Public Opinion Quarterly, 55,* 50–79.

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review, 107,* 261–288.

Cross, S. E., & Madson, L. (1997). Models of the self: Self-construal and gender. *Psychological Bulletin, 112,* 5–37.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience.* New York: Harper & Row.

Csikszentmihalyi, M., & LeFevre, J. (1989). Optimal experience in work and leisure. *Journal of Personality and Social Psychology, 56,* 815–822.

Cutler, S. E., Larsen, R. J., & Bunce, S. C. (1996). Repressive coping style and the experience and recall of emotion: A naturalistic study of daily affect. *Journal of Personality, 64,* 379–405.

Diener, E., & Larsen, R. J. (1984). Temporal stability and cross-situational consistency of affective, behavioral, and cognitive responses. *Journal of Personality and Social Psychology, 47,* 580–592.

Diener, E., Larsen, R. J., & Emmons, R. A. (1984, August). *Bias in mood recall in happy and unhappy persons.* Paper presented at the 92nd Annual Convention of the American Psychological Association, Toronto, Canada.

Dooling, D. J., & Christiaansen, R. E. (1977). Episodic and semantic aspects of memory for prose. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 428–436.

Dosher, B. A. (1984). Discriminating preexperimental (semantic) from learned (episodic) associations: A speed-accuracy study. *Cognitive Psychology, 16,* 519–555.

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology, 57,* 1082–1090.

Eagly, A. H., & Wood, W. (1991). Explaining sex differences in social behavior: A meta-analytic perspective. *Personality and Social Psychology Bulletin, 17,* 306–315.

Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related constructs. *Psychological Bulletin, 94,* 100–131.

Epstein, S. (1983). Aggregation and beyond: Some basic issues in the prediction of behavior. *Journal of Personality, 51,* 360–392.

Fazio, R. H. (1990). A practical guide to the use of response latency in social psychological research. In C. Hendrick & M. S. Clark (Eds.), *Review of personality and social psychology* (Vol. 11, pp. 74–97). Newbury Park, CA: Sage.

Feldman Barrett, L. (1997). The relationship among momentary emotion experience, personality descriptions, and retrospective ratings of emotion. *Personality and Social Psychology Bulletin, 23,* 1100–1110.

Feldman Barrett, L., Lane, R. D., Sechrest, L., & Schwartz, G. E. (2000). Sex differences in emotional awareness. *Personality and Social Psychology Bulletin, 26,* 1027–1035.

Feldman Barrett, L., Robin, L., Pietromonaco, P. R., & Eyssell, K. M. (1998). Are women the "more emotional" sex? Evidence from emotional experiences in social context. *Cognition and Emotion, 12,* 555–578.

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

Gasper, K., & Clore, G. L. (2000). Do you have to pay attention to your feelings to be influenced by them? *Personality and Social Psychology Bulletin, 26,* 698–711.

Hedges, S. M., Jandorf, L., & Stone, A. A. (1985). Meaning of daily mood assessments. *Journal of Personality and Social Psychology, 48,* 428–434.

Herr, P. M. (1986). Consequences of priming: Judgment and behavior. *Journal of Personality and Social Psychology, 51,* 1106–1115.

Herr, P. M., Sherman, S. J., & Fazio, R. H. (1983). On the consequences of priming: Assimilation and contrast effects. *Journal of Experimental Social Psychology, 19,* 323–340.

Herrmann, D. J., & Harwood, J. R. (1980). More evidence for the existence of separate semantic and episodic stores in long-term memory. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 467–478.

Josephs, R. A., Markus, H. R., & Tafarodi, R. W. (1992). Gender and self-esteem. *Journal of Personality and Social Psychology, 63,* 391–402.

Kahneman, D. (1999). Objective happiness. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 85–105). New York: Russell Sage Foundation.

Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better end. *Psychological Science, 4,* 401–405.

Klein, S. B., Babey, S. H., & Sherman, J. W. (1997). The functional independence of trait and behavioral self-knowledge: Methodological considerations and new empirical findings. *Social Cognition, 15,* 183–203.

Klein, S. B., Loftus, J., & Kihlstrom, J. F. (1996). Self-knowledge of an amnesic patient: Toward a neuropsychology of personality and social psychology. *Journal of Experimental Psychology: General, 125,* 250–260.

Klein, S. B., Loftus, J., Trafton, R. G., & Fuhrman, R. W. (1992). The use of exemplars and abstractions in trait judgments: A model of trait knowledge about the self and others. *Journal of Personality and Social Psychology, 63,* 739–753.

