



Published in final edited form as:

Soc Indic Res. 2017 January ; 130(1): 213–232. doi:10.1007/s11205-015-1172-6.

The Validity of the Day Reconstruction Method in the German Socio-Economic Panel Study

Ivana Anusic,

Michigan State University

Richard E. Lucas, and

Michigan State University

M. Brent Donnellan

Texas A & M University

Abstract

The Day Reconstruction Method is a useful tool for evaluating short-term changes in emotional experiences over a variety of daily situations. However, traditional method of collecting DRM data can be time-intensive for both researchers and participants. In this paper we provide evidence that a random-sampling approach to DRM assessment can provide useful data that are largely consistent with previous research that used the full version of the DRM. In a nationally representative sample of 2,303 people, we demonstrate that (1) there is variability in emotional ratings of episodes that replicates what has been found in prior studies, (2) correlations with global measures are typically small in magnitude ($< .30$), (3) correlations with personality are for the most part negligible, (4) correlations with global ratings of domain satisfaction are higher for domain-relevant situations, and (5) parents report more positive affect while providing care for their children when compared to other activities, and this effect can account for the observed differences in emotional experiences of parents and non-parents.

Although there are many ways to judge quality of life, one important perspective is a person's own subjective evaluation – a construct known as *subjective well-being (SWB)*. SWB has been the focus of considerable research, in part, because having a positive evaluation of their own lives is important to people (Diener, 2000). In addition, governments have become more and more interested in enhancing subjective perceptions of life quality in addition to traditionally used objective indicators such as the gross domestic product (Diener, Lucas, Schimmack, & Helliwell, 2009; Stiglitz, Sen, & Fitoussi, 2009). Indeed, some countries such as Canada, France, and the United Kingdom have developed programs that track national levels of well-being to help guide public policies (Samuel, 2009; Stratton, 2010; University of Waterloo, 2011). Given these important applications of SWB, it is important to assess the validity of measures of SWB.

Studies of SWB have typically used global measures that ask people to reflect on their lives as a whole and come up with a summary satisfaction score. This approach is useful when studying the effects of important life circumstances or personality characteristics on well-being. Alternatively, researchers may also be interested in studying short-term changes in emotional experiences, in particular the ways in which emotional experiences are linked to specific activities. For instance, the emotional experience of parents as they provide childcare has received much attention over the past years (see Nelson, Kushlev, English, Dunn, and Lyubomirsky). Global measures can provide interesting comparisons of general affective experiences and life satisfaction of parents and non-parents. However, these types of designs typically do not provide enough detail to evaluate whether, for example, parents are happier when spending time with their children than at other times (Nelson et al., 2013). For such questions, methods that assess people's emotional experiences repeatedly throughout the day are necessary.

One of the tools that have been developed to study activity-driven short-term changes in emotions is the Experience Sampling Method (ESM; Csikszentmihalyi & Larsen, 1987). In ESM participants are contacted several times a day at random intervals and asked to report on what they are doing and how they are feeling at that particular moment. ESM therefore provides researchers with real-time self-reports of emotional experiences. Although ESM is a useful tool for collecting momentary affective responses, it has important practical limitations. Participants are required to respond to a survey at multiple times throughout the day and this can be burdensome. In addition, it is sometimes impossible to answer the survey at the moment of contact (e.g., during an intense argument), leading to systematic exclusion of certain experiences that might contain important affective information. ESM can also be time intensive and costly for the researchers given the resources required to collect data from large samples of participants using specialized devices. These limitations make ESM impractical to use for most large scale studies of well-being such as those designed for informing public policies at the societal level. These factors might also make ESM impractical for many basic research studies where resources are also constrained.

In response to these limitations of the ESM, Kahneman et al. (2004) developed the Day Reconstruction Method (DRM) as a way of surveying people's daily activities and associated subjective emotional experiences. The DRM asks participants to reflect on their previous day and reconstruct the specific episodes that characterized their day. Specifically, people are asked about the duration of each episode, and what they did, who they were with, and how they felt during each episode. One advantage of the DRM over the ESM is that the DRM is completed in one sitting, resulting in fairly low burden on participants and researchers. Although the DRM is a less intensive method of tracking of situation-specific behaviors and emotions, research has shown that the DRM measures correspond closely to the data generated by ESM studies (Kahneman et al., 2004; Dockray et al., 2010).

Unique insights from research using the DRM

The DRM has become a useful tool for studying questions about time use and emotional experiences. For example, several studies have used the DRM to examine diurnal patterns of different emotions (Dockray et al., 2010; Kahneman et al., 2004; Stone et al., 2006). There is

a general agreement that positive emotions increase throughout the day. Negative emotions tend to decrease, but the strength of this effect is substantially smaller than for positive emotions (Dockray et al., 2010). Studies have also consistently shown that fatigue varies the most over the course of the day and follows a V-shaped pattern. Specifically, people's tiredness tends to decrease from the time they wake until about noon, and then it steadily increases until bedtime (Dockray et al., 2010; Kahneman et al., 2004; Stone et al., 2006).

Other studies have looked at differential effects of situations on experienced emotions for people with different personality and demographic characteristics. For example, Bakker et al. (2012) studied the association of doing work-related tasks in the evenings at home and found that it was more negatively associated with daily well-being of people higher in workaholism than those low in workaholism. In addition, the authors found that people high in workaholism benefited more from daily physical exercise than people low in workaholism. Oishi et al. (2011) found that location familiarity had different associations with positive affect for working and retired participants: retired participants reported more positive affect in familiar places, whereas working participants rated their positive affect higher in unfamiliar places. This study also showed cultural differences in the effects that interaction partners may have on our emotional experiences: Korean, but not American, participants reported more positive affect when interacting with a familiar rather than an unfamiliar person. Employment status has also been shown to play a role in experienced status of older adults: those who worked enjoyed their restful activities and weekends more than those who did not work (Tadic, Oerlemans, Bakker, & Veenhoven, 2013). In the same sample, Oerlemans, Bakker, and Veenhoven (2011) found that the association between time spent on social activities and happiness was stronger for people high in extraversion than those low in extraversion.

Studies employing the DRM can also detect contextual effects on experienced emotions. For instance, in a study of daily experiences of parents, Nelson et al. (2013) found that parents reported more positive affect and more meaning when taking care of their children than during other activities. Oerlemans et al. (2011) showed that restful activities moderated the effect of effortful activities on state happiness, such that people who spent more time doing effortful activities reported greater happiness if they also spent more time doing restful activities than if they spent less time resting. Thus, the DRM can meaningfully distinguish between different situations in people's days in a way that makes it possible to test theoretically-informed hypotheses about processes that underlie moment-to-moment changes in well-being.

Overall, it is clear that research using the DRM has generated some important insights into the nature of SWB. Experiential methods like the DRM allow researchers to shed light on unique questions and evaluate nuances and contextualized information. In short, using the DRM it is possible to collect detailed information about people's lives that simply cannot be gleaned from global measures of SWB.

