

*The Measurement of Subjective Well-Being in Survey Research*  
Vol. 16, No. 3, pp. 429–440

## USING THE DAY RECONSTRUCTION METHOD TO QUANTIFY TIME SPENT SUFFERING AMONG OLDER ADULTS WITH CHRONIC PAIN

Dylan M. Smith<sup>1</sup>

### ABSTRACT

The Day Reconstruction Method (DRM) is a structured diary designed to measure time use in a manner that is more valid than traditional written summary measures, but less burdensome than real-time electronic diary methods. The lower respondent burden and administration costs might make it feasible to utilize this method in large national surveys. Past studies using the DRM have generally focused on subjective mood states during different types of activities. In this study, we extended the DRM to also measure suffering from negative symptoms, such as pain and fatigue, in 122 older adults, most of whom suffer from chronic pain. Results indicated that the method was well tolerated in this population, with over 98% of the sample providing interpretable responses. Chronic pain respondents reported spending a considerable proportion of their day suffering from pain, fatigue, and depression; a much higher proportion than reported by healthy controls ( $p$ 's < .01). Time spent suffering was associated with lower income, negative mood, and lower life satisfaction and quality of life.

**Key words:** quality of life measurement, osteoarthritis, time-weighted experience.

## 1. Introduction

### 1.1. Overview of time-use measurement

To help inform public policy, there is a growing emphasis on capturing subjective appraisals of emotional experiences, as a supplement to traditional objective economic performance measures such as income, or to epidemiological measures of health such as longevity (Kahneman et al., 2004a, Dolan, 2008, Krueger et al., 2009). In particular, it has been proposed that a better sense of the “well-being” of a nation could be achieved if surveys examined how people spend their time (e.g., time spent at work, or engaged in leisure or social activities) and,

---

<sup>1</sup> Department of Preventive Medicine, HSC Level 3, Stony Brook University, Stony Brook, NY, USA 11794. E-mail: Dylan.smith@stonybrookmedicine.edu.

importantly, how people value their experiences during those activities. However, numerous concerns have arisen regarding the accuracy of traditional self-report measures that require respondents to remember and summarize their emotional experiences over some period of time, or ask respondents to estimate their general level of satisfaction with or liking for various activities (for reviews, see (Schwarz and Strack, 1999, Diener et al., 1999, Krueger et al., 2009). These concerns have led methodologists to consider ways of capturing subjective experiences that are less reliant on participants' ability to accurately remember subjective states, and/or to aggregate these experiences into a single summary score.

Ecological Momentary Assessment (EMA) refers to a class of methods designed to capture experiences as they occur, in everyday life, and thus avoid both reliance on memory, and context effects caused by artificial environments (e.g. a laboratory). In the most paradigmatic type of EMA, respondents report their current activities and provide subjective assessments of their emotions and experiences in real time, as they go about their daily lives (Shiffman and Stone, 1998). The usual method is for the respondent to wear an electronic device for a period of time, such as a week, that prompts them at various times throughout the day to respond to a brief survey. Answers are input directly into the device. Thus, from EMA data, researchers can compute average levels of variables of interest (thus avoiding the problem of relying on participants to aggregate their own experiences), and also explore peak and diurnal experiences. Data generated from these methods allows researchers to pair affective experiences with specific activities; by also considering time spent in different activities, across the population these data could be used to generate "national time accounts" (Krueger et al., 2009).

However, the cost of this kind of EMA procedure makes it prohibitive to "scale up" this method to the level of nationally representative surveys. Devices must be provided to respondents who must be trained in their use. Given the considerable respondent burden involved, response rates may be low—especially among some vulnerable or distressed groups—and participant compensation costs are likely to be substantial. Response rates may be especially low among people who have the most difficulty using the devices (e.g., those with certain disabilities), although technological innovation may eventually make this less of an issue (Smith, 2011).

## **1.2. Structured diary approach to subjective assessment**

The Day Reconstruction Method (DRM) was devised to assess subjective experiences in a manner that is specifically designed to avoid problems of many recall-based measures of time use while being more affordable and less burdensome than EMA methods (Kahneman et al., 2004b). In the DRM, participants follow a structured format in which they first divide a day into specific 'episodes' or events, and indicate the duration of the episode. They then describe those events in terms of the type of activity (e.g., commuting to work, having a meal, exercising), and provide a detailed rating of their affective state during the activity.

By first asking participants to recall the events of their day, and then provide ratings of those specific experiences, the DRM exploits the fact that while memory for ongoing experiences, like pain and mood, are flawed, memory for discrete events is much better (Robinson and Clore, 2002). Thus, it avoids (or at least reduces) some of the biasing factors noted above, such as the tendency to recall information that is congruent with peak or recent experiences, which are more easily remembered.

