Published in final edited form as:

Soc Indic Res. 2014 February 1; 115(3): 907–917. doi:10.1007/s11205-013-0240-z.

Validation of a Brief Yesterday Measure of Hedonic Well-Being and Daily Activities: Comparison with the Day Reconstruction Method

Christopher Christodoulou, PhD,

Department of Neurology, Stony Brook University, HSC T12-033 B, Stony Brook, NY 11794-8121, office: 631 444-8068, christopher.christodoulou@stonybrook.edu

Stefan Schneider, PhD, and

Department of Psychiatry, Stony Brook University, 125 Putnam Hall, Stony Brook, NY 11794-8790, 631 632-3114, Stefan.Schneider@stonybrook.edu

Arthur A. Stone, PhD

Department of Psychiatry, 125 Putnam Hall, Stony Brook, NY 11794-8790, 631 632-8833, arthur.stone@stonybrook.edu

Abstract

There has been increasing interest in the measurement of hedonic well-being (HWB), due in part to its broad implications in areas such as health and society. The day reconstruction method (DRM) is a validated technique assessing HWB and daily activities using instructions that help respondents recover their experiences from the previous day, thus reducing recall bias. Unfortunately, large-scale surveys are typically not able to implement the time-consuming DRM procedure; instead, they rely on single-item or very brief questionnaire assessment of HWB and time usage. Despite their wide use, brief questionnaires have rarely been compared to the DRM, which could provide validation of these short assessments. In the present study we compared these two questionnaire formats in 45 adults who completed both a DRM and a “hybrid” short form (HSF) questionnaire that included a very brief procedure to reconstruct yesterday’s events. Results were that the mean HSF ratings tended to overestimate HWB in comparison to the DRM, though effects were generally small to moderate. With regard to estimates of time spent on daily activities, the HSF also differed from the DRM, though with generally small to moderate effects. Correlations between estimates indicated that the HSF explained an average of almost half the variance in the DRM ratings for both HWB (r s ranging from .52 to .97) and time use (r s ranging from .43 to .85). In general, HSF ratings displayed considerable overlap with those of the DRM though the brief instructions apparently did not eliminate recall bias in the shorter questionnaire.

1. Introduction

Hedonic well-being (HWB) refers to the frequency and intensity of emotional experiences such as happiness, joy, stress, and worry that make a person’s life pleasant or unpleasant (Kahneman and Deaton 2010). A variety of disciplines have shown increasing interest in the accurate assessment of HWB, especially positive aspects of well-being (Seligman and Csikszentmihalyi 2000; Kahneman and Krueger 2006; Huppert et al. 2004; Krueger et al. 2009). Research has begun to delineate the neurobiological foundations of HWB (Davidson 2004) and to discern broad and important implications in areas such as health and society. In

Disclosures: AAS is a Senior Scientist with the Gallup Organization and a Senior Consultant with ERT, inc.

health research, positive affect has been found to predict response to illness (Cohen et al. 2003) and even survival among older men and women (Steptoe and Wardle 2011). In the economic and social arenas, there is a realization that traditional economic measures such as income provide an incomplete explanation of societal well-being (Easterlin 2001); Kahneman and Deaton 2010) and that appropriate measurement of HWB could serve as a useful complement to traditional economic indicators (Kahneman et al. 2004a; Seaford 2011).

One comprehensive method of evaluating HWB is the day reconstruction method (DRM) (Kahneman et al. 2004b; Dockray et al. 2010). This technique is used to assist individuals in recovering their experiences (time-usage) from yesterday. Participants are instructed to generate a sequential series of episodes of the previous day beginning with when they awoke and ending with when they went to bed (Kahneman et al. 2004b), which is similar to the procedure used in standard large scale time-use studies such as the American Time Use Survey (ATUS) (Herz and Devens 2001; Horrigan and Herz 2004). Participants are then asked to describe each episode and the feelings they experienced. Evoking the context of the previous day is intended to elicit specific memories and to reduce recall biases (Belli 1998). The DRM provides information regarding activities and time usage, which can then be associated with HWB ratings to give an indication of the relation between what one is doing and how one is feeling (Stone et al. 2006).