Challenges of the DRM method

Although the DRM is a useful method for investigation of processes that underlie changes in SWB, it can be a challenging method to administer in large samples. For instance, the DRM is still quite time-intensive as it can take over an hour to complete (e.g., Kahneman et al., 2004). To deal with this limitation, Miret et al. (2012) suggested randomly assigning participants to complete the DRM for only a portion of the day (e.g., morning, afternoon, evening). They demonstrated that the pooled set of these short surveys has similar properties as the full version of the DRM. However, events rarely fall neatly into such time-bounded categories and episodes that span multiple categories may lead to confusion (e.g., people may be at work throughout the morning and afternoon). An alternative method that has been adopted by national surveys such as the American Time Use Survey (2014) and the Panel Study of Income Dynamics (2014) involves a random episode sampling approach. In this case, participants list all episodes in the initial day reconstruction, but only provide details such as emotional ratings for three randomly selected episodes. This random-sampling approach offers a non-biased sampling of events while shortening the total completion time and burden on the participant.

Indeed, this approach might facilitate additional large scale data collections for researchers interested in studying the connections between daily activities, affect, and well-being. As it stands, there are relatively few available existing resources for these kinds of investigations. Fortunately, the German Socio-Economic Panel (GSOEP) has recently expanded data collection from a subsample of its panel that is geared toward testing and inclusion of more innovative methods, including the DRM (Richter & Schupp, 2012; Wagner, Frick, & Schupp, 2007). One of the main goals of this study is to introduce the DRM assessment in this dataset and evaluate how the random episode sampling approach in the nationally representative sample of German residents compares to the original study in which a full DRM was limited to a sample of working women.

Current study

In this paper we evaluate whether using only three rated episodes per day can provide similar information as the full DRM using a subsample of the nationally representative study of German residents. With reference to the original DRM study by Kahneman et al. (2004) that used a large sample of working women, we evaluate emotional ratings across different situations and diurnal rhythms of rated emotions. We extend these results by evaluating associations of the DRM ratings with global measures of well-being and personality ratings.

In addition, we look at predictions of domain-specific satisfaction variables with DRM measures in situations relevant to that domain and other situations. Fisher (2000) found that positive and negative emotions experienced at work were related to global evaluations of work satisfaction. In addition, there is a certain level of spillover of daily work satisfaction and affect at work to emotional experiences home, such that people who are more satisfied with their workday tend to report more positive and less negative affect at home (Judge & Ilies, 2004). However, because work satisfaction reflects satisfaction with a narrow domain of work, these judgments should reflect experiences at work to a greater extent than

experiences outside of work. Because the GSOEP includes questions about both global satisfaction with domains and assesses situation-specific emotions in the DRM, we were able to test the prediction that work satisfaction would be especially important for emotional experiences at work.

Finally, past work on emotional experiences of parents has turned up conflicting results. In a study of working women, Kahneman et al. (2004) found that childcare was associated with some of the lowest levels of positive affect, and high levels of negative affect and fatigue. On the other hand, Nelson et al. (2013) found that parents enjoyed activities involving children suggests that childcare may not be associated with negative experiences to the degree that the original study suggests. Because our study uses a relatively large national sample, we are able to test connections between childcare and emotional experiences and make overall comparisons of emotional experiences of parents and non-parents.

Method

Participants

Participants were the members of the German Socioeconomic Panel Innovation Sample (GSOEP-IS) surveyed in 2012 (Richter & Schupp, 2012). GSOEP-IS is a subsample of the larger GSOEP study of German households that has been ongoing since 1984 (Wagner et al., 2007). One of the goals of GSOEP-IS is to include more innovative survey questions and methods than are available in the core GSOEP, including the DRM. Like the larger GSOEP, the innovation sample is obtained using scientific sampling methods to approximate nationally representative sampling of German households. In 2012, GSOEP-IS surveyed approximately 2,300 households, with every person at least 16 years of age in the household being invited to take part in the survey. The DRM was administered to a subsample of 2,303 people (52% women) from 1,415 households, in-person by a trained interviewer. The average age of this sample was 51.8 years, $SD = 18.0$ years, range = 17 to 95 years.

Measures

DRM.—For the DRM portion of the survey, the participants were first asked to reconstruct their previous day by breaking it up into episodes. For each episode they specified starting and ending times, what they were doing, and whether the episode was pleasant or not pleasant. From this pool of episodes, three were randomly selected for detailed assessment. In particular, participants rated where they were (at home, at work, or somewhere else), who they were with, and how they felt. Options for episode categories, which included a broader range of activities than the original Kahneman et al. (2004) study, and interaction partners are reported in Table 1. On average participants took 12 minutes ($SD = 5$ minutes) to complete the DRM. The average number of listed episodes was 11.6 episodes ($SD = 4.0$), and the mean episode duration was 78 minutes ($SD = 103$ minutes).

For the three selected episodes, participants rated how strongly they experienced each of the following on a scale of 1 (not at all) to 7 (very much): happiness, anger, frustration, fatigue, mourning, worries, pain, enthusiasm, satisfaction, boredom, loneliness, stress, and a deeper meaning. We recoded the scale to range from 0 to 6 to make it comparable to Kahneman et

al., (2004). We created composites for positive affect (combining happiness, enthusiasm, and satisfaction; $\alpha = .81$) and negative affect (anger, frustration, mourning, worries, and stress; $\alpha = .82$). For each person we computed duration-weighted aggregate of positive affect, negative affect, fatigue, meaning, pain, boredom, and loneliness from the three rated episodes.

Life satisfaction.

Life satisfaction was assessed with a single item that asked participants “How satisfied are you with your life, all things considered.” Participants made ratings on an 11-point scale ranging from 0 (completely dissatisfied) to 10 (completely satisfied).

Global affect.

Participants were reported frequency of experiencing a number of feelings over the 4 weeks prior. The response scale ranged from 1 to 5 (Very rarely to Very often). We combined items that asked about experience of anger, worry, sadness, feeling rushed or pressed for time, and feeling down and gloomy into a single trait negative affect variable. We combined items that asked about frequency of feeling energetic, and calm and relaxed (reverse-scored) into a single variable we called energy. Additional items asked people to rate frequency of feeling happy, and frequency of feeling severe physical pain.

Domain satisfaction.

Participants also rated their satisfaction with several life domains. In this paper we focus on satisfaction with sleep and work. Ratings were made on an 11-point scale ranging from 0 (completely dissatisfied) to 10 (completely satisfied).

Personality.

The Big Five personality traits were assessed with a 15-item Big Five Inventory (BFI-S; Donnellan & Lucas, 2008; Hahn, & Gottschling, 2012). Participants rated the extent to which a series of statements applied to them on a 7-point scale (1 = does not apply to me at all; 7 = applies to me perfectly)¹ Average inter-item correlations for the Big Five scales were comparable to what was found using full 44-item BFI scales (Srivastava, Johan, Gosling, & Potter, 2003): .27 (vs. .30) for agreeableness, .36 (vs. .34) for conscientiousness, .40 (vs. .43) for extraversion, .35 (vs. .40) for neuroticism, and .36 (vs. .29) for openness to experience.

