The predictive value of daily vs. retrospective well-being judgments in relationship stability

Shigehiro Oishi [a,*], Helen W. Sullivan [b]

[a] Department of Psychology, University of Virginia, 102 Gilmer Hall, VA 22904-4400, USA
[b] Department of Psychology, University of Minnesota, USA

Abstract

The present study examined the role of daily and retrospective judgments of well-being and relationship satisfaction in relationship longevity. Participating couples completed a 14-day diary report of well-being and relationship satisfaction. After the daily diary survey, they evaluated the 14-day period. Participants also rated their global relationship satisfaction at that time. Retrospective judgments of daily well-being predicted later relationship status better than daily ratings of well-being did, whereas daily ratings of relationship domain satisfaction predicted later status better than retrospective judgments of daily relationship domain satisfaction did. Furthermore, *global* relationship satisfaction predicted later relationship status better than *daily* ratings of relationship domain satisfaction did. In general, global, summary judgments had a greater predictive value of future relationship status than did specific, daily judgments. Finally, synchronicity of daily fluctuations of well-being between partners predicted later relationship status.
© 2005 Elsevier Inc. All rights reserved.

Keywords: Subjective well-being; Well-being judgments; Close relationships

Global self-reports are believed to be vulnerable to various kinds of judgmental biases such as memory bias and current mood effect (Schwarz & Strack, 1999). Based on these findings, Kahneman (1999) argued that happiness should be measured by recording online hedonic tones, which he called *objective happiness,* rather than retrospective and global reports of happiness, which he called *subjective happiness.* Many researchers now recommend the use of experience sampling, event sampling, daily diary methods, and daily reconstruction method which reduce memory biases and other extraneous effects (e.g., Kahneman, Krueger, Schkade, Schwarz, & Stone, 2004; Stone, Shiffman, & DeVries, 1999). Although judgmental biases in well-being measures are well-documented, the utility of various measures of well-being is not yet well-known. As recognized by Kahneman (2000), we believe that judgmental biases themselves do not negate the importance of retrospective and global judgments in predicting future behavior. The present paper examines the utility of daily experiences vs. retrospective evaluation in an important behavioral outcome, namely, relationship stability.

Self-reports and judgments

Robinson and Clore (2002) classified self-reports into five types: online, retrospective, prospective, hypothetical, and time-inclusive reports. They argued that online reports are based mainly on experiential information and episodic memory, whereas retrospective, prospective, and time-inclusive reports are based on episodic memory (when available), situation-specific belief (e.g., "I get anxious during the mid-term period"), and identity-related belief (e.g., "I am an anxious person").

YJESP 1794

ARTICLE IN PRESS

No. of Pages 11; DTD=5.0.1

22 August 2005   Disk Used

Thomas (CE) / Prabakaran (TE)

2   *S. Oishi, H.W. Sullivan / Journal of Experimental Social Psychology xxx (2005) xxx–xxx*

Retrospective judgments are supposed to be based mainly on episodic memory if these judgments are made soon after the target event. As time passes, however, retrospective judgments become more and more dependent on semantic memory and beliefs. Similarly, when the time-frame is short (e.g., last 30 min), time-inclusive reports are based on experiences and episodic memory. However, as the time-frame becomes longer (e.g., last 2 weeks), they are based heavily on semantic memory and beliefs. Moreover, retrospective reports about a specific behavior (e.g., kindness to your partner) are different from reports about an abstract object (e.g., overall relationship quality) because an abstract object requires more integration in judgments (e.g., across times and domains) than does a specific target. In essence, the more integration judgments require, the more likely it is that these judgments become dependent on general beliefs.

Consistent with Robinson and Clore's (2002) belief model, Feldman Barrett (1997) showed that retrospective reports of the frequency of negative emotional experiences over the previous 90 days were predicted by neuroticism (which can be considered as an identity-belief about general negative emotional experiences), above and beyond the average of the actual negative emotional experiences (see Schimmack & Hartmann, 1997; for a case of repression; Christensen, Wood, & Feldman Barrett, 2003; for a case of self-esteem). Similarly, Robinson, Johnson, and Shields (1998) showed that although immediate emotional reactions to the experimental task were not different, men reported retrospectively having experienced pride and anger more than did women, and women reported retrospectively having experienced guilt and sympathy more than did men (see Oishi, 2002; Oishi & Diener, 2003 for similar cultural differences). These findings indicate that the time-inclusive retrospective judgments often depart from the actual frequency of emotional experiences, even when the retrospective judgments are made soon after the end of the experiences. A parallel difference also exists between specific and global reports. For instance, LaFrance and Banaji (1992) showed that gender differences in self-reported emotionality were more pronounced in global ratings than in specific ratings. Twenge and Campbell (2001) also found that self-esteem increased over time for the last 40 years when measured by global reports, whereas it did not change during the same period of time when measured by specific reports. Thus, even when the time-frame is identical, global ratings are likely to be more strongly influenced by general beliefs than are specific ratings.

In addition to the discrepancy between online and retrospective reports and between specific and global judgments, researchers have examined the utility of different types of judgments. In a series of ingenious studies, Kahneman, Fredrickson, Redelmeier, and colleagues have demonstrated that people's retrospective judgments and self-reported choice on a related task are often heavily influenced by the peak and the end experience (e.g., Fredrickson & Kahneman, 1993; Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993). For instance, Redelmeier and Kahneman (1996) demonstrated that retrospective reports about a colonoscopy were biased toward the peak and end experiences (i.e., the most painful moment and the pain in the end), and that their intention of repeating the same procedure was also predicted by these experiences, but not by the actual amount of time they were in pain. Recently, Wirtz, Kruger, Scollon, and Diener (2003) found that although the average online affective experience during spring break predicted participants' intention to repeat the vacation, the direct effect of the online affective experience was fully mediated by the retrospective judgment made at the end of the break.

There are three remaining issues in the previous research on the predictive value of retrospective judgments. First, unlike affective episodes previously examined (e.g., spring break), our daily affective experiences do not always have a clear beginning and end. We feel happy for someone and angry at someone with whom we have *ongoing* relationships (Berscheid & Reis, 1998). Thus, it is important to examine the predictive value of online experience vs. retrospective judgments in an ongoing, relationship context. Second, in the previous research the choice or decision was made soon after the episode ended and was hypothetical (e.g., "Would you take this same vacation over again?"). Thus, it is unclear whether participants in these experiments indeed repeated the same vacation or procedure when they encountered such a choice situation later. In the present research, we used relationship status 6 months after the assessment of daily well-being as the dependent variable to reduce such potential biases in the key decision task. Finally, because all the retrospective judgments in the previous research were global, summary judgments that required not only integration across times but also integration across domains (e.g., transportation, hotel), it was impossible to determine potential differences between specific and global retrospective judgments in predictive value. To this end, we asked participants to give retrospective reports on specific domains (e.g., "how satisfied are you with the way disagreement was resolved during the past 2 weeks?") as well as global targets (e.g., "how satisfied are you with the past 2 weeks?"). We were able to examine, therefore, the degree to which retrospective reports of specific relationship domains as well as global, summary reports about their lives during the 2-week period predicted the relationship status 6 months later.

