**"I feel better when…": An Analysis of the Memory-Experience Gap for Peoples'**

**Estimates of the Relationship between Health Behaviors and Experiences**

Andrew T. Gloster[a,b], Andrea H. Meyer[a], Cornelia Witthauer[a], Roselind Lieb[a], & Jutta Mata[a,c]

[a]University of Basel, Department of Psychology, Division of Clinical Psychology and Epidemiology, Missionsstrasse 62a, CH-4055 Basel

[b]University of Basel, Department of Psychology, Division of Clinical Psychology and Intervention Science, Missionsstrasse 62a, CH-4055 Basel

[c]University of Mannheim, Department of Psychology, Division of Health Psychology, Bismarckstrasse, 68161 Mannheim

Andrew Gloster: andrew.gloster@unibas.ch, +41 61 267 02 75

Andrea Meyer: andrea.meyer@unibas.ch, +41 61 267 06 55

Cornelia Witthauer: cornelia.witthauer@unibas.ch, +41 61 207 02 19

Roselind Lieb: roselind.lieb@unibas.ch, + 41 61 267 02 80

Jutta Mata: mata@uni-mannheim.de, +49 621 181 25 95

**Acknowledgement**

Abstract

**Objective**: People often overestimate how strongly behaviors and experiences are related. This memory-experience gap might have important implications for health care settings, which often require people to estimate associations such as "my mood is better when I exercise". This study examines how subjective correlation estimates between health behaviors and experiences relate to calculated correlations from online reports and whether subjective estimates are associated with engagement in actual health behavior. **Design**: Seven-month online study on physical activity, sleep, affect, and stress, with 61 online assessments. **Main Outcome Measures:** University students ($N$=168) retrospectively estimated correlations between physical activity, sleep, positive affect, and stress over the 7-month study period. **Results:** Correlations between experiences and behaviors (online data) were small ($r= -.12$-.14), estimated correlations moderate ($r= -.35$-.24). Correspondence between calculated and estimated correlations was low. Importantly, estimated correlations of physical activity with stress, positive affect, and sleep were associated with actual engagement in physical activity. **Conclusion**: Estimation accuracy of relations between health behaviors and experiences is low. However, association estimates could be an important predictor of actual health behaviors. This study identifies and quantifies estimation inaccuracies in health behaviors and points towards potential systematic biases in health settings, which might seriously impair intervention efficacy.

*Keywords:* daily diary methodology; correlation; estimation; belief; health behaviors

"I feel better when…": An analysis of the memory-experience gap for peoples' estimates of the relationship between health behaviors and experiences

People's memories of experiences and affect often differ from the actual occurrences. Such discrepancies are believed to reflect different sources of knowledge: a person's experience versus a person's belief about their experiences (Conner & Barrett, 2012; Robinson & Clore, 2002). The resultant memory-experience gap, defined as the discrepancy between the average of experienced affect and the overall retrospective evaluation of the experience, has been observed across multiple groups and behaviors (Miron-Shatz, Stone, & Kahneman, 2009). This gap routinely results in overestimating the frequencies or intensities of experiences and behaviors.

One situation in which persons' beliefs about experiences are frequently assessed and are of great importance is clinical settings. Health care providers commonly ask their patients about experiences relevant for their health and treatment and about the contexts and conditions in which these experiences occur. Studies that have compared retrospective recall to daily symptomatology clearly point to differential memory processes. At the most basic level, evidence exists that patients across various healthcare settings are largely inaccurate when recalling their own daily-recorded index data over as short a time frame as one week. For example, data collected in vivo and subjects' retrospective recall of this period have shown significant discrepancies for estimates of pain intensity (Stone, Broderick, Shiffman, & Schwartz, 2004), frequency of eating behaviors (Stein & Corte, 2003), use of coping behaviors (Stone et al., 1998), frequency of panic attacks (de Beurs, Lange, & Vandyck, 1992), timing of smoking lapses (Shiffman et al., 1997), presence of anxious cognitions (Marks & Hemsley, 1999), and depressive symptoms (Ben-Zeev & Young, 2010).

However, as pervasive as such inaccuracies between a person's estimated versus in-vivo assessed experiences are at the rote-recall level (i.e., frequency, duration), inaccuracies in the *relation* between phenomena (e.g., is exercise related to higher positive affect?) may be

even larger. Decades of studies have documented the belief of illusory correlations in a multitude of laboratory paradigms (e.g., Van Rooy, Vanhoomissen, & Van Overwalle, 2013), suggesting it to be a stable cognitive bias. Importantly, it remains relatively unexplored whether illusory correlations only apply to artificial laboratory stimuli, or whether they generalize to estimates of the relationship between self-generated day-to-day data, such as personal experiences and behaviors with which people are extremely familiar.