Results

Means

Table 1 shows means of rated emotions by activity, interaction partner, and place. Rank ordering of activities was generally consistent with the previous work using the DRM. To quantify the degree of similarity between ours and Kahneman et al.’s (2004) original study

¹The items started with the stem “I see myself as someone who...” and were as follows. Agreeableness: “Am sometimes too coarse with others”, “Able to forgive”, and “Friendly with others”; Conscientiousness: “Thorough worker”, “Tend to be lazy”, and “Carry out tasks efficiently”; Extraversion: “Am communicative”, “Am sociable”, and “Reserved”; Neuroticism: “Worry a lot”, “Somewhat nervous”, and “Deal well with stress”; Openness to experience: “Am original”, “Value artistic experiences”, and “Have lively imagination.”

we computed profile correlations of ratings of positive affect ($r = .74$), negative affect ($r = .84$), and fatigue ($r = .70$).² Replicating the findings of Kahneman et al. (2004), we found that positive affect dominated most activities. There were only 5% of experiences in which positive was absent, whereas negative affect was absent in 45% of rated episodes. Consistent with the original study, praying/working/meditation, intimate relations, and socializing with friends were among highest scoring activities for PA. We found that participating in sports was the activity associated with highest levels of PA; however this option not available in Kahneman et al.'s (2004) study. We should also note only few people gave emotion ratings for praying/working/meditation ($N = 14$) and intimate relations ($N = 4$), so ratings of these activities should be interpreted with caution. People reported some of the lowest positive affect for housework and commuting in both studies. The only activity with lower PA in this study was attending a doctor appointment, an option not included in the original study.

Negative affect was consistently highest for working/studying across both this and the original study. Using the phone, computer/internet, and shopping were also among the activities with highest reported NA. Activities that are generally done alone such as resting/napping, relaxing, and watching TV, tended to be associated with high ratings of fatigue, boredom, and loneliness in our study.

Ratings of pain were the highest for activities that typically involve responses to the experience of pain (doctor appointment and resting/napping) and for physical activities that may aggravate existing pain (sports, housework, and gardening). Meaning was highest for praying/worship/meditation, activities that involve effort and accomplishment (sports, working/studying, and gardening), and activities that involved interacting with others (taking care of one's children and socializing with friends).

There were also some notable differences across the studies. First, Kahneman et al. (2004) found that caring for children was among the least pleasurable activities. However, in our study it was one of the most pleasurable activities – it ranked relatively highly for positive affect and meaning, and 98% of people considered it an overall pleasant activity. On the other hand, caring for children also ranked relatively highly for ratings of negative affect and fatigue. A possible explanation is that although parents report feeling tired and experiencing relatively high levels of negative affect when taking care of their children, they also find meaning and experience high positive affect during these interactions. Another explanation is that some parents experience primarily positive emotions, whereas other parents experience primarily negative emotions while providing childcare. Further analyses of our data support the latter explanation: the correlation between positive and negative affect while taking care of children was small and negative ($r = -.11$) suggesting that parents generally do not tend to feel positive and negative emotions simultaneously.

Second, ratings of all emotions were overall lower in this study than Kahneman et al. (2004). For example, duration-weighted mean PA was 2.91 in our study but 3.89 in the original

²Profile correlations were computed on the subset of activities that overlapped across the studies. Exercise was excluded from Kahneman et al.'s study. From our study we excluded activities "caring for pets", "doctor appointment", "gardening", "getting ready", "reading", "sports", and "other. We used the average of "commuting for work" and "commuting for leisure" from our study to compare to the "commuting" activity in the original paper.

study. Ratings of negative affect and fatigue were also lower in our study (0.68 and 1.83 vs. 0.84 and 2.90). These differences could be due to differences in samples – our sample was approximately representative of German residents, whereas the original study used a sample of working women from the United States. We conducted additional analyses restricting our sample to women employed full- or part-time ($N = 496$), and obtained similar results as when using the full sample. In general, we found only very few and inconsistent gender differences in emotional ratings of episodes (see Table 1). Most notably, women reported higher overall levels of fatigue, but this difference was limited to only a couple of activities.

Finally, rates of reporting activities were overall smaller in our sample. For example, relaxing and resting/napping were reported by 30% and 29% of people in our study; in the original study 77% and 43% of people reported these activities in the DRM. Seventeen percent of people in our study, but 65% of women in Kahneman et al.'s (2004) study reported socializing with friends during the previous day. Proportion of people reporting praying/worship/meditation and intimate relations were negligible in our study (2% and 1%), but much more common in the original paper (23% and 11%). In some cases, such as for activities involving intimate relations, the discrepancies are possibly due to differences in data collection methodologies: data in this study were collected in face-to-face interviews, whereas the original study used a more anonymous paper-and-pencil questionnaire method. People are likely to underreport sensitive behaviors in a personal interview due to embarrassment (Tourangeau & Yan, 2007). Reasons for discrepancies are less clear for other activities, especially considering that there are also activities that were reported at the same rate (e.g., eating, watching TV, shopping).

Diurnal rhythms

To evaluate diurnal rhythms of rated motions we plotted means of emotional ratings over the course of the day in Figure 1. Kahneman et al. (2004) focused primarily on fatigue. In our data, fatigue showed most pronounced changes throughout the day. Consistent with the original study's findings we found a decrease and then an increase in fatigue over the course of the day, with the lowest levels occurring just before noon (although our pattern looked more like a W than V-shaped pattern). Kahneman et al. (2004) also demonstrated that fatigue levels differed across certain groups in meaningful ways. For example, they showed that people who slept less were more tired throughout the day compared to those who slept more. Information about the amount of sleep in the previous night was not available in our study; however we had ratings of overall sleep satisfaction. We plotted diurnal patterns of tiredness separately for those who rated their sleep satisfaction as 7 or less and those who rated it as 8 or more (median of sleep satisfaction = 7). Figure 2 shows these results. Although the difference between groups was somewhat smaller than in the original study, we found the same pattern – those who were less satisfied with sleep reported feeling more fatigue throughout the day. We found less consistency across studies for differences in fatigue across age groups. As Kahneman et al. (2004), we found that the youngest group started their day most fatigued and that the oldest group started the day least fatigued. However, in the original study the groups converged in the early afternoon whereas in our study that convergence was much less evident. For example, the oldest group was consistently the least fatigued group throughout the day.

Other variables showed much less variability over the course of the day. Most change was evident for positive affect and meaning, both of which increased throughout the day. However, these changes were relatively small: positive affect changed by 0.70 points and meaning changed by 0.42 points. Negative affect, loneliness, pain, and boredom declined only slightly throughout the day, with changes ranging from 0.12 to 0.30. These findings are consistent with previous work that has found more diurnal variability in positive emotions than in negative emotions (Clark, Watson, & Leeka, 1989; Stone et al., 2006).