Judgments of close relationships

In the close relationship literature, judgmental and memory processes have been extensively examined (see

YJESP 1794

22 August 2005   Disk Used

ARTICLE IN PRESS

No. of Pages 11; DTD=5.0.1

Thomas (CE) / Prabakaran (TE)

S. Oishi, H.W. Sullivan / Journal of Experimental Social Psychology xxx (2005) xxx–xxx   3

Berscheid & Reis, 1998, for review). For example, Murray and colleagues (e.g., Murray, Holmes, & Griffin, 1996, 2003) repeatedly showed that positive illusions about the partner are associated with high relationship satisfaction. Individuals who are satisfied with their relationship tend to view their romantic partner more favorably than the partner views himself or herself. There is a parallel phenomenon in memories of the partner's behaviors and the relationship. For instance, in a 4-year longitudinal study, Sprecher (1999) found that the couples who stayed in the same relationship throughout the study reported that their love, commitment, and satisfaction were increasing every year, although actual annual reports indicated otherwise (see also Frey & Karney, 2004; Karney & Frye, 2002; McFarland & Ross, 1987). Furthermore, beliefs about increases in love and commitment were associated with current feelings about the relationship and relationship stability. Spencher's findings highlight the critical role of beliefs about the relationship in relationship stability.

Robinson and Clore's (2002) belief model of self-reports suggests that retrospective reports about specific relationship behaviors are derived mainly from situation-specific beliefs (e.g., "I am witty when I am with my boyfriend"), whereas global retrospective reports are based on identity-beliefs (e.g., "My relationship is just OK"). Because an overall evaluation about the relationship and life is likely to be more relevant to the decision to stay together or break up than are specific evaluations, we hypothesized that global, summary judgments about the relationship would be a stronger predictor of the later relationship status than would specific retrospective judgments. Second, because many of the daily behaviors and feelings do not stay in long-term memory, we hypothesized that the average daily ratings would not be as strong a predictor of the later relationship status as would retrospective ratings.

Interdependence theory

In addition to testing the relative importance of daily vs. retrospective judgments, and specific vs. global judgments of well-being and relationship satisfaction, our second goal was to test Kelley et al.'s (1983) interdependence theory of close relationship in a novel way. High interdependence is characterized by frequent, strong, and diverse kinds of impact on each other for a long period of time. Berscheid, Snyder, and Omoto (1989) created a self-report scale that assesses the degree of interdependence. These researchers then successfully demonstrated that diversity and strength subscales predicted relationship dissolution 9 months later (see also Simpson, 1987). Yet another way of operationalizing the degree of interdependence between two people is to assess the degree to which one's daily well-being ratings

covary with their partner's daily well-being ratings (cf. Anderson, Keltner, & John, 2003). After all, if partners are highly interdependent, their daily well-being should fluctuate in a similar fashion. In other words, the higher the interdependence, the greater the covariation should be between two individuals' daily well-being ratings. Finally, the higher the degree of interdependence, the more stable the relationship should be.

In sum, the present research builds on and extends the previous research by testing a key theoretical issue in the well-being research in a daily, romantic relationship context, while simultaneously examining one of the most influential theories in close relationships using a new method.

Method

Participants

We recruited new dating couples via a student newspaper at the University of Minnesota. We defined a new dating couple as a couple who had been dating for less than 3 months. We recruited only new dating couples to minimize pre-existing between-couple differences in duration of the relationship, as it is known to be related to relationship quality and stability (e.g., Berscheid & Reis, 1998). Eighty-six heterosexual dating couples who met the criteria for participation responded to our newspaper advertisement. One-hundred and forty-two participants identified themselves as European American (82.6%), seven as Asian or Asian American (4.1%), three as African American, one as Hispanic American, four as other (15 did not specify). The average age was 20.72 years ($SD = 3.12$). Participants received $25 per person for their participation.

Procedure and materials

An experimenter met with each couple and explained the procedure of the study at the initial meeting. They were told that this study concerned daily satisfaction of newly dating couples. The participants completed a short questionnaire at the initial meeting (Time 1). We assessed the couple's global evaluations of relationship satisfaction by six items,[1] four of which came from Murray, Holmes, Dolderman, and Griffin (2000) 4-item relationship satisfaction scale (e.g., "I have a very strong relationship with my partner," "I am extremely happy with my current romantic relationship"). Participants rated them on a 9-point scale, $1 =$ not at all true to $9 =$ absolutely true ($\alpha = 0.82$). After the information ses-

---

[1] Two additional items were "I am very optimistic about the future of this relationship" and "I feel insecure about my relabation with my partner" (reversed item).

YJESP 1794

ARTICLE IN PRESS

22 August 2005   Disk Used

No. of Pages 11; DTD=5.0.1

Thomas (CE) / Prabakaran (TE)

4   *S. Oishi, H.W. Sullivan / Journal of Experimental Social Psychology xxx (2005) xxx–xxx*

sion, they were instructed to go to a designated website every night and complete a short daily survey individually for 14 consecutive days. The experimenter checked the database every morning and sent a reminder email to the participants who had not completed the daily survey the previous night. They were not allowed to make up a survey, except when they had no internet access (e.g., camping over the weekend). Compliance was excellent. On average, participants completed 13.47 daily surveys ($SD = 3.30$).

We assessed participants' daily well-being with two global items: "How was today?" on a 7-point scale, $1 =$ terrible to $7 =$ excellent, and "How satisfied are you with your life today?" on a 7-point scale, $1 =$ very dissatisfied to $7 =$ very satisfied ($\alpha = 0.86$[2]). In addition, they rated their satisfaction with 12 specific relationship domains everyday on the same 7-point scale: the partner, relationship with partner, sex life, partner's physical appearance, partner's social skills, partner's intelligence, support from partner, interactions with partner, time spent together with partner, the way disagreements were resolved, physical/ sexual intimacy with partner, and psychological intimacy with partner ($\alpha = 0.94$), and 6 non-relationship domains: academic performance, health, self, friends, professors/TAs, and weather ($\alpha = 0.70$).