LaFrance, M., & Banaji, M. (1992). Towards a reconsideration of the gender-emotion relationship. In M. S. Clark (Ed.), *Review of personality and social psychology* (Vol. 14, pp. 178–201). Newbury Park, CA: Sage.

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.

Levine, L. J. (1997). Reconstructing memory for emotions. *Journal of Experimental Psychology: General, 126,* 165–177.

Liberman, N., & Trope, Y. (1998). The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. *Journal of Personality and Social Psychology, 75,* 5–18.

Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes, 65,* 272–292.

Lorch, R. F., & Myers, J. L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 149–157.

Marsh, H. W., & Yeung, A. S. (1998). Top-down, bottom-up, and horizontal models: The direction of causality in multidimensional, hierarchical self-concept models. *Journal of Personality and Social Psychology, 75,* 509–527.

Martin, L. L., & Achee, J. W. (1992). Beyond accessibility: The role of processing objective in judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 195–216). Hillsdale, NJ: Erlbaum.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102,* 419–437.

McFarland, C., & Ross, M. (1987). The relation between current impressions and memories of self and dating partners. *Personality and Social Psychology Bulletin, 13,* 228–238.

McFarland, C., Ross, M., & DeCourville, N. (1989). Women's theories of menstruation and biases in the recall of menstrual symptoms. *Journal of Personality and Social Psychology, 57,* 522–531.

McNamara, T. P., & Altarriba, J. (1988). Depth of spreading activation revisited: Semantic mediated priming occurs in lexical decisions. *Journal of Memory and Language, 27,* 545–559.

Menon, G. (1993). The effects of accessibility of information on judgments of behavioral frequencies. *Journal of Consumer Research, 20,* 431–440.

Menon, G., Raghubir, P., & Schwarz, N. (1995). Behavioral frequency judgments: An accessibility-diagnosticity framework. *Journal of Consumer Research, 22,* 212–228.

Minsky, M. (1975). A framework for representing knowledge. In P. H. Winston (Ed.), *The psychology of computer vision* (pp. 211–277). New York: McGraw-Hill.

Mitchell, T. R., Thompson, L., Peterson, E., & Cronk, R. (1997). Temporal adjustments in the evaluation of events: The "rosy view." *Journal of Experimental Social Psychology, 33,* 421–448.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale, NJ: Erlbaum.

Neely, J. H., & Durgunoglu, A. Y. (1985). Dissociative episodic and semantic priming effects in episodic recognition and lexical decision tasks. *Journal of Memory and Language, 24,* 466–489.

Olasov, B., & Jackson, J. (1987). Effects of expectancies on women's reports of moods during the menstrual cycle. *Psychosomatic Medicine, 49,* 65–78.

Olson, J. M., Roese, N. J., & Zanna, M. P. (1996). Expectancies. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 433–465). New York: Guilford Press.

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

Porter, L. S., Marco, C. A., Schwartz, J. E., Neale, J. M., Shiffman, S., & Stone, A. A. (2000). Gender differences in coping: A comparison of trait and momentary assessments. *Journal of Social and Clinical Psychology, 19,* 480–498.

Redelmeier, D. A., & Kahneman, D. (1996). Patients' memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures. *Pain, 66,* 3–8.

Robinson, M. D. (2002). [Are reports of on-line feelings "veridical"?]. Unpublished raw data, North Dakota State University.

Robinson, M. D., & Clore, G. L. (in press). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin.*

Robinson, M. D., & Johnson, J. T. (1997). Is it emotion or is it stress?: Gender stereotypes and the perception of subjective experience. *Sex Roles, 36,* 235–258.

Robinson, M. D., Johnson, J. T., & Shields, S. A. (1998). The gender heuristic and the database: Factors affecting the perception of gender-related differences in the experience and display of emotions. *Basic and Applied Social Psychology, 20,* 206–219.

Robinson, M. D., & Ryff, C. D. (1999). The role of self-deception in perceptions of past, present, and future happiness. *Personality and Social Psychology Bulletin, 25,* 611–622.

Ross, M. (1989). The relation of implicit theories to the construction of personal histories. *Psychological Review, 96,* 341–357.

Rubin, D. C., & Wetzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review, 103,* 734–760.

Ruble, D. N., & Brooks-Gunn, J. (1979). Menstrual symptoms: A social cognition analysis. *Journal of Behavioural Medicine, 2,* 171–194.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding.* Hillsdale, NJ: Erlbaum.