Studies have demonstrated the reliability (Krueger and Schkade 2008) and validity of the DRM (Kahneman et al. 2004b; Stone et al. 2006; Dockray et al. 2010; Krueger et al. 2009). The data provided by the DRM compares favorably with the more time consuming process of experience sampling method (ESM) (Csikszentmihalyi and Larson 1987) or ecological momentary assessment (EMA) (See review by (Shiffman et al. 2008)), which involve repeated momentary assessments of one's current (momentary) HWB, often relying on handheld electronic devices that alert the individual to record their feelings at random points throughout the day during their daily activities. An important difference in the techniques is that EMA typically captures information about *moments* (at the time respondents are signaled) whereas the DRM captures information about *episodes*, thereby providing complete coverage of the day.

The DRM produces data that approximates EMA, and is less intrusive. However, it is still a time consuming process, requiring approximately 45 to 75 minutes to complete a standard version (Kahneman et al. 2004b). At a practical level this makes ESM/EMA and the DRM too burdensome for routine use in many large scale surveys and longitudinal investigations (Krueger et al. 2009). Such investigations often must rely on short questionnaires with a focus on single items or a small set of items to uncover the relation of HWB to various attributes and variables, for instance the relation of positive affect to aging (Stone et al. 2010).

Though short-form surveys (e.g., Gallup, Pew, etc.) are widely used, their relation to DRM data has rarely been investigated. One study that did compare short-form and DRM ratings for HWB, though not for daily activities, found that the short-form generated higher mean ratings of both positive and negative hedonic experiences than corresponding DRM experiences (Miron-Shatz et al. 2009). The authors interpreted this difference as a "memory-experience gap," by which memories for events are more intense than the average of the emotions they actually experienced. The DRM ratings are seen as more closely representing the experienced emotions than those from the short-form because the DRM is specifically designed to re-instantiate the context of the previous day, and its associated emotional experience.

The goal of the present investigation was to evaluate a “hybrid” short-form questionnaire that adopts some attributes of the DRM but would be more practical for use in large surveys. This hybrid short-form (HSF) includes a very brief reconstruction/re-instantiation of yesterday. Unlike the DRM, however, it does not collect HWB and time use data for individual episodes of the previous day, but rather collects general (non-episode specific) ratings of HWB (similar to other short-forms) and time spent on daily activities. The goal of this investigation was to assess the degree to which the HSF would provide data consistent with that obtained by the DRM, with the idea that a short-form could be used in large-scale survey research.

2. Methods

2.1. Participants

Participants were recruited through an online advertisement to the local university community seeking adult staff or faculty volunteers, 18 years and older, who were not students at the university, who were fluent in English, and who did not have uncorrected vision or hearing difficulties that might interfere with their participation. Each participant received \$50 for completion of the study.

2.2. Procedures

Individuals were tested in our university offices in small groups of usually two to five persons, depending upon their availability, though one individual was tested alone. Participants first received the HSF measures, which were administered orally in order to approximate telephone administration in a large-scale survey. Participants provided written responses to the HSF questions. Administration of the HSF began with the following comments to help participants with a brief reconstruction of the previous day's events. “I'd like you to think about YESTERDAY by trying to retrace your actions and locations – that is, what you were doing and where you were – throughout YESTERDAY. A good way to do this is by starting in the morning when you awoke and then moving to the next activity and so on. So, please spend 30 seconds or so doing this. Review your morning, then afternoon, and then evening.”

HWB adjectives were then introduced by telling participants they would now be asked for “some general impressions about yesterday.” Participants were asked “how” or “how much” they experienced the following HWB adjectives, which were rated on an 11 point scale ranging from 0 = “not at all” to 10 = “very much.” The items were: “Happy,” “Engaged,” “Satisfied,” “Worthwhile and Meaningful,” “Stress,” “Worried,” “Sad,” “Pain,” and “Wishing you were somewhere else.” With the exception of Happy and Worried, the items differ from those of the original DRM. The current items were predominantly adapted from a large daily survey, the Gallup Healthways Well-being Index (which included “Happiness,” “Stress,” “Worry,” “Sadness,” “Physical pain”) (Kahneman and Deaton 2010; Deaton 2012), with additional items included in order to reflect a broader conception of subjective well-being (i.e., “Engaged,” “Worthwhile and Meaningful,” “Wishing you were somewhere else,” “Satisfied”), including thought based components (White and Dolan 2009). Means for positive adjectives (Happy, Engaged, Satisfied, Worthwhile/Meaningful) and negative adjectives (Stress, Worried, Sad, and Wishing you were somewhere else) were calculated (Footnote: a dichotomous 0–1 format was also administered for many of the HWB adjectives and this data is available from the authors).