A behavior or experience never happens in isolation, but is rather embedded in a situational context comprised of other behaviors and experiences. Thus, understanding *relations* between such experiences and behaviors provides information on potential interdependencies between experiences and behaviors. The specific importance of understanding interdependencies or correlations in a health context lies in the fact that nearly all health care appointments require patients to offer judgments of their behavior in some sense or another (for example, "My pain is worse when…"; "I exercise more when…"; "My mood improves when…). Importantly, examinations of how individuals understand, attribute, and report on the relation between their own behaviors and experiences are scarce. Examining the covariation between two variables using both in vivo generated data as well as people's estimates of such a correlation offers a potentially important new perspective. In one such study, using an ecological momentary assessment (EMA) methodology, correlation between symptoms of obsessive-compulsive disorder and other behaviors showed that participants' estimations of correlations explained up to as much as 25% of the variance of the relationship between obsessive-compulsive symptoms and emotional states such as stress and anxiety. In contrast, the correlations based on patients' self-generated data explained only 4% of the variance (Gloster et al., 2008). If correlation misestimations are prevalent, it is important to determine what, if any, practical implications this has. That is, an absolute difference may still be relatively accurate. Therefore, it is crucial to simultaneously determine the extent of correlation misestimations, to assess what processes and variables are associated with such

5

discrepancies, and to explore potential implications. This is especially true in the domain of health reporting, where clinical and public health decisions are at stake.

The aim of the current study is to determine the presence, magnitude, and implication of potential illusory correlations across frequently documented targets and behaviors of health settings such as sleep, physical activity, stress, and affect (Housman & Dorman, 2005; Kirmayer, Robbins, Dworkind, & Yaffe, 1993; Schneiderman, Ironson, & Siegel, 2005). Thus, we explore, for example, whether someone believes that based on his or her experiences sleep and physical activity or stress and affect are associated (e.g., "when I sleep less, I exercise less"). In this study we refer to people's estimated correlations to describe the relations between experiences and behaviors that people subjectively judged at the end of data collection. We refer to daily diary-calculated correlations for the associations between experiences and behaviors that were calculated based on the online daily diary assessment over a 7-month study period.

Towards this end we asked participants to make judgments on variables with varying degrees of automaticity: behaviors whose engagement is under people's direct control (i.e., physical activity), behaviors that are habitual and subject to diurnal patterns (i.e., hours of sleep), as well as targets that are only indirectly influenced (i.e., stress & affect). Furthermore, these variables have been identified as important predictors for general health and have been shown to covary in natural settings (e.g., Affleck, Tennen, Urrows, & Higgins, 1994; Buman, Hekler, Bliwise, & King, 2011; Flueckiger, Lieb, Meyer, & Mata, 2014; Jacobs et al., 2007; Mata et al., 2012; Sadeh, Keinan, & Daon, 2004; Watson, 1988).

We examine the following research questions: (1a) Do people's estimated and daily-diary-calculated correlations differ? (1b) What is the relation between people's estimated and daily-diary-calculated correlations for health behaviors or experiences? (2) To explore the implications of people's estimated correlations, we tested both, whether people's estimated correlations as well as the correlations calculated based on the daily diaries are associated

6

with actual engagement in physical activity (calculated from the online assessments). Although exploratory, we hypothesized that actual physical activity would be associated with subjective estimates.

## Method

### Design and Procedure

Participants were recruited through advertisements during first-year introductory lectures at the University of [name omitted to maintain anonymity]. Particularly the introductory lectures in chemistry and physics are also obligatory for students from other areas such as pharmacy, biology, and teacher education. Thus, students of these areas were also recruited into the study. Students interested in participation gave their e-mail address to the research team and were invited via email to the online entry questionnaire. Participants received e-mail invitations to participate in a total of 61 short online-assessments.

Prompts to answer assessments were sent via email and could be answered after clicking on a link embedded within the email. Prompts during the first half of the study (early December to mid-May) occurred once every six days. This spacing was chosen to assess each day of the week equally and to reduce participant burden. In the second phase of the study, participants were prompted daily for 33 days (mid-May to mid-June 2012). The daily spacing was chosen to assess experiences and behaviors during the final exam period in more detail. At each prompt, participants entered their unique personal code. Participants were never asked to enter their e-mail address or any other information into the questionnaires that would allow them to be identified. On the 61st assessment day, participants were asked to estimate the correlation between physical activity, sleep, positive affect, and stress over the entire 7-month study period.

### Participants

A total of 323 students participated in the survey. On any given day, approximately 150 to 200 people participated in that day's assessment ($M$=183.2, $SD$=35.1 participants per

survey). Consistent with this, 168 completed the 61st assessment that included the correlation estimation. Participant characteristics are shown in Table 1. The 168 participants who estimated correlations over the 7-month study period responded to a total of 8278 assessments over the 61 assessment days. On average, each participant completed an average of $M$=49 assessments out of the 61 possible ($SD$=12.0, range 9–61).

All participants gave informed consent after receiving detailed information about the study. Students chose between receiving a financial compensation for participation or equivalent course credits. The study was approved by the local ethics committee, [name and reference number omitted to maintain anonymity].

