EXPERIENCE SAMPLING METHODS (ESM) IN ORGANIZATIONS: A REVIEW

BY

KATHLEEN A. YEARICK

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Arts in Psychology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Master's Committee:

Professor Daniel A. Newman, Chair
Professor James Rounds

ABSTRACT

We review research designs of ESM studies conducted in workplace settings (k = 167 samples). Eight ESM design features are summarized: sample size and number of observations, response rates, recruitment methods and incentives, survey timing factors (study duration, signal frequency, times of day), signaling strategies and reminder technologies, survey media, survey items (number of items, item sampling, constructs measured), and analytic strategies (lagged analyses, missing data treatment). Mean sample size was 93 and number of observations was 1,419. Average study duration was 10.13 days. Among studies that used multiple signals per day (56%), the average was 4.16 signals per day. 54% of studies did not report using incentives, 41% did not use reminders. Longer studies were more likely to provide incentives. Over time, online surveys are rising whereas paper-and-pencil surveys are disappearing. The average *between-persons response rate* was 63%, and *within-persons response rate* was 80%; although response rates were unrelated to incentives/design features. Interval-contingent signaling was most prevalent (59%), followed by signal-contingent signaling (19%). Event-contingent signaling was rare (4%). Few studies reported missing data treatments. Findings and implications are discussed.

ACKNOWLEDGEMENTS

Thanks to Angela Lee for collaborating with me on this project!

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

Organizational scholars have increasingly highlighted the dynamic nature of commonly studied constructs (e.g., Austin, Humphreys, & Hulin, 1989; Beal, Trougakos, Weiss, & Green, 2006; Beal, Weiss, Barros, & MacDermid, 2005; Dalal, Bhave, & Fiset, 2014; Ghiselli, 1956; Liu, Wang, Zhan, & Shi, 2009), and have adopted methods that capture within-persons changes and their consequences. In contrast to traditional cross-sectional designs, experience sampling methods (ESM) provide ways to record temporary, specific psychological states that spontaneously occur in the moment (Beal & Weiss, 2003; Bolger, Davis, & Rafaeli, 2003). Also known as ecological momentary assessment (EMA), daily diary studies, or event-sampling, ESM rests on the crucial ideas that lawful change occurs within individuals and that within-persons variance is not random error but in fact represents meaningful variance (Beal & Weiss, 2003). As seen in Figure 1, the number of ESM studies conducted in the workplace has increased (from one study in 1989 to twenty-five studies in 2014). A thorough understanding of norms and best practices in the implementation of ESM designs is important if we are to maximize the utility of this research method.

ESM study designs require many decisions on the part of the researcher, and the lack of clear norms may be preventing researchers from using ESMs to their full potential. Even scholars who routinely employ these methods may not recognize the full range of options available to them, and this lack of familiarity could hinder creation of optimal designs to fulfill particular studies' needs. Perhaps more importantly, researchers who have little or no experience with ESM research may not know how their own planned designs measure up to previous work, in terms of sample sizes, industries sampled, constructs measured, incentives provided, signaling

strategies, recruitment strategies, number and timing of daily surveys, between- and within-persons response rates, etc.

The purpose of the current paper is to increase awareness of ESM as a valuable tool to assess within-persons change in organizations, by summarizing ESM research practices. This paper extends previous introductions and reviews of the method in organizational settings (Beal & Weiss, 2003; Miner, Glomb, & Hulin, 2005; Ohly, Sonnentag, Niessen, & Zapf, 2010) by systematically analyzing 167 workplace studies that used ESM. Specifically, we conduct a comprehensive review of the qualitative and quantitative characteristics of ESM studies by outlining the eight major decisions required in ESM study design.

Before presenting this summary, we first provide more background information about conceptual and methodological aspects of ESM. Next, we describe two central beliefs that drive the implementation of ESM. Finally, we pose a series of research questions guiding our investigation.

Background of ESM

The use of ESM to study within-persons change originated in social and personality psychology (e.g., Csikszentmihalyi & Larson, 1987; Diener, Larsen, Levine, & Emmons, 1985; Wheeler & Reis, 1991; Zevon & Tellegen, 1982). ESM allows researchers to measure ongoing phenomena occurring in the natural environment with greater precision (Alliger & Williams, 1993; Hormuth, 1986), as well as short-term fluctuations on a daily basis (Ohly et al., 2010). Participants usually receive signals (e.g., from a beeper watch, a PDA [personal digital assistant], or email) one to seven times per day, for about one day to three weeks of study duration. These signals prompt immediate responses to survey items, and as a result participants provide responses at multiple time points. ESM is characterized by reduced memory bias, because

participants are typically asked to record psychological states "at this moment" rather than attempt to recall experiences in general or over an extended period of time, as required in a cross-sectional design (Alliger & Williams, 1993; Beal & Weiss, 2003; Ohly et al., 2010; Robinson & Clore, 2002).

ESM offers several advantages by acknowledging that variance previously assumed to be random error could in fact be meaningful *within-persons* or *intraindividual* variance. One purpose of ESM is to account for fluctuations over time by studying variables at the within-persons level. There is growing consensus among organizational scholars that subjective work experiences may be fleeting and changing (Alliger & Williams, 1993; Illies, Scott, & Judge, 2006; Sonnentag, Binnewies, & Mojza, 2008) and that criteria are dynamic (e.g., Austin et al., 1989; Dalal et al., 2014); and these dynamics call for the use of ESM. Indeed, empirical findings show that 56% of the variance in mood (Miner et al., 2005) and 62% of the variance in task performance (Dalal et al., 2014) occur at the within-persons level. Other constructs that exhibit within-persons variation include job affect (Fisher, 2000; Weiss, Nicholas, & Daus, 1999), work engagement and vigor (Sonnentag, 2003; Xanthopoulou, Bakker, Demerouti, & Schaufeli, 2009), counterproductive work behavior (CWB; Dalal, Lam, Weiss, Welch, & Hulin, 2009), and organizational citizenship behavior (OCB; Dalal et al., 2009; Ilies et al., 2006). By studying variables at the within-persons level, researchers can uncover new insights. For example, Dalal et al. (2014) found that—at the within-persons level of analysis—positive affect was more related to OCB, whereas negative affect was more related to CWB, thus arguing that CWB and OCB are two separate constructs. Additional research has also revealed more about the nature of criterion variables; for example, the within-persons level of OCB is related to several work variables

including mood regulation (Glomb, Bhave, Miner, & Wall, 2011), positive affect, and job satisfaction (Ilies et al., 2006).

Second, relationships assessed at the between-persons or *interindividual* level are independent from, and may differ greatly from, those studied at the within-persons or *intraindividual* level (Beal & Weiss, 2003; Fisher, 2003; Ostroff, 1993; Thorndike, 1939). Relationships at different levels of analysis can vary in terms of sign and magnitude (Bliese, Chan, & Ployhart, 2007; Dalal et al., 2014). For example, at the between-persons level, individuals who chronically exercise have lower blood pressure compared to those who rarely exercise; but at the within-persons level, exercisers have higher blood pressure in the moment of exercising than they do in moments of rest (Schwartz & Stone, 1998). Furthermore, the relationship between self-efficacy and task performance differs such that the two are positively related when studied at the between-persons level; but can be positively, negatively, or even unrelated at the within-persons level (Vancouver, Thompson, Tischner, & Putka, 2002). If relationships are similar at the between-persons and pooled within-persons levels, ESM studies can nonetheless reveal individual differences in the within-persons relations. Considering positive affect and negative affect for example, the between-persons correlation (-.30) and the average within-persons correlation (-.32) happen to be almost equal, but the within-persons relationship between positive and negative affect fluctuates greatly among individuals, ranging from strongly negative to moderately positive (Alliger & Willliams, 1993). In short, it is useful to examine relationships between variables at different levels of analysis, because doing so might reveal new phenomena.

Based on the empirical evidence cited above, we believe ESM is an important research method that allows organizational researchers to investigate dynamic processes at the within-

persons/state level, in addition to the between-persons/trait level, as well as mitigates retrospective recall bias. The present paper aims to provide descriptive norms about design choices in ESM by reviewing qualitative and quantitative features of studies conducted in the workplace. In the section below, we describe methodological characteristics of ESM studies and pose research questions to examine the average frequencies of design choices and outcomes.

Sample Size and Number of Observations

In contrast to the cross-sectional design, ESM requires a substantial commitment from participants by requesting responses at multiple time points. This burden might result in reduced sample sizes. However, small between-persons sample sizes do not prohibit inferences about pooled within-persons effects, particularly when the number of observations per participant is large and there is a large number of total data points (i.e., persons × occasions).

Research Question 1a: What is the average sample size (*N*) in ESM studies?

Research Question 1b: What is the average number of total observations in ESM studies?

Response Rates

Because participants have more responsibilities in an ESM study, response rates may be lower compared to cross-sectional studies. Additionally, there are multiple types of response rate in an ESM study, and we must distinguish the between-persons response rate from the within-persons response rate (Newman & Sin, 2007). As a further complication, there is more than one type of between-persons response rate that is typically reported in ESM studies. The various between-persons response rates in an ESM study can be defined as ratios involving the following three values: the number of potential participants contacted, the number of participants who responded to all or part of a survey, and the number of respondents included in the analyses by the researchers. First, we define the *between-persons response rate* as the number of individuals

who provided any response (including both full respondents and partial respondents), divided by the total number of people who were contacted with a survey invitation (Newman, 2014). Second, due to researchers' adopting different techniques to address missing data, some observations may have been excluded by the researcher. As such, *between-persons analysis rate* refers to the number of participants who are included in the final data analysis divided by the number who provided any response. Lastly, due to a variety of reasons, participants within a study often end up completing different numbers of daily/momentary surveys, so the *pooled within-persons response rate* is the number of daily/momentary surveys completed divided by the total number of daily/momentary signals sent, averaged across all individuals in the sample.