Following the presentation of the HWB items, the interviewer queried participants with regard to a series of daily activities, preceded by the question, “how did you spend time yesterday?” Following this general question was the following list of activities, each

requiring a “Yes” or “No” response: “Housework, including meal preparation, cleaning, bills,” “At work or working,” “Commuting,” “Childcare,” “Exercising, including walking,” “Watching TV,” “Leisure/Recreation activity, not including walking or TV watching,” and “With friends and/or family (awake time, including by phone, internet, or texting).” For each activity that a participant endorsed, they were then asked “How long?” Responses were coded as hours and minutes.

Participants then completed a written version of the DRM (Kahneman et al. 2004b). The DRM procedures began with life satisfaction and demographic questions. Participants were then asked to reconstruct their activities from yesterday in a diary. They were instructed to, “think of your day as a continuous series of scenes or episodes in a film. Give each episode a brief name that will help you remember it (for example, “commuting to work”, or “at lunch with B”, where B is a person or a group of people). Write down the approximate times at which each episode began and ended. The episodes people identify usually last between 15 minutes and 2 hours. Indications of the end of an episode might be going to a different location, ending one activity and starting another, or a change in the people you are interacting with.” (Kahneman et al. 2004b)(p. 1777) Participants were told that they would be using these notes to answer specific questions about these episodes, but that their notes were strictly personal and confidential and that they did not need to turn them in or show them to anyone. For each episode, the DRM asked questions about the same HWB adjectives used in the HSF (except that the DRM used “Stressed” instead of “Stress”), using the same scale of 0 = “not at all” to 10 “very much.”. Ratings for a HWB adjective were averaged across episodes for each person (with averages weighted by episode duration) to obtain an overall score for the day.

For each episode on the DRM participants were asked “what were you doing? (please check all that apply),” followed by the same list of activities used in the HSF. We calculated the overall duration of an activity for the day as the sum of the duration of all episodes during which the activity occurred. For episodes on which a participant selected more than one activity (e.g., for a single 60 minute episode one might select “commuting” and “with friends and/or family”), the full episode length (e.g., 60 minutes) was applied to each of the activities selected. We also examined an alternative method of calculating the duration of activities (i.e., evenly dividing the episode length by the number of activities selected), but results were similar with either approach and only the data for the full episode approach are presented in the results.

3. Results

A total of 49 participants were recruited for the study. Data for 4 individuals were excluded because they reported five or fewer episodes for the previous day, which we believed could distort the calculation of affect and time use estimates. Data for the remaining 45 participants were analyzed. Demographic characteristics of the predominantly female sample are provided in Table 1.

3.1. Hedonic Well-Being

T-tests comparing mean levels for each of the nine HWB ratings from the HSF versus DRM, showed statistically significant differences for five out of nine items ($p < 0.05$, using the sequentially rejective Bonferroni test) (Holm 1979), with the exception of Happy, Satisfied, Worthwhile/Meaningful, and Sad (Table 2). In each instance the HSF ratings were higher than those on the DRM. The difference was less than one point on the 0–10 scale for all adjectives except Stress/Stressed and Worried, which approached two points. Effect sizes (Cohen’s d) were generally small to moderate, ranging from 0.13 for Happy to 0.81 for Stress. HSF ratings were higher than those from the DRM for both positive and negative

items, and this tended to be stronger for negative items ($p = 0.045$). In addition to comparison of overall level differences, we examined the correlation between DRM and HSF for each adjective (see Table 2). All correlations were significant ($p < .01$), ranging from .52 (for worried) to .97 (for pain) in magnitude.