The DRM is designed to be self-administered and can be completed by most participants in a single sitting. Thus, it is much less burdensome and costly to field than the most rigorous EMA methods. Further, while the DRM was originally conceptualized as a way to pair activities with affect (positive and negative mood), its authors explicitly designed it to be adaptable to measure any type of subjectively felt experience.

### **1.3. Using the DRM to measure duration of suffering**

In considering using measures of subjective experience in national surveys, it has been suggested that it might be important to focus not only on emotional well-being, but on suffering, with the idea that alleviation of suffering might be a more suitable target for government intervention (Stone and Mackie, 2013). The adaptability of the DRM noted above makes it a potentially good candidate for quantifying suffering (Smith et al., 2008, Krueger and Stone, 2008). In this study, we adapted the DRM to capture experiences of pain and fatigue in a group of older adults, most of whom suffer from chronic knee pain.

Specifically, we tested the feasibility of an online version of the DRM, and computed time-weighted measures of suffering along the dimensions of pain, fatigue, and depression. We examined interrelations among these variables and tested associations with demographics and previously validated measures of subjective well-being. Finally, for demonstrative purposes we paired measures of suffering with specific activities as reported on the DRM to see whether some types of activities were generally associated with more pain, fatigue and depression.

To achieve these goals, we examined data from an observational study of older adults. The study was comprised of two waves of data collection, each involving a four-day observation period, and included numerous measures, including objective performance measures of physical function, Ecological Momentary Assessments (EMA) of pain and mood, and extensive measures of quality of life using traditional survey methods, in addition to the DRM. For this set of analyses, we focus on data from the DRM, and several survey measures, all at wave 1; other results from this study are reported elsewhere (Stone and Mackie, 2013, Smith, 2011).

## **2. Methods**

### **2.1. Population and sample**

124 participants in total, 77 of whom had chronic knee pain (most commonly due to osteoarthritis), and 47 of whom were healthy older adults, were enrolled in this study. Participants were recruited either through an informational letter and follow-up phone call to adults who are part of the Research Participant Program (RPP) at the University of Michigan's Geriatrics Center, or through the University of Michigan Health System's online registry of ongoing research studies. Participants were included if they were aged 50 and older, cognitively intact (i.e. score of  $\geq 24$  on the Mini Mental Status Exam (Folstein, 1975)), and English-speaking. To be included in the knee pain group, participants had to have at least mild knee pain (score of 5 or greater on the WOMAC pain subscale; (Bellamy et al., 1988)) over the past three months. Participants were excluded if they were non-ambulatory, had medical conditions or problems (other than knee pain) that interfered with daily activity performance or cause pain and fatigue (such as cardiopulmonary problems, neurological conditions, autoimmune diseases), or if they had a joint replacement or surgery of the knee or hip in the previous six months.

Participants were compensated up to \$160 for full participation in the study. Participation was voluntary and participants were free to withdraw from the study at any time without penalty. This research was approved and supervised by the Institutional Review Board at the University of Michigan.

### **2.2. Procedures**

We contacted potentially interested participants by phone for screening, to describe the study, and schedule an initial visit. At each of two time points, six to twelve months apart, data were collected at two separate visits. At the first visit, participants provided informed consent and demographic information, had their cognitive function assessed, and completed measures of medication use, comorbidities, and functional mobility. Participants were then electronically surveyed repeatedly over a four-day observation period to track symptoms and mood quality. On one of these days, chosen at random, participants completed an online version of the DRM (described below). Immediately after this observation period, participants returned for their second lab visit where they received payment and filled out additional questions about the observation period (including quality of life); a small number of participants who did not have internet access completed the DRM at this second lab visit.

## 2.3. Measures

We programmed an online version of the DRM, an instrument that provides a comprehensive assessment of experience-related affect throughout the day (Kahneman et al., 2004b). This instrument employs the aspects of a time-use study, in that respondents recall episodes of the previous day and construct a diary sequencing these episodes. Respondents then answer questions about the subjective experience associated with these episodes, as well as basic, objective questions about when and where episodes occurred, what specific tasks or activities respondents were engaged in, and with whom they were interacting. Questions probing subjective experience include four affect descriptors (happy, depressed, angry, enjoy) with response scales ranging from “not at all” (0) to “very much” (4).

Developers of the DRM designed it to be broadly adaptable to various types of self-report settings. In addition, the response scales used in the DRM can be easily modified to include domains relevant to chronic pain in older adults; thus, we added relevant items to the list of affect descriptors (pain, fatigue, physically active, physically limited).