Correlations with global measures

There is much interest in how momentary and global measures of well-being compare to each other both conceptually and empirically (Diener & Tay, 2014). Traditionally, subjective well-being is assessed using global measures (e.g., life satisfaction, global recall of affect). However, some researchers have called to define subjective well-being as a sum of moment-to-moment emotional experiences (e.g., Kahneman, 1999). In their review of the DRM, Diener and Tay (2014) have noted a need for more direct comparisons of the two types of measures. In our study we were able to do such comparisons because the GSOEP-IS includes assessment of global life satisfaction and recall of affect over the past month in addition to the DRM.

Table 2 shows correlations of duration-weighted aggregates of emotion ratings in the DRM and global measures of well-being. Frequent experience of positive affect throughout the day (i.e., DRM PA) was associated with higher life satisfaction, higher levels of trait happiness, and more energy. Both negative affect and fatigue reported in the DRM had notable correlations ($> .30$) with frequency of experience of negative affect over the past month. However, these associations were notably lower than what was obtained in large samples of college students by Anusic, Lucas, and Donnellan (2015) who found that DRM positive affect correlated on average at .32 (vs. .21 in this study) with a single-item life satisfaction measure and .54 with a trait measure of positive affect (vs. .23 for trait happiness and .21 with trait energy in this study). The association between DRM and trait measures of negative affect was also lower compared to past research that found an average correlation of .58 (vs. .33 in this study) between these two variables (Anusic et al., 2015). It is possible that differences between studies may reflect sample or age differences (e.g., college students may use their daily experiences to inform their overall life satisfaction ratings to a greater degree than the general population). More likely, these discrepancies may reflect higher degree of method effects in Anusic et al.'s (2015) study in which participants completed both the DRM and global measures fairly close together in a study designed to test different approaches to well-being assessment. In contrast, the design of the GSOEP may minimize demand characteristics and other method biases because it includes hundreds of questions that are not as clearly linked with one another. Indeed, once method effects were minimized by using self-rated DRM and informant-rated global well-being, correlations between the two types of measures were reduced in the Anusic et al. study.

Correlations of DRM negative affect and fatigue with chronic experience of pain were low ($r = .09$), in contrast to findings from previous studies that have found robust associations between trait levels of negative affect and pain (e.g., Robinson & Riley, 1998). The strongest

correlation of pain reported in DRM episodes was with chronic pain experienced in the month prior to the study ($r = .50$). In addition, people who reported most pain in their daily activities also reported lower life satisfaction and chronic energy levels, and higher tendency to experience negative affect in general. People who reported more loneliness in their daily episodes tended to be less satisfied with their lives and report less trait happiness.

Correlations with personality variables

Past research has suggested that people's personality dispositions tend to produce stability in their daily emotional experiences (e.g., Gray, 1987; Larsen & Ketelaar, 1991; Watson & Clark, 1984). However, the correlations between DRM-rated affect and the Big Five personality dimensions in our study tended to be low, generally lower in magnitude than .10 (see Table 3). The strongest relationship was observed between neuroticism and pain rated in DRM episodes, consistent with previous suggestions that neuroticism may predispose people to be more sensitive to bodily sensations and pain (Goubert, Crombez, & Van Damme, 2004). Overall, correlations with personality and DRM measures were smaller in magnitude than has been observed in previous research that found correlations of .46 between DRM negative affect and neuroticism, and .26 and $-.32$ between DRM positive affect and neuroticism and extraversion, respectively (Anusic et al., 2015). However, Anusic et al. obtained their correlations using a longitudinal latent variable approach with college studies that could isolate variance captured by the DRM that was stable over one month. Another reason for the discrepancy between the studies is that personality was assessed three years prior to the DRM measures in the GSOEP. Thus the obtained correlations essentially reflect the effect of stable aspects of personality on situation-specific affect. Regardless, the current study suggests that such effects are small, and that more temporally proximal features of one's personality are more strongly linked to situational affect.

Differential predictions of satisfaction with domains

We also investigated whether global measures could differentially predict DRM ratings in theoretically meaningful ways. Specifically, we evaluated the relationship between a global measure of work satisfaction and affect reported during activities and work and compared this to the relationship of global work satisfaction and affect during activities outside of the work. Because work is a specific type of situation with a range of activities that are specific to the work environment, we expected that global ratings of work satisfaction should predict emotional experiences at work but not outside of work.

To address these questions, we conducted nonlinear mixed effects analyses predicting DRM ratings from work satisfaction, activity location (work vs. not at work), and their interaction. We also allowed random variance in intercepts to capture between-person variability in experienced emotion. The results are shown in Table 4. During activities at work people reported experiencing less PA, and more NA, pain, and boredom than during other activities. The effect of work satisfaction on affect ratings was moderated by activity location. Although those who were more satisfied with their work reported slightly more PA, and slightly less NA, fatigue, pain, boredom, and loneliness during non-work activities, these associations were even more pronounced during activities at work. To provide an idea of the strength of this effect, Table 4 also shows correlations between work satisfaction and DRM

ratings for work and non-work activities. The average magnitude of correlation for activities at work was .27; for activities outside of work it was .07. Thus, global satisfaction with the domain of work has particularly strong associations with emotional experiences in situations relevant to the work domain.

Emotional experiences of parents

The relative ordering of activities involving caring for children in our study differed notably from what has been found by Kahneman et al. (2004). Our findings were consistent with suggestions of Nelson et al. (2013) that parents find spending time with their children enjoyable. However, it is also possible that whereas spending time with one's children leads to positive emotional experiences, providing childcare does not. In our study we were able to test this prediction.

We conducted a series of multilevel models predicting positive affect from various variables and controlling for age. The results of these models can be found in Table 5. First, we replicated Nelson et al.'s (2013) finding that parents experience more positive affect when spending time with their children than during other activities, and that there were no significant differences in this effect between women and men. A possible reason for this finding is that some activities that do not involve children are day-to-day activities that are not necessarily pleasant, but that people do not think of when they consider whether they are happy when spending time with their kids (e.g., work). In order to explore this idea we restricted our selection of activities that did not involve children, and found that even when we left out activities at work or those that involved getting ready and commuting for work. The effect of children as companions remained: parents experienced more positive affect when with their kids than at other times. Another idea we explored was that parents may be happier when spending time with kids but not while providing childcare. In contrast with this prediction, we found that parents reported more positive affect for activities that were classified as childcare than for other activities (the top 5 other activities were: eating, cooking, watching TV, housework, and getting ready). Overall, parents reported more positive affect than non-parents, but this difference disappeared when we compared parents' ratings of activities that did not involve children. The difference between experienced positive affect between parents and non-parents appears to be due to parents spending time with their children.

Discussion

The main goal of our study was to test the viability of using a random sampling DRM in a nationally representative long-time running study of German residents. We did this by comparing our results to the results of the original DRM study by Kahneman et al. (2004) based on a sample of working women. We advanced previous DRM studies by demonstrating differential relationships between emotional experiences and work satisfaction for activities at different locales (work vs. other), and by testing additional predictions about the relationship between emotional experiences and parental status.