3.2. Time Use Activities

The HSF ratings significantly differed from the DRM on three of the eight types of activities: Housework, Commuting, and Leisure/Recreation (Table 3); in each case, the mean duration was shorter on the HSF than on the DRM. The mean time estimates for the two questionnaire formats were within one hour for all activities except Housework, and Leisure/Recreation. Effect sizes for the mean differences were generally small to moderate, and ranged from 0.09 for Work to 0.71 for Housework and Leisure/Recreation. Correlations between time estimates derived from HSF and DRM ranged from $r = 0.43$ for TV to $r = 0.85$ for Work, with most between $r = 0.50$ and $r = 0.70$ (all r s significant at $p < .01$)

3.3. Pattern of relations among HWB and time use variables for the DRM and HSF

We next examined associations between the HWB ratings and the time spent on each of the activities during the day, separately for the DRM and the HSF. Table 4 presents the generally modest correlations. We explored whether the correlations for the HSF differed from those found for the DRM using a test of differences between dependent (i.e., correlated) correlations. With few exceptions, the sets of correlations were not found to differ ($ps > .10$). In an analogous examination of the pattern of intercorrelations among HWB items for the DRM and HSF, respectively (Table 5), the two methods again showed similar patterns of intercorrelations. For each method, HWB items of similar valence (e.g., positive with positive or negative with negative) tended to correlated positively, while those with opposing valences displayed negative correlations.

4. Discussion

Results indicate that estimates of HWB from the HSF differed from those of the DRM, though effect sizes tended to be modest with the HSF ratings generally within one point of those for the DRM on an 11 point scale. Correlations between HWB ratings from the HSF and DRM all reached significance, with the strength of the correlations indicating that HSF ratings explained an average of almost half the variance in the DRM ratings.

Like the DRM, the HSF ratings of negatively valenced HWB items tended to be lower than those of positive items. To the extent that HSF ratings of HWB differed from the DRM, they did so in a predictable fashion, with the HSF providing higher means across most HWB items. This pattern of results is similar to that of an earlier study comparing the two questionnaire formats (Miron-Shatz et al. 2009). Like that earlier study, the present investigation found that the overestimation for negative items was greater than it was for positive items.

Administration of the HSF always preceded the DRM, so it is possible that some sort of order effect could be responsible for the HSF overestimation. However, the earlier investigation counterbalanced administration of a short-form questionnaire and DRM, and found a mean elevation of short form HWB ratings regardless of order (Miron-Shatz et al. 2009). It is notable that the present investigation showed results similar to the earlier study, even though the earlier one did not instruct participants to reconstruct their day prior to collecting short-form ratings. This similarity suggests that the brief instructions for the HSF were not particularly effective in re-instantiating the context of the previous day, and its associated emotional experience, as compared to the more extensive instructional set in the

DRM. Overestimation of symptom ratings has been observed with other forms of recall measures as well, such end of day recall of pain and fatigue as compared to mean EMA ratings (Schneider et al. 2011).

Estimates of time spent on activities showed some differences between the HSF and DRM. The effect size of the differences ranged from small to moderate, with HSF estimates within one hour of those for the DRM for the majority of activities. Correlations between time use estimates from the HSF and DRM were uniformly significant, with HSF ratings explaining an average of close to half the variance in DRM ratings. Importantly, the HSF displayed a pattern of relationships between time use and HWB that was quite similar to that of the DRM, and evidenced some face validity, including a positive correlation between work and stress, a negative correlation between watching television and worthwhile/meaningful, and a negative correlation between spending time with friends/family, and wishing you were somewhere else. In addition, correlations amongst HWB items for the HSF were similar to those found for the DRM, with both methods tending to display positive correlations among HWB items of similar affective valence and negative correlations among items of opposite valence.

The generalizability of the present findings is limited by the relatively small size of the sample as well as its demographic composition, which included an over-representation of women and persons with higher levels of income and education. The findings would clearly benefit from replication in a larger, more representative sample. In addition, the study was limited by its primary focus on HWB rather than other aspects of subjective well-being (e.g., eudaimonic approach that measures how well a person is functioning in their environment and realizing their potential (Huppert et al. 2009)), though two items (“Engaged,” “Worthwhile and Meaningful”) tap some aspects of a broader understanding of subjective well-being.

In general, the results show that HSF ratings display considerable overlap with those of DRM; perhaps enough to make valid statements at a group level, if not at the level of individuals. However, the brief instructions used in the HSF did not appear to lessen the tendency of recall measures to overestimate HWB. (Miron-Shatz et al. 2009; Schneider et al. 2011) One possible modification of the HSF procedure would be to present the activity items first as a way to use participants’ processing of the activity items to improve their re-instantiation at the time when their HWB ratings are made. Any attempt to modify the the HSF must balance the goal of better eliciting memories for past events with the need for brevity inherent in large scale surveys. In its current form the procedure is brief, as we have collected ancillary data finding that the nine HWB ratings could be administered to a sample of 40 to 85 year old participants in an average of less than one minute, with administration of the eight activity items requiring an addition of only one and a half more minutes on average (data available from the authors).