At visit 2, positive and negative affect over the previous four days was measured using the positive affect/negative affect schedule (Watson et al., 1988). This instrument includes a list of 20 different feelings and emotions. Survey respondents are asked to indicate to what extent they have felt each of these feelings or emotions during the past week using a scale ranging from “Very slightly or Not at all” (1) to “Extremely” (5). Items include such feelings and emotions as “Interested,” “Excited,” “Nervous,” and “Uncertain about things.”

An overall QOL question asked participants to choose a number between 0 and 100 that best represents their current QOL (0 \_ *the worst imaginable QOL*; 100 \_ *the best imaginable QOL*). Overall life satisfaction was measured with the Satisfaction with Life Scale (Diener et al., 1985). This instrument comprises statements about respondents’ general feelings and attitudes concerning their life, such as “In most ways my life is close to my ideal” and “So far I have gotten the important things I want in life.” Respondents indicate how much they agree or disagree with these statements on a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*).

## 3. Results

### 3.1. Response rate and demographics

Of 123 participants, 122 completed the DRM and provided affect ratings of at least three episodes (75 knee pain and 47 healthy controls). Participants divided

their previous day into an average of 8.4 episodes (ranging from 3 to 16 per participant), thus providing a total of 1043 separate episode ratings of mood and symptoms. Episodes were approximately 90 minutes in length, on average, although they ranged in length from 2 minutes to over 10 hours. The sample was 60%; the average age was 59 years. Participants were female and 71% White, 24% African American.

### 3.2. Proportion of time in activities while experiencing suffering

We examined “suffering” along three dimensions: pain, fatigue, and depression. To compute time weighted proportion scores for pain, we first identified any episode in which the participant reported experiencing pain. Next, we summed the length of all painful episodes for each participant, and computed the proportion of the length of that episode to the total length of all episodes for that participant (61, or 6% of the episodes had missing or non-interpretable length estimates; for these we utilized mean imputation). The same procedures were used for fatigue and depression. Finally, we computed a composite variable that estimated the proportion of episodes spent suffering at all—that is, from either fatigue, pain, or depression. As these are proportions, the possible range is from 0 to 1.

Table 1 shows the mean, observed range, and standard deviation for each dimension of suffering, and composite suffering for the total sample, and then broken down by group (chronic pain versus healthy control). Chronic pain participants reported spending more of their time in pain, feeling fatigued, and feeling depressed than did health controls. Regression analyses confirmed that these group differences were significant, after controlling for age and gender (all three  $p$ 's < .01).

**Table 1.** Means, standard deviations, and observed ranges of proportion of time spent suffering for the full sample and by group.

	Total sample (n=122)	Chronic pain sample (n=75)	Non-pain sample (n=47)
Proportion of time with pain (0-1)	.60 (.45)	.79 (.30)	.21 (.34)*
Proportion of time with fatigue (0-1)	.61 (.54)	.71 (.36)	.45 (.37)*
Proportion of time with depression (0-1)	.29 (.38)	.38 (.42)	.15 (.07)*
Proportion of time any symptom (0-1)	.76 (.33)	.90 (.23)	.55 (.36)*

*Note.* The significance tests in the table indicate whether pain participants differed from non-pain participants in regression models that controlled for age and gender.

\* $p < .01$ .

Next, we examined associations with demographics and previously validated measures of subjective well-being and quality of life. Table 2 depicts these correlations. As shown, time spent in pain or depressed was associated with lower income, lower life satisfaction, more negative affect, and lower overall quality of life. Time spent fatigued was associated with more negative affect, and lower quality of life. Although not depicted in Table 2, we also examined these associations within each group and did not observe notable differences in the patterns.

**Table 2.** Correlations between different measures of suffering and with demographics and measures of well-being (n=122).

	Proportion pain	Proportion fatigue	Proportion depressed	Age	Income	Negative Mood	Life Sat.	Quality of life
Prop. pain	--							
Prop. fatigue	.59**	--						
Prop. Depressed	-.40**	.44**	--					
Age	-.10	-.12	-.28**	--				
Income	-.26**	-.05	-.28**	.17 <sup>a</sup>	--			
Negative Mood	.35**	.30**	-.43**	-.10	-.19*	--		
Life Sat.	-.29**	-.14	-.45**	.29**	-.51**	-.42**	--	
Quality of Life	-.27**	-.20*	-.43**	.21*	.50**	-.41**	.67**	--

Note: Two participants failed to return for wave 2 and thus did not provide responses for mood, life satisfaction, or quality of life; for correlations involving these variables, n = 120. Three additional participants declined to respond to the income question; for correlations involving this variable, n= 117.