Research Question 2a: What is the average *between-persons response rate* (i.e., number of participants who provided any response / number of people contacted) in ESM studies?

Research Question 2b: What is the average *between-persons analysis rate* (i.e., number of participants included in data analysis / number of participants who provided any response) in ESM studies?

Research Question 2c: What is the average *pooled within-persons response rate* (i.e., number of daily surveys completed / number of possible daily surveys) in ESM studies?

Recruitment and Incentives

Because ESM studies require a higher degree of time commitment from participants, obtaining and retaining participants can be difficult. Recruitment strategies in ESM studies can be similar to those in cross-sectional designs and include posting paper flyers or online advertisements, contacting organizations to survey employees, or recruiting working students in school. Incentives can potentially help motivate people to participate in studies and complete

multiple surveys, and vary in form and monetary value. Rewards given in cross-sectional studies such as money, gift cards, and lottery drawings for prizes are often thought to be useful, but perhaps larger or different types of incentives are needed in ESM studies to compensate for the greater burden.

Research Question 3a: What are the frequencies of recruitment strategies used in ESM studies?

Research Question 3b: What are the frequencies of incentives used in ESM studies?

Research Question 3c: Do recruitment strategies and incentives relate to response rates in ESM studies?

Survey Timing Factors

ESM studies are characterized by surveys administered at multiple time points, so researchers must also make decisions about the timing of the surveys. It can be difficult to determine the frequency and duration of surveys needed to adequately capture the targeted within-person phenomenon, which can also vary depending on which constructs are studied. Some psychological processes (e.g., creativity [Amabile, Barsade, Mueller, & Staw, 2005]) may require several months to unfold, whereas others (e.g., affect [Daniels, Boocock, Glover, Hartley, & Holland, 2009]) can transpire in less than a day. Due to researchers' limited resources and concerns over participant burden, there may be a tradeoff between study duration (i.e., number of days) and the signal frequency (i.e., number of daily/momentary surveys) in an ESM study. For example, researchers who plan a three-week long study may choose to obtain responses only once per day, whereas researchers who adopt a one-week long design may sample participants four times a day. Another feature to consider is the time(s) of day at which surveys are scheduled. There are many options, such as before arriving at work, in the morning or afternoon

during work hours, before bed at night, or based on personal work shifts for jobs with irregular schedules (e.g., nurses, flight attendants). Moreover, the time of survey delivery could depend on the constructs of interest, such that different constructs are measured at different times. For instance, recovery measured at bedtime is related to improved affect measured the next morning (Sonnentag, Binnewies, & Mojza, 2008), flow in the afternoon is related to decreased burnout and increased work engagement at bedtime (Demerouti, Bakker, Sonnentag, & Fullagar, 2011), and positive reflection intervention in the afternoon is related to decreased stress and improved health at bedtime (Bono, Glomb, Shen, Kim, & Koch, 2013). Lastly, researchers typically want to obtain trait-level measures of study variables with a one-time baseline survey, which can be administered either before or after the within-persons ESM surveys.

Research Question 4a: What is the average study duration (i.e., number of days) of ESM studies?

Research Question 4b: What is the average signal frequency (i.e., number of daily/momentary surveys) in ESM studies?

Research Question 4c: How often are daily/momentary surveys scheduled at different times of day (i.e., morning, noon, early afternoon, late afternoon, night, other)?

Research Question 4d: Is the baseline survey (i.e., trait-level measures) typically administered before or after the daily/momentary surveys?

Signaling Strategies and Reminder Technologies

In order to deliver surveys to participants, researchers need to decide on the *signaling strategy* and *reminder technology* to use. Signaling strategy refers to the method by which signals alert participants to provide responses to a daily survey, and includes three major types: event-contingent, interval-contingent, and signal-contingent. Event-contingent signaling

describes recording responses based on the occurrence of specific events (Beal & Weiss, 2003; Wheeler & Reis, 1991), such as interpersonal conflicts with coworkers. Interval-contingent signaling refers to recording responses at regular, predetermined times (Wheeler & Reis, 1991), such as every day at 10:00 A.M. and 2:00 P.M. Signal-contingent signaling is defined as recording responses at random times throughout the day (Wheeler & Reis, 1991), usually when a reminder is provided by the researchers via email or PDA alert. These three signaling strategies have their respective advantages and disadvantages, and each can be more suitable for certain research questions. For example, event-contingent signaling is useful for low frequency behaviors such as CWB (Beal & Weiss, 2003), but participants may forget to complete surveys in the absence of a physical reminder (e.g., a beep). On the other hand, interval-contingent signaling can be more convenient because participants are aware of the exact times at which they should complete surveys. This could increase response rates; but because participants come to expect the signals, there may be problems of habituation and reactivity (Beal & Weiss, 2003). In signal-contingent signaling, if participants miss the random signals, then there may be increased missing data concerns. However, signal-contingent signaling may provide a more comprehensive sample of times across each day. An ESM feature closely related to signaling strategy (event-, interval-, or signal-contingent) is reminder technology, or the method used to send the signal to participants. Researchers can remind participants to record responses in several ways, including PDAs, email, beepers, and more.

Research Question 5a: What are the average frequencies of signaling strategies (i.e., event-, interval-, and signal-contingent) used in ESM studies?

Research Question 5a: What are the average frequencies of reminder technologies (e.g., PDA, email, and beeper) used in ESM studies?

Survey Media

The next feature to consider is the survey medium, or format, by which participants record responses. Survey media generally used in traditional cross-sectional research designs consist of paper-and-pencil surveys and online surveys. These survey formats can also be implemented in ESM studies, and an additional format includes PDAs. For ESM studies that adopt a paper-and-pencil format, it is questionable whether this method constitutes momentary assessment if there is a time lag before responses are recorded and the time lag is unknown. An advantage of PDAs and online surveys is that time stamps can be recorded to determine whether participants are compliant in reporting responses within appropriate time frames. Given the technological advances in recent years, there may have been a shift from prevalence in paper-and-pencil and PDAs to online surveys and smartphone applications.

Research Question 6: What are the average frequencies of survey media (e.g., paper-and-pencil, PDA, and online) used in ESM studies?

Survey Items

ESM enables researchers to investigate psychological processes at the within-persons level, but are there particular constructs that are frequently examined in ESM studies? Another issue is survey length, or the number of items in a survey. Researchers often seek the optimal balance between adequate construct coverage and reasonable burden for participants. ESM studies typically administer two types of surveys: the baseline survey that is administered at one point in time to measure trait-levels of variables, and the daily/momentary survey that captures within-persons responses at multiple time points. Survey length may be more of a concern for the daily survey than for the baseline survey, because the daily survey must be completed many times. Also, because participants are repeatedly exposed to and respond to the same items,

reactivity and habituation might occur (Bolger et al., 2003). A possible solution to reactivity is *item sampling*, in which *different* subsets of items measuring the *same* construct are selected from a larger item bank for each daily survey in addition to a fixed set of items (Miner et al., 2005). In this way, participants are exposed to some new items in each momentary survey, which is intended to prevent reactivity and priming effects.

Research Question 7a: How many items on average are in the baseline survey in ESM studies?

Research Question 7b: How many items on average are in daily/momentary surveys in ESM studies?

Research Question 7c: Which constructs are commonly studied at the within-persons level in ESM studies?

Research Question 7d: What is the frequency of using item sampling in ESM studies?

Analysis Properties

During data analysis, ESM researchers have the option to conduct lagged analyses to study the relationship between a variable at one time and another variable at a later time. Lagged analyses are believed to lend greater support for making causal inferences, compared to a cross-sectional or concurrent analyses. As with all longitudinal designs, missing data are inevitable, and methods to address this problem include maximum likelihood estimation, listwise deletion, pairwise deletion, and imputation (Enders, 2010).

Research Question 8a: What is the frequency of conducting lagged analyses in ESM studies?

Research Question 8b: What are the frequencies of missing data techniques used in ESM studies?

CHAPTER 2: METHOD

Literature Search and Inclusion Criteria

We conducted a keyword search in PsycINFO and ProQuest Dissertations databases through January 2015 to locate relevant studies. We also searched the Society for Industrial and Organizational Psychology (SIOP) conference programs for unpublished studies from 2011-2015. Keywords included *diary study*, *ecological momentary assessment* (EMA), *experience sampling method* (ESM), and *within-persons* in combination with common organizational topics, such as job performance, organizational citizenship behavior, and work-family conflict (see Appendix A for a complete list of the search terms). To be eligible for inclusion, a study had to implement an ESM design (i.e., to prompt participants at multiple time points using a signaling strategy) in a workplace setting (i.e., sample of employed participants). Our search yielded 156 studies (130 published and 26 unpublished), including 167 independent samples.

Coding

For each independent sample, we coded sample size (*N)*, number of total observations, *between-persons response rate* (including both full respondents and partial respondents, divided by the number of people contacted with a survey invitation; Newman, 2014), *between-persons analysis rate* (number of respondents included in the analyses, divided by number of full and partial respondents), *pooled within-persons response rate* (total number of observations divided by number of possible observations [persons × occasions], sometimes referred to as the "compliance rate" or "overall response rate"), recruitment strategy, incentive type, study duration (number of days), signal frequency (number of surveys per day), time(s) of daily surveys, timing of baseline survey, signaling strategy, reminder technology, survey medium, number of items in

baseline survey, number of items in daily/momentary surveys, constructs studied at the within-persons level, use of item sampling, use of lagged analyses, and missing data treatment.

