MEASURING HAPPINESS

The strength of association between the ESM and DRM when excluding the carryover effect

Supervisor: M. Hendriks
Second reader: Dr. M.J. Burger

Master's thesis submitted in partial fulfillment of the requirements for the degree of Master of Science

July 22, 2016

By Suzanne Hielckert
416071

Erasmus School of Economics | Behavioral Economics

ABSTRACT

Knowledge about what makes us happy allows for the assisting of individuals in making more informed decisions. Self-reporting of happiness is an often-used method for acquiring data on people's well-being. The Experience Sampling Method (ESM) is an established subjective well-being measure, though it can be burdensome for participants. Another method, the Day Reconstruction Method (DRM), was introduced in 2004 by Kahneman and colleagues (2010) to alleviate this burden. This method requires participants to reconstruct their day retrospectively, and therefore may memory biases come into play. It is therefore important to research the validity of the DRM as compared to the ESM. Previous studies have attempted to compare the two methods by having participants provide both DRM and ESM happiness ratings for the same day, and found that the DRM provides reliable estimates of individual's happiness. In such research design, however, there may be a carryover effect from the ESM to the DRM happiness ratings, due to the fact that both reports are filed on the same day, causing the previously given ESM ratings to pollute the DRM reports. Hence, this carryover effect may overestimate the strength of association between the two types of happiness ratings. This study therefore seeks to explore the strength of association between the ESM and DRM ratings when excluding this carryover effect. To that end, 128 individuals were asked to file DRM and ESM happiness reports on alternate days for one week. The two types of ratings are compared while controlling for all other factors, and the partial correlation is found to be 0.636, which is significantly lower than found by previous comparison studies. Furthermore, several moderators are examined to explore the strength of association between the ESM and DRM, among which timing, activities, location and social setting. The adjusted means of the ESM and DRM happiness ratings differ significantly during each hour of the day, and for about 30% of the examined activities. Additionally, evidence was found for the conservatism bias, stating that affect is less intense when recalled than when actually experienced. We conclude that there is indeed evidence for the carryover effect in previous studies, leading to an overestimation of the strength of association. Although the DRM is still found to be a reliable measure of experienced well-being, one should be wary of the existing discrepancy between the ESM and DRM happiness ratings.

Table of Contents

ACKNOWLEDGEMENTS

I wish to thank the following people for their contribution to this master's thesis; Martijn Hendriks, my supervisor, for his patient guidance, constructive suggestions and editorial input; Dr. M.J. Burger, my second reader, for his willingness to read and evaluate my work; and Joris Kleyheeg for his support with the statistical analysis of my research.

I would also like to extend my thanks to the Happiness Research Organisation, for the effort they put into designing the application for gathering the data, without which this thesis would not have been written.

Finally, I wish to thank my parents, Robert Hielckert and Clara Waardenburg, for their enthusiasm and their endless willingness to test the application.

1. INTRODUCTION

People's well-being is traditionally measured by economic indicators, such as per capita GDP at the societal level, and income at the individual level. However, these indicators fail to capture other aspects that people value in life, such as a sense of safety, belongingness, esteem, and self-actualization (Maslow, 1943). There are researchers who claim that one reaches happiness if these needs are fulfilled, while unfulfilled needs are a source of unhappiness (Durayappah, 2010).

Data about people's happiness is important as it allows for assessing the well-being of certain groups of people, or society as a whole. These data can therefore assist individuals in making more informed decisions. Additionally, life evaluation metrics are used in many research disciplines, such as economics, psychology, sociology and public policy (Frey & Stutzer, 2002; Seligman & Csikszentmihalyi, 2000; Stiglitz et al., 2009). Furthermore, there are objective benefits to happy people, as happiness leads to increased levels of – for instance – altruism, health, creativity, productivity and income (De Neve et al., 2013; Diener & Seligman, 2004). Subjective well-being, or the self-reporting of happiness, is an often-used method for acquiring data on people's well-being (Tay et al., 2014).

One method of happiness assessment is an evaluative measure, which asks questions comparable to: "Taking all things together, how happy are you with your life?". Respondents can then rate how happy they are on a scale from, for instance, extremely unhappy to extremely happy (OECD, 2013). It is important to note, however, that there are many factors that affect peoples' happiness ratings, such as current mood, and the order and framing of questions (Diener et al, 2013). Additionally, several studies concluded that there are certain cognitive biases affecting the subjective well-being ratings. These biases include, for instance, the peak-end-rule and recall bias (Diener, 1984; Durayappah, 2010). Although these biases are present in subjective well-being measures, extensive research has evidenced that evaluative well-being measures are sufficiently valid and reliable to be adopted in policy evaluations and scientific research, when the data is collected with caution (Diener et al., 2013; OECD, 2013).

Currently, there are two often-used methods to measure subjective well-being besides global happiness reporting that attempt to minimize cognitive biases. The first method is the Experience Sampling Method (ESM), and the second is the day reconstruction method (DRM), as developed by Kahneman et al. (2004). Both methods capture experienced well-being, whereas life evaluations, in contrast, capture memorized well-being. The ESM asks participants' attention multiple times a day, where they have to report how they are feeling, what they are doing, where they are, and who they are with**.** This method provides maximized ecological valid result as people report their mood and emotions while they are actually experiencing them, which minimizes recall bias and other cognitive heuristics (Shiffman et al., 2008). It can be a large burden for the

participants, however, to have their activities disrupted six times daily. Additionally, this method does not necessarily give a good overview of the full day, as the participants report only six small samples during the day, rather than reviewing the entire day (Kahneman et al., 2004). The DRM was developed to relieve the participants' burden by only requiring a mood report once a day. This method asks participants to divide the previous day in episodes of activities, and describing for each episode the same aspects as for the ESM: what they were doing, on which location, with whom they were, and how they felt during that episode. The order of these questions, however, is less optimal than for the ESM reports. In ESM reports, the participants are asked first how they are feeling, and subsequently asked what they are doing, where they are and whom is with them. In DRM reports, participants are first asked what they were doing, where they were and with whom, and only after that asked how happy they were during that episode. Hence, the participants are first reminded of the situation they were in before reporting their happiness ratings, which may pollute their happiness reporting. The DRM is fast becoming a key instrument in assessing individuals' happiness levels (Cohen et al., 2003; Kahneman et al., 2006; Knabe et al., 2010; Hendriks et al., 2016. Several studies state that people can reconstruct their previous day well, and that the cognitive biases therefore do not distort the presented image (Robinson & Clore, 2002; Kahneman et al., 2004; Dockray et al., 2010; Diener & Tay, 2014). However, the general literature onrecall bias suggests that people cannot remember everything in detail, and that our memory can be unreliable (Stone et al., 2000). Cognitive biases associated with recall may therefore distort the image provided by DRM.

To date, there are a few studies that investigate the association between ESM and DRM ratings (Kahneman, 2004; Krueger & Schkade, 2008; Dockray et al., 2010; Kim et al., 2013; Tay et al., 2014). These studies report that DRM is a reliable adaptation of ESM, as there seems to be a high correlation between the happiness ratings of the two methods. The abovementioned comparison studies require participants to give both ESM and DRM ratings concerning the same day. As mentioned by Diener and Tay (2014), it is likely that there are carry-over effects when participants are using both ESM and DRM daily (Diener & Tay, 2014; Hendriks et al., 2014). That is, the happiness reports in the ESM condition are likely to influence the ratings given on a later time in the DRM condition. This carry-over effect may be due to confirmation bias, which is the tendency to recall information that is in line with one's prior beliefs (OECD, 2013), or people may use their previous ESM answer as an anchor for rating their happiness later in the DRM condition. Hence, the ESM and DRM ratings may be dispersed more than these comparison studies suggest.

This study seeks to measure the differences in ratings between these two methods without carry-over effects, by asking participants to report mood ratings in the ESM and DRM conditions on alternate days for one week. This way, individuals will not be vulnerable to particularly consistency bias, because the

DRM rating will concern a different day than the given ESM ratings. The research question of this study is therefore:

- What is the strength of association in happiness ratings between Experience Sampling Method and Day Reconstruction Method when eliminating the carry-over effect?

However, the validity of DRM ratings, as compared to the ESM ratings, may depend on the contex, such as recall period, features of the episode, and characteristics of participants (Kahneman & Krueger, 2006; OECD, 2013; Shiffman et al., 2008). Furthermore, the ESM and DRM ratings may be more aligned in the evening than in the morning, or may be more or less aligned for enjoyable activities than for unpleasant activities. An improved understanding of these relationships could help formulate when ESM is a more appropriate method than DRM, and when DRM ratings suffice despite of possible cognitive biases. Hence, in order to approach the main research question, this paper will explore five sub-questions concerning possible moderators of the strength of association between the two methods.

- What is the (partial) correlation between ESM and DRM happiness ratings?
- What is the strength of association between ESM and DRM happiness ratings across time?
- Does the strength of association between ESM and DRM happiness ratings differ between activities?
- Does the strength of association between ESM and DRM happiness ratings differ between locations?
- Does the strength of association between ESM and DRM happiness ratings differ between social settings?

The remaining part of this paper proceeds as follows: chapter two begins by laying out the theoretical framework for the research. Chapter three is concerned with the methodology for this study. The fourth chapter presents the findings of the research, focusing on the main research question and the three sub questions raised in this introductory chapter. The last chapter discusses the findings and concludes.

2. THEORETICAL FRAMEWORK

2.1 Happiness and subjective well-being

Largely, there are two main aspects that are assumed to determine one's happiness level: life satisfaction and the hedonic component, which is related to affect (Alibpour et al., 2012; Kahneman & Krueger, 2006; Ryan & Deci, 2001). Life satisfaction entails how satisfied one is in life in general, and is constructed retrospectively. The hedonic component focuses on affect, or, more specifically, the frequency of experiencing positive and negative emotions. According this approach, one should avoid negative affects (e.g. pain) and attain positive affects (e.g. pleasure) when pursuing happiness (Fisher, 2010; Ryan & Deci, 2001; Sandvik et al., 1993; Waterman et al., 2008).

Regarding the measurement of happiness, direct reports of subjective well-being measures are typically used. Both the ESM and the DRM are self-report measures, and thus subjective as well. Intuitively, this type of measure makes sense, as the nature of happiness is subjective (Kahneman & Krueger, 2006; Alibpour et al., 2012). Additionally, it has been demonstrated that these direct reports of subjective well-being are highly correlated with visible signs of happiness, such as smiling behavior and physiological responses. This finding, together with other statistical tests, suggests that the subjective well-being measure for happiness shows high construct validity (Kahneman & Krueger, 2006; Alibpour et al., 2012; OECD, 2013).