\*\* = p < .01

\* = p < .05

a = p < .10

### 3.3. Suffering by type of activity

Next, for demonstrative purposes, we paired pain and fatigue scores to specific activities listed under each episode. To examine whether a given type of activity was associated with more pain, we subtracted the participants' average pain level from the pain score reported during the activity. That is, for each individual we computed a score indicating whether a given activity was generally

associated with greater or less pain than that generally experienced by that individual (this was done to avoid confounding across participants; i.e., participants with more pain may be less likely to work).

For the total group, pain was highest while at work (average difference score = .08,  $n = 50$ ) and while exercising (average difference score = .23,  $n = 48$ ). Fatigue was highest while watching television (.18,  $n = 91$ ), during childcare (.09,  $n = 43$ ), while exercising (.09), and while relaxing (.13,  $n = 96$ ). When we examined activities within each group, patterns were generally similar.

## **4. Discussion**

### **4.1. Overview of key findings**

These data provide initial support for the idea that a structured diary of time use, the Day Reconstruction Method, can be adapted to measure dimensions of suffering across a day in older adults with chronic illness. The online instrument was well tolerated by participants, 98% of whom were able to complete the instrument and provide affect ratings of at least three episodes. Findings indicated that pain was prevalent. Knee arthritis participants reported pain 85% of the time spent in their episodes throughout the day; however, pain was also reported 21% of the time in episodes reported by healthy older adults. Feeling fatigued was also common, though much more so in the knee pain group. Depression was the least prevalent among the three measures examined, occurring 15% of the time in the healthy group and 38% of the time in the pain group. Across both groups, proportion of time spent suffering was generally associated with lower income, negative affect, and lower life satisfaction and quality of life.

The DRM produces much more detailed information about time spent in different activities than do traditional summary survey measures, and there is some theoretical and empirical basis to assume that this information is more accurate as well. Past research comparing the DRM to the “gold standard” approach of repeated experience sampling has generally found relatively close agreement—closer than that observed using traditional survey methods (Kim et al., 2013, Stone and Mackie, 2013).

For descriptive purposes, we also paired ratings of pain and fatigue with specific types of activities. Few clear patterns emerged; however, this may be a function of the relatively modest sample size (compounded by the fact that not all participants were engaged in all activity types). Nonetheless, we did observe some indication that pain levels were higher when participants were at work and especially while exercising. Given that participants on average spent much more time at work than exercising, work pain may have been a bigger overall component of total time in pain. However, given the sample size we performed no significance test and made no claim regarding the replicability of the pattern we observed. In larger, more representative studies this technique could be quite

useful in identifying the types of activities that most contribute to time spent suffering. The format of the DRM also allows for weighting by the length of activity, further refining its ability to identify activities that contribute the most to the overall experience of pain.

Future studies can also expand upon the general approach pursued here. Additional dimensions of suffering can be examined (e.g., stress and anxiety, other physical symptoms including disability limitations, etc.) in different subpopulations (e.g., other health conditions, different racial/ethnic groups). In addition, the data generated from the DRM allows for other analytic approaches, including weighting by intensity of the experience of pain, for example, or utilizing a more strict criteria to define suffering (e.g., restricting to the highest pain levels, rather than including any pain level).

#### **4.2. Caveats and limitations**

The DRM is designed to be more accurate than traditional written summary measures, and it provides estimates of time spent engaged in various activities, and affect experienced during those activities; however, the time data should be interpreted with caution. Obviously, these are estimates based on potentially flawed memories of how the day went, but in addition the format of the DRM makes interpretation of these estimates somewhat ambiguous, for at least two reasons. First, participants are allowed to define how many episodes to split the day into, and they often leave gaps in time. Second, while they may report experiencing a symptom during a given episode, they may not have experienced it for the entire duration of that episode. For these reasons, rather than try to estimate how many minutes each participant reported experiencing a negative symptom, we computed a time-weighted proportion of episodes in which a symptom was experienced. This can be interpreted as how much time the participants spent in episodes with, for example, pain, as a proportion of the total time spent in all episodes the participant felt were sufficiently meaningful to list in reconstructing their day. As such, it leaves open the possibility that suffering occurred at other points in the day not listed as episodes.

Third, while the DRM is much less burdensome and expensive to administer than experience sampling methods using electronic diaries, it still requires nearly 45 minutes for most participants to complete the full instrument. Thus, it may be impractical for use in many national surveys. However, researchers have successfully fielded a brief version of the DRM in at least one large survey study, by having respondents rate only a subsample of the episodes respondents listed (Krueger and Stone, 2008).