Recruitment strategy

We coded types of recruitment strategy that were used by researchers to obtain study participants: (1) online or hard copy advertisements, (2) researchers' contacting organizations, (3) working students from class, (4) drawn from a larger study, (5) other (i.e., not one of the four strategies listed above), or (6) multiple recruitment strategies (i.e., a combination of some of the first four strategies).

Incentive type

Type of incentive offered to participants as compensation was coded: (1) none/unreported, (2) money, (3) gift card, (4) lottery (i.e., entered into a drawing for prizes such as money [e.g., Fuller et al., 2003; Miner & Glomb, 2010] or gift cards [e.g., Kim, 2012]), (5) other (i.e., not one of the four types listed above), or (6) multiple (i.e., a combination of some of the first four types). When incentive type was money, the US dollar amount was recorded. When incentive type was lottery, the prize was also recorded.

Survey timing factors

For the time of day when surveys were administered, six major categories were coded: (1) morning (i.e., 7:00 A.M. to noon, including before work), (2) noon, (3) early afternoon (i.e., 1:00 P.M. to 4:00 P.M.), (4) late afternoon (i.e., 4:00 P.M. to 6:00 P.M., including end of the workday and right after work), (5) night (i.e., 6:00 P.M. to 10:00 P.M., including early evening and before bed), or (6) other (e.g., random signals throughout the day, irregular work hours such as those of nurses and flight attendants). Unlike most of our coded variables, the time(s) of daily surveys contained non-mutually exclusive categories; that is, each sample coded could contain

any combination of categories (e.g., morning, noon, and night). Additionally, we coded whether each sample received a baseline survey measuring baseline information chronologically: (1) before the ESM phase of the study, (2) after the ESM phase, or (3) both before and after the ESM phase.

Signaling strategy

Type of signaling strategy, the procedure by which participants were prompted to complete daily surveys, was coded as: (1) event-contingent, (2) interval-contingent, (3) signal-contingent, (4) multiple (i.e., more than one signaling strategy; most often interval- and signal-contingent signaling [e.g., Reicherts & Pihet, 2000]), (5) other (e.g., researchers hand-delivered each survey onsite [Wang et al., 2013]), or (6) none (i.e., no signaling strategy was specified).

Reminder technology

The means by which researchers notified participants to complete a daily survey were coded: (1) none, (2) text message, (3) email, (4) PDA, (5) beeper, (6) smartphone application, (7) paper reminder, (8) watch, (9) multiple (i.e., more than one reminder technology), or (10) other (e.g., researcher was on-site to give reminders [Grech, Neal, Yeo, Humphreys & Smith, 2009]).

Survey medium

Type of survey medium, the method by which study participants provided daily survey responses, was coded into six categories: (1) paper-and-pencil, (2) PDA, (3) online, (4) smartphone application, (5) phone calls, or (6) multiple (i.e., more than one survey medium).

Item sampling

We coded dichotomously (i.e., “yes” or “no”) whether studies measured the same items across time points. For studies that did not measure the same items across time points, we

recorded whether studies used (1) item sampling or (2) measured different constructs across time points.

Constructs

To address the question of the constructs, or topics, that have been studied at the within-persons level in ESM studies, 38 categories of constructs were coded dichotomously (i.e., "yes" or "no") to indicate whether a sample measured the particular construct. As presented in Table 7, these 38 categories were subsequently collapsed into nine broad construct categories for simplicity. Because studies often measured multiple constructs, the construct categories are not mutually exclusive.

Analytic strategies

The use of lagged analyses was coded as a dichotomous variable (i.e. "yes" or "no"). The missing data treatment used for each sample was coded as one of the following five categories: (1) listwise deletion, (2) pairwise deletion, (3) deleting outliers, (4) single imputation, (5) multiple imputation, or (6) unreported (i.e., did not indicate how missing data were handled).

Sample characteristics

We coded whether the study was published ("yes" or "no"), the year it was published (or the year it was completed, for unpublished studies), and the country in which the study was conducted. We also coded the sample's employment status (i.e., full-time employees, primarily full-time employees, part-time employees, working students, or MBA students), job, industry, mean organizational tenure, mean job tenure, percent male, percent Caucasian, and the number of organizations from which the sample was drawn (i.e., a single organization versus multiple organizations).

Interrater Reliability

The first two authors independently coded all the studies, and the interrater reliability indices are displayed in Tables 1 and 2. For numerical variables, the average interrater reliability (r = .97; ICC1 = .97) was very high. The average interrater reliability for categorical variables (κ = .75) indicates moderate to high agreement (Landis & Koch, 1977). However, we note that reliability indices below .70 (i.e., for recruitment strategy, time(s) of daily surveys, and some construct categories) represent variables that are both subjective in nature and often described ambiguously in primary studies. For the categorical variables, because reliability estimates were lower, interrater reliability was recalculated after interrater interpretational discrepancies in construct definitions were identified and corrected (see Appendix C for a list of each coder discrepancy and how it was resolved). This resulted in an average interrater reliability of .84 for the categorical variables.

Analyses

For numerical variables, we calculated mean, standard deviation, median, minimum, maximum, and quartile values. For categorical variables, we calculated frequencies and percentages. We note that not all studies reported information for every variable (e.g., recruitment strategy, timing of baseline survey, and survey medium).

Additionally, we examined whether response rates differed based on ESM design features. First, we conducted t tests to compare the response rates (i.e., *between-persons response rates, between-persons analysis rates,* and *pooled within-person response rates*) between samples that received incentives and samples for which there were no (or unreported) incentives. We also compared response rates between samples that received reminders and those that did not receive reminders. Furthermore, we analyzed the standardized mean differences (i.e., Cohen's *d*) in response rates in these samples: incentives vs. no (or unreported) incentives, reminders vs. no

reminders. Next, we conducted *t* tests to compare the signal frequency as well as the study duration between samples that received incentives and those in which there were no (or unreported) incentives. Then, we conducted ANOVAs to compare samples that used different recruitment strategies, signaling strategies, reminder technologies, and survey media. Finally, we calculated the correlations between response rates and numerical variables, including monetary incentive amount, study duration, signal frequency, number of items per daily survey, and year that study was published/completed.

CHAPTER 3: RESULTS

Results for numerical variables are displayed in Table 3, and results for categorical variables are displayed in Tables 4-7.

Sample Characteristics

Figure 2 illustrates the frequencies of the countries in which ESM studies were conducted; studies were most commonly conducted in the US (43%; $k = 63$). Frequencies and percentages of sample types are displayed in Figure 3, which indicates that the majority of studies was conducted on full-time employees (69%; $k = 101$). As presented in Tables 4 and 5, employees were most commonly sampled from multiple jobs (62%; $N = 93$), multiple industries (46%; $k = 72$), and multiple organizations (68%; $k = 110$) rather than from a single job, industry, or organization. Additional descriptive statistics are reported in Table 3. Samples' average job tenure was 6 years ($SD = 5$) and average organizational tenure was 9 years ($SD = 5$). Samples were 58% female and 73% Caucasian.

Sample Sizes

The mean sample size (N) was 93 participants ($SD = 99$, $Mdn = 69$), answering Research Question 1a. Sample size and year of study were positively related ($r = .17$; $k = 167$, $p < .05$), indicating the sample sizes have grown over time. The mean number of observations was 1,419 ($SD = 1{,}951$, $Mdn = 650$), addressing Research Question 1b (see Table 3).

Response Rates

To answer Research Questions 2a, 2b, and 2c, the mean *between-persons response rate* was 63.30% ($SD = 22.97\%$, $Mdn = 68\%$), the mean *between-persons analysis rate* was 79.02% ($SD = 19.05\%$, $Mdn = 86\%$), and the mean *pooled within-persons response rate* was 79.87% ($SD = 11.49\%$, $Mdn = 82\%$; see Table 3). Next, we compared response rates across different design

conditions (described in the sections below), but found no statistical effects of study design features on response rates (see Appendix B).

*Incentives vs. no/unreported incentives***.** There was no difference in *between-persons response rate* between samples that received incentives (M = 60.81%) and samples for which there were no (or unreported) incentives (M = 62.23%, $t(85.41) = .26$, $p = .79$; $d = .05$). Similarly, there was no difference in *between-persons analysis rate* between samples with incentives (M = 78.81%) and no (or unreported) incentives (M = 74.99%, $t(79.73) = -.76$, $p = .45$; $d = -.08$). And there was no difference in *pooled within-person response rate* for samples with incentives (M = 79.16%) versus no (or unreported) incentives (M = 74.67%, $t(55.00) = -.98$, $p = .33$; $d = -.34$).

Reminders vs. no reminders. The *pooled within-persons response rate* was no different between samples that received reminders (M = 79.43%) compared to samples that did not receive reminders (M = 84.38%, $t(3.59) = .86$, $p = .44$; $d = .74$; note this test has low statistical power because the number of studies in the *no reminder* condition was only $k = 5$; see Appendix B).

Recruitment strategies. Recruitment strategies had no influence on the *between-persons response rate* (one-way ANOVA $F(5,72) = .26$, $p = .94$), *between-persons analysis rate* ($F(5,71) = 1.08$, $p = .38$), or the *pooled within-persons response rate* ($F(4,40) = .15$, $p = .97$).

Signaling strategies. The *pooled within-persons response rates* were similar in studies that used signal-contingent (M = 77.94%) and interval-contingent signaling (M = 81.82%; $F(4,46) = 1.37$, $p = .26$). Studies that used event-contingent signaling did not provide information on the *pooled within-persons response rate*.

Reminder technologies. Reminder technologies had no influence on the *between-persons response rate* ($F(8,76) = 1.33$, $p = .25$), *between-persons analysis rate* ($F(8,74) = .47$, $p = .87$), or the *pooled within-persons response rate* ($F(8,42) = 1.63$, $p = .15$).