Subjective well-being can be reported both in real-time and retrospectively. Real-time evaluations are correlated with retrospective reports: individuals usually report the same emotions as in real-time – positive or negative (OECD, 2013). The focus in this study lies on experienced well-being measures, rather than positive and negative affect, as the nature of both ESM and DRM are based entirely on such experienced well-being measures (Kahneman et al., 2004). Even though it is correlated with real-time evaluations, retrospective evaluations may provide a distorted image of the experience for different reasons. To fully understand why, it is useful to make a distinction between experienced utility, how people feel during the event, and remembered utility, how people remember the experience once it is over (Kahneman & Krueger, 2006). Remembered utility requires individuals to weigh their feelings during an experience and aggregate them in a certain way. In a perfect world, these individuals would simply sum their real-time utilities to construct an overall evaluation of an event or their life. In reality, however, these retrospective evaluations are subject to several cognitive biases. These biases arise from heuristics that allow us to quickly make sense of our surroundings without carefully having to consider every detail (Kahneman et al., 1982). For instance, these retrospective reports are affected by context cues, such as current weather: sunny days increase these reports significantly (OECD, 2013; Schwarz & Clore, 1983). Additionally, there are a few cognitive biases that

one should be wary of due to the fact that they are both subjective well-being measures. For instance, social desirability plays a role in self-report data: the conscious or subconscious tendency to select the response that is most likely to be conform social norms (Uchida & Kitayama, 2009). Furthermore, one should be wary of response biases, which can rise in many forms such as the tendency to agree, or disagree with questionnaire items irrespective of the question. Or the tendency to solely choose the most extreme or by contrast merely moderate response categories (OECD, 2013).

The next section will give a brief account of the cognitive biases relevant to recall and memory. See the OECD (2013) report for biases present in subjective wellbeing measures in general.

2.2 ESM and DRM

The experience sampling method is designed to collect data on people's affects in real time, and requires people to repeatedly report their behavior and experience of that behavior. More specifically, subjects are prompted at set hours, around six times each day, that they have to describe what they are doing, where they are doing that, who they are with and their current mood. Typically, subjects give mood reports for several days or longer, in order to get an overview of how events and their evaluations vary in time and context (Shiffman et al., 2008). The ESM sampling can be either random, or tailored specifically to the research purpose. For instance, when determining the craving for nicotine after quitting smoking, one could require mood reporting during instances when the subject used to smoke. The main advantage of the ESM, and other real-time data collection methods, is that it eliminates biases involved with (mood) memory and recall. The subjects are reporting on their current mood and activity, which indicates high ecological validity: the findings are generalizable to real-world settings as the data is collected in such setting (Shiffman et al., 2008). A disadvantage to ESM is that it arguably gives no accurate overview of the full day, as the subjects report only on – more or less – six instances during the day, rather than accounting for the full 24 hours (Kahneman et al., 2004). The implication of this disadvantage is that uncommon or short are most likely not recorded. Another negative aspect of this method is the necessity to constantly carry a device, which may seem like both expensive and a strain. However, since nearly everyone has a mobile phone these days, people do not have to wear an additional device on them. A smartphone App can solve both problems. An additional advantage of a smartphone App downloaded on subjects' own cellphones is that the Hawthorne effect, which is the effect that people know they are being watched has on their behavior, is reduced (Hendriks et al., 2014). Perhaps the largest disadvantage of the ESM is that being disrupted from their activities frequently during the day can be burdensome for subjects. This disadvantage makes it difficult to implement a largescale

experiment using this method (Kahneman & Krueger, 2006). This disadvantage is especially problematic for research into well-being at work: regularly looking on one's phone is often negatively valued or even prohibited.

An alternative to ESM that creates less disruption of subjects' daily life is the day reconstruction method, which was introduced by Kahneman et al. (2004). This method was designed to resemble ESM, but in a less burdensome manner, although still more time-consuming than global happiness measures (Tay et al., 2014). Participants are asked to first divide the previous day in episodes of activities. For every episode is then recorded how the person was feeling, what he was doing, if there were an interaction partner and where he was. Another advantage to DRM, besides that is less disruptive of subjects' daily activities, is that the data provides a clear overview of affect per activity or situation (Diener & Tay, 2013). The following section will discuss biases that may be present in the DRM and therefore may distort the image of the truly experienced subjective well-being resulting from this method.

2.3 The (potential) disadvantages of the DRM

A disadvantage to the DRM is that subjects divide the day into episodes of activity. Work, for instance, is one episode, even though mood may drastically change during an 8 hour working day (Diener & Tay, 2013). A second disadvantage is that aggregating data can be difficult, as the different subjects often do not divide their days in the same number of episodes. This effect stems from the fact that people think differently about time, and thus may divide their day into smaller or larger episodes (Zimbardo & Boyd, 1999).

Additionally, DRM relies on recall-based data, for which is known that memory and recall heuristics bias the results. DRM is designed to minimize the effect of those heuristics, as it has a relatively short recall period – a maximum of 24 hours. The question is whether the systematic biases that arise from memory and recall are present in such short period as well, or whether their effect is indeed limited. If these biases are indeed present in DRM data, they form a disadvantage to the method. Since ESM uses merely real-time data, these heuristics solely affect DRM data, to a greater or lesser extent.

There are several cognitive biases that arise from (mood) memory and recall. Table 1 presents an overview of these biases. Firstly, a bias involved in recalling emotions the peak-end rule (Durayappah, 2010; Kahneman, 2000). This bias entails that when reflecting on an episode, subjects' mood ratings are strongly affected by their mood towards the end of the episode, and the most intense emotion experienced during the episode (Durayappah, 2010; Kahneman, 2000). Secondly, duration neglect is affecting recall-based data. People tend to fail accounting for the duration of experiencing a certain affect. Specifically, several studies reported that the duration of the emotion is hardly correlated to the

evaluation of the experience, even though it is common sense that people prefer a longer period of happiness to a shorter one, and a shorter period of pain to a longer one (Schreiber and Kahneman 2000, Kahneman, 2000; Redelmeier & Kahneman, 1996). Thirdly, recall-based data is concerned with the availability heuristic. This bias entails that the ease with which events or affects are retrieved is heavily influenced by readily available examples in one's mind. Mood-congruent recall is based on this heuristic: subjects are more likely to remember experiences for which the affect is consistent with their current mood (Shiffman et al., 2008). For instance, people that are currently happy have increased accessibility of happy memories, and vice versa for sad people. Hence, subjects tend to overestimate certain emotions and underestimate others that do not match their current state of mind. A more general implication of this heuristic is that researchers should be very careful for priming effects when designing a survey or smartphone application: survey context, such as the layout and question order, can affect the response severely as it may prime the subjects to think in a certain direction (Kahneman & Krueger, 2006; OECD, 2013). Fourthly, recall-based data is influenced by the fading affect bias, which states that negative experiences fade out in people's memories quicker than information with positive associations with the same intensity. In terms of recall, the implication is that people are generally less likely to remember negative experiences than positive ones, because they simply are more salient in our memory (Gibbons et al., 2011; Walker et al., 2003). Fifthly, context effects are found to have an effect on retrospective mood reporting. Current weather, for instance, influences ratings: people report to be significantly happier on sunny days (Tay et al., 2014).

Sixthly, emotions in memory are often less extreme than when experiencing them. This is called the conservatism bias, stating that affects are recorded as more extreme in real-time than retrospectively, both for frequency and intensity (Fischhoff et al., 1977; Daniel & Hirshleifer, 1998).

These potential biases may significantly distort DRM-ratings – either negatively or positively, which may cause serious limitations for drawing inferences on such data. Given these potential biases, it is necessary to determine the comparability of the ESM and DRM happiness ratings.

Table 1: Potential mechanisms toward recall bias in the DRM

Bias	*Explanation*
Peak-end-rule	People tend to evaluate an experience on its most intense emotion (peak) and on the emotions felt at the end of the episode (end).
Duration neglect	People tend to fail to incorporate the length of the episode of a certain affect when evaluating an experience.

Availability heuristic	The ease of recalling a certain experience or affect is heavily affected by one's current mood and situation.
Fading affect bias	People tend to forget information associated with negative emotions more easily than positive ones
Context effects	Memory is dependent on current context, - e.g. weather – such that past experiences that were in a similar context are more easily remembered.
Conservatism bias	Affect is generally less intense in memory than when experienced in real-time.

2.4 Comparison studies

Thus far, there are several studies that have attempted to determine whether data based on recall, or specifically the day reconstruction method, brings about the same results as the experience sampling method despite its dependence on memory.

Kahneman et al. (2004), upon introducing the day reconstruction method, begun to reveal the utility of DRM by collecting DRM ratings from 909 working women, and compared these with established ESM data with similar affect categories. They concluded that the results from their data collection were comparable to the established ESM data set. First, the "tiredness" ratings were extremely similar for both datasets; both V-shape reaching a low around 12pm. Second, the correlations of affects were high, although higher for positive emotions (0.7) than for negative ones (0.4). Taking these findings together, the DRM appears a good approximation of ESM ratings. A limitation of this study is, however, that the DRM and ESM ratings were reported by two different samples and at different moments in time.

Dockray and colleagues (2010) attempted to validate the DRM by comparing these reports to ESM ratings. Specifically, they examined 94 working women, asking them to give six ESM ratings and one DRM rating per 24 hours of one working and one leisure days – hence, a total of 2 days. The participants could rate their mood in several affect categories, namely happiness, stress, tiredness, anger and frustration. Consequently, they compared the average DRM episode rating of certain hour to ESM rating of that time. They result that the correlations between ESM and DRM ratings vary between 0.52 and 0.79, upon which they conclude that both intensity and variation of affect are similar for ESM and DRM, and that DRM thus is a valuable measure for everyday experience when ESM is too costly or impractical.

A problem with this study, and others with similar research designs and similar results (e.g. Hedges et al., 1985) is that ESM and DRM reports were provided for the same episodes. It is likely that the ESM mood reports earlier influence the DRM report the of that day (Diener & Tay, 2013; Hendriks et al.,

2014). Firstly, the subjects pay more attention to their emotions due to the ESM reports than they would if they had not been participating, which increases the accuracy of the image provided by the DRM ratings. Secondly, the reported ESM ratings are likely to still be in the subjects' memory when they are rating their moods in the DRM report (ibid). The consistency bias may come into play then, which is the tendency to appear rational ensure consistency with previously reported answers (OECD, 2013). Research has suggested that the consistency bias arises especially when participants have to retrospectively construct their attitudes and behaviors – which are key elements in the DRM (Podsakoff et al., 2003). Another possible source of contamination is the confirmation bias, which is the tendency to favor or recall information – or in this case, affects – that are consistent with one's prior beliefs. Research has suggested that information that is inconsistent with one's self-image is less likely to be encoded in the memory (Swann et al., 1999). This may suggest that, when prior beliefs about one's day are shaped in the ESM ratings, DRM ratings may be heavily affected. Hence, the DRM ratings may be seriously affected by the ESM reports if DRM ratings are collected concerning the same day. These potential hazards of the carryover effect to DRM ratings are also listed in table 2.

The current study more critically examines the following hypothesis specified and tested in previous research by excluding the carry-over effect of ESM evaluations to the DRM evaluation.