We also assessed the perceptions of the self and the partner using a 10-item scale, which was culled from Murray et al.'s (1996) 22-item interpersonal qualities scale. Participants were asked to indicate how well each of the 10 specific behaviors described their partner that day when they were with him/her on a 9-point scale ranging from $1 =$ not at all characteristic to $9 =$ completely characteristic of him/her. Positive behaviors of the partner were kind/affectionate, witty/fun, responsive to your needs, sociable, and patient ($\alpha = 0.94$), and negative behaviors of the partner were critical/judgmental, controlling/dominant, moody/irritable, distant, and jealous ($\alpha = 0.87$). They were also asked to indicate how well each of the 10 behaviors described themselves when they were with their partner that day using the same 9-point scale ($\alpha = 0.94$ for positive, 0.87 for negative behaviors of the self).

When participants completed the 2-week daily surveys, they met with the experimenter and completed a short questionnaire (Time 2). Included in the Time 2 questionnaire were two global well-being items that correspond to the two daily well-being items: "How were the last two weeks?" and "How satisfied are you with your life during the past two weeks?" rated on the same 7-point scales ($\alpha = 0.83$). Participants also rated their satisfaction with the 12 specific relationship domains

($\alpha = 0.94$) and the six non-relationship domains ($\alpha = 0.69$) during the 2-week period. They also completed the perceptions of the self and the partner during the 2-week period on the same 10-item scales ($\alpha = 0.92$ for the positive behaviors of the self, .88 for the negative behaviors of the self, .88 for the positive behaviors of the partner, .88 for the negative behavior of the partner). We also assessed their global relationship satisfaction again using the same scale used at Time 1 ($\alpha = 0.89$).

Out of the original 86 couples, 7 couples (8.14%) did not complete the daily and/or Time 2 questionnaire. Six months later, participants were contacted via email and were asked their relationship status. Out of the 79 couples who completed the daily, Time 1, and Time 2 questionnaires, 65 couples indicated their relationship status 6 months after the completion of the Time 2 questionnaire. Out of the 65 couples, 48 couples (73.8%) were still in the same relationship and 17 (26.2%) were not in the same relationship 6 months later.

Results

Were retrospective judgments biased?

Table 1 shows average daily ratings and retrospective ratings for men and women separately, and paired t tests comparing average daily with retrospective ratings for men and women separately. With the exception of daily well-being (e.g., "how was today?"), which already involves aggregation across various domains, retrospective judgments were significantly higher than the average daily ratings for both men and women. Separate analyses for men and women are, however, suboptimal because dependency between men and women in our data (i.e., dyadic nature of our data structure) is not taken into account. Thus, we tested whether retrospective judgments made at the end of the daily study significantly deviated from the average of actual daily ratings for each variable with multilevel random coefficient models (MRCM) using hierarchical linear analysis (HLM 5.04, Raudenbush, Bryk, Cheong, & Congdon, 2001). Our data consist of three levels: within-individual (repeated measures, Level 1), between-individual, within-couple (Level 2), and between-couple (Level 3).[3] This analysis is conceptually equivalent of the paired t-test reported in Table 1, but also takes into account the within-couple dependency between men and women, and the nested nature of our data. Similar to

---

[2] Cronbach's α for daily measures reported in text were computed at the level of day across all the participants.

[3] Although in a typical multilevel random coefficient model daily ratings would be modeled at Level 1 and retrospective ratings would be modeled at Level 2, in this analysis both were modeled at Level 1. The main reason was that the central question here is concerned with the mean differences between the retrospective and the average daily judgments, not with the degree of covariation between them.

YJESP 1794	**ARTICLE IN PRESS**	No. of Pages 11; DTD=5.0.1
22 August 2005 Disk Used		Thomas (CE) / Prabakaran (TE)

S. Oishi, H.W. Sullivan / Journal of Experimental Social Psychology xxx (2005) xxx–xxx 5

Table 1
Means (*SD*s) of actual and recalled daily well-being, daily positive and negative behaviors of the self and the partner, and relationship satisfaction and global life satisfaction before and after the daily diary study for men and women

	Actual	Recalled	Paired *t*	*r*
Men				
Daily well-being	5.62 (0.56)	5.62 (0.87)	0.05	0.77**
Daily relation domain satisfaction	5.91 (0.69)	6.11 (0.82)	−3.77**	0.82**
Daily positive behavior of Self	6.62 (1.15)	7.01 (1.25)	−4.10**	0.80**
Daily negative behavior of self	2.17 (1.12)	2.47 (1.34)	−3.21**	0.82**
Daily positive behavior of partner	6.75 (1.07)	7.34 (1.04)	−6.45**	0.76**
Daily negative behavior of partner	2.17 (1.10)	2.38 (1.35)	−2.52*	0.87**
	Time 1	Time 2		
Global relationship satisfaction	7.89 (1.09)	7.79 (1.22)	0.73	0.52**
Satisfaction with life scale	26.58 (5.26)	27.78 (4.45)	−2.44*	0.70**
Women				
Daily well-being	5.55 (0.57)	5.57 (0.84)	−0.19	0.61**
Daily relation domain satisfaction	5.99 (0.58)	6.22 (0.70)	−4.99**	0.79**
Daily positive behavior of self	6.78 (1.17)	7.32 (1.07)	−5.40**	0.75**
Daily negative behavior of self	2.07 (1.09)	2.37 (1.33)	−4.00**	0.90**
Daily positive behavior of partner	7.09 (1.12)	7.59 (1.09)	−5.38**	0.77**
Daily negative behavior of partner	1.83 (0.94)	2.05 (1.14)	−2.84**	0.84**
	Time 1	Time 2		
Global relationship satisfaction	7.95 (0.96)	8.02 (1.08)	−0.58	0.60**
Satisfaction with life scale	27.21 (4.92)	27.95 (4.38)	−1.59	0.68**

Note. *N* = 80 couples.
Daily well-being was rated on a 1–7 point scale. Other scales were rated on a 1–9 point scale. *R* indicates a correlation coefficient between actual and recalled ratings.
  * *p* < .05.
  ** *p* < .01.

Barnett, Marshall, Raudenbush, and Brennan (1993) married couple study, we created two dummy codes at Level 1. Each average daily rating was given 1 and a retrospective rating was given 0 for the first dummy code, $D1$; an average daily rating was given 0 and a retrospective rating was given 1 for the second dummy code, $D2$. This full-rank dummy coding was used to simulate a multivariate regression, in which a pair of outcome variables can be tested simultaneously (Gonzalez & Griffin, 2002). The Level 1 model was as follows: $Y = \beta_1 * D_1 + \beta_2 * D_2 + r$. There is no intercept in Level 1 because in this model when D_1 and D_2 were both zero (although in reality there is no such case), "*y*" should be zero, too. Level 2 model was as follows: $\beta_1 = \gamma_{01} + u \cdot \beta_2 = \gamma_{02}$. In Level 2, γ_{01} indicates the mean of daily ratings, whereas γ_{02} indicates the mean of a retrospective rating. Because there were no gender differences in any of the variables in Table 1, we did not include sex at Level 2. We did not include any predictors at Level 3 because our hypotheses here were not concerned with between-couple differences, but instead concerned with the difference between γ_{01} and γ_{02}.