Survey media. Survey media had no influence on the *between-persons response rate* ($F(4,72) = 1.24$, $p = .30$), *between-persons analysis rate* ($F(5,73) = .77$, $p = .58$), or the *pooled within-persons response rate* ($F(5,44) = 1.13$, $p = .36$).

Numerical variables. For all three response rates, there was no relation with monetary incentive amount, study duration, signal frequency (number of signals per day), number of items on the momentary or daily survey, or year (see Table 8).

Recruitment and Incentives

Figure 4 displays the frequencies and percentages of type of recruitment strategy. To answer Research Question 3a, the most commonly used recruitment strategy was researchers' contacting organizations (58%; $k = 87$), followed by online or hard copy advertisements (17%; $k = 26$), being part of a larger study (12%; $k = 18$), working students from class (6%; $k = 8$), other strategies (5%; $k = 8$), and multiple strategies (2%; $k = 3$).

Frequencies and percentages of type of incentive are presented in Figure 5. To answer Research Question 3b, most studies did not report providing incentives (54%; $k = 90$). When studies did provide incentives, money was the most popular (26%; $k = 44$), with an average US Dollar amount of \$57.33 per person (SD = \$41.82, Mdn = \$50.00).

Survey Timing Factors

On average ESM studies lasted 10.13 days ($SD = 9.88$, $Mdn = 8$), addressing Research Question 4a. Figure 13 depicts the distribution of studies by study duration. Studies that provided incentives lasted longer ($M = 13.4$ days) than those that provided no (or unreported) incentives

(*M* = 6.7 days; *t*(85.71) = -3.75, *p* < .001). Participants completed an average of 2.90 surveys (*SD* = 3.13, *Mdn* = 2) per day, answering Research Question 4b (see Table 3). Among the subset of studies that used multiple signals per day (56% of studies), the average signal frequency per day was 4.16. Figure 14 illustrates the distribution of signal frequency. Signal frequency did not differ in samples that provided incentives and those with no (or unreported) incentives (*t*(130.89) = 1.25, *p* = .22).

Figure 6 displays the frequencies and percentages of the time of day when daily/momentary surveys were administered. To answer Research Question 4c, the most popular survey time was late afternoon (35%; *k* = 59), followed by morning (28%; *k* = 46), night (26%; *k* = 44), other (21%; *k* = 35), early afternoon (19%; *k* = 32), and noon (5%; *k* = 8).

In regard to Research Question 4d, the vast majority of samples (88%; *k* = 120) administered a baseline survey *before* the ESM phase, whereas a few (4%; *k* = 5) administered a baseline survey after the ESM phase, and some (8%; *k* = 11) included a baseline survey both before and after the ESM phase.

Signaling Strategy and Reminder Technology

Figure 7 illustrates the frequencies and percentages of signaling strategies. To answer Research Question 5a, interval-contingent signaling (59%; *k* = 99) was most prevalent, followed by signal-contingent signaling (19%; *k* = 31), no signaling (i.e., as in paper-and-pencil diary studies; 10%; *k* = 17), multiple (7%; *k* = 12), event-based signaling (4%; *k* = 7), and other signaling strategies (1%; *N* = 1).

Frequencies and percentages of reminder technologies are presented in Figure 8. To address Research Question 5b, many studies did not use any reminders (41%; *k* = 69), and when studies did use reminders, email (23%; *k* = 38) and PDA (20%; *k* = 33) were the most common.

Other reminder technologies included beeper (4%; *k* = 7), other (4%; *k* = 7), watch (2%; *k* = 4), text message (2%; *k* = 3), multiple (2%; *k* = 3), paper (1% *k* = 2), and smartphone application (1%; *k* = 1; Foo, Uy & Baron, 2009).

Survey Medium

Figure 9 displays the frequencies and percentages of survey media used. Paper-and-pencil (*k* = 62; 40%) was the most popular, followed by PDAs (27%; *k* = 42) and online surveys (26%; *k* = 41). Additional survey formats included multiple (3%; *k* = 5), phone call (3%; *k* = 4), and smartphone application (1%; *k* = 2). Figure 10 illustrates the use of the three most common survey media across time, revealing an overall increase in all three formats, along with a spike in online surveys corresponding to a drop in paper-and-pencil and PDA surveys starting in 2013.

Survey Items

Baseline surveys contained an average of 33 items (*SD* = 43, *Mdn* = 20), answering Research Question 7a, whereas daily surveys contained an average of 21 items (*SD* = 15, *Mdn* = 18), answering Research Question 7b (see Table 3). To answer Research Question 7c, most samples administered the same items at each time point (72%; *k* = 121), very few used item sampling (5%; *k* = 8), and one-in-four measured different constructs across time (23%; *k* = 38).

Table 6 and Figure 11 display the frequencies and percentages of the nine construct categories studied at the within-persons level. Results for the 38 specific constructs are presented in Table 7. To answer Research Question 7d, affect/emotion/mood was most commonly studied (60%; *k* = 100), followed by stress (45%; *k* = 75), work behaviors (37%; *k* = 61), health (35%; *k* = 58), situational factors (32%; *k* = 53), job attitudes (29%; *k* = 48), individual differences (20%; *k* = 33), motivation (16%; *k* = 27), and non-work attitudes/activities (11%; *k* = 19).

Analysis Strategies

With regard to Research Question 8a, the majority of samples did not report procedures used to treat missing data (65%; $k = 108$), whereas many samples reported using listwise deletion (33%; $k = 55$), and only a few used ad hoc single imputation methods (1%; $k = 2$; i.e., imputation from spouse's responses, and earlier score carried forward), pairwise deletion (1%; $k = 2$), or deleted outliers (1%; $k = 1$).[1]

[1] In the software SAS (PROC MIXED) and in HLM, the default missing data routines involve maximum likelihood missing data estimation; as such, some of the studies with unreported missing data treatments might have nonetheless been using state-of-the-art missing data strategies (Newman, 2014).

CHAPTER 4: DISCUSSION

Experience sampling methods have been increasingly used by organizational researchers to study dynamic constructs and within-persons variation, so a comprehensive review of methodological features should be useful. In fact, extrapolating the recent trend suggests that use of ESM may continue to grow in organizational research. It seems there are very few psychological constructs in the workplace that would *not* benefit from being examined at the within-persons level of analysis. The present study systematically summarized quantitative and qualitative aspects of ESM studies conducted in workplace settings over the past 20 years. Based on our findings, we offer norms, trends, recommendations, and critical comments related to the common practices of organizational ESM researchers.

First, because ESM studies took place most frequently in the US, knowledge of cross-cultural variation in ESM parameters is likely to be limited. Nations differ in various cultural dimensions, including individualism-collectivism and power distance (Hofstede, 1980; Hofstede & Hofstede, 2010), which could potentially influence psychological states measured via ESM. In the current review, we found that over three quarters of ESM studies occurred in individualist cultures (i.e., Germany, Netherlands, US, UK, Australia, and Switzerland); thus more studies would need to be conducted in collectivist cultures (e.g., China, Greece, Mexico) in order to gain a better understanding of how ESM results generalize cross-culturally.

The majority of ESM samples was obtained from multiple jobs, multiple industries, and multiple organizations, which calls to attention a potential ambiguity in the levels of analysis of estimated parameters. For example, in multi-organization samples, individual-level (between-persons) effects are confounded with between-organizations effects, which can create an ecological fallacy (Bliese, 2000; Newman, Joseph, & Feitosa, 2015). Due to the fact that ESM

multilevel analysis oftentimes reveals meaningful and novel relationships between variables at the between-persons level of analysis (including cross-level interactions), ESM researchers should make a greater effort to sample participants within a single job, industry, or organization in order to avoid confounding person-level variance with organization-level variance.

Next, the average sample size in ESM studies was 93, whereas the average number of observations was 1,419. This is equal to approximately 15 observations per participant. These findings have important implications for statistical power, or the ability to detect an effect of a certain magnitude with a specific degree of confidence, in within-persons studies. To extend the argument, we note that beyond simply reducing the probability of false negative conclusions, statistical power can also enhance the replicability of findings (e.g., Francis, 2012; Ioannidis, 2005; Simmons, Nelson, & Simonsohn, 2011). In order to increase power to detect within-persons effects, the number of observations per participant should be increased. This suggests that a small *N* (i.e., < 93) may suffice as long as the number of total observations is large (i.e., > 1,419). On the other hand, for assessing between-persons effects, *N* = 93 may yield inadequate statistical power (Cohen, 1992). Finally, cross-level effects include direct effects and interaction effects, and general rules of thumb argue for a sample size of at least 30 in both the lower and upper levels (i.e., 30-30 rule; e.g., Kreft & de Leeuw, 1998). In contrast, recent research demonstrates that the power to detect cross-level interaction effects is influenced by: (a) the magnitude of direct cross-level effects, (b) the standard deviation of Level 1 slope coefficients, and (c) both the average lower and upper level sample sizes (Mathieu, Aguinis, Culpepper, & Chen, 2012). In particular, Matheiu et al. (2012) advocate for a focus on larger lower level samples rather than a greater number of upper level units (i.e., 3:2 premium of Level 1 sample size versus Level 2 sample size; Mathieu et al., 2012, p. 959). In contrast, for ESM studies, the

ratio of the average within-persons sample size to the average number of persons is 15:93 ≈ 1:6, which is much smaller than Mathieu et al.'s recommended ratio of 3:2 for detecting cross-level interactions. In summary, ESM studies seem well-designed for detecting within-persons effects, but are not optimally designed for detecting cross-level interactions, in terms of statistical power.

For response rates obtained in ESM studies, we observed an average of a 63% *between-persons response rate*. This is strikingly similar to the average response rates found by Anseel et al. (2010, 52%) and Roth and BeVier (1998, 57%), both of whom summarized between-persons response rates primarily from cross-sectional studies in the organizational sciences. In other words, between-persons response rates for ESM studies are no lower than between-persons response rates for non-ESM studies.