H0: DRM ratings are on average not statistically significantly different from ESM ratings

Table 2: potential problems of the carry-over effect between ESM and DRM ratings on the same day

Potential problem	*Explanation*
Attention	Previously given ESM report increases attention to one's mood during the day, causing increased accuracy in DRM report
Saliency	People may still remember their ESM mood rating for a specific time or activity, affecting their DRM report.
Consistency bias	People like to be consistent in their answers, which may lead to distorting the DRM report based on their memory of the ESM reports earlier that day
Confirmation bias	Information that is inconsistent with self-image, or view on specific experience, is less strongly encoded in memory.

3. METHODS

3.1 Sample

Participants in this study were recruited according to convenience sampling. The participating individuals were told that the research involved questions about their demographics, and the tracking of their happiness. Additionally, they were instructed that their answers would remain confidential, and they could opt out at any moment. To deter opting out, the subjects were told that they could win a small prize if they listed their e-mail addresses upon completion. A total of 332 participants downloaded the application and filled in the baseline questionnaire. However, following Dockray et al. (2010), subjects with chronic illnesses were excluded from the sample. Moreover, to examine the within-subject difference between ESM and DRM ratings, subjects were included solely if they had filed both ESM and DRM reports. A total of 128 participants fulfilled these requirements and were included in the analysis. Together, they have filed a total of 3519 reports. The number of episode ratings per participant ranges from 4 to 89, with an average of 27.5 and standard deviation of 19.06. These reports include a total of 963 ESM episode ratings and 2556 DRM episode ratings. The average number of DRM episodes per day is 9.91.

3.2 Procedure

As recommended by the OECD (2013), participants are asked to fill in a baseline questionnaire prior to commencing the ESM and DRM assessments. This survey starts by asking subjects about their current overall happiness and life satisfaction, which they can answer on a 11-point scale. Subsequently, they are asked to provide some basic information about themselves, such as age, gender, religion, education level, family status, job status, and household income. The survey then goes on to ten statements for which the subjects have to decide to what degree they are applicable to their personal situation on a 7-point scale – ranging from 'strongly disagree' to 'strongly agree'. Statements include, for instance, "I am trustworthy, disciplined" or "I am friendly, warm". Together, the answers to these statements provide a personal score for the big five personality traits: conscientiousness, openness, extraversion, emotional stability, and agreeableness.

After the baseline questionnaire, the participants are asked to give ESM and DRM ratings on alternate days, for a total of seven days. Half of the subject start with the ESM assessment on the first day, whereas the other half start with DRM ratings. This division of participants in either group is determined randomly. This study uses the same set of menus for both ESM and DRM reports, with the same wording for mood, activities, locations, and social setting. This decision was

made because it allows for clear comparison of the DRM and ESM ratings. This way, differences in ratings that stem from interpretation are limited. Additionally, usage of the same 11-point happiness scale in both conditions facilitates comparison as well. The order of the questioning is different for the two methods, however. During the ESM assessment, the participant is first asked how he is feeling, and then what he is doing, where he is and whether there is an interaction partner. During the DRM assessment, the participant is first asked what he has done that day, where he was and with whom, and only then asked how he was feeling during each episode.

For both methods, a single-item question for the assessment of happiness was used rather than the affect balance scale for several reasons. Firstly, because it is unclear how to correctly weigh each affect to get a valid aggregated mood overview (Hendriks et al., 2014; White & Dolan, 2009). By using a single-item question ("how do/did you feel?") the participants can weigh their feelings themselves and decide what emotions were more important than others for their happiness rating. Secondly, the single-item scale is found to be strongly correlated with multi-item results, while the single-item scale is less burdensome for the participants (Hendriks et al., 2014; Knabe et al., 2010).

This research design differs from previous studies that have compared ESM and DRM ratings, in that DRM and ESM assessments are used on alternate days, rather than give both ratings each day (Dockray et al., 2010). The design is different in order to eliminate the carryover effect between the two ratings.

3.3 ESM Assessments

The ESM assessment prompts participants to report their happiness six times a day on a 11-point scale ranging from extremely unhappy to extremely happy. Specifically, they are asked "How do you currently feel?"[1] Furthermore, subjects are shown multiple menus, on which they indicate what they are doing, where they are, and who is with them. Below, figures 1 provide examples of what the ESM report in the application.

[1] In Dutch: hoe voelt u zich op dit moment?

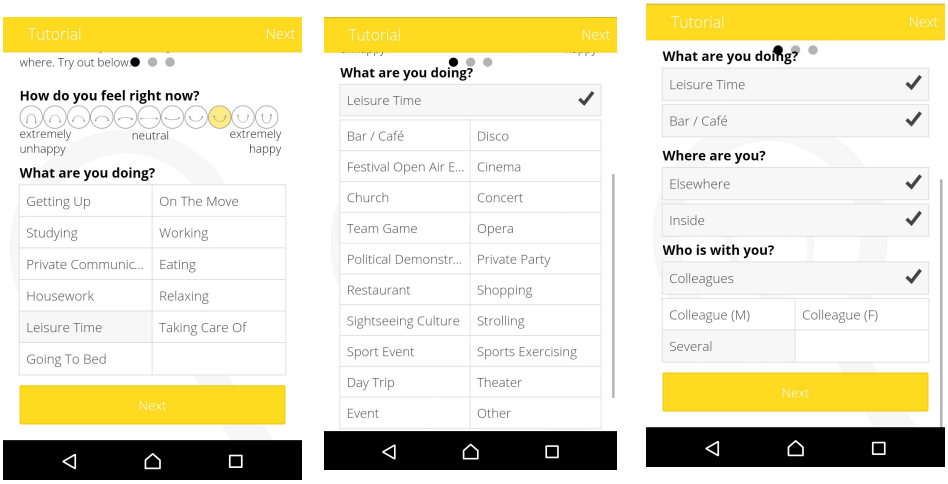


Figure 1: example ESM report

3.4 DRM Assessments

The DRM assessment requires participants to first reconstruct their day by dividing it into episodes. The notification for the DRM assessment prompted the subjects' attention at 9PM or later in the evening, depending on the subjects' application settings. In the DRM assessment, the subjects could indicate their activity, location and interaction partner in the same set of menus as in the ESM condition. Each episode lasts at least 15 minutes, and ends whenever the situation changes – that is, when the location, activity or interaction partner have changed. In order to proceed to the ratings per episode, the whole day must be accounted for. When the activities, locations and interaction partners are specified for whole day, the subject is asked to rate their happiness during each episode on a 11-point scale, ranging from extremely unhappy to extremely happy. Specifically, they were asked: "Now, rate how happy you were during the individual episodes."[2]

Subsequently, the participants are asked how happy they had felt during the whole day, and what kind of day it was – work or leisure day. Additionally, there is an optional field where participants can enter notes. Below, figure 2 provide an example of an empty and a filled in DRM diary.

[2] Author's translation. In Dutch: Geef nu aan hoe gelukkig u was gedurende de individuele gebeurtenissen.

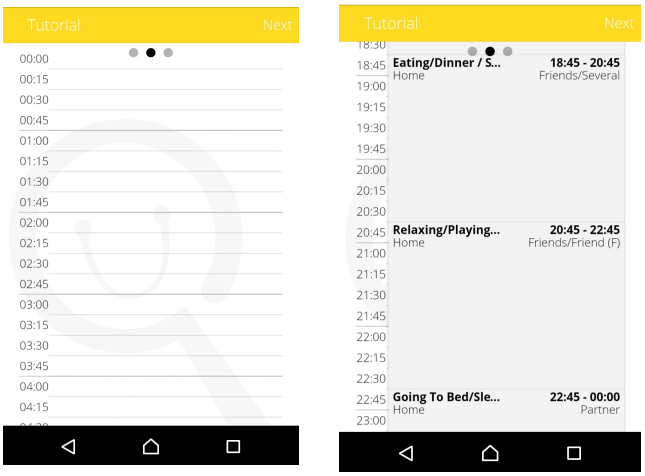


Figure 2: empty DRM report

3.5 Statistical Analysis

The difference between this study's design and Dockray et al.'s (2010) is that in their study there are corresponding DRM and ESM evaluations for the same period, whereas in this study they are provided for different days. Hence, the ESM and DRM reports cannot be matched in the same fashion as in Dockray et al. (2010), who linked the timing of the DRM happiness rating to its corresponding ESM rating in that hour on the same day.

Even though the menus for activity, location and social setting are the same between ESM and DRM assessment, there is a difference in reporting: the ESM ratings consist of six different points in time each day, whereas the DRM ratings are continuous. In order to compare ratings, we statistically control for all other possible causes of dispersion of the two types of ratings, such as location, activity, individual characteristics, timing and social setting. Additionally, timing was reduced to 24 one-hour timing categories. For instance, an ESM rating reported at 4.37pm is positioned in the 4PM-5PM period. The timing of the DRM ratings is chosen as follows. For sleeping, the happiness rating upon waking up is used: if a participant slept until 7.30AM, this episode was registered in the 7AM-8AM period. For all other activities, the average time is chosen: for an activity that lasted from 8.30PM until 10.30PM, it is registered as 9.30PM and therefore in the 9PM-10PM period.

A basic OLS model is estimated to approach the main question on the strength of association between the ESM and DRM ratings. Plus, three extensions for this model are constructed that are linked to sub-questions 2, 3 4, and 5. In these extensions, which were formulated to answer the main research question, interaction effects are created between the ESM/DRM dummy and respectively timing, activities, locations and social settings. For each model, the standard errors are clustered in the analysis to ensure that the observed variances concern within individual residuals, rather than between individuals. This way, the total variation

will not be underestimated. The standard errors are clustered by adding the command 'margins' to the regression in STATA. Below, table 3 provides an overview of the models. These interaction effects are created to determine the significance of the differences in mean ESM and DRM happiness ratings. Knowing this significance of difference allows us to estimate the strength of association between the ESM and DRM ratings across timing, activities, locations and social settings. Furthermore, several control variables have been included in the models. First, individual dummies are included, to control for the person fixed effects. Second, date dummies are included to control for the calendar fixed effects, since the participants have provided happiness reports on 12 different days. Third, the results from Dockray et al. (2010) indicated that happiness ratings are significantly affected by what type of day it is: a leisure day or a working day. Therefore, a dummy for leisure days is included to control for this effect. Together, this information will support our tentative conclusion concerning the strength of association between ESM and DRM ratings in general.

The first sub-question concerns the partial correlation between ESM and DRM happiness ratings. To find this correlation, the adjusted means for each participant for both their ESM and DRM ratings are calculated. These adjusted means are the mean happiness rates when controlling for clustered standard errors, activities, locations, social setting, timing, date, and type of day. The correlation between these adjusted means of the whole sample is subsequently computed to answer the first sub-question. All analyses are carried out using STATA v.14.