Schimmack, U. (1996). *The relationship between extraversion/neuroticism*

and positive/negative affect: A meta-analysis. Unpublished manuscript, Free University, Berlin, Germany.

Schimmack, U. (1997). *Frequency judgments of emotion: How accurate are they, and how are they made?* Unpublished dissertation, Free University, Berlin, Germany.

Schimmack, U., & Diener, E. (1997). Affect intensity: Separating intensity and frequency in repeatedly measured affect. *Journal of Personality and Social Psychology, 73,* 1313–1329.

Schwarz, N., & Bless, H. (1992). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 217–245). Hillsdale, NJ: Erlbaum.

Shelton, J. R., & Martin, R. C. (1992). How semantic is automatic semantic priming? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1191–1210.

Shields, S. A. (1987). Women, men, and the dilemma of emotions. In P. Shaver & C. Hendrick (Eds.), *Review of personality and social psychology* (Vol. 7, pp. 229–250). Newbury Park, CA: Sage.

Shields, S. A. (1991). Gender in the psychology of emotion: A selective research review. In K. T. Strongman (Ed.), *International review of studies on emotion* (Vol. 1, pp. 227–245). New York: Wiley.

Shields, S. A. (1995). The role of emotion beliefs and values in gender development. In N. Eisenberg (Ed.), *Review of personality and social psychology* (Vol. 15, pp. 212–232). Thousand Oaks, CA: Sage.

Shoben, E. J., Wescourt, K. T., & Smith, E. E. (1978). Sentence verification, sentence recognition, and the semantic–episodic distinction. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 304–317.

Slade, P. (1984). Premenstrual emotional changes in normal women: Fact or fiction? *Journal of Psychosomatic Research, 28,* 1–7.

Spence, J. T., Helmreich, R., & Stapp, J. (1975). Ratings of self and peers on sex role attributes and their relation to self-esteem and conceptions of masculinity and femininity. *Journal of Personality and Social Psychology, 32,* 29–39.

Stangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin, 111,* 42–61.

Stapel, D. A., Koomen, W., & van der Pligt, J. (1997). Categories of category accessibility: The impact of trait concept versus exemplar priming on person judgments. *Journal of Experimental Social Psychology, 33,* 47–76.

Stone, A. A., Hedges, S. M., Neale, J. M., & Satin, M. S. (1985). Prospective and cross-sectional mood reports offer no evidence of a "Blue Monday" phenomenon. *Journal of Personality and Social Psychology, 49,* 129–134.

Stone, A. A., Shiffman, S. S., & DeVries, M. W. (1999). Ecological momentary assessment. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 26–39). New York: Russell Sage Foundation.

Strack, F. (1992). The different routes to social judgments: Experiential versus informational strategies. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 249–276). Hillsdale, NJ: Erlbaum.

Sulin, R. A., & Dooling, D. J. (1974). Intrusion of thematic idea in retention of prose. *Journal of Experimental Psychology, 103,* 255–262.

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization–minimization hypothesis. *Psychological Bulletin, 110,* 67–85.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103,* 193–210.

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

Trafimow, D., Triandis, H. C., & Goto, S. G. (1991). Some tests of the distinction between the private self and the collective self. *Journal of Personality and Social Psychology, 60,* 649–655.

Tulving, E. (1984). Precis of elements of episodic memory. *Behavioral and Brain Sciences, 7,* 223–268.

Tulving, E. (1993). What is episodic memory? *Current Directions in Psychological Science, 2,* 67–70.

Watson, D. (2000). *Mood and temperament.* New York: Guilford Press.

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

Watson, D., & Walker, L. M. (1996). The long-term stability and predictive validity of trait measures of affect. *Journal of Personality and Social Psychology, 70,* 567–577.

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1997). Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness. *Psychological Bulletin, 121,* 331–354.

Widiger, T. A., & Settle, S. A. (1987). Broverman et al. revisited: An artifactual sex bias. *Journal of Personality and Social Psychology, 53,* 463–469.

Williams, J. E., & Bennett, S. M. (1975). The definition of sex stereotypes via the adjective check list. *Sex Roles, 1,* 327–337.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin, 116,* 117–142.

Winkielman, P., Knäuper, B., & Schwarz, N. (1998). Looking back in anger: Reference periods change the interpretation of emotion frequency questions. *Journal of Personality and Social Psychology, 75,* 719–728.

Received September 5, 2000
Revision received December 1, 2001
Accepted January 17, 2002 ■