References

- Belli RF. The structure of autobiographical memory and the event history calendar: potential improvements in the quality of retrospective reports in surveys. [Review]. *Memory*. 1998; 6(4): 383–406. [PubMed: 9829098]
- Cohen S, Doyle WJ, Turner RB, Alper CM, Skoner DP. Emotional style and susceptibility to the common cold. [Comparative Study, Research Support, Non-U.S. Gov't Research Support, U.S. Gov't P.H.S.]. *Psychosom Med*. 2003; 65(4):652–657. [PubMed: 12883117]
- Csikszentmihalyi M, Larson R. Validity and reliability of the Experience-Sampling Method. [Research Support, Non-U.S. Gov't Research Support U.S. Gov't P.H.S.]. *J Nerv Ment Dis*. 1987; 175(9):526–536. [PubMed: 3655778]

- Davidson RJ. Well-being and affective style: neural substrates and biobehavioural correlates. [Research Support U.S. Gov't PHS. Review]. *Philos Trans R Soc Lond B Biol Sci*. 2004; 359(1449):1395–1411. [PubMed: 15347531]
- Deaton A. The financial crisis and the well-being of Americans. *Oxford Economic Papers-New Series*. 2012; 64(1):1–26.
- Dockray S, Grant N, Stone AA, Kahneman D, Wardle J, Steptoe A. A Comparison of Affect Ratings Obtained with Ecological Momentary Assessment and the Day Reconstruction Method. *Soc Indic Res*. 2010; 99(2):269–283. [PubMed: 21113328]
- Easterlin RA. Income and happiness: Towards a unified theory. *Economic Journal*. 2001; 111(473): 465–484.
- Herz D, Devens RM. The American time-use survey. *Industrial Relations*. 2001; 40(3):526–529.
- Holm S. A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*. 1979; 6(2):65–70.
- Horrigan M, Herz D. Planning, designing, and executing the BLS American time-use survey. *Monthly Labor Review*. 2004; 127(10):3–19.
- Huppert FA, Baylis N, Keverne B. Introduction: why do we need a science of well-being? *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*. 2004; 359(1449):1331–1332.
- Huppert FA, Marks N, Clark A, Siegrist J, Stutzer A, Vitterso J, et al. Measuring Wellbeing Across Europe: Description of the ESS Well-being Module and Preliminary Findings. *Soc Indic Res*. 2009; 91(3):301–315.
- Kahneman D, Deaton A. High income improves evaluation of life but not emotional wellbeing. *Proceedings of the National Academy of Sciences of the United States of America*. 2010; 107(38): 16489–16493. [PubMed: 20823223]
- Kahneman D, Krueger AB. Developments in the measurement of subjective well-being. *Journal of Economic Perspectives*. 2006; 20(1):3–24.
- Kahneman D, Krueger AB, Schkade D, Schwarz N, Stone A. Toward national wellbeing accounts. *American Economic Review*. 2004a; 94(2):429–434.
- Kahneman D, Krueger AB, Schkade DA, Schwarz N, Stone AA. A survey method for characterizing daily life experience: the day reconstruction method. [Research Support, Non-U.S. Gov't Research Support U.S. Gov't, Non-P.H.S. Research Support U.S. Gov't P.H.S.]. *Science*. 2004b; 306(5702): 1776–1780. [PubMed: 15576620]
- Krueger, AB.; Kahneman, D.; Schkade, D.; Schwarz, N.; Stone, AA. National time accounting: The currency of life. In: Krueger, AB., editor. *Measuring the subjective well-being of nations: National accounts of time use and well-being*. Chicago: University of Chicago Press; 2009. p. 9-86.
- Krueger AB, Schkade DA. The Reliability of Subjective Well-Being Measures. *J Public Econ*. 2008; 92(8–9):1833–1845. [PubMed: 19649136]
- Miron-Shatz T, Stone A, Kahneman D. Memories of Yesterday's Emotions: Does the Valence of Experience Affect the Memory-Experience Gap? *Emotion*. 2009; 9(6):885–891. [PubMed: 20001131]
- Schneider S, Stone AA, Schwartz JE, Broderick JE. Peak and end effects in patients' daily recall of pain and fatigue: a within-subjects analysis. [Research Support, N.I.H., Extramural]. *J Pain*. 2011; 12(2):228–235. [PubMed: 20817615]
- Seaford C. Policy: time to legislate for the good life. *Nature*. 2011; 477(7366):532–533. [PubMed: 21956313]
- Seligman ME, Csikszentmihalyi M. Positive psychology. An introduction. *Am Psychol*. 2000; 55(1): 5–14. [PubMed: 11392865]
- Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. [Review]. *Annu Rev Clin Psychol*. 2008; 4:1–32. [PubMed: 18509902]
- Steptoe A, Wardle J. Positive affect measured using ecological momentary assessment and survival in older men and women. [Research Support N.I.H., Extramural Research Support, Non-U.S. Gov't]. *Proc Natl Acad Sci U S A*. 2011; 108(45):18244–18248. [PubMed: 22042845]