[Table 1 near here]

**Measures**

**Entry and Exit Questionnaire.** Sociodemographic information was collected through the entry questionnaire (see Table 1).

**Day-Level Questionnaire.** Daily positive affect was assessed with the three positive affect items of the German version of the pleasantness scale (Roecke, 2006), that is, happy, delighted/joyful, and content. The pleasantness scale is based on items from the Positive Affect Negative Affect Scale (Watson, Clark, & Tellegen, 1988) and has shown sufficient variability and predictive validity in another university student sample (Flueckiger et al., 2014). Participants indicated the extent to which they were experiencing each emotion during the last 24 hours on a seven-point Likert scale. Reliability of the three items was $\alpha = .97$, calculated based on the mean values of the three affect items per person over the 7-month study period.

Physical activity was evaluated with the Godin Leisure Time Questionnaire (Godin & Shephard, 1997), adapted to the daily survey format. Participants reported physical activity during the last 24 hours in number of minutes engaged in mild, moderate, and strenuous activity. Following Godin & Shepard (1997), the total physical activity score was calculated

as minutes of light activity x 3 + minutes of moderate intensity x 5 + minutes of strenuous activity x 9. Thus, the daily activity score represents the total daily leisure activity weighted by duration and intensity. Higher scores reflect more intense levels and/ or longer duration of physical activity.

Hours slept were assessed with one item taken from the Pittsburg Sleep Quality Index (PSQI; Buysse, Reynolds, Monk, Berman, & Kupfer, 1989), asking participants how long (hours:minutes) they had slept the previous night.

Experienced stress was assessed by asking participants to report which hassles they had experienced during the morning, afternoon, and evening (the question structure was based on the Day Reconstruction Method (Kahneman, Krueger, Schkade, Schwarz, & Stone, 2004) and their rating of each hassle's intensity. For the analyses, we used the number of reported hassles per 24-hour period.

Correlation estimation was measured by asking participants to estimate the relation between health behaviors and/ or experiences on a scale from −100 to +100 (adapted from Gloster et al., 2008). The number of hours participants slept, their amount of physical activity, positive affect, and stressful events were paired in all possible combinations, resulting in six estimation questions such as, "What is the relation between the number of hours you slept per night and the amount of physical activity over the study period, that is, from early December to mid-June?".

**Statistical Analysis**

People's estimated correlations were directly accessible for analyses whereas Daily-diary-calculated correlations were calculated by correlating the two health behaviors/experiences with each other for each individual. All analyses were conducted using R (R Core Team, 2014) or MPlus (Muthén & Muthén, 1998-2012); the type of analysis conducted is described in the respective Results section below.

<div align="center">

**Results**

</div>

**Research Question 1a: Do people's estimated and daily-diary-calculated correlations differ?**

To examine this question, we compared people's estimated and daily-diary-calculated mean levels of correlation using t-tests for paired samples. Independent of their sign (i.e., positive or negative), the mean level of estimated correlations was almost always higher than the mean level of daily-diary-calculated correlations (see Table 2). These differences were always significant, with small to large effect sizes. The largest differences were for the variable pairs sleep-positive affect, sleep-physical activity, and stress-positive affect, suggesting that the size of estimated correlations including sleep and positive affect might be particularly overestimated. Importantly, four of the correlations based on daily-diary data were close to zero ($r \leq .10$), yet only one estimated association reflected this, but was in the opposite direction (i.e., Stress-Physical Activity). See also Figure 1.

[Table 2 near here]

[Figure 1 near here]

**Research Question 1b: What is the relation between people's estimated and daily-diary-calculated correlations?**

Acknowledging that although the direct comparison showed differences, they estimated and daily-dairy correlations might still be associated, we assessed how closely the correlation estimates of the two assessment methods were related. For this step we assessed the association between people's estimated correlation and their daily-diary-calculated correlation measures. Daily-diary-calculated correlation measures were thereby determined using multilevel structural equation models. To this end we first set up a linear mixed model for each pair of health behaviors/experiences that contained a random slope coefficient, denoting subject specific associations for the daily-diary-calculated data. The estimates of these slopes are typically named empirical Bayes estimates and these were then used as measure of daily-diary-calculated correlation. Empirical Bayes estimates are more precise and

more stable than estimates obtained from ordinary linear regression analysis (Singer & Willett, 2003). Thus we correlated the empirical Bayes estimates of the slopes with people's estimated correlations (Table 3). Although associations between the two methods were always positive, they were never significant. In sum, across all participants and most variable pairs, the size of estimated correlations was weakly, if at all, related to the size of daily-diary-calculated correlations.