To test the difference between a daily average and a retrospective rating, we compared the baseline model that allows γ_{01} and γ_{02} to be different with the model that constrains γ_{01} and γ_{02} to be the same (i.e., daily average = retrospective rating). This is conceptually equivalent to the model comparison approach in structural equation modeling (SEM). Analogous to a χ^2 difference in SEM, a significant χ^2 coefficient here indicates that the constrained model is a worse fit than the baseline model. In the present context, the significant χ^2 coefficient indicates a significant mean difference between the average daily and retrospective ratings. When we examined the daily and retrospective ratings of positive behaviors of the self, using the aforementioned HLM model, we found a significant difference between them, $\chi^2(1) = 8.79$, $p < .01$. As seen in the relevant descriptive statistics in Table 1, when making retrospective judgments, our participants thought that they exhibited more positive behaviors (e.g.., kind) toward their partner than when they made actual daily ratings. Similarly, when they made retrospective judgments, our participants thought that their partner exhibited more positive behaviors toward them than when they rated their partner every day, $\chi^2(1) = 14.67$, $p < .01$. Our participants also reported a higher level of relationship domain satisfaction (e.g., sex, intimacy) at the time of retrospective judgments than during the corresponding daily period, $\chi^2(1) = 9.69$, $p < .01$. Interestingly, our participants also indicated at the time of retrospective judgments that they exhibited more negative behaviors (e.g., critical) toward their partner during the 2-week period than when they rated their behaviors every day,

YJESP 1794

ARTICLE IN PRESS

No. of Pages 11; DTD=5.0.1

22 August 2005   Disk Used

Thomas (CE) / Prabakaran (TE)

6   *S. Oishi, H.W. Sullivan / Journal of Experimental Social Psychology xxx (2005) xxx–xxx*

Table 2
The relative predictive value of daily vs. retrospective ratings on later relationship status: HLM Bernoulli model analyses

	Separate		Simultaneous	
	β (SE)	t value	β (SE)	t value
Daily well-being	0.79 (0.34)	2.35*	−0.66 (0.65)	−1.01
Retro daily well-being	0.88 (0.27)	3.22**	1.27 (0.49)	2.58*
Daily relation domain satisfaction	1.04 (0.35)	2.94**	1.14 (0.55)	2.08*
Retro relation domain satisfaction	0.75 (0.28)	2.66*	−0.05 (0.49)	−0.10
Daily positive behavior of self	0.33 (0.16)	2.15*	0.17 (0.25)	0.66
Retro positive behavior of self	0.33 (0.16)	2.10*	0.21 (0.25)	0.81
Daily negative behavior of self	−0.10 (0.15)	−0.70	−0.07 (0.31)	−0.23
Retro negative behavior of self	−0.08 (0.13)	−0.62	−0.03 (0.28)	−0.11
Daily positive behavior of partner	0.31 (0.16)	1.96*	0.09 (0.26)	0.33
Retro positive behavior of partner	0.36 (0.17)	2.17*	0.30 (0.27)	1.09
Daily negative behavior of partner	0.01 (0.17)	0.05	0.13 (0.34)	0.37
Retro negative behavior of partner	−0.02 (0.14)	−0.16	−0.12 (0.29)	−0.40
Daily relation domain satisfaction	1.04 (0.35)	2.94**	0.50 (0.43)	1.16
T2 global relationship satisfaction	0.74 (0.22)	3.33**	0.59 (0.26)	2.25*

Note. Relationship status was coded as follows: broke up = 0, stayed together = 1. "β" denotes an unstandardized regression coefficient. In the separate analyses, daily or retrospective rating was the only predictor. In the simultaneous analyses, both daily and retrospective ratings were predictors. Level 1 model was as follows: Status = $\beta_0 + \beta_1*$daily(male) $+ \beta_2*$daily(female) $+ \beta_3*$retro(male) $+ \beta_4*$retro(female) + error. Because constraining β_1 and β_2, and β_3 and β_4 to be the same resulted in no significant change in overall fit of the model, we constrained β_1 and β_2 to be the same as β_3 and β_4, respectively, in the analyses shown above.

* $p < .05$.
** $p < .01$.

$\chi^2(1) = 4.57$, $p < .05$. Consistent with previous research on social judgments (e.g., Kahneman, 1999; Ross, 1989), therefore, participants' retrospective judgments about relationship behaviors significantly deviated from the actual daily ratings, with the exception of daily negative behavior of the partner, $\chi^2(1) = 2.22$, $p = .13$, and daily well-being, $\chi^2(1) = 0.10$, *n.s.*[4]

What predicted relationship status 6 months later?

Next, we explored which daily ratings would predict the relationship status 6 months later. Because the dependent variable here is relationship status 6 months after the completion of the daily diary study, which is the same for both members of the couple, the data structure is quite different from the previous analyses. We used a Bernoulli model of HLM because the dependent variable here was dichotomous. We first predicted relationship status from the daily report by male participants and the daily report by female participants. Thus, Level 1 model was as follows: status = $\beta_0 + \beta_1*$daily (male) $+ \beta_2*$daily(female) + r. We tested sex differences by comparing the two models: the baseline model in which β_1 and β_2 were allowed to be different and the constrained model, in which β_1 and β_2 were constrained to be the same. We also repeated the same analyses using the ret-

rospective judgments. There were no sex differences in any variables. The first column of Table 2, labeled "Separate" describes the results from these analyses. When examined separately, both daily ratings and retrospective judgments of well-being, relationship domain satisfaction, positive behaviors of the self and the partner predicted later relationship status. In contrast, neither daily ratings nor retrospective judgments of negative behavior of the self and the partner predicted later relationship status.

We then tested the critical issue of the relative predictive values of daily vs. retrospective ratings. This time we included both daily and retrospective ratings at Level 1: status = $\beta_0 + \beta_1*$daily(male) $+ \beta_2*$daily(female) $+ \beta_3*$ retro (male) $+ \beta_4*$retro(female) + r. Again, we were able to constrain β_1 and β_2, and β_3 and β_4 to be the same without compromising the overall fit of the model (this was the case for all the analyses below). Consistent with our hypothesis, the retrospective judgment of daily well-being (i.e., global retrospective judgment) was a significant predictor of later relationship status, whereas the average daily well-being (i.e., global daily judgment) was not (see the "Simultaneous" column of Table 2). Given that the average daily well-being was a significant predictor of later relationship status when the retrospective report was not included in the equation, and that daily well-being was significantly associated with the retrospective report (see Table 1), the present findings indicate that the direct effect of the average daily well-being was mediated by the retrospective report of daily well-being.