- Stone AA, Schwartz JE, Broderick JE, Deaton A. A snapshot of the age distribution of psychological well-being in the United States. [Research Support N.I.H. Extramural]. *Proc Natl Acad Sci U S A*. 2010; 107(22):9985–9990. [PubMed: 20479218]
- Stone AA, Schwartz JE, Schkade D, Schwarz N, Krueger A, Kahneman D. A population approach to the study of emotion: diurnal rhythms of a working day examined with the Day Reconstruction Method. [Research Support N.I.H. Extramural Research Support, Non-U.S. Gov't Research Support U.S. Gov't, Non-P.H.S.]. *Emotion*. 2006; 6(1):139–149. [PubMed: 16637757]
- White MP, Dolan P. Accounting for the Richness of Daily Activities. *Psychol Sci*. 2009; 20(8):1000–1008. [PubMed: 19549079]

Table 1

Demographic characteristics of sample

Demographic Characteristic	Sample (N = 45)
Age (mean, SD) *	41.4 (12.6)
Female %	77.8%
Race	
White %	86.7%
African American %	2.2%
Asian %	4.4%
Native American %	2.2%
Other %	4.4%
Ethnicity (Hispanic) %	8.9%
Married %	53.3%
Education	
High school graduate %	8.9%
Some college %	13.3%
College graduate or more %	51.1%
Advanced degree %	26.7%
Family Income	
Less than \$20,000 %	2.2%
\$20,000–34,999	8.9%
\$35,000–49,999	15.6%
\$50,000–\$74,999 %	24.4%
Greater than \$75,000 %	48.9%

*
(n = 44), age not provided by one participant

Table 2

Hedonic well-being (HWB) as measured by hybrid short-form (HSF) and day reconstruction method (DRM), and correlation between the two measurement approaches.

HWB	HSF Mean (SD)	DRM Mean (SD)	Effect Size of Difference (Cohen's d)	p value of Difference *	Correlation	P value of Correlation *
Happy	7.04 (2.00)	6.79 (2.00)	0.13	0.269	0.70	<0.001
Engaged	7.42 (2.06)	6.58 (2.13)	0.40	0.040	0.53	0.002
Satisfied	7.40 (1.88)	6.84 (1.86)	0.30	0.068	0.68	<0.001
Worthwhile/ Meaningful	7.53 (2.02)	7.06 (2.01)	0.23	0.246	0.61	<0.001
Stress	4.58 (2.89)	2.62 (1.85)	0.81	<0.001	0.80	<0.001
Worried	3.00 (2.60)	1.70 (1.56)	0.61	<0.001	0.52	0.002
Sad	1.38 (2.11)	0.97 (1.33)	0.23	0.246	0.64	<0.001
Wish You Were Somewhere Else	3.29 (3.25)	2.34 (2.38)	0.33	0.006	0.85	<0.001
Pain	1.09 (2.16)	0.74 (1.86)	0.17	<0.001	0.97	<0.001

* corrected p values using sequentially rejective Bonferroni test (Holm 1979)

Table 3

Time use activities (hours) as measured by hybrid short-form (HSF) and the full episode day reconstruction method (DRM), and correlation between the two measurement approaches.