[Table 3 near here]

Potentially, participants need not know the exact size of the correlation between two variables. Rather, it might be sufficient to be *relatively* accurate, that is, to estimate that a correlation between two variables is positive, negative, or zero. Therefore, next, we categorized the estimated correlations for all variable pairs simply as positive (r > 0), negative (r < 0), or zero (r = 0). If knowing simply the direction of the correlation (positive, negative, no correlation) were sufficient, then the daily-diary-calculated correlations would correspond: positively estimated relations should also be positive, those in the negative category should be negative, and those in the zero category should be zero or close to zero. This possibility was tested with a one-factorial analysis of variance with categorized estimated correlations as factor and daily-diary-calculated correlations as outcome. However, for all six variable-pairs, the sign of the daily-diary-calculated correlations generally did not correspond to the sign of the estimated categories (i.e., daily-diary-calculated correlations in the positive estimation category were not generally positive; see Table 4). Moreover, the daily-diary-calculated correlations also did not differ between the three categories of estimated correlations (Table 4). In sum, whether a participant estimated a correlation pair to be positive, negative, or zero was not reflected in the daily-diary-calculated correlations. Note that the results were similar when the criterion for categorization was changed (e.g., positive (r > .10), negative (r < −.10).

[Table 4 near here]

**Research Question 2: Are people's correlation estimates associated with physical activity behavior?**

Physical activity is the one variable in this data set that was both under the person's direct control (unlike stress and positive affect) and that was not subject to diurnal rhythms (unlike sleep). As such, it is a discreet behavior that was most reliant on motivation components and hence its association with the other variables may be easier for people to estimate. We used correlations to determine whether participants' correlation estimates for the variable pairs predicted the *actual* physical activity reported in the daily-diary data. Average physical activity reported during the daily-diary assessment (i.e., throughout the semester) was positively related to the estimated correlations between physical activity/sleep, physical activity/stress, and short of being significantly related to physical activity/positive affect (see Table 5a). As a means of comparison, we also examined whether correlations derived from daily-diary data predicted the *actual* physical activity. In contrast to participants' estimates, the data-derived correlations were not associated with actual physical activity in any of the pairs (see Table 5b).

[Table 5a and 5b near here]

## Discussion

The findings of this study suggest that people's estimated correlations were little or not related to daily-diary-calculated correlations between health behaviors and experiences. Importantly, people's estimated correlations between physical activity with sleep and stress were associated with *actual* engagement in physical activity – independent of a very low or nonexistent correspondence between estimated and daily-diary-calculated correlations.

The findings suggest that estimated correlations between behaviors and experiences remain illusory and a cognitive bias, also in a natural context. That is, their low correspondence does not appear to be an artifact of the laboratory setting, where several previous studies on this phenomenon were conducted (Meiser & Hewstone, 2010). Previous

research on cognitive biases, for example on overconfidence, has suggested that bias mostly occurs in artificial laboratory settings in which task difficulty is different from a natural context: People have little experience with their performance in the laboratory task context and misjudge it. Importantly, overconfidence can disappear in a natural task context (Gigerenzer, Hoffrage, & Kleinbolting, 1991). Therefore, one of the goals of this study was to examine whether a similar effect could be found for illusory correlations in a natural setting with variables participants are extremely familiar with – their own health behaviors and experiences. However, at least in this current data set, such an influence of context could not be detected.

To our knowledge, this is also the first study on illusory correlations in a health behavior setting. The current findings are in line with previous findings in the domain of social cognition and stereotype formation (Van Rooy et al., 2013), which suggest that humans assume relations and patterns where there are actually none. The overestimation observed in these health behaviors were also consistent with the overestimations of correlation observed in patients diagnosed with obsessive-compulsive disorder with respect to the relation between symptoms and emotional states (Gloster et al., 2008).

Importantly, people's estimated correlations that higher physical activity is associated with longer sleep duration, less stress, and to a lesser degree higher positive affect – despite being mostly illusory – were nevertheless associated with actual engagement in physical activity. In contrast, correlations calculated based on the daily diaries were not associated with actual engagement in physical activity. This suggests that the implicit representation of an association between physical activity and other behaviors or experiences could be one of the underlying motives or rationales for engaging in this health behavior. That is, "I will engage in physical activity now because I *believe* it will improve my sleep and decrease my stress". This is in line with observations that belief may be more strongly related to some health decisions than actual experience (Conner & Barrett, 2012; Houtveen & Oei, 2007). The

estimated correlations between physical activity and the other three target variables only predicted physical activity, the behavior under participants' direct control. Nevertheless, it is important to be mindful that the causal direction of this association remains open. For example, it is possible that people sleep longer after exercising or that they exercise more after sleeping longer. Future research is needed to help determine which association is referenced when participants estimate the relationship between experiences.

This study extends documentation of the experience-memory gap (Kahneman et al., 2004; Miron-Shatz et al., 2009) in health behaviors. Although different sources of knowledge are likely emphasized in each assessment method (i.e., experience vs. correlation estimate) and each is important, we concur with others that overreliance on questionnaires is scientifically and clinically limiting (Conner & Barrett, 2012; Ebner-Priemer & Trull, 2009; FDA. 2006). Although retrospectively estimated correlations are necessary and important, caution should be exercised when interpreting these data, particularly if inferences are to be generalized to daily experiences. This is especially true when people are required to engage in relatively complicated cognitive tasks for which they likely rely on simple rules rather than calculating correlations according to the respective mathematical formula. An accurate or even relatively accurate understanding of one's own behavior may simply be too difficult a task for most people, especially if large volumes of information over long time frames are used or in answering questions about what they estimated to correlate with targeted symptoms. These results further suggest that future research needs to better understand how and under which conditions various actual behaviors relates to retrospective estimations thereof.