Next, we tested the predictive value of daily relationship domain satisfaction and the retrospective judgment

---

[4] It should be noted that the significant departure from the actual daily ratings does not mean that retrospective judgments were inaccurate (see Funder & Colvin, 1991). In terms of the rank-order of individuals on the perceptions of the self and partner, there was a large degree of stability (see Table 1).

YJESP 1794

ARTICLE IN PRESS

22 August 2005   Disk Used

No. of Pages 11; DTD=5.0.1

Thomas (CE) / Prabakaran (TE)

S. Oishi, H.W. Sullivan / Journal of Experimental Social Psychology xxx (2005) xxx–xxx   7

of daily relationship domain satisfaction. The results were the complete opposite of the analysis with daily well-being. Namely, the average daily relationship domain satisfaction (i.e., daily, specific judgment) was a significant predictor of later relationship status, whereas the retrospective judgment (i.e., retrospective, specific judgment) was not (see Table 2). We then tested the predictive value of daily and retrospective ratings of positive behaviors of the self (i.e., daily, specific vs. retrospective, specific judgments). Although, respectively, each of these variables predicted later relationship status, neither of them was significant when both were simultaneously included. Similarly, although, respectively, each of the daily and retrospective rating of positive behavior of the partner predicted later relationship status, neither was significant when both were simultaneously included.

Because we also assessed global relationship satisfaction at the end of the daily diary study, we were able to examine the relative predictive value of daily relationship domain satisfaction (i.e., daily, specific judgment) vs. global relationship satisfaction (i.e., global, summary judgment), which presents an interesting test for the relative importance of specific vs. global summary judgments. Although daily relationship domain satisfaction was a strong predictor of later relationship status in the previous analysis (against the retrospective report of daily relationship domain satisfaction), daily relationship domain satisfaction was no longer a predictor of later relationship status when global relationship satisfaction at Time 2 was included (see Table 2). Thus, the direct effect of daily relationship domain satisfaction, which assessed daily satisfaction with specific relationship domains, on later relationship status was mediated by global relationship satisfaction, which is a summary judgment about the relationship.

Did peak and end experiences predict the later relationship status?

We then examined whether the peak (highest rating over 14 days) and the end (the final day rating) experiences predicted later relationship status. As seen in Table 3, peak daily well-being, relationship domain satisfaction, and positive behavior of the self predicted relationship status 6 months later. In contrast, none of the end reports predicted later relationship status. Because previous research showed the relative predictive power of the peak and the end experience over the average ratings (e.g., Fredrickson & Kahneman, 1993), we conducted the

Table 3
The relative predictive value of peak/end ratings vs. daily ratings on later relationship status: HLM Bernoulli model analyses

	Separate		Simultaneous	
	β (SE)	t value	β (SE)	t value
Peak daily well-being	0.91 (0.37)	2.48*	0.65 (0.55)	1.18
Daily well-being			0.32 (0.50)	0.64
Peak relation domain satisfaction	0.86 (0.36)	2.42*	−0.34 (0.63)	−0.54
Daily relation domain satisfaction	1.28 (0.48)	2.67*		
Peak positive behavior of self	0.39 (0.20)	1.97*	0.10 (0.35)	0.29
Daily positive behavior of self			0.26 (0.27)	0.96
Peak negative behavior of self	−0.03 (0.09)	−0.29	0.14 (0.20)	0.67
Daily negative behavior of self			−0.30 (0.33)	−0.91
Peak positive behavior of partner	0.28 (0.19)	1.49	−0.07 (0.32)	−0.23
Daily positive behavior of partner			0.36 (0.26)	1.39
Peak negative behavior of partner	−0.00 (0.09)	−0.00	−0.01 (0.21)	−0.06
Daily negative behavior of partner			0.03 (0.37)	0.07
End daily well-being	0.01 (0.18)	0.03	−0.36 (0.31)	−1.15
Daily well-being			0.98 (0.41)	2.37*
End relation domain satisfaction	0.15 (0.22)	0.68	−0.17 (0.33)	−0.52
Daily relation domain satisfaction	0.86 (0.60)	1.44		
End positive behavior of self	0.11 (0.15)	0.76	0.28 (0.25)	1.13
Daily positive behavior of self			−0.28 (0.33)	−0.86
End negative behavior of self	−0.03 (0.13)	−0.25	0.00 (0.30)	0.01
Daily negative behavior of self			−0.05 (0.39)	−0.13
End positive behavior of self	0.03 (0.14)	0.22	0.10 (0.19)	0.52
Daily positive behavior of self			−0.13 (0.22)	−0.58
End negative behavior of partner	−0.05 (0.13)	−0.36	−0.31 (0.31)	−0.99
Daily negative behavior of partner			0.40 (0.43)	0.93

Note. **$p < .01$. Relationship status was coded as follows: broke up = 0, stayed together = 1. "β" denotes an unstandardized regression coefficient. In the separate analyses, peak/end rating was the only predictor. In the simultaneous analyses, both peak/end and daily/retrospective ratings were predictors. Level 1 model was as follows: status = $\beta_0 + \beta_1*$peak/end(male) + β_2*peak/end(female) + β_3*daily(male) + β_4*daily(female) + error. Because constraining β_1 and β_2, and β_3 and β_4 to be the same resulted in no significant change in overall fit of the model, we constrained β_1 and β_2 to be the same as β_3 and β_4, respectively, in the analyses shown above. Results for daily ratings in separate analyses are reported in Table 2.
*$p < .05$.

YJESP 1794

22 August 2005   Disk Used

ARTICLE IN PRESS

No. of Pages 11; DTD=5.0.1

Thomas (CE) / Prabakaran (TE)

8   *S. Oishi, H.W. Sullivan / Journal of Experimental Social Psychology xxx (2005) xxx–xxx*

equivalent test of this. To our surprise, none of the peak reports predicted later relationship status, above and beyond the average daily ratings (see the "Simultaneous" column in Table 3). It should also be noted that retrospective ratings of daily well-being predicted later relationship status, above and beyond the end report, $\beta = 0.98$, $t = 2.37$, $p < .05$, indicating the predictive value of retrospective judgments of daily well-being was not due to the closeness in time to when later relationship status was measured.