Table 3: Overview of models in analysis

	Model 1.0	*Model 1.1*	*Model 1.2*	*Model 1.3*	*Model 1.4*
Dummy ESM/DRM	X	X	X	X	X
Individual dummies	X	X	X	X	X
Activity dummies	X	X	X	X	X
Location dummies	X	X	X	X	X
Social setting dummies	X	X	X	X	X
Timing dummies	X	X	X	X	X
Date dummies	X	X	X	X	X
Leisure day dummy	X	X	X	X	X
*Timing*Dummy ESM/DRM*		X			

*Activity*Dummy ESM/DRM*	X		
*Location*Dummy ESM/DRM*		X	
*Social setting*Dummy ESM/DRM*			X

4. RESULTS

4.1 Descriptive statistics

The sample characteristics are presented in table 4. The average ESM rating is 6.73, whereas the average DRM rating is 6.83. Furthermore, the standard error for the ESM ratings was 0.05, whereas the standard error for the DRM ratings was 0.03. Participants were on average 38 years old, and around 90% of them was born in the Netherlands. The bulk of the participants is in a relationship of is married, and around 45% of the subjects has children. The majority of the sample has an educational level of higher vocational or university, and about 20% of the participants was a student at the time of the data collection. The participating subjects rate their overall happiness and life satisfaction quite high: respectively 7.29 and 7.45 on a scale from 0 to 10. The data points were collected during 12 consecutive days.

Table 4: Descriptive statistics

Variable	*Mean (SD)/%*
Age in years	38.3 (12.05)
Gender (% male)	26.6
Born outside NL	9.5
Household situation	
Single	26.2
In relationship	30.1
Married, living together	38.6
Married, living separately	1.1
Divorced	4.0
Has children	45.8
Is Atheist	49.9
Education level	
High school	7.1
Intermediate vocational	6.3
Higher vocational	38.2
University	48.4
Has job (% yes)	65.6
Student	18.9
Monthly income	
< €1,500	40.7
€1,500-€3,000	47.2

> €3,000	8.8
Personality (1-7)	
Extraversion	4.99 (1.43)
Conscientiousness	5.57 (1.09)
Openness	5.59 (0.99)
Agreeableness	4.43 (0.89)
Emotional stability	4.81 (1.42)
Overall happiness (0-10)	7.29 (1.55)
General life satisfaction (0-10)	7.45 (1.17)

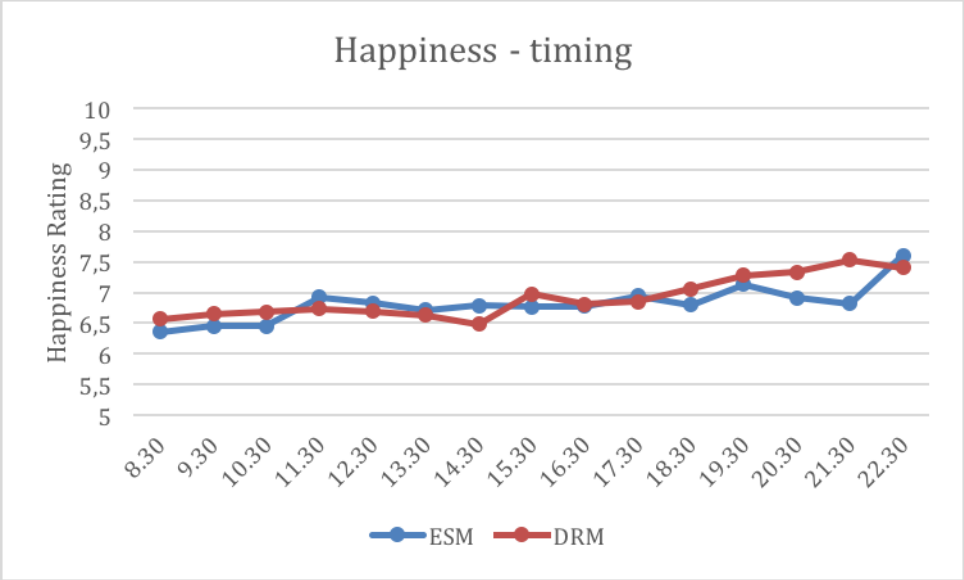
4.2 Partial correlation

In order to answer the first sub-question, concerning the partial correlation between the ESM and DRM happiness ratings, the adjusted means for each participant has been computed. As mentioned in the procedure section, these adjusted means account for clustered standard errors, activities, locations, social setting, timing, date, and type of day. These two average ratings for all 128 subjects is presented in table 6 in the appendix. It is The correlation between all ESM ratings and all DRM ratings is 0.636. According to Evans (1996), this can be considered as a strong positive relation between the two types of ratings. The results of the basic model regression are presented in table 6 of the appendix of this article.

4.3 Moderators

The second sub-question is what the strength of association between DRM and ESM ratings over time. In order to answer this question, the adjusted means for both the ESM and DRM happiness ratings were computed for each hour. These adjusted means were based on OLS model 1.1. Figure 1 illustrates the mean ESM and DRM ratings for 8AM until 11PM. The timeframes were averaged or clarity purposes – that is, the time category 8AM-9AM is noted as 8.30AM in this graph. Timeframes before 8AM and after 11PM were not presented in the graph due to a lack of observations. It can be seen that happiness ratings are relatively low in the morning, and slowly increase during the day. Furthermore, the figure shows that the DRM ratings are not systematically higher or lower than the ESM ratings at the corresponding timeframe. In the morning and in the evening, DRM ratings are on average higher. However, between 11AM and 5PM, ESM ratings are generally higher. The adjusted mean ESM and DRM ratings and the significance of their difference is presented in table 5.

Figure 1: adjusted means of DRM and ESM happiness ratings over the day.



In order to answer the other sub-questions concerning activities, locations and social settings as moderators for the strength of the ESM/DRM association, the differences between adjusted mean DRM and ESM ratings was computed. In order to control for the existing difference in average ESM and DRM rating – 6.73 for the ESM and 6.83 for the DRM – 0.1 is added to all ESM happiness ratings in the sample. This way, the difference between the two ratings does not stem from an inherent difference, but rather from the difference in rating for that specific activity, location or social setting. Table 5 presents these adjusted means – based on the OLS models 1.3, 1.4, and 1.5 – of the ESM and DRM happiness ratings for each category, location and social setting. When comparing these results to the results without adding the value 0.1, which are presented in table 7 in the appendix, it becomes clear that the results do not change significantly.

Firstly, the activities will be discussed. There is a significant difference between ESM and DRM ratings for transportation. For the activities traveling by car and by bicycle, the ESM ratings are significantly higher than their DRM counterparts. For studying, the ESM rating is significantly higher than the DRM rating as well. The activity work was divided in two parts: paid and unpaid work. For unpaid work, the difference between the two ratings is not significant, whereas the ESM rating for paid work was significantly higher than the DRM rating. The activity private communication, however, has a significantly higher DRM rating. The absolute difference between these two average ratings is relatively high (5.83 for ESM and 7.33 for DRM). The activity food was divided into breakfast, lunch, and dinner. From these four categories, only breakfast yielded a significant result, with a higher ESM than DRM rating. Turning now to the activity housekeeping, tidying up, cleaning and other housekeeping chores all three have significantly higher ESM than DRM ratings. Additionally, the ESM rating for

multimedia is significantly higher than the DRM as well. Hence, almost all significant differences had higher ESM than DRM ratings. Solely the activity private communication has a significantly higher DRM rating. Also, in sum, the activities with a significant difference between ratings are almost all in the categories waking up, studying, transport, working, housekeeping, eating and private communication. The categories relaxing, leisure, caring for someone and going to sleep together yielded solely one significant result. Secondly, the social setting as a moderating effect in the ESM/DRM association was assessed. The computation of the significance of the difference between the adjusted ESM and DRM means did not uncover any significant results. Thirdly, comparing the adjusted means for ESM and DRM ratings for the different locations, yielded solely one significant result. The average DRM rating for being at home was significantly higher than the ESM rating.

Table 5: Adjusted means, their difference, and the significance of their difference. The number of observations per variable are noted as well. The value 0.1 was added to all ESM ratings to control for inherent differences between the ESM and DRM ratings. The variables are included in this table solely when the adjusted means could be estimated for both ESM and DRM ratings. *=p<0.10, **=p<0.05, ***=p<0.01

Type of event	***Average ESM rating***	***Average DRM rating***	***Mean difference***	***Significance of difference***	***N***
	N=963	**N=2556**			
Timing					
8AM-9AM	6.46	6.57	-0.11	***	252
9AM-10AM	6.55	6.65	-0.10	***	212
10AM-11AM	6.94	6.68	0.26	***	216
11AM-Noon	7.02	6.73	0.29	***	201
Noon-1PM	6.93	6.69	0.24	***	233
1PM-2PM	6.81	6.63	0.18	***	219
2PM-3PM	6.89	6.48	0.41	***	181
3PM-4PM	6.87	6.97	-0.10	***	187
4PM-5PM	6.88	6.81	0.07	***	204
5PM-6PM	7.04	6.85	0.19	***	237
6PM-7PM	6.9	7.06	-0.16	***	240
7PM-8PM	7.23	7.28	-0.05	***	255
8PM-9PM	7.01	7.33	-0.32	***	240
9PM-10PM	6.92	7.53	-0.61	***	102
10PM-11PM	7.70	7.40	0.30	***	45
11PM-midnight	9.20	7.46	1.74		16
Midnight-1AM	7.67	7.01	0.66	***	67
1AM-2AM	8.38	6.93	1.45	***	20
2AM-3AM	9.33	6.99	2.34	***	23
6AM-7AM	7.01	6.10	0.91	***	105
7AM-8AM	6.34	6.40	-0.06	***	191

Activities					
Waking up	6.02	6.114	-0.09		317
Transport	7.06	6.54	0.52	***	329
Car	7.28	6.54	0.74	***	156
Bicycle	7.27	6.75	0.52	**	84
Other	6.61	6.33	0.28		89
Studying	6.57	5.93	0.64	**	99
Working	6.84	6.60	0.24		480
Paid	6.93	6.66	0.27	**	370
Unpaid	6.57	6.38	0.19		110
Private communication	5.83	7.33	-1.50	*	82
Eating	7.13	7.13	0.00		471
Breakfast	7.03	6.69	0.34	**	130
Lunch	7.14	7.10	0.04		106
Dinner	7.3	7.38	-0.08		235
Housekeeping	6.67	6.42	0.25		292
Tidying up	6.79	6.13	0.66	**	47
Groceries shopping	5.31	6.57	-1.26		58
Preparing food	6.77	6.88	-0.11		63
Cleaning	6.72	5.97	0.75	***	57
Other	6.73	6.27	0.46	**	67
Relaxing	7.19	7.25	-0.06		553
Cuddling	7.54	8.80	-1.26		25
Multimedia	7.42	7.31	0.11	*	295
Napping	6.18	7.24	-1.06		36
Sex	8.68	8.78	-0.10		28
Playing games	7.20	7.08	0.12		46
Other	6.83	7.17	-0.34		123
Leisure	7.31	7.52	-0.21		258
Bar/Cafe	8.00	7.96	0.04		41
Shopping	7.42	7.29	0.13		27
Hiking	8.14	7.78	0.36		25
Playing sports	7.62	7.71	-0.09		52
Other	7.04	7.43	-0.39		113
Caring for someone	6.37	6.27	0.10		231
Going to bed	6.02	6.53	-0.51		353
Social setting					
Alone	6.57	6.54	0.03		1428
Work-related	7.86	7.33	0.53		53
Children	6.76	6.89	-0.13		292

Friends	7.63	7.78	-0.15		244
Colleagues	7.01	6.84	0.17		296
Siblings	6.53	7.01	-0.48		23
Parents	6.86	6.98	-0.12		43
Partner	8.18	7.62	0.56		612
Other family members	7.05	6.96	0.09		433
Other	6.78	6.42	0.36		95
Locations					
At home	6.71	6.80	-0.09	*	2167
At work	7.01	6.86	0.15		386
Elsewhere - Inside	7.24	6.87	0.37		446
Elsewhere - Outside	6.89	6.94	-0.05		321
Transport	6.89	6.67	0.22		199

5. DISCUSSION

This study set out to determine what the strength of association is between ESM and DRM happiness ratings, when eliminating the carryover effect, which may have distorted previous studies that had their subjects provide both ESM and DRM reports for the same day. This research focused on within-person variance between the two methods, to eliminate the possibility that the observed variance between the two ratings is caused by participant heterogeneity. The partial correlation between the mean ESM and DRM ratings per participating individual is 0.636, which is a relatively strong correlation. However, the correlation found in this study is lower than the correlation Dockray et al. (2010) found after adjustment for attenuation, which was between 0.71 and 0.90 for different circumstances. This result confirms our expectation that the ESM and DRM are less strongly associated than Dockray et al. (2010) concluded, and that their overestimation of this correlation is likely due to a carryover effect between the two ratings.