The observed phenomenon likely has more practical and potentially detrimental effects in some settings than in others. It is precisely for this reason that we targeted health behaviors, where effective treatment presumably requires at least minimally accurate reports from the patient. Based on these results, health practitioners should be aware of potential

inaccuracies when assessing even common bivariate associations ("I sleep better when…").

Given that the memory-experience gap has been found to be greater for negative mood than

positive mood (Miron-Shatz et al., 2009) and that reporting of negative affect is common in

healthcare settings, particularly in mental health care, future research should explicitly test

whether the memory-experience gap is especially pronounced in a health context. This could

be particularly troubling given the nearly ubiquitous reliance on such judgments in the

research and clinical care of all disorders. On the other hand, beliefs about associations, even

inaccurate beliefs, may be equally important in predicting actual behaviors in some domains.

Which degree of inaccuracy or belief impacts the predictive validity across various domains

remains an important empirical question.  Inaccurate patient estimation has the potential to

compromise treatment efficacy (Haynes, Leisen, & Blaine, 1997). Irrespective of the clinical

implications, the scientific understanding of these behaviors requires attention to both actual

correlations as measured in situ and retrospectively estimated relations.

The mechanisms that account for the discrepancy between actual experiences (e.g., physical

activity) and estimated experiences (e.g., memory of physical activity and how it related to other

psychological states) is not currently clear, but likely due to numerous influences. Two different lines

of research might be particularly interesting for better understanding this phenomenon. (1) Contextual

aspects of the actual occurrence that are likely outside the awareness of the person during the

experience may contribute to the discrepancy. Ekkekakis (2003) suggested that exercise is related to

affect via both, cognitive factors (e.g., physical self-efficacy) and interoceptive (e.g., muscular or

respiratory) cues. At low intensity, cognitive factors predominantly influence affect, for example, a

person feels happy that they eventually did go out and enjoys what they are doing. At very high

intensities, interoceptive cues predominantly determine the affect state. The closer the body gets to

functional limits, the higher negative affect and lower positive affect. This can be seen as a cue to

lower exercise intensity or stop it. In line with this are findings that positive affect is often below

baseline – at least temporarily – after high or very high doses of exercise (see Reed & Ones, 2006, for

an overview and meta-analysis). Thus, our memory of the relation between affect and physical activity

is likely affected by the fact that this relation is not linear but rather dependent on the dosage. At the same time, this variability makes it even more important to study various bouts of physical activity and to understand how these experiences affect our subjective representation of the relation between physical activity and affect. Further, (2) memory consolidation and recall processes likely also contribute to the discrepancy. A large body of research has shown that a person's prediction about emotional reactions to future events (i.e., affective forecasting) is generally inaccurate (e.g., Wilson & Gilbert, 2005). At the same time, a recent study on anticipated and consummatory pleasure of daily activities showed that people were generally accurate about their predictions (Wu et al., 2016). However, the mechanisms helping us to understand when these predictions are accurate and when not, need yet to be examined.

This study's strength is the utilization of an assessment period that stretched over several months in a naturalistic setting and targeted a large sample of behaviors reflecting participants' individual lives and choices. The study also had several limitations. First, this study examined university students' health behaviors only. This participant group may differ from older participants or participants who are actively seeking health care. Second, participants' subjective estimates about the causal direction of the associations were not explicitly assessed. Third, longer or shorter time frames may have rendered different results. Various lengths of assessment periods (e.g., one day vs. two semesters) and intensity of assessments (e.g., continuous assessment vs. daily reports) each bring advantages and disadvantages with respect to participant fatigue and ability to capture fluctuating states. A particular advantage of the long (i.e., two semester) assessment period used in this study is that it eliminates noise from highly salient but seldom occurring events (e.g., work promotion). Nevertheless, appropriate caution should be applied when interpreting these results. Finally, whereas the estimation task was purposefully standardized in order to promote experimental control and was correctly understood and applied by the participants, this procedure differs from "naturally" generated estimations. Future research is needed to

develop additional methods of assessing estimations as well as documenting advantages and disadvantages with respect to "naturally" generated estimations.

Future research needs to examine whether and how clinical decisions are detrimentally affected by misestimations of associations between health behaviors and experiences. Researchers need to better control effects of illusory correlation estimates because the magnitude of discrepancies between estimation and actual occurrences may be equal to the effects of intervention studies and limit researchers' and practitioners' ability to capture active processes (Conner & Barrett, 2012; Frewen, Allen, Lanius, & Neufeld, 2012). Therefore, future research should also determine how daily-diary assessment can practically augment daily practice. We recommend that practitioners query only a limited, recent time frame and to capitalize on people's beliefs regarding the estimates of associations – particularly about physical activity. That is, it may be beneficial for physicians to query about the believed effects of physical activity in addition to frequency. Finally, additional affective states relevant in clinical and health contexts such as feeling fit, energized, calm, indices of well-being, etc. should be explored.