Despite differences in methodology (i.e., interview vs. paper and pencil questionnaire), sample characteristics (representative vs. working women), DRM method (random sampling

of episodes vs. all episodes), the results of this were generally similar to the results of Kahneman et al.'s (2004) original study. This increases our confidence in the validity of the random-sampling DRM. However, these results also differed in some important ways from previous studies, and these differences raise important questions about cross-cultural differences and the generalizability of theoretically important results from Kahneman et al.'s original study.

This study, like the original study, revealed that activities that involved sport participation, praying/worship/meditation, and socialization with friends tended to be associated with the highest positive affect and meaning, and lowest negative affect, loneliness, and boredom. The rank-ordering of situations based on their emotional ratings was generally consistent across the studies, giving confidence to the idea that people are to some degree consistent in how they respond to particular situations (Rauthmann et al., 2014).

The highest rating of negative affect was obtained for ratings of activities that involved working, commuting for work, and using the phone. Interestingly, people also reported relatively high levels of deep meaning during work activities. Thus, although people tended to experience relatively high levels of negative emotions at work, they were also able to find meaning in their work. In contrast, feelings of loneliness, boredom, and fatigue were most evident during activities that involved relaxing, resting/napping, and watching TV, and these activities were also associated with relatively low levels of meaning.

The patterns of diurnal rhythms in our study were generally consistent with other studies despite our use of random sampling of episodes. We found that fatigue showed most variation throughout the day, and replicated the pattern of decrease in fatigue from waking to mid-day, followed by increase until bedtime that has been found in other studies (Dockray et al., 2010; Kahneman et al., 2004; Stone et al., 2006). We also found meaningful patterns of between-person variation in fatigue that reflects people's overall satisfaction with sleep. Those who were more satisfied with sleep reported less fatigue throughout the day compared to those who were less satisfied. We also found that people who were 50 years or older tended to report less fatigue over the course of the day, compared to younger age groups. This finding partially replicates Kahneman et al.'s (2004) study that found that fatigue levels of different age groups converge in the afternoon and remain at similar levels for the remaining of the day.

Although scholars have suggested that daily aggregates of DRM should capture overall quality of life, we found only modest correlations between the two types of measures. Replicating findings of a previous study, DRM-rated positive affect correlated most strongly with global measures of life satisfaction, happiness, and energy, whereas DRM-rated negative affect correlated most strongly with global ratings of negative affect (Anusic et al., 2015). However, the observed correlations tended to be modest, with highest correlations falling in the .2 to .3 range. These results are consistent with the idea that the DRM and global measures capture different aspects of well-being (Tay, Chan, & Diener, 2014).

Overall, the two measurement methods converged the most strongly for ratings of pain ($r = .50$). People experienced more pain over the previous month also reported feeling pain during

the daily activities reported in the DRM. Daily ratings of pain were also associated with lower life satisfaction and energy, and higher negative affect. Experience of pain plays an important role in both subjective judgments of life quality and objective outcomes (McNamee & Mendolia, 2014). For fatigue, in addition to its association with lower energy levels, we found that it was associated with higher negative affect. In contrast, loneliness was primarily associated with lower life satisfaction and happiness ratings.

In contrast to theoretically meaningful correlations between DRM and global measures of well-being, we found very low correlations between daily aggregates of DRM affect and personality traits. A possible reason for such low correlations is that personality was assessed three years prior to the DRM. Alternatively, these results suggest that broad personality factors have low predictive utility for specific emotions from a single situation, although they may predict consistent patterns of thoughts, feelings, and behaviors (e.g., Epstein, 1979).

Despite low correlations with some of the criterion variables, we demonstrated that the DRM assessed in this sample can show meaningful differences across situations. Specifically, we showed that global work satisfaction was related to ratings of positive and negative affect, fatigue, pain, boredom, and loneliness to a greater extent at work than outside of work. People who are more satisfied with their work feel better at work, but this positivity does not extend to other activities. This provides further evidence of the validity of the DRM assessment because it shows that the correlations between global measures and DRM measures do not simply reflect method effects (e.g., just global positivity or negativity in responding). This is particularly impressive given that the questions about job satisfaction were embedded in a separate questionnaire that included many other questions, and so demand characteristics are not a likely explanation for these findings.

Our findings concerning parents were consistent with Nelson et al.'s (2013) idea that parents enjoy spending time with their children more so than other activities. Interestingly, this effect does not appear to be specific to purely leisure activities – parents reported higher positive affect during activities involving childcare than other activities during which they spent time with their children. Overall, although parents reported more positive affect during their activities than non-parents, we found that this difference primarily reflected positivity during activities that involved children. Thus, these data suggests that parenting is a positive experience that is associated with greater positive affect in the days of parents.

In sum, our results suggest that the random episode sampling approach used in the GSOEP-IS as well as some other large studies can capture similar information as the full DRM. The main benefit of the random episode approach is that it takes less time to complete and thus can be useful for a wider range of research designs. In addition, having the DRM embedded within the larger GSOEP survey, allows scholars to take advantage of the broad scope of questions included in this panel study to explore diverse questions about people's daily experiences.

Acknowledgments

This research was supported by NIA grant AG040715 awarded to Richard E. Lucas and M. Brent Donnellan.