Synchronicity in daily well-being: A test of the interdependence theory

Based on the interdependence theory of interpersonal relationships (Kelley et al., 1983), we hypothesized that the higher the interdependence, the greater the covariation should be between two individual's daily well-being. We tested this hypothesis using 2-level HLM[5]. At Level 1 (within-couple), day t well-being of each male participant was predicted from their female partner's day t well-being: day t male's well-being $= \beta_0 + \beta_1*$ female partner's day t well-being $+ r$. Female partner's daily well-being was centered around her own mean over 14 days. Thus, β_0 indicates male partner's daily well-being on a day when female partner's daily well-being was her average. β_1 indicates the degree of psychological interdependence because it reflects a change in male's daily well-being associated with female's change in her daily well-being. At Level 2 (between-couple), Level 1 intercept, β_0, and regression coefficient, β_1, were predicted from later relationship status: $\beta_0 = \gamma_{00} + \gamma_{01}*status + u$; $\beta_1 = \gamma_{10} + \gamma_{11}*status + u$. Status was coded as 0 if the couple broke up within 6 months and 1 if the couple was still together 6 moths later. This coding leads γ_{00} and γ_{10} to indicate average β_0 and β_1, respectively, for the couples who broke up within 6 months.

Among the couples who broke up within 6 months, the degree to which female's daily well-being was associated with her male partner's daily well-being was positive and marginally different from zero, $\gamma_{10} = 0.12$, $t = 1.78$, $p = .07$. Consistent with our hypothesis, this within-couple covariation of daily well-being was significantly larger among the couples who were still in the relationship 6 months later than among the couples who broke up within the following 6 months, $\gamma_{11} = 0.17$, $t = 2.03$, $p < .05$. On a day when a female partner's daily

well-being was 1 point higher than her typical day, male partner's daily well-being was 0.12 higher than his typical day among the couples who broke up within 6 months. On a day when a female partner's daily well-being was 1 point higher than her typical day, male partner's daily well-being was 0.29 (from 0.12 + 0.17) higher than his typical day among the couples who stayed together (see Fig. 1).

We repeated this analysis on other key variables. Somewhat surprisingly, the size of within-couple covariation did not differ between the two types of couples in terms of daily positive behaviors of the self, $\gamma_{11} = 0.06$, $t = 0.65$, n.s., daily negative behavior of the self, $\gamma_{11} = -0.05$, $t = -1.01$, n.s., daily relationship domain satisfaction, $\gamma_{11} = -0.03$, $t = -0.25$, n.s., or daily non-relationship domain satisfaction, $\gamma_{11} = 0.04$, $t = 0.47$, n.s.

Finally, we examined whether the degree to which a female participant's daily relationship domain satisfaction was associated with her partner's overall daily well-being was different across couples, depending on their later relationship status using the 2-level HLM analysis. Level 1 (within-couple) model was as follows: male's day t well-being $= \beta_0 + \beta_1*$ female's day t relationship domain satisfaction $+ r$. Level 1 intercept, β_0, and regression coefficient, β_1, were predicted from later relationship status at Level 2 (between-couple): $\beta_0 = \gamma_{00} + \gamma_{01}*status + u$; $\beta_1 = \gamma_{10} + \gamma_{11}*status + u$. Among the couples who broke up within 6 months, male participant's daily well-being was marginally associated with female partner's daily relationship domains satisfaction, $\gamma_{10} = 0.19$, $t = 1.89$, $p = .06$. More important, the degree of within-couple covariation was marginally larger among the
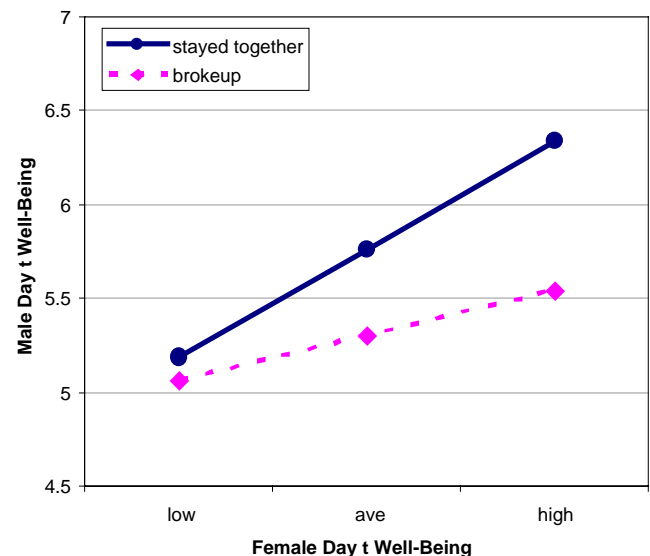
footnote

---

[5] In addition to the 2-level model discussed here, we also conducted a 3-level model, in which daily ratings were modeled at Level 1 (within-individual), sex was modeled at Level 2 (within-couple), and later relationship status was modeled at Level 3 (between-couple). The 3-level showed the same moderation effect of the relationship stability found in the 2-level analysis. For simplicity, we present the results from the 2-level model. Likewise, a 3-level analysis on the relationship domain satisfaction and daily well-being revealed the significant moderation effect of the relationship stability.

Fig. 1. Synchronicity of daily well-being among couples who stayed in the relationship and couples who broke up within 6 months. *Note.* "Low" in *X*-axis indicates two points below each individual's own daily average, whereas "high" indicates two points higher than each individual's own daily average.

YJESP 1794

22 August 2005   Disk Used

ARTICLE IN PRESS

No. of Pages 11; DTD=5.0.1

Thomas (CE) / Prabakaran (TE)

S. Oishi, H.W. Sullivan | Journal of Experimental Social Psychology xxx (2005) xxx–xxx   9

couples who stayed together than among the couples who broke up within 6 months, $\gamma_{11} = 0.22$, $t = 1.81$, $p = .07$. In other words, on a day when female's daily relationship domain satisfaction was 1 point higher than her typical day, her partner's daily well-being was 0.19 higher than his typical daily well-being among the couples who broke up within 6 months. In contrast, on a day when female's daily relationship domain satisfaction was 1 point higher than her typical day, her partner's daily well-being was 0.41 (from 0.19 + 0.22) higher than his typical day among the stable couples (see Fig. 2). In short, the female partner's daily relationship domain satisfaction appears to affect the male partner's overall daily well-being (and vice versa) more strongly among stable couples than among unstable couples.