In sum, this study documents a new dimension of the memory-experience gap, namely illusory correlations between retrospectively estimated correlations and behaviors and experiences reported in the moment. The current study extends knowledge about illusory correlations to the health context, identifying important implications for health professionals and patients. Whereas this may be further evidence that our memories "favor prudence over accuracy" (Kahneman et al., 2004; Miron-Shatz et al., 2009) and may make evolutionarily sense, it may not be without cost when aiming to understand and positively influence health.

<center>**References**</center>

Affleck, G., Tennen, H., Urrows, S., & Higgins, P. (1994). Person and contextual features of daily stress reactivity - Individual-differences in relations of undesirable daily events with mood disturbance and chronic pain intensity. *Journal of Personality and Social Psychology, 66*, 329-340. doi: 10.1037/0022-3514.66.2.329

Ben-Zeev, D., & Young, M. A. (2010). Accuracy of hospitalized depressed patients' and healthy controls' retrospective symptom reports an experience sampling study. *Journal of Nervous and Mental Disease, 198*, 280-285. doi: 10.1097/NMD.0b013e3181d6141f

Buman, M. P., Hekler, E. B., Bliwise, D. L., & King, A. C. (2011). Exercise effects on night-to-night fluctuations in self-rated sleep among older adults with sleep complaints. *Journal of Sleep Research, 20*, 28-37. doi: 10.1111/j.1365-2869.2010.00866.x

Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index - a new instrument for psychiatric practice and research. *Psychiatry Research, 28*, 193-213. doi: 10.1016/0165-1781(89)90047-4

Conner, T. S., & Barrett, L. F. (2012). Trends in ambulatory self-report: The role of momentary experience in psychosomatic medicine. *Psychosomatic Medicine, 74*, 327-337. doi: 10.1097/PSY.0b013e3182546f18

de Beurs, E., Lange, A., & Vandyck, R. (1992). Self-monitoring of panic attacks and retrospective estimates of panic - Discordant findings. *Behaviour Research and Therapy, 30*, 411-413. doi: 10.1016/0005-7967(92)90054-K

Ebner-Priemer, U. W., & Trull, T. J. (2009). Ecological momentary assessment of mood disorders and mood dysregulation. *Psychological Assessment, 21*, 463-475. doi: 10.1037/a0017075

Ekkekakis, P. (2003). Pleasure and displeasure from the body: Perspectives from exercise. *Cognition & Emotion, 17*, 213-239. doi: http://dx.doi.org/10.1080/02699930302292

Flueckiger, L., Lieb, R., Meyer, A. H., & Mata, J. (2014). How health behaviors relate to academic performance via affect: An intensive longitudinal study. *Plos One, 9*. doi: 10.1371/journal.pone.0111080

Frewen, P. A., Allen, S. L., Lanius, R. A., & Neufeld, R. W. J. (2012). Perceived causal relations: Novel methodology for assessing client attributions about causal associations between variables including symptoms and functional impairment. *Assessment, 19*, 480-493. doi: 10.1177/1073191111418297

Gigerenzer, G., Hoffrage, U., & Kleinbolting, H. (1991). Probabilistic mental models - a Brunswikian theory of confidence. *Psychological Review, 98*, 506-528. doi: 10.1037/0033-295x.98.4.506

Gloster, A. T., Richard, D. C. S., Himle, J., Koch, E., Anson, H., Lokers, L., & Thornton, J. (2008). Accuracy of retrospective memory and covariation estimation in patients with obsessive-compulsive disorder. *Behaviour Research and Therapy, 46*, 642-655. doi: 10.1016/j.brat.2008.02.010

Godin, G., & Shephard, R. (1997). Godin leisure-time exercise questionnaire. *Med Sci Sports Exerc, 29*, 36-38.

Haynes, S. N., Leisen, M. B., & Blaine, D. D. (1997). Design of individualized behavioral treatment programs using functional analytic clinical case models. *Psychological Assessment, 9*, 334-348.

Housman, J., & Dorman, S. (2005). The Alameda County study: a systematic, chronological review. *Journal of Health Education, 36*, 302-308.

Houtveen, J. H., & Oei, N. Y. L. (2007). Recall bias in reporting medically unexplained symptoms comes from semantic memory. *Journal of Psychosomatic Research, 62*, 277-282. doi: 10.1016/j.jpsychores.2006.11.006

Jacobs, N., Myin-Germeys, I., Derom, C., Delespaul, P., van Os, J., & Nicolson, N. A. (2007). A momentary assessment study of the relationship between affective and adrenocortical stress responses in daily life. *Biological Psychology, 74*, 60-66. doi: 10.1016/j.biopsycho.2006.07.002

Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., & Stone, A. A. (2004). A survey method for characterizing daily life experience: The day reconstruction method. *Science, 306*, 1776-1780. doi: 10.1126/Science.1103572

Kirmayer, L. J., Robbins, J. M., Dworkind, M., & Yaffe, M. J. (1993). Somatization and the recognition of depression and anxiety in primary care. *American Journal of Psychiatry, 150*, 734-741.