References

- Anusic I, Lucas RE, & Donnellan MB (2015). Comparing the reliability and validity of global self-report measures of subjective well-being to day reconstruction measures. Unpublished manuscript.
- Budowski M, Tillmann R, Zimmermann E, Wernli B, Scherpenzeel A, & Gabadinho A (2001). The Swiss Household Panel 1999–2003: Data for research on micro-social change. *ZUMA-Nachrichten*, 49, 100–125.
- Clark LA, Watson D, & Leeka J (1989). Diurnal variation in the positive affects. *Motivation and Emotion*, 13, 205–234.
- Csikszentmihalyi M, & Larsen RE (1987). Validity and reliability of the experience-sampling method. *The Journal of Nervous and Mental Disease*, 175, 526–536. [PubMed: 3655778]
- Diener E (2000). Subjective well-being: The science of happiness and a proposal for a national index. *American Psychologist*, 55, 34–43. [PubMed: 11392863]
- Diener E, Lucas R, Schimmack U, & Helliwell J (2009). *Well-being for public policy*. New York, NY: Oxford University Press.
- Dockray S, Grant N, Stone AA, Kahneman D, Wardle J, Steptoe A (2010). A comparison of affect ratings obtained with ecological momentary assessment and the day reconstruction method. *Social Indicators Research*, 99, 269–283. [PubMed: 21113328]
- Diener E, & Tay L (2014). Review of the Day Reconstruction Method (DRM). *Social Indicators Research*, 116, 255–267.
- Donnellan MB, & Lucas RE (2008). Age differences in the Big Five across the life span: Evidence from two national samples. *Psychology and Aging*, 23, 558–566. [PubMed: 18808245]
- Epstein S (1979). The stability of behavior: On predicting most of the people much of the time. *Journal of Personality and Social Psychology*, 37, 1097.
- Fisher CD (2000). Mood and emotions while working: Missing pieces of job satisfaction? *Journal of Organizational Behavior*, 21, 185–202.
- Goubert L, Crombez G, & Van Damme S (2004). The role of neuroticism, pain catastrophizing and pain-related fear in vigilance to pain: A structural equations approach. *Pain*, 107, 234–241. [PubMed: 14736586]
- Gray JA (1987). The neuropsychology of emotion and personality In Stahl SM & Iversen SD (Eds.), *Cognitive neurochemistry* (pp. 171–190). Oxford, England: Oxford University Press.
- Hahn E, & Gottschling J (2012). Short measurement of personality : Validity and reliability of the GSOEP Big Five Inventory (BFI-S). *Journal of Research in Personality*, 46, 355–359.
- Haisken-DeNew JP, & Frick R (2005). *Desktop companion to the German Socio-Economic Panel (SOEP)*. Berlin: German Institute for Economic Research.
- Judge TA, & Ilies R (2004). Affect and job satisfaction: A study of their relationship at work and at home. *Journal of Applied Psychology*, 89, 661–673. [PubMed: 15327352]
- Kahneman D (1999). Objective happiness In Kahneman D, Diener E, & Schwarz N (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 3–25). New York, NY: Russell Sage.
- Kahneman D, Krueger AB, Schkade DA, Schwarz N, & Stone AA (2004). A survey method for characterizing daily life experience: The Day Reconstruction Method. *Science*, 306, 1776–1780. [PubMed: 15576620]
- Larsen RJ, & Ketelaar T (1991). Personality and susceptibility to positive and negative emotional states. *Journal of Personality and Social Psychology*, 61, 132–140. [PubMed: 1890584]
- Lucas RE, & Donnellan MB (2011). Personality development across the life span: Longitudinal analyses with a national sample from Germany. *Journal of Personality and Social Psychology*, 101, 847–861. [PubMed: 21707197]
- McNamee P, & Mendolia S (2014). The effect of chronic pain on life satisfaction: Evidence from Australian data. *Social Science & Medicine*, 121, 65–73. [PubMed: 25306411]
- Nelson SK, Kushlev K, English T, Dunn EW, & Lyubomirsky S (2013). In defense of parenthood: Children are associated with more joy than misery. *Psychological Science*, 24, 3–10. [PubMed: 23201970]

- Oerlemans WGM, Bakker AB, & Veenhoven R (2011). Finding the key to happy aging: A day reconstruction study of happiness. *Journal of Gerontology: Psychological Sciences*, 66B, 1–10.
- Oishi S, Kurtz JL, Miao FF, Park J, Whitchurch E (2011). The role of familiarity in daily well-being: Developmental and cultural variation. *Developmental Psychology*, 47, 1750–1756. [PubMed: 21875185]
- Rauthmann JF, Gallardo-Pujol D, Guillaume EM, Todd E, Nave CS, Sherman RA, ... Funder DC (2014). The Situational Eight DIAMONDS: A taxonomy of major dimensions of situation characteristics. *Journal of Personality and Social Psychology*, 107, 677–718. [PubMed: 25133715]
- Richter D, & Schupp J (2012). SOEP Innovation Sample (SOEP-IS): Description, structure, and documentation. Retrieved from http://www.diw.de/documents/publikationen/73/diw_01.c.407141.de/diw_sp0463.pdf
- Robinson ME, & Riley JL (1998). The role of emotion in pain In Gatchel RJ & Turk DC (Eds.), *Psychosocial factors in pain* (pp. 74–89). New York, NY: Guilford Press.
- Samuel H (2009). Nicolas Sarkozy wants to measure economic success in 'happiness'. *The Telegraph*. Retrieved from <http://www.telegraph.co.uk/news/worldnews/europe/france/6189530/Nicolas-Sarkozy-wants-to-measure-economic-success-in-happiness.html>
- Srivastava S, John OP, Gosling SD, & Potter J (2003). Development of personality in early and middle adulthood: Set like plaster or persistent change? *Journal of Personality and Social Psychology*, 84, 1041–1053. [PubMed: 12757147]
- Stiglitz JE, Sen A, & Fitoussi J (2009). Report by the commission on the measurement of economic performance and social progress. Commission on the Measurement of Economic Performance and Social Progress. Retrieved from http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf
- Stone AA, Schwartz JE, Schwarz N, Schkade D, Krueger A, & Kahneman D (2006). A population approach to the study of emotion: Diurnal rhythms of a working day examined with the day reconstruction method. *Emotion*, 6, 139–149. [PubMed: 16637757]
- Stratton A (2010). Happiness index to gauge Britain's national mood. *The Guardian*. Retrieved from <http://www.theguardian.com/lifeandstyle/2010/nov/14/happiness-index-britain-national-mood>
- Summerfield M (2010). HILDA user manual—Release 9. Melbourne, Australia: Melbourne Institute of Applied Economic and Social Research, University of Melbourne.
- Tadic M, Oerlemans WGM, Bakker AB, & Veenhoven. (2013). Daily activities and happiness later in life: The role of work status. *Journal of Happiness Studies*, 14, 1507–1527.
- Tay L, Chan D, & Diener E (2014). The metrics of societal happiness. *Social Indicators Research*, 117, 577–600.
- Taylor MF, Brice J, Buck N, & Prentice-Lane E (2009). *British Household Panel Survey user manual: Volume A. Introduction, technical report and appendices*. Colchester, UK: University of Essex.
- Tourangeau R, & Yan T (2007). Sensitive questions in surveys. *Psychological Bulletin*, 133, 859–883. [PubMed: 17723033]
- University of Waterloo (2011). CIW releases Canada's first-ever index to measure national wellbeing. Retrieved from <http://uwaterloo.ca/applied-health-sciences/news/ciw-releases-wellbeing-index>
- Wagner GG, Frick JR, & Schupp J (2007). The German Socio-Economic Panel Study (SOEP): Scope, evolution and enhancements. *Schmollers Jahrbuch*, 127, 139–169.
- Watson D, & Clark LA (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin*, 96, 465–490. [PubMed: 6393179]
- Watson D, & Pennebaker JW (1989). Health complaints, stress, and distress: Exploring the central role of negative affectivity. *Psychological Review*, 96, 234–254. [PubMed: 2710874]