Finally, this is a non-representative clinical sample of limited size; thus, we cannot assert that the proportion of time spent in pain observed here will generalize to broader populations. As noted, the small sample size makes it difficult to study how type of activity relates to pain and fatigue.

### **4.3. Conclusion**

Using the Day Reconstruction Method, we observed that some older adults spend a substantial proportion of their time experiencing symptoms such as pain, fatigue, and depression, and that time spent with these symptoms is negatively associated with quality of life. Taken together, these findings highlight the adaptability of the DRM, and demonstrate its potential usefulness for capturing suffering, in addition to emotional well-being.

### **Acknowledgements**

This research was supported by a grant from the National Institute on Aging (R21AG030177). I wish to thank Catherine Hytner, William Lopez, and Rachel Tocco for their assistance in data collection and coding. I also wish to thank Susan Murphy for her contributions to the design of the parent study that made it possible to generate these data.

**REFERENCES**

- BELLAMY, N., BUCHANAN, W. W., GOLDSMITH, C. H., CAMPBELL, J., STITT, L. W.,(1988). Validation study of WOMAC: A health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *The Journal of Rheumatology*, 15, 1833–1840.
- DIENER, E., EMMONS, R. A., LARSEN, R. J., GRIFFIN, S.,(1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, 49, 71–75.
- DIENER, E., SUH, E. M., LUCAS, R. E., SMITH, H. L.,(1999). Subjective well-being: Three decades of progress. *Psychological Bulletin*, 125, 276–302.
- DOLAN, P.,(2008). Developing methods that really do value the ‘Q’ in the QALY. *Health Economics, Policy and Law*, 3, 69–77.
- FOLSTEIN, M., FOLSTEIN S. E., MCHUGH, P. R., (1975). "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*, 12, 189–198.
- KAHNEMAN, D., KRUEGER, A., SCHKADE, D., SCHWARZ, N.,STONE, A. A.,(2004a). Toward National Well-Being Accounts. *AEA Papers and Proceedings*, 94, 429–434.
- KAHNEMAN, D., KRUEGER, A. B., SCHKADE, D., SCHWARZ, N., STONE, A. A.,(2004b). A survey method for characterizing daily life experience: The day reconstruction method (DRM). *Science*, 306, 1776–1780.
- KIM, J., KIKUCHI, H., YAMAMOTO, Y.,(2013). Systematic comparison between ecological momentary assessment and day reconstruction method for fatigue and mood states in healthy adults. *British Journal of Health Psychology*, 18, 155–167.
- KRUEGER, A. B., KAHNEMAN, D., SCHKADE, D., SCHWARZ, N., STONE, A. A.,(2009). National Time Accounting: The Currency of Life. In: KRUEGER, A. B. (ed.) *Measuring the Subjective Well-Being of Nations: National Accounts of Time Use and Well-Being*. Chicago: University of Chicago Press.
- KRUEGER, A. B., STONE, A. A.,(2008). Assessment of Pain: A community-based diary survey in the USA. *The Lancet*, 371, 1519–1525.
- ROBINSON, M. D., CLORE, G. L.,(2002). Episodic and semantic knowledge in emotional self-report: Evidence for two judgment processes. *Journal of Personality and Social Psychology*, 83, 198–215.

- SCHWARZ, N., STRACK, F.,(1999). Reports of subjective well-being: judgmental processes and their methodological implications. In: KAHNEMAN, D., DIENER, E., SCHWARZ, N. (eds.) *Well-Being The foundations of hedonic psychology*. New York.
- SHIFFMAN, S., STONE, A. A.,(1998). Ecological momentary assessment: A new tool for behavior medicine research. In: KRANTS, D. S. (ed.) *Technology and Methods in Behavioral Medicine*. Mahwah: Lawrence Erlbaum Associates.
- SMITH, D. M.,(2011). Measures of time-use and well-being in survey research. *Forum for Health Economics and Policy*, 14, 1–14.
- SMITH, D. M., BROWN, S. L., UBEL, P. A.,(2008). Mispredictions and misrecollections: Challenges for subjective outcome measurement. *Disability and Rehabilitation*, 30, 418–424.
- STONE, A. A., MACKIE, C., (eds.),(2013). *Subjective Well-Being: Measuring Happiness, Suffering, and Other Dimensions of Experience*, Washington, D.C.: The National Academies Press.
- WATSON, D., CLARK, L. A., TELLEGEN, A.,(1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality & Social Psychology*, 54, 1063–1070.