The current study further found a 79% *between-persons analysis rate*, meaning that only 79% of the respondents had their data included in the analyses; probably due to listwise deletion of partial respondents. The practice of listwise deletion—i.e., analyzing less than 100% of the respondents—is increasingly recognized as a suboptimal missing data strategy, which produces unnecessary missing data bias (Enders, 2010; Newman, 2014). We recommend that future ESM researchers should employ the default missing data routines in SAS (PROC MIXED) and HLM (i.e., maximum likelihood missing data routines), rather than simply deleting the data that are provided by the partial respondents.

Finally, the current study found an 80% *pooled within-persons response rate*, on average. We are encouraged to learn that the within-persons response rate is as high as 80% in organizational ESM studies, because a high within-persons response rate suggests less missing data bias in the estimation of within-persons parameters.

Interestingly, the mean levels of response rates (both between-persons and within-persons) did not differ based on incentives, reminders, signaling strategy, reminder technology, or survey medium. Similarly, response rates were not correlated with monetary incentive amount, study duration, signal frequency, number of items on daily surveys, or year of publication. It is unclear why these response rates are not influenced or related to features of ESM design as one might expect. However, we note that these results should be taken with caution because not all ESM studies contained information about the three types of response rates, so the power to statistically detect differences was limited. We urge future ESM researchers to clearly report response rates, as this information can meaningfully change results. Further, the lack of an association between incentives and response rates is consistent with past findings (e.g., Anseel et al., 2010; Roth & BeVier, 1998). We caution the reader that lack of correlation might *not* imply lack of causation in the current case (i.e., we did not observe a correlation between incentives and response rates). For instance, it is possible that researchers would offer more incentives in situations where they believed *a priori* that responses would be difficult to get. If this were true, then any would-be positive correlation between incentives and response rates would be suppressed, due to a selection effect (i.e., if scenarios with low expected response rates were selected to receive incentives). In such cases, it is possible that incentives might actually enhance response rates, but no correlation would be observed, because the low-response groups were selectively chosen to get incentives.

The average study duration with an ESM design was 10 days, with about 3 signals per day. However, we note that these frequencies were not normally distributed (i.e., are positively skewed; see Figure 13 and Figure 14), such that 75% of studies lasted 10 days or fewer, with 60% of studies administering 2 signals or fewer per day. Interestingly, signal frequency was not

related to the presence of incentives or monetary incentive amount, but the average study duration was longer for samples that provided an incentive. In other words, researchers were less likely to pay participants for short duration studies.

The most common survey medium in ESM studies was paper-and-pencil, but there are serious concerns about its suitability for capturing momentary states. The major disadvantage of paper-and-pencil surveys is the lack of time record or time stamp to ensure participants' compliance with timing rules. For example, if surveys are supposed to be completed at 10:00 A.M., it is impossible to determine whether participants actually recorded responses at that time. If participants fail to comply with the prescribed timeline in paper-and-pencil surveys, it negates one of the main goals of ESM studies, because results will be susceptible to recall biases that the ESM method aims to overcome. In contrast, online surveys tend to be accessed more quickly because employees are usually close to a computer or mobile device. Due to this, it is questionable whether paper-and-pencil surveys should be used to record momentary states.

Missing data tends to be a problem in ESM studies, but the vast majority of researchers do not report the way in which missing data were treated. This is alarming because there are a number of strategies to deal with missing data and each has different influences on data analysis and final results. Of the few ESM studies that did report missing data treatment, most used listwise deletion with various exclusion rules such as excluding participants with fewer than one third, one half, etc., of all possible responses (e.g., Biron & van Veldhoven, 2012; King, Mohr, Peddie, Jones, & Kendra, 2014). Alternatively, some ESM studies excluded participants who did not comply with time restrictions and completed surveys at inappropriate times (e.g., Edmonson, Shaffer, Chaplin, Burg, Stone, & Schwartz, 2013; Harris & Daniels, 2005). Other ESM studies reported excluding data due to a combination of reasons including participants' poor compliance,

exhibiting no variance (e.g., Binnewies & Wornlein, 2011), failing to respond to certain items (e.g., Fay & Sonnentag, 2012), stopping their diary entries (Hulsheger, Alberts, Feinholdt, & Lang, 2013), and more. A consistent missing data treatment in ESM studies is necessary in order to compare findings across studies. Listwise deletion seems to be the most popular technique, but in reality only creates more missing data by eliminating the responses that were provided by partial respondents. Instead, multiple imputation or maximum likelihood estimation would be better alternatives because they are less biased and more powerful (Enders, 2010; Newman, 2014). We urge researchers using ESM to clearly report missing data techniques.

Although ESM is a valuable tool in organizational research, there are some samples that may be difficult to study via ESM due to the inherent characteristics of certain jobs. ESM requires participants to briefly stop engaging in work tasks in order to complete momentary surveys at multiple time points. Whereas most office workers are able to do this, employees such as bus drivers, construction workers, or medical doctors cannot immediately stop working in the middle of a task to respond to a signal. This limits the samples to which extant ESM findings can be generalized, which researchers should keep in mind. Perhaps there are adjustments to the ESM design that can be made to accommodate unique professions. For example, flight attendants who participate in ESM studies complete surveys after each flight rather than at regular intervals (e.g., Xanthopoulou, Baker, Heuven, Demerouti, & Schaufeli, 2008).

Another issue with ESM studies is that most measurement instruments were developed and validated in samples at the between-persons level, but are being adopted in momentary surveys to support within-persons inferences. Some researchers have modified items or response options to be more suitable for momentary responses. For instance, items measuring organizational citizenship behavior are usually based on a 1-5 agreement or frequency scale, but

response options have been changed to dichotomous "yes/no" with the logic that it is rare for employees to report engaging in these behaviors multiple times per day (e.g., Dalal et al., 2009; Glomb et al., 2011). Although this is a logical adjustment, these modified scales must be validated, ideally at the within-persons level, in order to be used properly. When a between-persons scale is used at the within-persons level, researchers assume constructs are analogous or isomorphic across the two levels (Kozlowski & Klein, 2000), but this is problematic in the absence of evidence of isomorphism (see Bliese, Chan, & Ployhart, 2007; Tay, Woo, & Vermunt, 2014). Because existing between-persons scales are likely to continue to be used to obtain momentary responses, researchers should examine the extent to which isomorphism exists, and perhaps develop new scales or scale formats for use at the within-persons level (cf. Dyer, Hanges, & Hall, 2005).

Lastly, *item sampling* is a strategy to address reactivity and habituation in ESM studies (Miner et al., 2005), but is rarely used. The main strength of item sampling is to keep participants engaged in filling out multiple surveys and to avoid retest/priming effects by presenting some new items at each time point. In this way, participants will not be able to expect the exact items that will appear on surveys, which in theory will help with the accuracy of their responses. However, item sampling also has several weaknesses, such as difficulty in establishing comparability of different items used at different time points to measure the same construct. Perhaps this is why item sampling is used so rarely in ESM studies.

In sum, it is our hope that the present paper elucidates the common practices used by ESM researchers in organizations, as well as the time trends in the usage of various design features. Understanding these norms should make it easier for researchers interested in within-persons phenomena to design new studies that are methodologically stronger than past ESM

designs. We also hope this knowledge of methodological norms and trends in organizational ESM research will encourage scholars to consider optimal tradeoffs in future ESM designs.

TABLES

Table 1. Interrater Reliability (Cohen's Kappa) for Categorical Variables.

Variable	Kappa	Corrected
All Variables	**.75**	**.84**
All Variables (except constructs)	**.86**	**.86**
Missing Data Treatment	1.00	--
Recruit	.68	--
Incentives	.94	--
Time	.60	--
Signaling Strategy	.82	--
Reminder Technology	.85	--
Survey Medium	.95	--
Country	.96	--
Sample	.75	--
Job	.96	--
Industry	.97	--
Number of Orgs	.78	--
Published	.91	--
Timing of Baseline survey	.76	--
Item Sampling	.91	--
Lagged Analyses	.91	--
Constructs - Overall	**.75**	**.83**
positive affect/emotion/mood	.71	.73
negative affect/emotion/mood	.86	.86
general affect/emotion/mood	.76	.73
emotional labor/regulation	.86	.86
leadership	1.00*	1.00*
job characteristics	.55	.62
work events	.76	.83
workload	.87	.87
sleep	.66*	.66*
burnout	.90	.90
general health	.82	.82
beliefs about self	.70	.77
personality	.65	.61
job satisfaction	.84	.87
organizational justice	.85	.85
perceived support	.57	.72

Table 1 (cont.)

work engagement	.54	.75
flow	.83	.91
mindfulness	1.00*	1.00*
learning	.80*	.80*
recovery	.85	.91
approach/avoidance	.50*	.50*
general motivation	.48	.78
family	.44	.91
life satisfaction	.66	1.00*
activities outside work	.94	.94
conflict/constraint	.58	.80
demands	.94	.97
stress/strain	.75	.88
work family conflict	.69	.89
time pressure	.72	.77
coping	.91	.91
performance	.81	.88
CWB	.94	.94
OCB	.83	.83
creativity	.87	.94
general work behavior	.42	.80

Note. The second column displays interrater reliability calculated between two raters. The third column displays post-hoc corrections of interrater reliability, corrected through discussion (see Appendix C). * denotes low base rate (i.e., both raters had fewer than 5 counts of the construct).

Table 2. Interrater Reliability (Intraclass Correlation Coefficient) for Quantitative Variables.