Furthermore, research into the two types of ratings on different moments during the day provides an overview of the strength of association over time. The correlation here seems to be merely moderate. The difference in adjusted mean happiness rating between the two methods is significant for each timeframe with sufficient observation. Although the significance of difference is consistent, one method is not systematically higher or lower than the other: the DRM ratings are higher during some hours, and the ESM ratings at other times. However, even though the difference in rating is statistically significantly different, the ratings are not that far apart absolute measures. During the day, the two methods are farthest apart in the timeframe 9PM-10PM, with 0.71 points apart on a scale from 0 to 10. This finding is in contrast to our expectation about the two ratings being stronger associated as the day passes. Additionally, this difference in rating may be insufficient to refrain from using the DRM. Also, it was expected that the strength of association between the DRM and ESM ratings was stronger in the evening than in the morning, because the recall period for the DRM report is smaller in the evening. In contrast, the largest difference between the two average ratings was at night. Hence, it seems that differences in recall period with a maximum of 24 hours do not necessarily affect the strength of association between ESM and DRM.

Research into the strength of association between ESM and DRM ratings with respect to different activities provides mixed results. For about 30% of the activities in the analysis with sufficient observations have significantly different ESM than DRM ratings. All but one of these significant results have a higher ESM than DRM rating. Interestingly, most of the activities with a significant difference between the two ratings are relatively unpleasant. The categories studying, working, housekeeping and commuting, which yielded the most significant activities, are assumed to be viewed less entertaining than leisure and relaxing. Private communication, which may be more pleasant than the abovementioned

categories, was the sole activity with a significant rating difference where the DRM rating was higher than its ESM counterpart. This may indicate that the DRM yields higher ratings for pleasant activities, conditional on the fact that the ratings are significantly different.

Considering how the strength of association between ESM and DRM ratings varies across social settings, several interaction partners have been examined. This part of the analysis did not uncover any statistically significant results. When examining whether the location of the episode affects the strength of association between ESM and DRM ratings, it was found that the two ratings for being at home differ significantly, where the DRM rating is generally higher than the ESM rating. The other four locations yielded sufficiently similar reports.

It is interesting to note that the standard deviation for DRM happiness ratings is on average smaller than for ESM happiness ratings – 0.05 for ESM and 0.03 for DRM. This result provides evidence for the conservatism bias, which states that emotions are less intense in one's memory than when experienced in real-time.

Taken together, these results suggest that the DRM reports do provide similar happiness ratings as the ESM reports. The partial correlation is lower than suggested by Dockray et al. (2010), however, and the two types of ratings can differ significantly across timing and activities. Therefore, it is advisable to carefully consider the circumstances in which the DRM will be used as a substitute for the ESM. This is the first study to investigate the strength of association between the two methods while excluding the carryover effect.

Notwithstanding its contributions, there are also limitations to this research study. First, even though it was attempted to reach an audience as wide as possible, the sample in this study was a convenience sample rather than a random sample. Due to the fact that solely people that were willing or interested in participating in the research are in the sample, the results may not be generalizable to the whole population. Second, the application with which the data was collected did not function perfectly. Some subjects did not receive any notifications, whereas others received notifications solely on the first couple of days. Without this bug, there would have been more data-points per individual, allowing for higher validity of the data. A reminder-notification that is built in the application would also stimulate participants to complete the whole seven days. Third, even though there was no carryover effect from the ESM to the DRM ratings, the fact that individuals in the study knew they were participating in a happiness research may have influenced the DRM ratings. The participants are likely to have been more aware of the emotions they experienced during the day, leading to more accurate DRM reports than otherwise would be the case.

Despite these limitations, this study does extend our knowledge of the validity of DRM ratings as compared to ESM ratings. In conclusion, by having participants file the ESM and DRM happiness ratings on alternate days, the carryover effect is eliminated, allowing for a more reliable comparison between

the two methods. The examination of these two types of happiness ratings indicates that there indeed was a carryover effect in previous comparison studies, leading to an overestimation of the strength of association. Further analysis suggested that although there are significant differences between the two types of happiness ratings across time during the day and activities, the ESM and DRM do provide similar results. The findings of this study provide a valuable addition to the existing literature about the ESM and DRM as tools to explore people's daily activities and emotions, and suggest that although for some situations the strength of association between the two is sufficiently high, one should be wary when using the DRM as a substitute for the ESM when researching happiness.

6. REFERENCES

Alipour, A. (2012). What is happiness. *Interdisciplinary Journal of Contemporary Research in Business 3*(660), 2044-4621

Bergner M, & Rothman ML. (1987). Health status measures: an overview and guide for selection. *Annual Review Of Public Health, 8*, 191-210.

Cohen, S., Doyle, W.J., Turner, R., Alper, C.M., & Skoner, D.P. (2003). Sociability and susceptibility to the common cold. *Psycholocial Science, 14*(5)*,* 389-395

Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under-and overreactions. *The Journal of Finance, 53*(6), 1839-1885.

De Neve, J.-E., Diener, E., Tay, L., & Xuereb, C. (2013) The objective benefits of subjective well-being. In Helliwell, J., Layard, R., & Sachs, J., eds. *World Happiness Report 2013.* New York: UN Sustainable Development Solutions Network.

Diener, E. (1984). Subjective well-being. *Psychological Bulletin, 95,* 542-575

Diener, E. (2000). Subjective well-being. *American Psychologist, 55*, pp. 34–43

Diener, E., Inglehart, R., & Tay, L. (2013). Theory and Validity of Life Satisfaction Scales. *Social Indicators Research*, *112*(3), 497-527. doi:10.1007/s11205-012-0076-y

Dockray, S., Grant, N., Stone, A.A., Kahneman, D., Wardle, J. & Steptoe, A. (2010). A comparison of effect ratings obtained with ecological momentary assessment and the day reconstruction method. *Social Indicators Research, 99*(2), 269-283

Diener, E., & Seligman, M. E. P. (2004). Beyond Money: Toward an Economy of Well-Being. *Psychological Science In The Public Interest, 5*(1), 1-31.

Diener, E., & Suh, E. M. (Eds.). (2000). *Cultural and subjective well-being*. Cambridge, MA: MIT Press.

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin*, *125*(2), 276-302

Diener, E. & Tay, L. (2014). Review of the day reconstruction method (DRM). *Social Indicators Research, 116*(1), 255-267

Di Tella, R. and R. MacCulloch (2006). Some uses of happiness data in economics. *Journal of Economic Perspectives 20*, 25–46.

Durayappah, A. (2010). The 3p model: A general theory of subjective well-being. *Journal of Happiness Studies*

Eich, E., Reeves, J.L., Jaeger, B. and Graf-Radford, S.B., (1985). Memory for pain: relation between past and present pain intensity, *Pain, 23,* 375-379.

Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences*. Pacific Grove, CA: Brooks/Cole Publishing.

Finkenauer, C. & Rimé, B. (1998). Socially shared emotional experiences vs. emotional experiences kept secret; differential characteristics and consequences. *Journal of social en Clinical Psychology, 17*(3), 295-318

Fisher, C. D. (2010). Happiness at Work. *International Journal Of Management Reviews*, *12*(4), 384-412

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance, 3*(4), 552-564

Freeman, R. (1978). Job Satisfaction as an Economic Variable. *American Economic Review. 68*(2), pp. 135– 41.

Frey, B.S. & Stutzer, A. (2002), What can economists learn from happiness research? *Journal of Economic Literature 40*, 402–435.

Gibbons, Jeffrey A.; Lee, Sherman A.; Walker, W. Richard (2011). The fading affect bias begins within 12 hours and persists for 3 months. *Applied Cognitive Psychology 25* (4): 663–672

Godoy, R., Reyes-García, V., Gravlee, C. C., Huanca, T., Leonard, W. R., McDade, T. W. (2009). Moving beyond a snapshot to understand changes in the well-being of native Amazonians: Panel evidence (2002-2006) from Bolivia. *Current Anthropology, 50*(4), 560-570.

Hedges, S. M., Jandorf, L., & Stone, A. A. (1985). Meaning of daily mood assessments. *Journal Of Personality And Social Psychology*, *48*(2), 428-434.

Helliwell, J., Layard, R. & Sachs, J. (eds.) (2012) *World happiness report.* The Earth Institute, Columbia University, New York, USA.

Hendriks, M., Ludwigs, K. & Veenhoven, R. (2014). Why are locals happier than immigrants? The role of daily life. *Social Indicators Research,* 1-28

Izquierdo C. (2005). When "health" is not enough: societal, individual and biomedical assessments of well-being among the Matsigenka of the Peruvian Amazon. *Social Science & Medicine,* 61(4), 767-783.

Juster, F. T., Stafford, F. P., & University of Michigan (Ann Arbor, Mich). (1985). *Time, goods, and well-being.* Ann Arbor, Mich.: University of Michigan.

Kahneman D. and A. Krueger (2006). Developments in the Measurement of Subjective Well- Being. *Journal of Economic Perspectives* 20(1), 3–24.

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

Kahneman, D., Krueger, A.B., Schkade, D.A., Schwartz, N. & Stone, A.A. (2006). Would you be happier if you were richer? A focusing illusion. *Science, 312*(5782), 1908-1910

Kahneman, D., & Riis, J. (2005). Living, and thinking about it: Two perspectives on life. In F. Huppert, N. Baylis, & B. Keverne (Eds.), *The science of well-being* (pp. 285–304). Oxford, UK: Oxford University Press.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases.* New York, NY: Cambridge University Press.