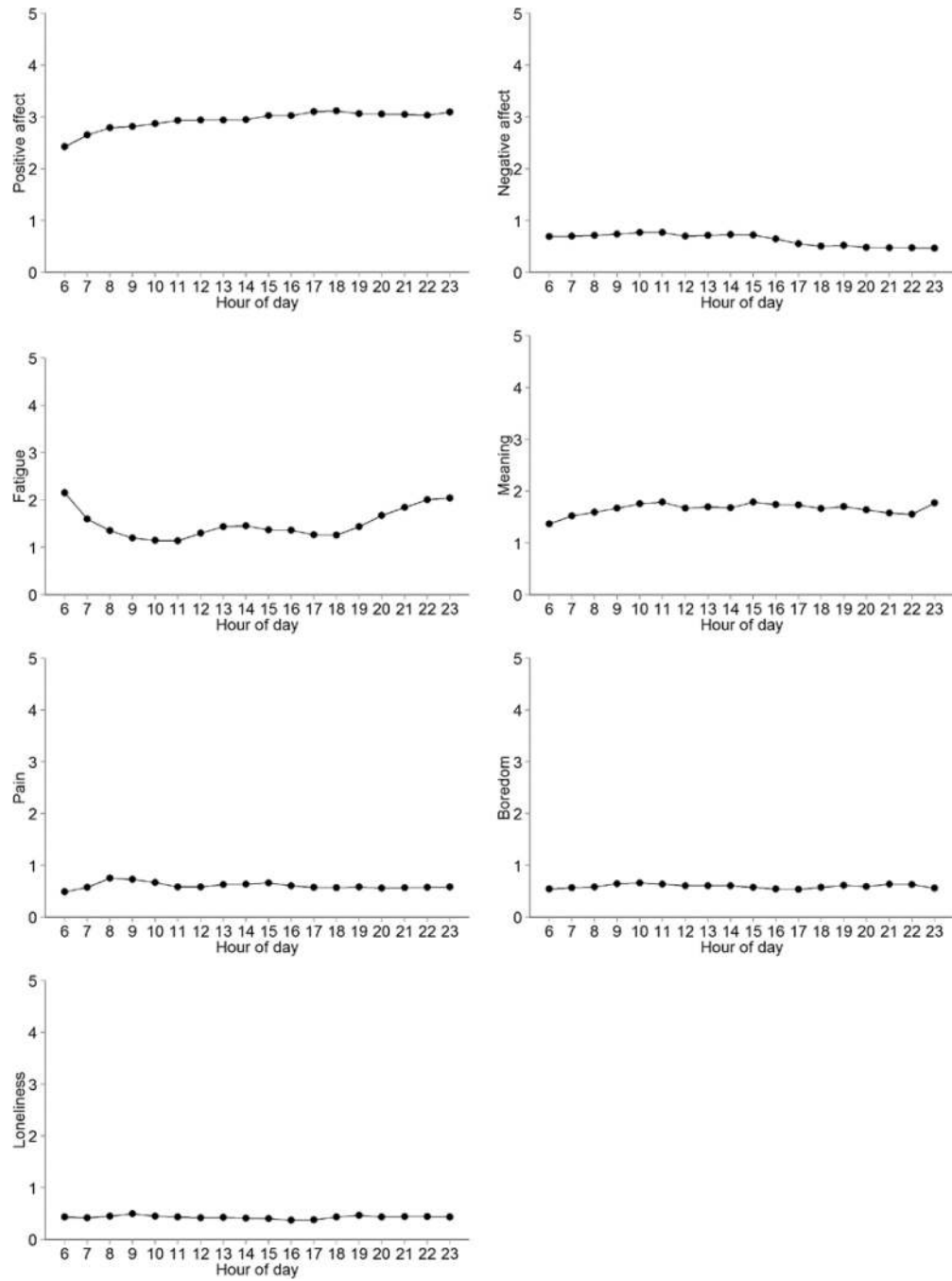


Figure 1.
Diurnal patterns for DRM-rated emotions.

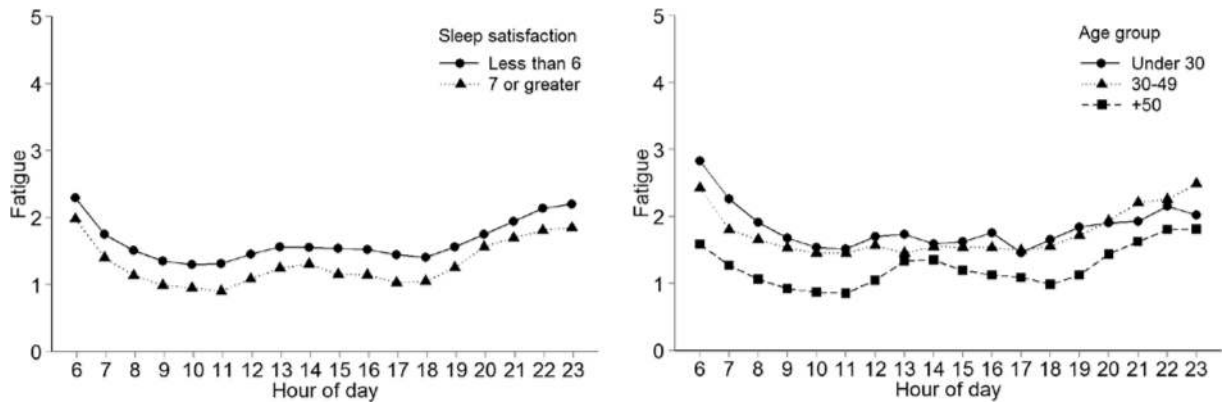


Figure 2. Diurnal patterns for fatigue by age group (left panel) and satisfaction with sleep (right panel).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Mean affect ratings by episode.

	Mean affect rating										Proportion of pleasant (vs. unpleasant)	Mean hours/day	Number of people rating	Proportion of sample reporting	
	Positive affect	Negative affect	Fatigue	Meaning	Boredom	Loneliness	Pain								
What were you doing?															
Sports	4.48 ^W	0.28	0.62	2.57	0.27	0.22	0.81					1.93	79	.12	
Praying/worship/meditation	4.38	0.34	0.50	5.23	0.14	0.07	0.07					0.89	14	.02	
Intimate relations	4.33	0.15	1.25	0.00	0.00	0.00	0.50					1.54	4	.01	
Socializing with friends	4.05	0.38	1.18	2.09	0.43	0.24	0.39					3.07	113	.17	
Caring for pets	3.78	0.34	1.01	1.81	0.32	0.34	0.68					0.97	75	.11	
Caring for children	3.73	0.62	1.45	2.32 ^M	0.42	0.28	0.38					1.86	162	.16	
Other	3.38	0.69	1.19 ^W	2.27	0.50	0.39	0.72					2.04	364	.41	
Gardening	3.25	0.39	0.68	2.04	0.23 ^M	0.19	0.73					1.93	71	.10	
Relaxing	3.12	0.47	2.10	1.61	0.56	0.60	0.60					1.64	238	.30	
Eating	3.08	0.37	1.11	1.39	0.46	0.41	0.59					0.60	976	.89	
Reading	3.06	0.30 ^M	1.04	1.82	0.25	0.28	0.62					1.05	171	.24	
On the phone	2.89	0.90	0.98	1.66	0.27	0.37	0.43					0.64	59	.12	
Cooking	2.88	0.41	1.26	1.38	0.47	0.42 ^M	0.56					0.59	436	.52	
Computer/internet	2.87 ^M	0.52	1.10 ^W	1.44	0.63	0.37	0.37					1.70	192	.24	
Working/studying	2.79	1.13	1.49	2.22	0.64	0.34	0.55					5.79	384	.41	
Watching TV	2.74	0.43	1.79	1.10	0.80	0.57	0.63 ^W					2.27	581	.72	
Commuting (leisure)	2.70	0.52	1.52	1.24	0.60	0.42	0.48					1.36	190	.22	
Resting/napping	2.69	0.39	3.04	1.23	0.77	0.57	0.98					1.54	204	.29	
Getting ready	2.35	0.42	2.10	1.17	0.45	0.45	0.67					0.38	760	.88	
Shopping	2.26	0.54	1.00 ^W	0.91	0.58	0.26	0.58					1.29	220	.31	
Commuting (work/study)	2.21	0.76	1.93	1.27	0.67	0.46	0.48					0.98	390	.37	
Housework	2.14	0.51	1.26	1.20	0.65 ^M	0.49	0.78					1.27	433	.50	