Time Use Activities	HSF Mean (SD)	DRM Mean (SD)	Effect Size of Difference (Cohen's d)	p value of Difference *	Correlation	P value of Correlation *
Housework	1.33 (1.26)	2.65 (2.29)	0.71	<0.001	0.50	0.002
Work	5.26 (3.93)	4.89 (4.28)	0.09	0.870	0.85	<0.001
Commuting	0.96 (0.97)	1.47 (1.50)	0.40	0.018	0.69	<0.001
Childcare	1.53 (3.31)	1.00 (2.14)	0.19	0.320	0.81	<0.001
Exercising	0.35 (0.48)	0.42 (0.82)	0.10	0.994	0.60	<0.001
TV	2.00 (1.27)	2.17 (1.88)	0.11	0.994	0.43	0.003
Leisure/Recreation	1.26 (1.68)	3.38 (3.90)	0.71	<0.001	0.52	<0.001
Friends/Family	4.72 (3.99)	3.83 (4.08)	0.22	0.315	0.70	<0.001

* corrected p values using sequentially rejective Bonferroni test (Holm 1979)

Table 4

Correlations between time use and hedonic well-being items for the DRM and for the HSF

		Housework	Work	Commuting	Childcare	Exercise	TV	Leisure/ Recreation	Friends/ Family
Happy	DRM	-.22	-.22	.04	-.05	-.09	-.18	-.04	.03
	HSF	-.08	-.30	.07	.06	.14	-.05	.13	.30
Engaged	DRM	-.20	.20	.18	.14	-.14	-.18	-.16	.10
	HSF	-.08	.06	.15	.18	.11	-.30	-.10	.18
Satisfied	DRM	-.14	-.04	.13	-.02	.05	-.34	-.09	.04
	HSF	.13	-.16	.16	.10	.25	-.30	.08	.14
Worthwhile/Meaningful	DRM	-.04	.11	.15	.02	.01	-.31	-.24	-.03
	HSF	.16	.06	.05	.17	.02	-.30	-.09	.19
Stressed	DRM	.04	.48	-.01	.17	-.06	-.15	-.19	-.22
	HSF	.28	.42	.00	.12	.19	-.03	-.23	-.23
Worried	DRM	.12	.09	-.08	.16	.02	-.03	-.15	-.05
	HSF	.19	.04	.04	.28	.25	.29	-.11	.02
Sad	DRM	-.01	.05	-.18	-.08	-.05	-.03	-.05	-.26
	HSF	-.14	.05	.02	-.15	.03	.29	-.04	-.22
Wish You Were Somewhere Else	DRM	.15	.39	-.17	-.09	.07	-.24	-.21	-.30
	HSF	.17	.38	-.07	-.04	.08	-.07	-.17	-.37
Pain	DRM	.13	-.06	-.21	-.15	-.07	-.02	-.02	-.27
	HSF	.15	-.01	-.09	-.16	-.15	.22	.00	-.12

Bolded numbers represent correlations for the DRM and HSF that differ at $p < 0.10$.

Table 5

Correlations between hedonic well-being items for the DRM and for the HSF.

		Engaged	Satisfied	Worthwhile/ Meaningful	Stressed	Worried	Sad	Wish You Were Somewhere Else	Pain
Happy	DRM	.47	.82	.64	-.46	-.64	-.66	-.63	-.41
	HSF	.45	.58	.48	-.57	-.33	-.60	-.53	-.52
Engaged	DRM		.60	.56	.06	-.20	-.34	-.27	-.40
	HSF		.58	.66	.01	-.16	-.47	-.09	-.44
Satisfied	DRM			.85	-.32	-.49	-.52	-.39	-.35
	HSF			.77	-.37	-.35	-.44	-.29	-.47
Worthwhile/ Meaningful	DRM				-.07	-.38	-.40	-.17	-.19
	HSF				-.23	-.35	-.40	-.23	-.31
Stressed	DRM					.58	.45	.62	.25
	HSF					.61	.21	.67	.24
Worried	DRM						.70	.53	.29
	HSF						.20	.34	.12
Sad	DRM							.52	.63
	HSF							.13	.35
Wish You Were Somewhere Else	DRM								.59
	HSF								.43

Bolded numbers represent correlations for the DRM and HSF that differ at $p < 0.10$.