Marks, M., & Hemsley, D. (1999). Retrospective versus prospective self-rating of anxiety symptoms and cognitions. *Journal of Anxiety Disorders, 13*, 463-472. doi: 10.1016/S0887-6185(99)00015-8

Mata, J., Thompson, R. J., Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Gotlib, I. H. (2012). Walk on the bright side: physical activity and affect in major depressive disorder. *Journal of Abnormal Psychology, 121*, 297-308. doi: 10.1037/a0023533

Meiser, T., & Hewstone, M. (2010). Contingency learning and stereotype formation: Illusory and spurious correlations revisited. *European Review of Social Psychology, 21*, 285-331. doi: 10.1080/10463283.2010.543308

Miron-Shatz, T., Stone, A., & Kahneman, D. (2009). Memories of yesterday's emotions: Does the valence of experience affect the memory-experience gap? *Emotion, 9*, 885-891. doi: 10.1037/a0017823

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

R Core Team. (2014). R: A language and environment for statistical computing, from http://www.R-project.org/

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin, 128*, 934-960. doi: 10.1037//0033-2909.128.6.934

Roecke, C. (2006). *Intraindividual variability in positive and negative affect: Age-related and individual differences in magnitude and coupling with cognitive performance.* Freie Universität Berlin. Retrieved from http://www.diss.fu-berlin.de/diss/receive/FUDISS_thesis_000000002464

Reed, J. & Ones, D. S. (2006). The effect of acute aerobic exercise on positive acitvated affect: A meta-analysis. *Psychology of Sport and Exercise, 7,* 477-514.

Sadeh, A., Keinan, G., & Daon, K. (2004). Effects of stress on sleep: The moderating role of coping style. *Health Psychology, 23*, 542-545. doi: 10.1037/0278-6133.23.5.542

Schneiderman, N., Ironson, G., & Siegel, S. D. (2005). Stress and health: Psychological, behavioral, and biological determinants. *Annual Review of Clinical Psychology, 1*, 607-628. doi: 10.1146/annurev.clinpsy.1.102803.144141

Shiffman, S., Hufford, M., Hickcox, M., Paty, J. A., Gnys, M., & Kassel, J. D. (1997). Remember that? A comparison of real-time versus retrospective recall of smoking lapses. *Journal of Consulting and Clinical Psychology, 65*, 292-300. doi: 10.1037/0022-006x.65.2.292.A

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford, UK: Oxford University Press.

Stein, K. F., & Corte, C. M. (2003). Ecologic momentary assessment of eating-disordered behaviors. *International Journal of Eating Disorders, 34*, 349-360. doi: 10.1002/eat.10194

Stone, A. A., Broderick, J. E., Shiffman, S. S., & Schwartz, J. E. (2004). Understanding recall of weekly pain from a momentary assessment perspective: absolute agreement, between- and within-person consistency, and judged change in weekly pain. *Pain, 107*, 61-69. doi: 10.1016/j.pain.2003.09.020

Stone, A. A., Schwartz, J. E., Neale, J. M., Shiffman, S., Marco, C. A., Hickcox, M., . . . Cruise, L. J. (1998). A comparison of coping assessed by ecological momentary assessment and retrospective recall. *Journal of Personality and Social Psychology, 74*, 1670-1680. doi: 10.1037/0022-3514.74.6.1670

U.S. Department of Health and Human Services Food and Drug Administration (FDA). (2006). *Guidance for industry: Patient-reported outcome measures: Use in medical product development to support labeling claims*.

Van Rooy, D., Vanhoomissen, T., & Van Overwalle, F. (2013). Illusory correlation, group size and memory. *Journal of Experimental Social Psychology, 49*, 1159-1167. doi: 10.1016/j.jesp.2013.05.006

Watson, D. (1988). Intraindividual and interindividual analyses of positive and negative affect - Their relation to health complaints, perceived stress, and daily activities. *Journal of Personality and Social Psychology, 54*, 1020-1030. doi: 10.1037/0022-3514.54.6.1020

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect - The Panas Scales. *Journal of Personality and Social Psychology, 54*, 1063-1070. doi: 10.1037/0022-3514.54.6.1063

Wilson, T. D., & Gilbert, D. T. (2005). Affective forecasting. Knowing what to want. *Current Directions in Psychological Science, 14*, 131-134.