Kim, J., Kikuchi, H. & Yamamoto, Y. (2013). Systematic comparison of between ecological momentary assessment and day reconstruction method for

fatigue and mood states in healthy adults. *British Journal of Health Psychology, 18,* 155-167

Kim, M. S., Kim, H. W., Cha, K. H., & Lim, J. (2007). What makes Koreans happy? Exploration on the structure of happy life among Korean adults. *Social Indicators Research,* 82(2), 265-286.

Knabe, A., Ratzel, S., Schob, R. & Weimann, J. (2010). Dissatisfied with life but having a good day: time-use and well-being of the unemployed. *The Economic Journal, 120,* 867-889

Krueger, A.B. & Schkade, D.A. (2008). The reliability of subjective well-being measures. *Journal of Public Economics, 92,* 1833-1845

Maslow, A.H. (1943) "A Theory of Human Motivation." *Psychological Review 50*(4), 370-396.

Mayo, F.H. (1983). Personality traits and the retrieval Of positive and negative memories. *Personality and Individual Differences, 4,* 465-471.

Oishi, S., Diener, E., Choi, D. W., Kim-Prieto, C., & Choi, I. (2007). The dynamics of daily events and well-being across cultures: When less is more. *Journal of Personality and Social Psychology, 93*, 685–698

OECD (2013), *OECD Guidelines on Measuring Subjective Well-being*, OECD Publishing. http://dx.doi.org/10.1787/9789264191655-en

Podsakoff, P.M., MacKenzie, S.B. & Lee, J. (2003). Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology, 88*(5), 879-903

Redelmeier, D.A., & Kahneman, D. (1996). Patients' memories of painful medical treatments: real-time and retrospective evaluations of two minimally invasive procedures. *Pain, 66*(1), 3-8.

Robinson, M.D. & Clore, G.L. (2002). Belief and feeling: evidence for an accessibility model of emotional self-report. *Psychological Bulletin, 128*(6), 934-960

Ryan, R. M., & Deci, E. L. (2001). On Happiness and Human Potentials: A Review of Research on Hedonic and Eudaimonic Well-Being. *Annual Review Of Psychology*, *52*(1), 141-166

Sandvik, E., Diener, E., & Seidlitz, L. (1993). Subjective Well-Being: The Convergence and Stability of Self-Report and Non-Self-Report Measures. *Journal Of Personality*, *61*(3), 317-342

Schimmack, U., Oishi, S., & Diener, E. (2002). Cultural influences on the relation between pleasant emotions and unpleasant emotions: Asian dialectic philosophies or individualism-collectivism. *Cognition & Emotion, 16*(6), 705–719

Schreiber, C. A., & Kahneman, D. (2000). Determinants of the remembered utility of aversive sounds. *Journal of Experimental Psychology: General, 129*, 27–42.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45,* 513–523.

Seligman, M.E.P. (2002). *Authentic happiness: using the new positive psychology to realize your potential for lasting fulfillment*. New York: Free Press

Seligman, M.E.P. Csikszentmihalyi, M. (2000). Positive Psychology. An Introduction. *American Psychologist. 55* (1) 5-14

Shiffman, S., Stone, A.A. & Hufford, M.R. (2008). Ecological momentary assessment. *Annual review of clinical psychology, 4,* 1-32

Steptoe A, Deaton A, & Stone AA. (2015). Subjective wellbeing, health, and ageing. *Lancet (London, England)*, *385*(9968), 640-8

Stone, A.A., Bachrach, C.A., Jobe, J.B., Kurtzman, H.S. & Cain, V.S., eds. (2000). *The Science of Self-Report: Implications for Research and Practice*. Mahwah, NJ: Erlbaum

Stone, A., Schwartz, J., Broderick, J., & Shiffman, S. (2005). Variability of Momentary Pain Predicts Recall of Weekly Pain: A Consequence of the Peak (or Salience) Memory Heuristic. *Personality And Social Psychology Bulletin, 31*(10), 1340-1346.

Swann, W.B., Pelham, B.W. & Krull, D.S. (1989), Agreeable Fancy or Disagreeable Truth? Reconciling Self-Enhancement and Self-Verification, *Journal of Personality and Social Psychology* (American Psychological Association) *57*(5), 782–791

Tay, L., Chan, D. & Diener, E. (2014). The metrics of societal happiness. *Social Indicators Research, 117,* 577-600

Thomas, David L.; Diener, Ed (1990). Memory accuracy in the recall of emotions. *Journal of Personality and Social Psychology, 59*(2): 291–297

Uchida, Y., & Kitayama, S. (2009). Happiness and unhappiness in East and West: Themes and variations. *Emotion, 9*(4), 441-456

Walker, W. Richard; John J. Skowronski; Charles P. Thompson (2003). Life is pleasant – and memory helps keep it that way. *Review of General Psychology (Educational Publishing Foundation) 7* (2): 203–210

White, M. P., & Dolan, P. (2009). Accounting for the richness of daily activities. *Psychological Science, 20*, 1000–1008

Zimbardo, P. G., & Boyd, J. N. (1999). Putting time in perspective: A valid, reliable individual-differences metric. *Journal of Personality and Social Psychology, 77*, 1271–1288.

7. APPENDIX

Table 6: adjusted means for each participant and total partial correlation

Individual	*Adjusted means ESM rating*	*Adjusted means DRM rating*	*Partial correlation*
1	6.50	7.75	0.636
2	6.34	6.15	
3	6.61	6.60	
4	6.38	6.62	
5	5.72	5.17	
6	7.95	6.81	
7	6.32	7.11	
8	5.88	6.63	
9	6.25	6.39	
10	8.18	7.21	
11	5.56	6.71	
12	8.32	7.61	
13	5.46	6.31	
14	6.84	7.40	
15	6.60	6.66	
16	8.03	8.15	
17	6.48	6.14	
18	7.04	7.13	
19	6.69	7.25	
20	9.00	7.75	
21	7.08	8.32	
22	6.15	6.84	
23	5.93	7.01	
24	7.15	7.15	
25	6.15	6.69	
26	5.19	6.08	
27	6.73	8.72	
28	5.62	6.06	
29	5.90	6.33	
30	6.42	6.06	
31	6.77	6.38	
32	6.37	6.14	
33	6.49	6.20	
34	4.69	6.74	
35	6.19	5.66	
36	9.26	8.43	
37	7.48	7.51	
38	4.37	7.26	

39	7.16	6.94
40	7.29	6.49
41	6.41	7.33
42	6.54	5.76
43	8.42	8.17
44	7.13	6.81
45	6.82	6.10
46	6.34	6.71
47	6.25	5.10
48	8.18	7.84
49	3.05	4.66
50	6.90	7.14
51	8.06	7.22
52	5.32	5.57
53	7.51	7.25
54	7.01	7.41
55	5.29	5.61
56	7.08	6.99
57	5.02	6.20
58	6.65	4.88
59	7.81	7.89
60	7.96	7.79
61	8.38	8.87
62	6.79	6.66
63	7.48	6.18
64	6.00	5.59
65	7.13	7.10
66	4.66	6.33
67	6.42	8.03
68	7.52	7.18
69	5.47	3.12
70	7.12	7.09
71	7.10	6.74
72	7.20	6.80
73	7.70	8.93
74	8.17	7.82
75	6.29	5.99
76	7.59	7.87
77	5.96	5.50
78	7.56	7.22
79	5.03	4.78
80	6.65	6.75
81	6.31	8.08

82	8.42	7.15
83	3.34	6.01
84	6.84	6.91
85	7.33	6.93
86	6.76	6.61
87	6.42	6.08
88	8.26	8.21
89	6.80	7.87
90	7.01	6.86
91	8.16	6.77
92	6.59	7.00
93	6.14	6.75
94	7.74	6.24
95	6.10	5.96
96	6.58	6.19
97	6.81	6.79
98	5.37	3.99
99	6.99	7.67
100	5.67	7.18
101	7.81	7.42
102	6.57	7.19
103	7.92	7.68
104	6.87	7.39
105	6.91	7.05
106	7.81	7.74
107	7.01	6.92
108	8.02	7.59
109	8.31	8.36
110	6.54	6.48
111	8.17	8.51
112	7.12	7.96
113	7.39	5.55
114	6.60	5.32
115	7.04	7.73
116	6.34	6.72
117	7.07	7.21
118	8.90	9.37
119	6.45	6.59
120	7.11	7.24
121	7.10	7.20
122	8.07	7.01
123	6.90	6.62
124	6.20	6.80

125	6.52	6.35
126	6.44	6.53
127	6.64	6.58
128	7.08	7.23

Table 7: Original adjusted means, their difference, and the significance of their difference. The number of observations per variable are noted as well. The variable is included in this table solely when the adjusted means could be estimated for both ESM and DRM ratings. *=p<0.10, **=p<0.05, ***=p<0.01

Type of event	***Average ESM rating***	***Average DRM rating***	***Mean difference***	***Significance of difference***	***N***
	N=963	**N=2556**			
Timing					
8AM-9AM	6.36	6.57	-0.21	***	252
9AM-10AM	6.45	6.65	-0.2	***	212
10AM-11AM	6.84	6.68	0.16	***	216
11AM-Noon	6.92	6.73	0.19	**	201
Noon-1PM	6.83	6.69	0.14	**	233
1PM-2PM	6.71	6.63	0.08	**	219
2PM-3PM	6.79	6.48	0.31	**	181
3PM-4PM	6.77	6.97	-0.2	***	187
4PM-5PM	6.78	6.81	-0.03	***	204
5PM-6PM	6.94	6.85	0.09	***	237
6PM-7PM	6.80	7.06	-0.26	***	240
7PM-8PM	7.13	7.28	-0.15	***	255
8PM-9PM	6.91	7.33	-0.42	***	240
9PM-10PM	6.82	7.53	-0.71	***	102
10PM-11PM	7.60	7.40	0.2	**	45
11PM-midnight	9.10	7.46	1.64		16
Midnight-1AM	7.57	7.01	0.56	*	67
1AM-2AM	8.28	6.93	1.35		20
2AM-3AM	9.23	6.99	2.24		23
6AM-7AM	6.91	6.10	0.81		105
7AM-8AM	6.24	6.40	-0.16	***	191
Activities			0		
Waking up	5.92	6.114	-6108.08		317
Transport	6.96	6.54	0.42	*	329
Car	7.18	6.54	0.64	***	156
Bicycle	7.17	6.75	0.42	**	84
Other	6.51	6.33	0.18		89