	Mean affect rating										Proportion rated as pleasant (vs. unpleasant)	Mean hours/day	Number of people rating	Proportion of sample reporting
	Positive affect	Negative affect	Fatigue	Meaning	Boredom	Loneliness	Pain							
Doctor appointment	1.34 ^M	0.73	0.84	1.43	0.76	0.32	1.35	.65	1.50	37	.07			
Who were you with?														
Friends	3.72	0.44	1.19	2.01	0.45	0.23	0.50	.96	2.56	236	-			
Parents/relatives	3.40	0.60	1.40	1.54	0.55	0.32	0.53	.97	1.63	166	-			
Your children	3.39	0.57	1.58	1.92	0.40 ^M	0.27	0.41	.96	1.36	392	-			
Supervisor/teacher/trainer	3.23	1.23	1.86	2.72	0.71	0.27	0.45	.92	7.06	74	-			
Partner/spouse	3.14	0.43	1.36 ^W	1.58	0.50	0.28	0.63	.96	1.35	1096	-			
Other	3.07	0.63	1.11	2.39	0.47	0.29	0.71	.89	2.43	169	-			
Clients/customers/students	2.93	1.14	1.56	2.30	0.79	0.25	0.22	.84	5.85	86	-			
Colleague	2.76	1.17	1.64	2.08	0.56	0.32	0.51	.89	5.49	281	-			
Nobody	2.55	0.49	1.57 ^W	1.30	0.58 ^M	0.54	0.64	.93	1.08	1895	-			
Where were you														
At home	2.83	0.44	1.53 ^W	1.44	0.53 ^M	0.45	0.65	.95	1.14	2146	-			
Somewhere else	2.80	0.59	1.37	1.52	0.55	0.34	0.51	.91	1.60	1101	-			
At work	2.80	1.14	1.51 ^W	2.01	0.65	0.39	0.57	.87	5.49	335	-			
Duration-weighted mean	2.91	0.59	1.44 ^W	1.60	0.58	0.42	0.62	-	-	-	-			
% time > 0	95%	55%	53%	53%	28%	20%	25%	-	-	-	-			

Note: Superscripts M and W indicate significant ($p < .05$) gender differences in emotional ratings:

W = women rated higher than men;

M = men rated higher than women.

Correlations of duration-weighted aggregates of DRM affect ratings and global well-being measures.

Table 2.

	PA	NA	Fatigue	Meaning	Pain	Boredom	Loneliness
Life satisfaction	.21	-.16	-.10	.04	-.22	-.15	-.22
Happy	.23	-.13	-.07	.02	-.15	-.11	-.23
Negative affect	-.11	.33	.31	.05	.20	.16	.18
Energy	.21	-.19	-.22	.04	-.22	-.12	-.16
Pain	-.05	.09	.09	.05	.50	.05	.14

Notes: PA = positive affect, NA = negative affect. Trait negative affect is a composite of angry, worried, sad, rushed, and gloomy. Trait energy is a composite of calm (reverse scored) and energetic. Correlations greater in magnitude than .20 are bolded.

Table 3.

Correlations of DRM-rated affect with the Big Five personality variables.

	Agreeableness	Conscientiousness	Extraversion	Neuroticism	Openness to experience
PA	.05	.02	.05	-.05	.06
NA	-.09	-.06	-.04	.09	.01
Fatigue	-.08	-.04	-.01	.12	.00
Meaning	-.05	-.05	.00	.02	.05
Pain	-.03	.00	-.06	.14	-.02
Boredom	-.12	-.13	-.08	.07	-.04
Loneliness	-.08	-.06	-.07	.12	-.03

Note: PA = positive affect, NA = negative affect.

Table 4. Interactions of work satisfaction and work (vs. not work) activities in predicting of DRM ratings.

	PA	NA	Fatigue	Meaning	Pain	Boredom	Loneliness
Intercept	2.39* (0.12)	0.89* (0.07)	2.26* (0.15)	1.51* (0.17)	0.82* (0.10)	0.81* (0.09)	0.63* (0.08)
At work	-1.12* (0.30)	1.60* (0.16)	0.44 (0.37)	-0.12 (0.35)	0.56* (0.19)	0.65* (0.21)	0.38 (0.19)
Work satisfaction	0.06* (0.02)	-0.05* (0.01)	-0.06* (0.02)	0.00 (0.02)	-0.04* (0.01)	-0.04* (0.01)	-0.03* (0.01)
Work satisfaction X At work	0.16* (0.04)	-0.12* (0.02)	-0.12* (0.05)	0.09 (0.05)	-0.06* (0.02)	-0.09* (0.03)	-0.06* (0.03)
Correlations with work satisfaction							
At work	.36	-.32	-.29	.09	-.27	-.32	-.24
Not at work	.07	-.13	-.07	.00	-.07	-.07	-.07

Notes: PA = positive affect, NA = negative affect. Correlations between DRM affect ratings and work satisfaction differed at work vs. not at work for all ratings except meaning ($p < .05$). Work variable was coded 1 for activities done at work, 0 otherwise.

Table 5.

Results of models predicting positive affect from parenting-related variables.

Model	Intercept	Kids	Gender	Kids x Gender	Childcare	Parent	Age	N observations	N people
Parents: activities with kids vs. other activities	2.69* (0.10)	0.61* (0.08)	-	-	-	-	-0.01 (0.01)	1,531	521
Interactions with gender	2.69* (0.10)	0.63* (0.08)	-0.06 (0.06)	0.12 (0.08)	-	-	-0.01 (0.01)	1,531	521
Exclude work	2.68* (0.10)	0.61* (0.08)	-	-	-	-	-0.01 (0.01)	1,415	520
Exclude work, commuting to work, getting ready	2.82* (0.11)	0.46* (0.09)	-	-	-	-	-0.01 (0.01)	1,099	501
Parents, activities with kids: childcare vs. other activities	3.04* (0.16)	-	-	-	0.40* (0.14)	-	-0.01 (0.01)	415	285
Parents vs. non-parents	2.79* (0.03)	-	-	-	-	0.14* (0.07)	0.00 (0.00)	6,578	2,231
Parents (activities not involving children) vs. non-parents	2.78* (0.03)	-	-	-	-	-0.03 (0.07)	0.00 (0.00)	6,034	2,203