Discussion

We examined the predictive value of daily vs. retrospective judgments in relationship stability using the 14-day diary method. This study revealed several intriguing phenomena. First, participants overestimated daily relationship domain satisfaction, positive behaviors of the self and the partner when making retrospective judgments. Somewhat surprisingly, participants also overestimated daily *negative* behaviors of the self and the partner (see Table 1). This suggests that participants did not simply view the past through rosy lenses (Mitchell, Thompson, Peterson, & Cronk, 1997), or were motivated to view the past in an egotistical way (Greenwald, 1980). Rather, this pattern of results signals judgment errors
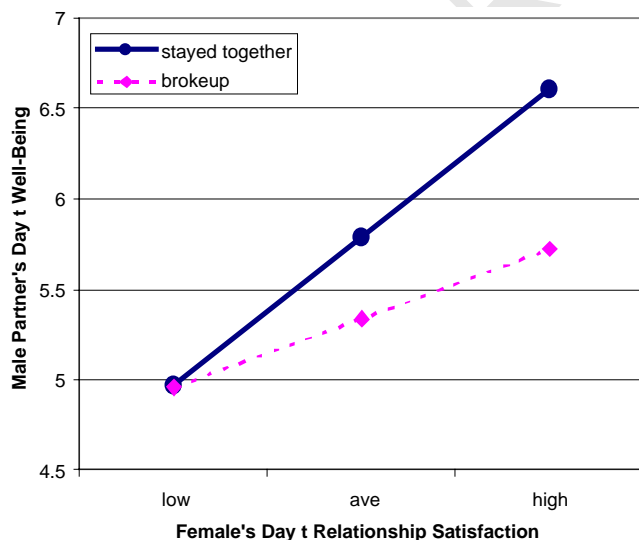

Fig. 2. The degree to which female's daily relationship domain satisfaction predicted their male partner's daily well-being among couples who stayed in the relationship and couples who broke up within 6 months. *Note.* "Low" in *X*-axis indicates two points below each individual's own daily average, whereas "high" indicates two points higher than each individual's own daily average.

when forming retrospective judgments on specific behaviors or domains over an extended period of time, which require aggregation of the relevant information across time. It is interesting to note that participants did not overestimate their daily well-being. This might be because daily well-being (e.g., "how satisfied are you with your life today?") requires integration of various life domains, whereas daily ratings of behaviors (e.g., "how kind were you toward your partner today?") and domains (e.g., "how satisfied are you with the way the disagreement was resolved today?") requires much less integration of information. Whereas daily ratings of specific behaviors and domains are based chiefly on episodic memory, daily well-being judgments might be based on semantic memory and identity-belief as much as episodic memory.

Second, we found that retrospective daily well-being (retrospective, global judgment) had a stronger predictive value for later relationship status than did actual daily well-being (daily, global judgment). In contrast, daily relationship domain satisfaction (daily, specific judgment) was a stronger predictor of later relationship status than were retrospective judgments (retrospective, specific judgment presumably based on a situation-specific belief). It is also interesting that global relationship satisfaction (global judgment presumably based on a general belief) predicted later status better than daily relationship domain satisfaction did (Table 2). Finally, in terms of daily behaviors, although daily and retrospective reports predicted later relationship status, respectively, together neither of them predicted it. This suggests that neither daily nor retrospective reports of specific relationship behaviors have a clear advantage in their predictive power. Taken together, it seems safe to conclude that it is not retrospective judgments per se that are superior to daily judgments in predicting future relationship status. Rather, it is global, summary judgments that are superior to specific, daily judgments. This is an important contribution of the present research because previous research (e.g., Wirtz et al., 2003) did not distinguish between retrospective judgments and global judgments. In terms of Robinson and Clore's (2002) belief model of self-reports, then, the present findings suggest that general identity-based beliefs predict important behavioral outcomes better than situation-specific beliefs do.

Third, as predicted, we found that within-couple covariation of daily well-being predicted relationship status 6 months later. Among the stable couples, on a day a female participant had a good day, her male partner tended to evaluate the day as good, as well; among the couples who broke up, on a day a female participant had a good day, her male partner sometimes reported that the day was not good. Thus, synchronicity of daily well-being is an important indicator of interdependence between partners, and a powerful predictor of future

YJESP 1794

ARTICLE IN PRESS

22 August 2005   Disk Used

No. of Pages 11; DTD=5.0.1

Thomas (CE) / Prabakaran (TE)

10   *S. Oishi, H.W. Sullivan / Journal of Experimental Social Psychology xxx (2005) xxx–xxx*

relationship status. It is interesting that synchronicity of daily behaviors and other daily ratings were not related to future relationship status. Thus, it appears that synchronicity of specific behavior (e.g., witty, critical) and specific relationship domain satisfaction (e.g., sex) is not necessary for relationship stability. Rather, it is synchronicity of daily well-being (or overall evaluation) that is critical for relationship stability.

In a related vein, we also found that a female participant's daily relationship domain satisfaction was more strongly associated with her partner's overall daily well-being among couples who stayed together than among couples who broke up within 6 months. These between-couple differences in within-couple covariation indicate that individuals in a stable relationship weighed their partner's relationship domain satisfaction more heavily in evaluating their overall daily well-being than did those in an unstable relationship (or that one partner's well-being affected the other partner's relationship domain satisfaction). This is consistent with the value-as-a-moderator model of well-being (Oishi, Diener, Suh, & Lucas, 1999), which assumes that well-being judgments are based on satisfaction in personally important domains. Extending previous research on this model (e.g., Oishi, Schimmack, & Colcombe, 2003), the present research showed that the basis of well-being judgments can predict an important future behavior such as relationship status.

Although the present findings are informative, they should be interpreted with the following limitations in mind. First, the data presented here were limited to young, heterosexual couples in an early stage of a romantic relationship. This might explain one unexpected finding: the negative behaviors of the self and the partner did not predict future relationship status. In contrast to our findings, Gottman and Levenson (1992) found that the frequency of negative behaviors (e.g., criticism) as well as positive behaviors during the interaction predicted marital status 4 years later.[6] It is plausible that at an early stage of a romantic relationship, feelings of love and passion are dominant forces, and therefore, positive aspects of the relationship outweigh negative aspects. It is important in the future to test whether the types of behaviors that predict future relationship status shift from positive to negative over the course of a relationship. Second, we did not find any end effect often found in the previous research (Kahneman et al., 1993). This might be largely because daily well-being and relationship satisfaction do not consist of a single affective episode. Recall that in Kahneman et al.'s (1993) experiment, the target experience was a single affective episode which had a clear beginning and end. Day 14 report was the arbitrary end of the daily survey. The end effect should be much stronger when there is a clear ending in the affective experience under study. Thus, our results do not negate the importance of peak and end experience in general. Finally, our participants were mostly European Americans. Previous research found cultural differences in online vs. retrospective judgments of well-being (e.g., Oishi, 2002; Oishi & Diener, 2003). Thus, the generalizability of the present findings needs to be tested in the future.

In conclusion, the present findings demonstrate the power of global, summary judgments over daily, specific ratings in predicting an important life outcome in an ongoing, relationship context. This suggests that other important life outcomes, such as group and organizational membership (e.g., job turn-over), might be better predicted by global satisfaction judgments rather than by daily or online hedonic measures. Despite recent critiques of global self-reports of well-being (Kahneman, 1999; Schwarz & Strack, 1999), the present research illustrates that global self-reports have an important utility in predicting future behaviors of great significance, at least on par with daily or specific reports. Equally important, the present study indicates that synchronicity of daily well-being within couples is a reliable predictor of relationship stability and an index of interdependence. As demonstrated here, the investigation of judgment processes in an ongoing relationship enriches the knowledge about social judgments, well-being, and close relationships.