Studying	6.47	5.93	0.54	**	99
Working	6.74	6.60	0.14		480
Paid	6.83	6.66	0.17	**	370
Unpaid	6.47	6.38	0.09		110
Private communication	5.73	7.33	-1.6	***	82
Eating	7.03	7.13	-0.1		471
Breakfast	6.93	6.69	0.24	**	130
Lunch	7.04	7.10	-0.06		106
Dinner	7.20	7.38	-0.18		235
Housekeeping	6.57	6.42	0.15		292
Tidying up	6.69	6.13	0.56	**	47
Groceries shopping	5.21	6.57	-1.36		58
Preparing food	6.67	6.88	-0.21		63
Cleaning	6.62	5.97	0.65	***	57
Other	6.63	6.27	0.36	**	67
Relaxing	7.09	7.25	-0.16		553
Cuddling	7.44	8.80	-1.36		25
Multimedia	7.32	7.31	0.01	*	295
Napping	6.08	7.24	-1.16		36
Sex	8.58	8.78	-0.2		28
Playing games	7.10	7.08	0.02		46
Other	6.73	7.17	-0.44		123
Leisure	7.21	7.52	-0.31		258
Bar/Cafe	7.90	7.96	-0.06		41
Shopping	7.32	7.29	0.03		27
Hiking	8.04	7.78	0.26		25
Playing sports	7.52	7.71	-0.19		52
Other	6.94	7.43	-0.49		113
Caring for someone	6.27	6.27	0		231
Going to bed	5.92	6.53	-0.61		353
Social setting			0		
Alone	6.47	6.54	-0.07		1428
Work-related	7.76	7.33	0.43		53
Children	6.66	6.89	-0.23		292
Friends	7.53	7.78	-0.25		244
Colleagues	6.91	6.84	0.07		296
Siblings	6.43	7.01	-0.58		23
Parents	6.76	6.98	-0.22		43
Partner	8.08	7.62	0.46		612

Other family members	6.95	6.96	-0.01		433
Other	6.68	6.42	0.26		95
Locations			0		
At home	6.61	6.80	-0.19	*	2167
At work	6.91	6.86	0.05		386
Elsewhere - Inside	7.14	6.87	0.27		446
Elsewhere - Outside	6.79	6.94	-0.15		321
Transport	6.79	6.67	0.12		199

*Table 8: Regression output of the basis OLS model 1.0. Reference categories: Timing: midnight – 1AM; Activity: Getting up; Social setting: alone; Location: home. *=p<0.10, **=p<0.05, ***=p<0.01*

Variable	*B-coefficient*	*SE*
Constant	0.862**	0.342
Dummy DRM/ESM	0.031	0.088
Leisure day	0.215**	0.085
Timing		
1AM-2AM	0.351	0.412
2AM-3AM	0.277	0.331
3AM-4AM	-0.445	0.503
4AM-5AM	-0.026	0.369
5AM-6AM	-0.270	0.269
6AM-7AM	-0.161	0.241
7AM – 8AM	0.074	0.214
8AM – 9AM	0.121	0.203
9AM – 10AM	0.049	0.201
10AM – 11AM	0.219	0.214
11AM – noon	0.132	0.213
Noon – 1PM	0.062	0.206
1PM – 2PM	0.021	0.208
2PM – 3PM	-0.061	0.237
3PM – 4PM	0.284	0.197
4PM – 5PM	0.198	0.201
5PM – 6PM	0.228	0.205
6PM – 7PM	0.224	0.207
7PM – 8PM	0.446**	0.195
8PM – 9PM	0.318	0.197

9PM – 10PM	0.370	0.239
10PM – 11PM	0.614***	0.223
11pm-midnight	0.637**	0.321
Activity		
Transportation – Car	0.431**	0.149
Transportation – Bicycle	0.755***	0.212
Transportation -Other	0.168	0.182
Studying	0.169	0.350
Paid work	0.302	0.155
Unpaid work	0.174	0.237
Private communication	0.764**	0.304
Eating – breakfast	0.538***	0.119
Eating – lunch	0.842***	0.145
Eating - dinner	0.817***	0.151
Eating – Snack	0.907***	0.225
Housekeeping – Tidying up	0.172	0.193
Housekeeping – Groceries	0.155	0.182
Housekeeping – Cooking	0.588***	0.180
Housekeeping – Cleaning	0.130	0.280
Housekeeping –Other	0.152	0.216
Relaxing - Multimedia	0.990***	0.117
Relaxing – Cuddling	1.602***	0.387
Relaxing – Sex	2.192***	0.270
Relaxing – Napping	0.326	0.249
Relaxing – Gaming	0.699**	0.198
Relaxing – Other	0.627***	0.172
Leisure – Bar/café	1.164***	0.220
Leisure –Shopping	1.230***	0.217
Leisure –Hiking	1.617***	0.330
Leisure – Sports	1.199***	0.188
Leisure – Other	0.879***	0.140
Caring for someone	0.120	0.193
Going to bed	0.384***	0.133
Locations		
At work	0.016	0.085
Elsewhere – Inside	0107	0.072
Elsewhere – Outside	-0.059	0.081
Elsewhere – Transportation	-0.026	0.079
Social setting		
Friend	0.791***	0.130
Colleagues	0.490***	0.167
Other work-related	0.890***	0.24
Siblings	0.110	0.168

Parents	0.078	0.285
Children	0.264**	0.118
Partner	0.288**	0.113
Other family	0.124	0.120
Other	-0.033	0.176
Individual dummies	YES	
Date dummies	YES	
N individuals	128	
N observations	3519	
R-squared	0.369	

*Table 9: Regression output of the OLS-model 1.1: Timing interaction effects. Reference categories: Timing: midnight – 1AM; Activity: Getting up; Social setting: alone; Location: home. *=p<0.10, **=p<0.05, ***=p<0.01*

Variable	*B-coefficient*	*SE*
Constant	6.546***	0.305
Dummy DRM/ESM	3.698	0.319
Leisure day	0.182*	0.105
Timing		
1AM-2AM	0.274	0.416
2AM-3AM	0.223	0.313
3AM-4AM	-0.450	0.507
4AM-5AM	-0.024	0.372
5AM-6AM	-0.279	0.269
6AM-7AM	-0.185	0.245
7AM – 8AM	0.073	0.217
8AM – 9AM	0.149	0.208
9AM – 10AM	0.128	0.220
10AM – 11AM	0.138	0.222
11AM – noon	0.048	0.218
Noon – 1PM	0.003	0.214
1PM – 2PM	-0.036	0.221
2PM – 3PM	-0.189	0.263
3PM – 4PM	0.361	0.205
4PM – 5PM	0.186	0.211
5PM – 6PM	0.177	0.211
6PM – 7PM	0.289	0.214
7PM – 8PM	0.488**	0.198
8PM – 9PM	0.430**	0.201

9PM – 10PM	0.569**	0.250
10PM – 11PM	0.615**	0.243
11pm-midnight	0.390*	0.229
Activity		
Transportation – Car	0.406***	0.146
Transportation – Bicycle	0.729***	0.207
Transportation -Other	0.137	0.177
Studying	0.033	0.326
Paid work	0.291	0.148
Unpaid work	0.164*	0.233
Private communication	0.744**	0.314
Eating – breakfast	0.516***	0.119
Eating – lunch	0.851***	0.142
Eating - dinner	0.772***	0.146
Eating – Snack	0.871***	0.245
Housekeeping – Tidying up	0.136	0.192
Housekeeping – Groceries	0.150	0.181
Housekeeping – Cooking	0.576***	0.176
Housekeeping – Cleaning	0.083	0.270
Housekeeping –Other	0.124	0.215
Relaxing - Multimedia	0.963***	0.114
Relaxing – Cuddling	1.674***	0.280
Relaxing – Sex	2.189***	0.276
Relaxing – Napping	0.317	0.257
Relaxing – Gaming	0.672***	0.203
Relaxing – Other	0.619***	0.169
Leisure – Bar/café	1.182***	0.252
Leisure –Shopping	1.238***	0.215
Leisure –Hiking	1.641***	0.323
Leisure – Sports	1.162***	0.193
Leisure – Other	0.862***	0.138
Caring for someone	0.123	0.205
Going to bed	0.378***	0.132
Locations		
At work	0.003	0.083
Elsewhere – Inside	0.114	0.076
Elsewhere – Outside	-0.069	0.080
Elsewhere – Transportation	-0.040	0.078
Social setting		
Friend	0.833***	0.142
Colleagues	0.495***	0.272
Other work-related	0.864***	0.181
Siblings	0.123	0.125

Parents	0.104	0.258
Children	0.248**	0.117
Partner	0.304**	0.148
Other family	0.196	0.169
Other	-0.034	0.192
Interaction effects		
*Dummy ESM*Midnight-1AM*	-3.411***	0.367
*Dummy ESM*1AM-2AM*	-2.314***	0.581
*Dummy ESM*2AM-3AM*	-1.664***	0.434
*Dummy ESM*6AM-7AM*	-2.990***	0.512
*Dummy ESM*7AM – 8AM*	-3.866***	0.460
*Dummy ESM*8AM – 9AM*	-3.861***	0.383
*Dummy ESM*9AM – 10AM*	-3.884***	0.296
*Dummy ESM*10AM – 11AM*	-3.395***	0.364
*Dummy ESM*11AM – noon*	-3.421***	0.377
*Dummy ESM*Noon – 1PM*	-3.492***	0.371
*Dummy ESM*1PM – 2PM*	-3.483***	0.370
*Dummy ESM*2PM – 3PM*	-3.311***	0.374
*Dummy ESM*3PM – 4PM*	-3.853***	0.343
*Dummy ESM*4PM – 5PM*	-3.629***	0.364
*Dummy ESM*5PM – 6PM*	-3.495***	0.375
*Dummy ESM*6PM – 7PM*	-3.866***	0.380
*Dummy ESM*7PM – 8PM*	-3.764***	0.345
*Dummy ESM*8PM – 9PM*	-4.063***	0.375
*Dummy ESM*9PM – 10PM*	-4.383***	0.428
*Dummy ESM*10PM – 11PM*	-3.769***	0.448
Individual dummies	YES	
Date dummies	YES	
N individuals	128	
N observations	3519	
R-squared	0.380	

*Table 10: Regression output of the OLS-model 1.2: Activity interaction effects. Reference categories: Timing: midnight – 1AM; Activity: Getting up; Social setting: alone; Location: home. *=p<0.10, **=p<0.05, ***=p<0.01*

Variable	*B-coefficient*	*SE*
Constant	6.487**	0.293
Dummy DRM/ESM	-0.480	0.292
Leisure day	0.216**	0.085
Timing		
1AM-2AM	0.350	0.441
2AM-3AM	0.269	0.326
3AM-4AM	-0.517	0.487
4AM-5AM	-0.097	0.360
5AM-6AM	-0.254	0.261
6AM-7AM	-0.256	0.236
7AM – 8AM	0.099	0.207
8AM – 9AM	0.174	0.200
9AM – 10AM	0.099	0.203
10AM – 11AM	0.275	0.214
11AM – noon	0.174	0.213
Noon – 1PM	0.121	0.206
1PM – 2PM	0.076	0.205
2PM – 3PM	-0.051	0.239
3PM – 4PM	0.332	0.196
4PM – 5PM	0.257	0.200
5PM – 6PM	0.256	0.210
6PM – 7PM	0.261	0.210
7PM – 8PM	0.481**	0.198
8PM – 9PM	0.368*	0.200
9PM – 10PM	0.458*	0.233
10PM – 11PM	0.611***	0.230
11pm-midnight	0.567*	0.306
Activity		
Transportation – Car	0.285*	0.160
Transportation – Bicycle	0.653***	0.212
Transportation -Other	0.037	0.221
Studying	-0.286	0.427
Paid work	0.156	0.172
Unpaid work	0.074	0.240
Private communication	1.149**	0.280
Eating – breakfast	0.471***	0.115
Eating – lunch	0.8766***	0.166
Eating - dinner	0.809***	0.163
Eating – Snack	0.640**	0.322
Housekeeping – Tidying up	-0.058	0.224