Variable	ICC(1)
All Variables	.97
Sample Size (*N*)	.94
Number of Observations	1.00
Between-persons response rate	1.00
Between-persons analysis rate	.99
Pooled within-persons response rate	.99
Number of days	1.00
Number of surveys per day	.97
Organizational tenure	1.00
Job tenure	1.00
Percent male	.89
Percent Caucasian	1.00
Year of study	.98
Number of items in baseline survey	.98
Number of items in daily survey	.79

Note. Interrater reliability was calculated between two raters. ICC(1) = intraclass correlation coefficient, which in this case is equal to Pearson's r between two raters.

Table 3. Descriptive Statistics of Quantitative Variables.

Variable	Min	Lower	Median	Upper	Max	M	SD
Sample Size (*N*)	7	47	69	106	1,020	93.34	98.65
Number of observations	12	360	650	1,419	11,471	1,419.13	1,951.10
Between-persons response rate	15	45	68	79	100	63.30	22.97
Between-person analysis rate	5	70	86	92	100	79.02	19.05
Pooled within-persons response rate	52	70	82	87	100	79.87	11.49
Number of days	1	5	8	10	95	10.13	9.88
Number of surveys per day	1	1	2	4	30	2.90	3.13
Organizational tenure	3	5	8	10	27	8.53	5.20
Job tenure	1	3	5	8	18	6.48	4.72
Percent male	0	23	40	57	100	42.04	24.46
Percent Caucasian	0	59	80	91	100	72.90	24.58
Number of items in baseline survey	1	10	20	35	253	32.54	40.09
Number of items in daily survey	2	10	18	25	88	20.71	14.59

Note. Min = minimum value; Lower = lower quartile; Max = maximum value; Upper = upper quartile; M = mean; SD = standard deviation.

Table 4. Frequency of Jobs Sampled in ESM Study Samples.

Frequency	Job
93	multiple jobs
6	nurses
5	call center workers, managers
4	teachers
3	physicians, police officers, software programmers
2	accountants, customer service workers, engineers, flood control workers, HR staff
1	academic staff members, cheerleading camp instructors, correctional officers, doctoral and post-graduate students, emergency workers, entrepreneurs, flight attendants, hospital staff, interior designers, IT consultants, medical coders, navy patrol vessel crew members, police service call center workers, professors, railway controllers, sales representatives, self-employed workers, servers

Note. The first column displays the number of samples drawn from a particular job. The second column displays the jobs sampled once, twice, etc.

Table 5. Frequency of Industries Sampled in ESM Study Samples.

Frequency	Industry
72	multiple industries
22	educational services
15	health care and social assistance
12	service-providing
8	information
6	professional and business services
4	accommodation and food services, manufacturing
3	social assistance
2	motor vehicle and parts dealer
1	administrative and support services, arts/entertainment/recreation, financial services, food services and drinking places, performing arts/spectator sports, retail trade, telecommunications, transportation and warehousing

Note. The first column displays the number of samples drawn from a particular industry. The second column displays the industries sampled once, twice, etc. Industry names were obtained from the US Bureau of Labor Statistics.

Table 6. Frequency of ESM Study Samples that Measured Each Construct.

Broad Construct Category	Frequency (# of samples)	% of all Samples
affect/emotion/mood	100	.60
Stress	75	.45
work behaviors	61	.37
Health	58	.35
situational factors	53	.32
job attitudes	48	.29
individual differences	33	.20
motivation	27	.16
non-work attitudes, perceptions, and activities	19	.11

Note. Total # samples = 167. The first column lists broad categories into which constructs from ESM study samples are organized. Constructs are measured at the within-person level (i.e., daily surveys). The second column lists the number of samples that measured the construct. The third column lists the percentage of samples that measured a given construct (i.e., the second column divided by 167).

Table 7. Construct Category Frequencies and Percentages.

Broad Construct Category	Construct Category	Frequency	Frequency of Broad Category	% of Broad Category
affect/emotion/mood	general affect/emotion/mood	51	100	.51
	negative affect/emotion/mood	36	100	.36
	positive affect/emotion/mood	30	100	.30
	emotional labor/regulation	22	100	.22
situational factors	work events	31	53	.58
	job characteristics	23	53	.43
	workload	8	53	.15
	leadership	2	53	.04
health	burnout	38	58	.66
	general health	16	58	.28
	sleep	5	58	.09
individual differences	beliefs about self	17	33	.52
	personality	8	33	.24
job attitudes	job satisfaction	21	48	.44
	work engagement	16	48	.33
	perceived support	8	48	.17
	organizational justice	6	48	.13
motivation	general motivation	23	27	.85
	recovery	11	27	.41
	flow	6	27	.22
	learning	3	27	.11
	approach/avoidance	3	27	.11
	mindfulness	2	27	.07
non-work attitudes,	activities outside work	9	19	.47
perceptions, and activities	family	6	19	.32
	life satisfaction	5	19	.26
stress	stress/strain	27	75	.36
	conflict/constraint	17	75	.23
	demands	17	75	.23
	work family conflict	10	75	.13
	time pressure	8	75	.11
	coping	5	75	.07
work behaviors	general work behavior	22	61	.36
	performance	21	61	.34

Table 7 (cont.)

counterproductive work behavior	18	61	.30
organizational citizenship behavior	13	61	.21
creativity	8	61	.13

Note. The first column lists broad categories into which constructs from ESM studies are organized. Constructs are measured at the within-person level (i.e., daily surveys). The second column lists 38 specific construct categories that make up the 9 broad categories. The third column lists the number of times construct categories were measured (i.e. how many samples measured the category). The fourth column lists the number of times broad construct categories were measured (i.e. how many samples measured the broad category). The fifth column indicates the percentage of samples that measured a construct category relative to the number of the samples that measured its broad category (i.e., the third column divided by the fourth column).

Table 8. Response Rate Correlations.

	Between Person Response Rate	Between Person Analysis Rate	Pooled Within-Persons Response Rate
Monetary Incentive Amount	-.27 / 25	.27 / 26	.26 / 19
Study Duration (number of days)	.13 / 83	.10 / 81	-.09 / 51
Signal Frequency (# surveys/day)	.01 / 78	.12 / 76	.12 / 49
Number of Daily Survey Items	.06 / 82	.11 / 79	.00 / 50
Year	-.03 / 85	-.07 / 83	-.05 / 51

Note: Each correlation is followed by its sample size (*k*). Pairwise deletion was used. None of the correlations above are significant.

FIGURES

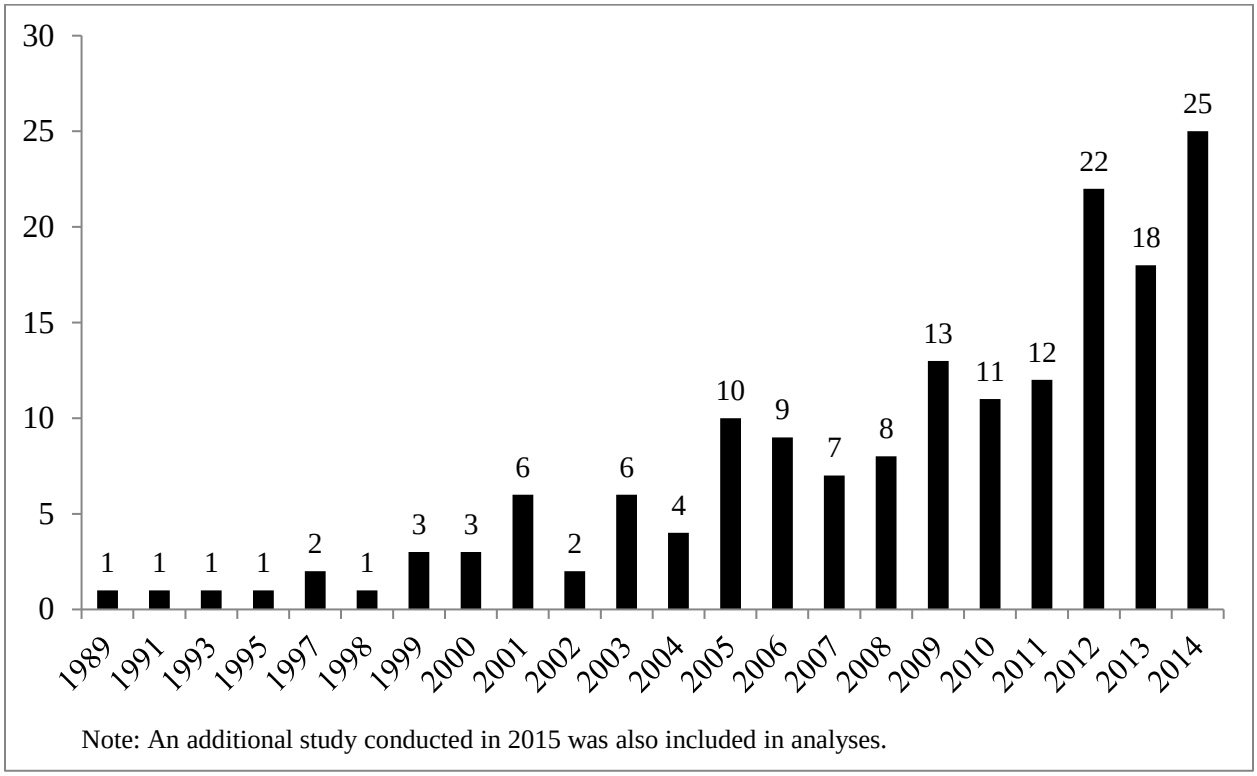

Figure 1. Number of ESM studies conducted from 1989-2014.