Wu, H., Mata, J., Furman, D. J., Whitmer, A. J., Gotlib, I. H., & Thompson, R. J. (2016). Anticipatory and consummatory pleasure and displeasure in major depressive disorder: An experience sampling study. *Journal of Abnormal Psychology. Advance Online Publication*. doi: 10.1037/abn0000244

Table 1

*Sample Characteristics*

|  | (N=168) |
| --- | --- |
| Gender (women %) | 81.5% |
| Age (*M, SD*) | 20.6 (2.3) |
| Minutes of sleep/ night (*M, SD*) | 447 (39) |
| Physical activity (*M, SD*) | 629 (365) |
| Positive affect (*M, SD*) | 4.8 (1.1) |
| Stress (*M, SD*) | 2.0 (1.7) |

*Note.* Physical activity in minutes weighted by intensity; stress represents sum of hassles per day.

Table 2

*Means, Standard Deviations, and statistical difference values for person's estimated versus daily-diary-calculated correlations*

| Correlation | Estimated association M (SD) | Daily-diary-calculated association1 M (SD) | Paired t-test; difference estimated vs. daily-diary-calculated correlation | Effect size (d) |
|---|---|---|---|---|
| Physical Activity & Positive Affect | 18.5 (3.6)** | 14.14 (1.4)** | $t(153) = 1.17$; $p = 0.24$ | 0.13 |
| Sleep & Positive Affect | 24.37 (2.8)** | 3.77 (1.5)* | $t(153) = 6.71$; $p < .001$ | 0.81 |
| Sleep & Physical Activity | 10.76 (2.9)** | -4.29 (1.6)* | $t(153) = 4.66$; $p < .001$ | 0.55 |
| Stress & Positive Affect | -36.2 (3.9)** | -10.52 (1.7)** | $t(152) = -5.47$; $p < .001$ | 0.68 |
| Stress & Physical Activity | -3.72 (2.9) | 5.06 (1.6)** | $t(152) = -2.40$; $p=0.011$ | 0.26 |
| Stress & Sleep | -16.34 (3.2)** | -7.91 (1.4)** | $t(152) = 2.16$; $p = 0.010$ | 0.26 |

*Note.* ** $p < .001$, * $p < .05$ for one-sampled *t*-tests examining whether a correlation was significantly different from zero.

Table 3

*Association between person's estimated and daily-diary-calculated correlations*

| Correlation | z-value[1] | *p*-value |
| --- | --- | --- |
| Physical Activity & Positive Affect | 0.04 | 0.816 |
| Sleep & Positive Affect | 2.01 | 0.134 |
| Sleep & Physical Activity | 2.81 | 0.218 |
| Stress & Positive Affect | 1.75 | 0.148 |
| Stress & Physical Activity | 1.14 | 0.206 |
| Stress & Sleep | 1.09 | 0.108 |

*Note.* [1]*z*-value for the correlation between people's estimated correlation and daily-diary-calculated correlation based on multilevel structural equation models.

Table 4

*Do daily-diary-calculated correlations correspond to person's estimated correlations, if the person's correlation estimates are categorized as positive, negative or zero?*

| Correlation | Estimated correlation negative | Estimated correlation zero | Estimated correlation positive | Statistical difference between categories |
|---|---|---|---|---|
| Physical Activity & Positive Affect | .15 | .13 | .14 | $F$ (2,163) =0.11 $p$ = .89 |
| Sleep & Positive Affect | -.02 | .01 | .05 | $F$ (2,163) = 1.528 $p$ = .22 |
| Sleep & Physical Activity | -.10 | -.01 | -.05 | $F$ (2,164) = 2.313 $p$ = .10 |
| Stress & Positive Affect | -.12 | -.01 | -.05 | $F$ (2,159) = 1.915 $p$ = .15 |
| Stress & Physical Activity | .04 | .05 | .06 | $F$ (2,160) = 0.10 $p$ = .90 |
| Stress & Sleep | -.09 | -.05 | -.07 | $F$ (2,160) = 0.87 $p$ = .42 |

*Note.* F-values are based on an omnibus test.

Table 5a

*Are people's estimated correlations about physical activity with positive affect, stress, or sleep associated with actual reported physical activity?*

| Pearson's estimated Correlation | Daily-diary data | *Corr. coeff.* | *t*-value | *p*-value |
|---|---|---|---|---|
| Physical Activity & Positive Affect | physical activity | .14 | 1.80 | 0.074 |
| Sleep & Physical Activity | physical activity | .25 | 3.29 | 0.001 |
| Stress & Physical Activity | physical activity | .16 | 2.14 | 0.034 |

Table 5b

*Are people's correlations calculated based on the daily diaries about physical activity with positive affect, stress, or sleep associated with actual reported physical activity?*

| Pearson's correlation based on daily diaries | Daily-diary data | *Corr. coeff.* | *t*-value | *p*-value |
|---|---|---|---|---|
| Physical Activity & Positive Affect | physical activity | .08 | 1.39 | 0.167 |
| Sleep & Physical Activity | physical activity | .04 | 0.70 | 0.483 |
| Stress & Physical Activity | physical activity | -.06 | -0.96 | 0.340 |