Housekeeping – Groceries	0.241	0.165
Housekeeping – Cooking	0.571***	0.204
Housekeeping – Cleaning	-0.178	0.247
Housekeeping –Other	-0.001	0.246
Relaxing - Multimedia	0.919***	0.135
Relaxing – Cuddling	2.048***	0.348
Relaxing – Sex	2.223***	0.288
Relaxing – Napping	0.970**	0.445
Relaxing – Gaming	0.630**	0.263
Relaxing – Other	0.721***	0.191
Leisure – Bar/café	1.152***	0.252
Leisure –Shopping	1.162***	0.206
Leisure –Hiking	1.496***	0.348
Leisure – Sports	1.180***	0.227
Leisure – Other	1.007***	0.149
Caring for someone	0.035	0.224
Going to bed	0.385***	0.136
Locations		
At work	0.002	0.081
Elsewhere – Inside	0.103	0.074
Elsewhere – Outside	-0.044	0.082
Elsewhere – Transportation	0.015	0.082
Social setting		
Friends	0.773***	0.132
Colleagues	0.477***	0.163
Other work-related	0.901***	0.231
Siblings	0.083	0.166
Parents	0.056	0.282
Children	0.260**	0.115
Partner	0.276**	0.110
Other family	0.138	0.118
Other	0.005	0.184
Interaction effects		
*Dummy ESM*Getting up*	0.387	0.333
*Dummy ESM*Transportation – Car*	1.259***	0.379
*Dummy ESM*Transportation – Bicycle*	1.076***	0.462
*Dummy ESM*Transportation - Other*	0.781	0.506
*Dummy ESM*Studying*	1.154**	0.485

*Dummy ESM*Paid work*	0.763**	0.324
*Dummy ESM*Unpaid work*	0.722	0.534
*Dummy ESM*Private communication*	-1.074*	0.619
*Dummy ESM*Eating – breakfast*	0.854**	0.398
*Dummy ESM*Eating – lunch*	0.628	0.452
*Dummy ESM*Eating - dinner*	0.386	0.330
*Dummy ESM*Eating – Snack*	1.074***	0.397
*Dummy ESM*Housekeeping – Tidying up*	1.143**	0.540
*Dummy ESM*Housekeeping – Groceries*	-0.732	0.723
*Dummy ESM*Housekeeping – Cooking*	0.362	0.567
*Dummy ESM*Housekeeping – Cleaning*	1.156***	0.390
*Dummy ESM*Housekeeping –Other*	0.954	0.447
*Dummy ESM*Relaxing - Multimedia*	0.574*	0.341
*Dummy ESM*Relaxing – Cuddling*	-0.729=8	0.507
*Dummy ESM*Relaxing – Sex*	0.277	0.762
*Dummy ESM*Relaxing – Napping*	-0.535	0.551
*Dummy ESM*Relaxing – Gaming*	0.600	0.390
*Dummy ESM*Relaxing – Other*	0.149	0.427
*Dummy ESM*Leisure – Bar/café*	0.563	0.578
*Dummy ESM*Leisure – Shopping*	0.580	0.579

*Dummy ESM*Leisure – Hiking*	0.881	0.763
*Dummy ESM*Leisure – Sports*	0.404	0.400
*Dummy ESM*Leisure – Other*	0.093	0.349
*Dummy ESM*Caring for someone*	0.394	0.544
Individual dummies	YES	
Date dummies	YES	
N individuals	128	
N observations	3519	
R-squared	0.384	

*Table 11: Regression output of the OLS-model 1.2: Location interaction effects Reference categories: Timing: midnight – 1AM; Activity: Getting up; Social setting: alone; Location: home. *=p<0.10, **=p<0.05, ***=p<0.01*

Variable	*B-coefficient*	*SE*
Constant	6.474**	0.299
Dummy DRM/ESM	0.219	0.177
Leisure day	0.219**	0.084
Timing		
1AM-2AM	0.358	0.417
2AM-3AM	0.275	0.331
3AM-4AM	-0.447	0.508
4AM-5AM	0.028	0.362
5AM-6AM	-0.276	0.267
6AM-7AM	-0.080	0.240
7AM – 8AM	0.068	0.213
8AM – 9AM	0.118	0.202
9AM – 10AM	0.037	0.202
10AM – 11AM	0.221	0.215
11AM – noon	0.116	0.212
Noon – 1PM	0.051	0.204
1PM – 2PM	0.023	0.208
2PM – 3PM	-0.055	0.234
3PM – 4PM	0.283	0.196
4PM – 5PM	0.199	0.202
5PM – 6PM	0.230	0.205
6PM – 7PM	0.222	0.206

7PM – 8PM	0.437**	0.195
8PM – 9PM	0.320	0.196
9PM – 10PM	0.361	0.239
10PM – 11PM	0.618***	0.220
11pm-midnight	0.649**	0.322
Activity		
Transportation – Car	0.435***	0.149
Transportation – Bicycle	0.758***	0.213
Transportation -Other	0.179	0.185
Studying	0.100	0.350
Paid work	0.291*	0.153
Unpaid work	0.174	0.239
Private communication	0.762**	0.303
Eating – breakfast	0.532***	0.117
Eating – lunch	0.853***	0.144
Eating - dinner	0.823***	0.152
Eating – Snack	0.901***	0.223
Housekeeping – Tidying up	0.170	0.194
Housekeeping – Groceries	0.161	0.185
Housekeeping – Cooking	0.596***	0.178
Housekeeping – Cleaning	0.135	0.281
Housekeeping –Other	0.158	0.215
Relaxing - Multimedia	0.992***	0.118
Relaxing – Cuddling	1.596***	0.298
Relaxing – Sex	2.196***	0.272
Relaxing – Napping	0.308	0.248
Relaxing – Gaming	0.696***	0.199
Relaxing – Other	0.636***	0.171
Leisure – Bar/café	1.149***	0.218
Leisure –Shopping	1.248***	0.217
Leisure –Hiking	1.611***	0.318
Leisure – Sports	1.171***	0.193
Leisure – Other	0.878***	0.140
Caring for someone	0.140	0.194
Going to bed	0.390***	0.134
Locations		
At work	-0.059	0.100
Elsewhere – Inside	-0.030	0.095
Elsewhere – Outside	-0.070	0.089
Elsewhere – Transportation	-0.114	0.092
Social setting		
Friends	0.807***	0.130
Colleagues	0.508***	0.169

Other work-related	0.922***	0.223
Siblings	0.107	0.169
Parents	0.094	0.287
Children	0.290**	0.117
Partner	0.298**	0.115
Other family	0.146	0.121
Other	-0.011	0.177
Interaction effects		
*Dummy ESM*At home*	-0.301*	0.176
*Dummy ESM*At work*	-0.050	0.229
*Dummy ESM*Elsewhere-Inside*	0.170	0.215
*Dummy ESM*Elsewhere-outside*	-0.259	0.229
Individual dummies	YES	
Date dummies	YES	
N individuals	128	
N observations	3519	
R-squared	0.371	

*Table 12: Regression output of the OLS-model 1.2: Social setting interaction effects Reference categories: Timing: midnight – 1AM; Activity: Getting up; Social setting: alone; Location: home. *=p<0.10, **=p<0.05, ***=p<0.01*

Variable	*B-coefficient*	*SE*
Constant	6.456**	0.304
Dummy DRM/ESM	0.054	0.199
Leisure day	0.219**	0.084
Timing		
1AM-2AM	0.351	0.415
2AM-3AM	0.275	0.333
3AM-4AM	-0.452	0.503
4AM-5AM	0.017	0.372
5AM-6AM	-0.267	0.265
6AM-7AM	-0.160	0.243
7AM – 8AM	0.078	0.215
8AM – 9AM	0.132	0.203
9AM – 10AM	0.059	0.203
10AM – 11AM	0.229	0.214
11AM – noon	0.139	0.215
Noon – 1PM	0.070	0.206

1PM – 2PM	0.029	0.208
2PM – 3PM	-0.050	0.237
3PM – 4PM	0.295	0.198
4PM – 5PM	0.203	0.203
5PM – 6PM	0.235	0.205
6PM – 7PM	0.242	0.209
7PM – 8PM	0.450**	0.197
8PM – 9PM	0.331*	0.198
9PM – 10PM	0.387	0.238
10PM – 11PM	0.612***	0.225
11pm-midnight	0.640**	0.320
Activity		
Transportation – Car	0.429***	0.147
Transportation – Bicycle	0.763***	0.213
Transportation -Other	0.189	0.183
Studying	0.075	0.353
Paid work	0.294*	0.157
Unpaid work	0.173	0.237
Private communication	0.768**	0.306
Eating – breakfast	0.535***	0.119
Eating – lunch	0.849***	0.145
Eating - dinner	0.816***	0.149
Eating – Snack	0.897***	0.232
Housekeeping – Tidying up	0.184	0.195
Housekeeping – Groceries	0.161	0.182
Housekeeping – Cooking	0.599***	0.179
Housekeeping – Cleaning	0.130	0.279
Housekeeping –Other	0.145	0.219
Relaxing - Multimedia	0.995***	0.116
Relaxing – Cuddling	1.623***	0.281
Relaxing – Sex	2.169***	0.273
Relaxing – Napping	0.330	0.251
Relaxing – Gaming	0.712***	0.200
Relaxing – Other	0.627***	0.168
Leisure – Bar/café	1.166***	0.219
Leisure –Shopping	1.241***	0.219
Leisure –Hiking	1.620***	0.334
Leisure – Sports	1.219***	0.189
Leisure – Other	0.872***	0.142
Caring for someone	0.114	0.192
Going to bed	0.373***	0.133
Locations		
At work	0.007	0.084

Elsewhere – Inside	0.110	0.074
Elsewhere – Outside	-0.066	0.083
Elsewhere – Transportation	-0.032	0.077
Social setting		
Friends	0.835***	0.144
Colleagues	0.429**	0.216
Other work-related	0.855***	0.223
Siblings	0.135	0.181
Parents	-0.063	0.306
Children	0.302**	0.136
Partner	0.327**	0.126
Other family	0.122	0.143
Other	-0.138	0.207
Interaction effects		
*Dummy ESM*Alone*	-0.016	0.209
*Dummy ESM*Other*	0.350	0.388
*Dummy ESM*Other work-related*	0.411	0.535
*Dummy ESM*Friends*	0.411	0.282
*Dummy ESM*Colleagues*	-0.182	0.275
*Dummy ESM*Siblings*	0.141	0.281
*Dummy ESM*Parents*	-0.373	0.473
*Dummy ESM*Children*	0.824	0.294
*Dummy ESM*Partner*	-0.161	0.235
Individual dummies	YES	
Date dummies	YES	
N individuals	128	
N observations	3519	
R-squared	0.371	