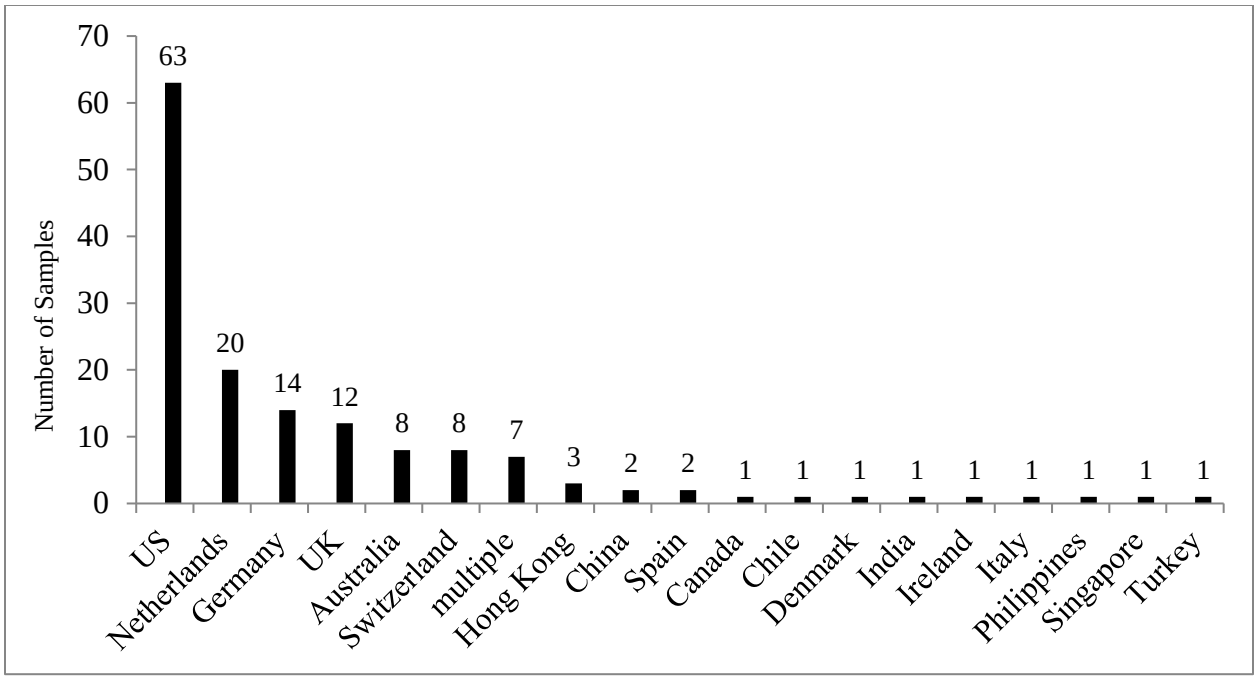

Figure 2. Countries from which samples are drawn.

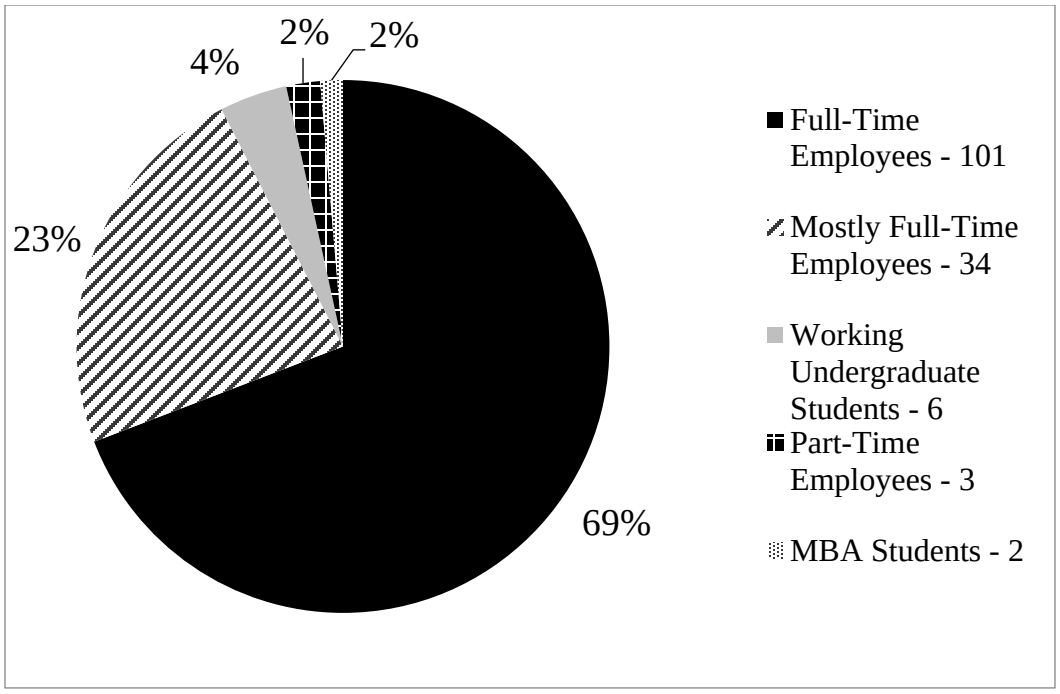

Figure 3. Sample types: Frequencies and percentages.

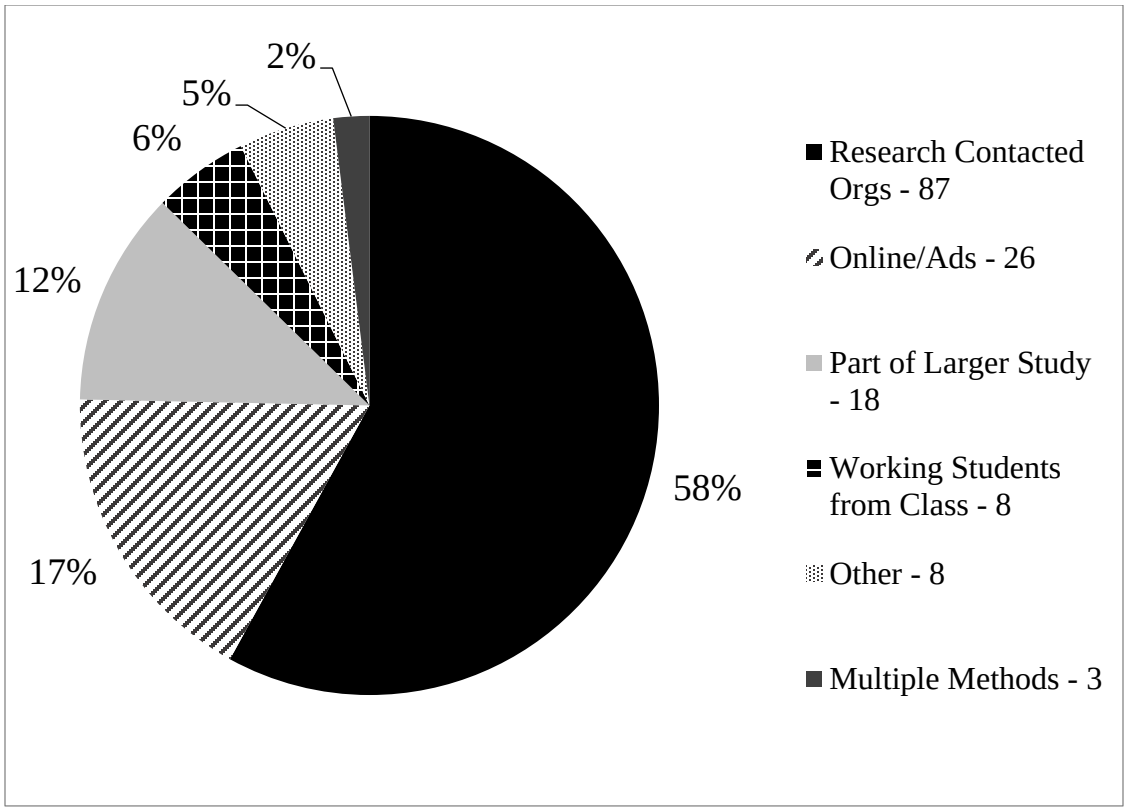

Figure 4. Recruitment strategies: Frequencies and percentages. (k= 150 studies reported this information).

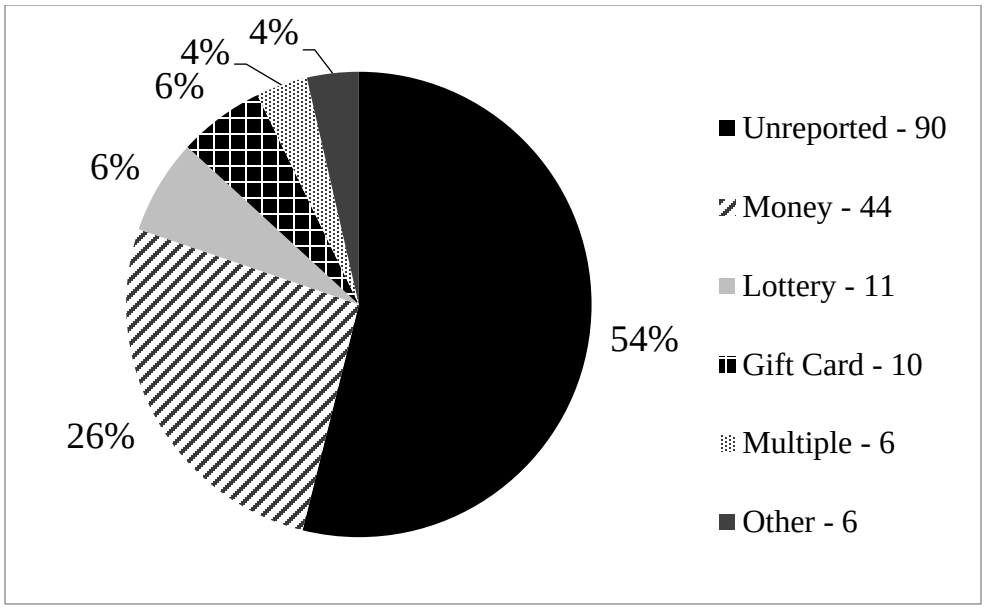

Figure 5. Incentive types: Frequencies and percentages.