Uncited reference

Reis and Gable (2000).

Acknowledgments

---

[6] When we conducted the analyses using the ratio of positive vs. negative behavior of the self and the partner, as opposed to analyzing them separately, these new variables did not predict later relationship status. This is largely because positive behaviors predicted later status, but negative behaviors did not. In other words, we did not find the effects found by Gottman and Levenson (1992), even when we analyzed the data using the positive vs. negative ratio.

References

Anderson, C., Keltner, D., & John, O. P. (2003). Emotional convergence between people over time. *Journal of Personality and Social Psychology, 84*, 1054–1068.

YJESP 1794

ARTICLE IN PRESS

No. of Pages 11; DTD=5.0.1

22 August 2005   Disk Used

Thomas (CE) / Prabakaran (TE)

S. Oishi, H.W. Sullivan / Journal of Experimental Social Psychology xxx (2005) xxx–xxx   11

Barnett, R. C., Marshall, N. L., Raudenbush, S. W., & Brennan, R. T. (1993). Gender and the relationship between job experiences and psychological distress: A study of dual-earner couples. *Journal of Personality and Social Psychology, 64*, 794–806.

Berscheid, E., & Reis, H. T. (1998). Attraction and close relationships. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 2, fourth ed, pp. 193–281). Boston, MA: McGraw-Hill.

Berscheid, E., Snyder, M., & Omoto, A. M. (1989). The relationship closeness inventory: Assessing the closeness of interpersonal relationships. *Journal of Personality and Social Psychology, 57*, 792–807.

Christensen, T. C., Wood, J. V., & Feldman Barrett, L. (2003). Remembering everyday events through the prism of self-esteem. *Personality and Social Psychology Bulletin, 29*, 51–62.

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

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

Frey, N. E., & Karney, B. R. (2004). Revision in memories of relationship development: Do biases persist over time. *Personal Relationships, 11*, 79–97.

Funder, D. C., & Colvin, C. R. (1991). Explorations in behavioral consistency: Properties of persons, situations, and behaviors. *Journal of Personality and Social Psychology, 60*, 773–794.

Gonzalez, R., & Griffin, D. (2002). Modeling the personality of dyads and groups. *Journal of Personality, 70*, 901–924.

Gottman, J. M., & Levenson, R. W. (1992). Marital processes predictive of later dissolution: Behavior, physiology, and health. *Journal of Personality and Social Psychology, 63*, 221–233.

Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist, 35*, 603–618.

Kahneman, D. (1999). Objective happiness. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-Being: The foundations of hedonic psychology* (pp. 3–25). New York: Russell Sage Foundation.

Kahneman, D. (2000). New challenges to the rationality assumption. In D. Kahneman & A. Tversky (Eds.), *Choices, values, and frames* (pp. 758–774). 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.

Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., & Stone, A. A. (2004). A survey method for characterizing daily life experience: The day reconstruction method. *Science, 306*, 1776–1780.

Karney, B. R., & Frye, N. E. (2002). "But we've been getting better lately": Comparing prospective and retrospective views of relationship development. *Journal of Personality and Social Psychology, 82*, 222–238.

Kelley, H. H., Berscheid, E., Christensen, A., Harvey, J. H., Huston, T. L., Levinger, G., et al. (1983). *Close relationships*. San Francisco: Freeman.

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.

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

Murray, S. L., Holmes, J. G., & Griffin, D. W. (1996). The benefits of positive illusions: Idealization and the construction of satisfaction in close relationships. *Journal of Personality and Social Psychology, 70*, 79–98.

Murray, S. L., Holmes, J. G., Dolderman, D., & Griffin, D. W. (2000). What the motivated mind sees: Comparing friends' perspectives to married partners' views of each other. *Journal of Experimental Social Psychology, 36*, 600–620.

Murray, S. L., Holmes, J. G., & Griffin, D. W. (2003). Reflections on the self-fulfilling effects of positive illusions. *Psychological Inquiry, 14*, 289–295.

Oishi, S. (2002). Experiencing and remembering of well-being: a cross-cultural analysis. *Personality and Social Psychology Bulletin, 28*, 1398–1406.

Oishi, S., & Diener, E. (2003). Culture and well-being: The cycle of action, evaluation and decision. *Personality and Social Psychology Bulletin, 29*, 939–949.

Oishi, S., Diener, E., Suh, E., & Lucas, R. E. (1999). Value as a moderator in subjective well-being. *Journal of Personality, 67*, 157–184.

Oishi, S., Schimmack, U., & Colcombe, S. (2003). The Contextual and systematic nature of life satisfaction judgments. *Journal of Experimental Social Psychology, 39*, 232–247.

Raudenbush, S., Bryk, A., Cheong, Y. F., & Congdon, R. (2001). *HLM5: hierarchical linear and nonlinear modeling*. Chicago, IL: Scientific Software International.

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.

Reis, H. T., & Gable, S. L. (2000). Event-sampling and other methods for studying everyday experience. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 190–222). New York: Cambridge University Press.

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin, 128*, 934–960.

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.

Schimmack, U., & Hartmann, K. (1997). Interindividual differences in the memory representation of emotions: Exploring the cognitive processes in repression. *Journal of Personality and Social Psychology, 73*, 1064–1079.

Schwarz, N., & Strack, F. (1999). Reports of subjective well-being: Judgmental processes and their methodological implications. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-Being: The foundations of hedonic psychology* (pp. 61–84). New York: Russell Sage.

Simpson, J. A. (1987). The dissolution of romantic relationships: Factors involved in relationship stability and emotional distress. *Journal of Personality and Social Psychology, 53*, 683–692.

Sprecher, S. (1999). "I love you more today than yesterday": Romantic partners' perceptions of changes in love and related affect over time. *Journal of Personality and Social Psychology, 76*, 46–53.

Stone, A. A., Shiffman, S. S., & DeVries, M. W. (1999). In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-Being: The foundations of hedonic psychology*. New York: Russell Sage Foundation.

Twenge, J. M., & Campbell, W. K. (2001). Age and birth cohort differences in self-esteem: A cross-temporal meta-analysis. *Personality and Social Psychology Bulletin, 5*, 321–344.

Wirtz, D., Kruger, J., Scollon, C. N., & Diener, E. (2003). What to do on spring break. The role of predicted, online, and remembered experience in future choice. *Psychological Science, 14*, 520–524.