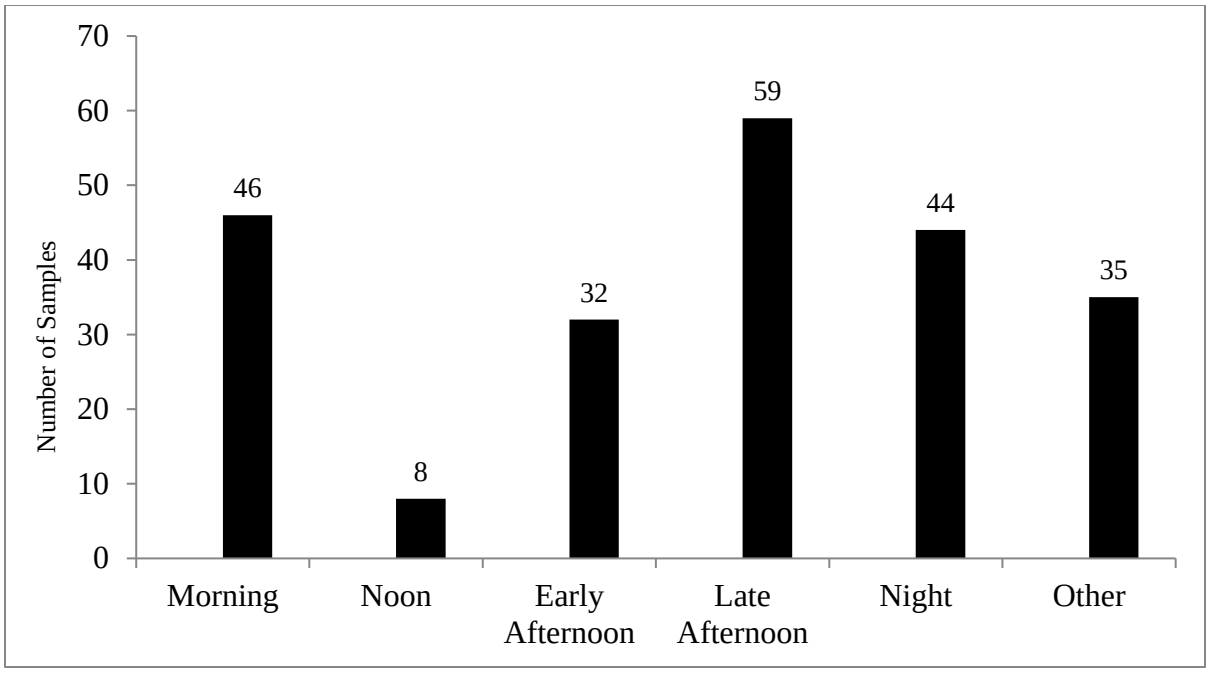

Figure 6. Times of day when participants were surveyed. (k= 136 studies reported this information).

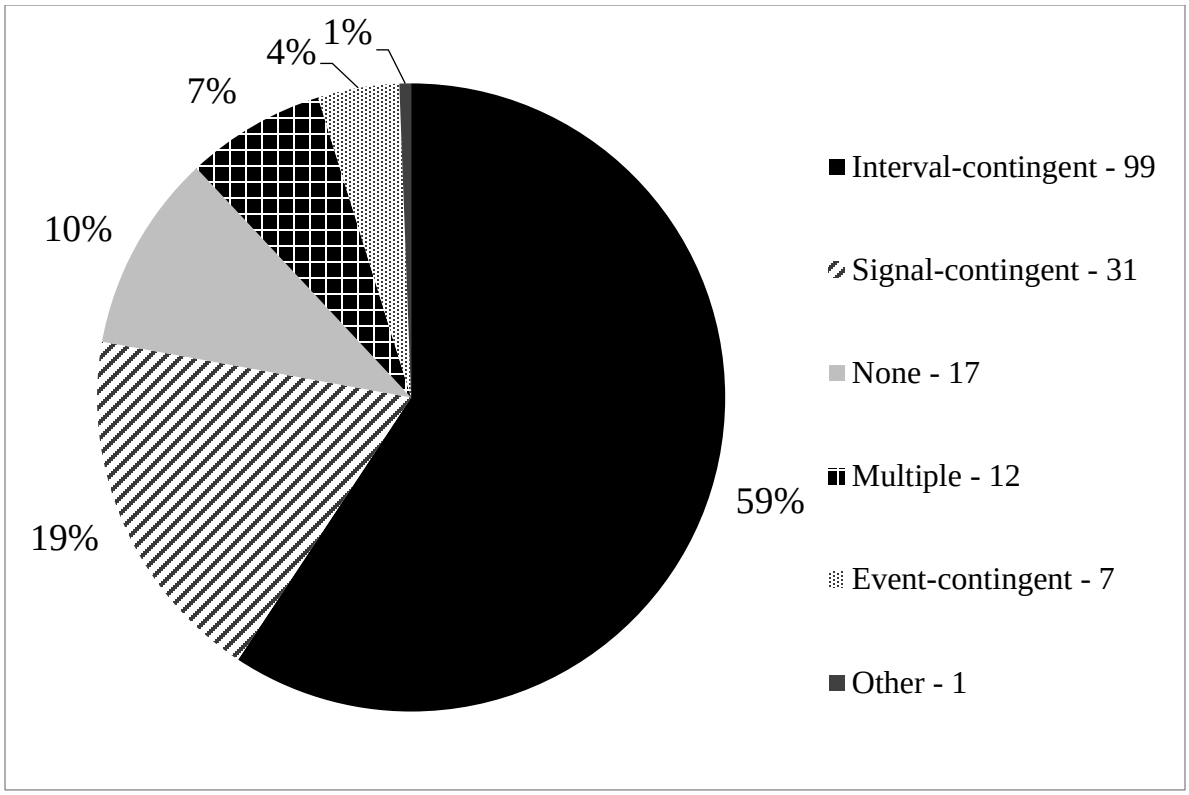

Figure 7. Signaling strategies: Frequencies and percentages.

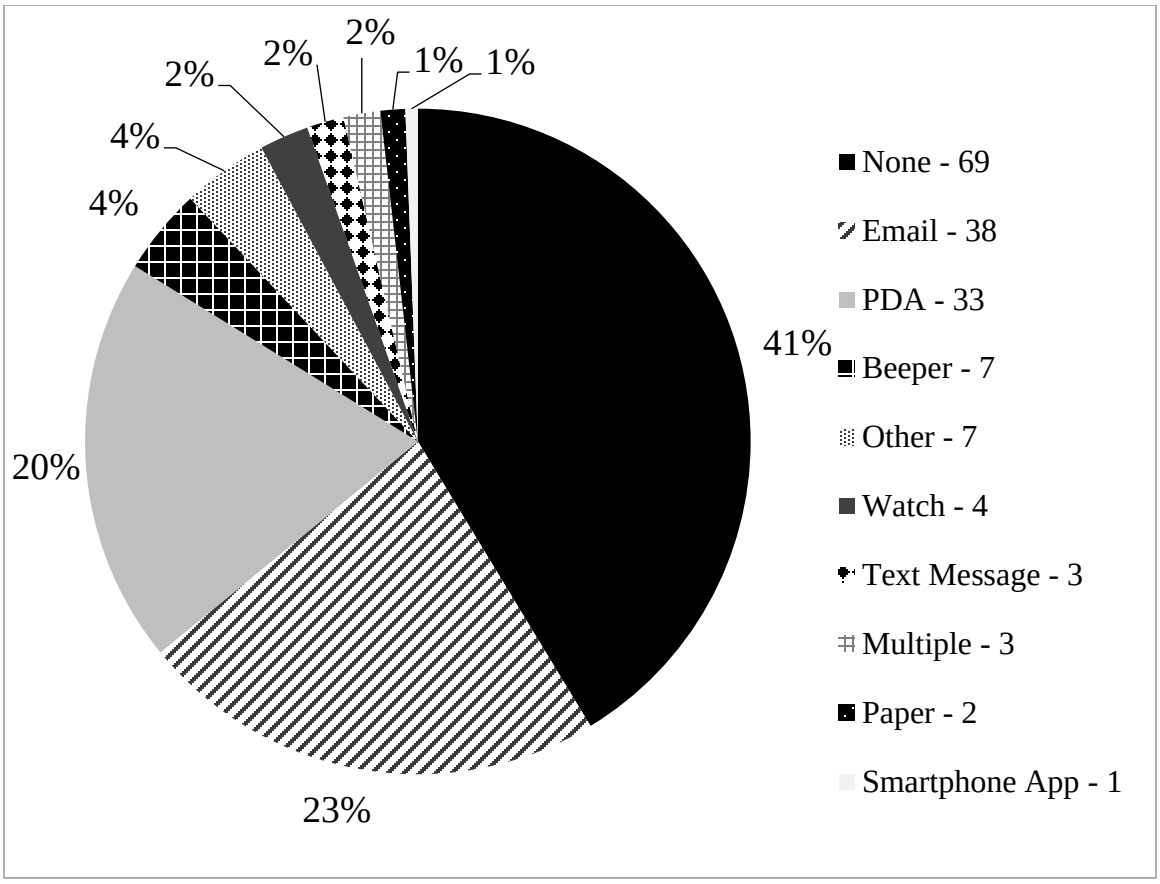

Figure 8. Reminder technologies: Frequencies and percentages.

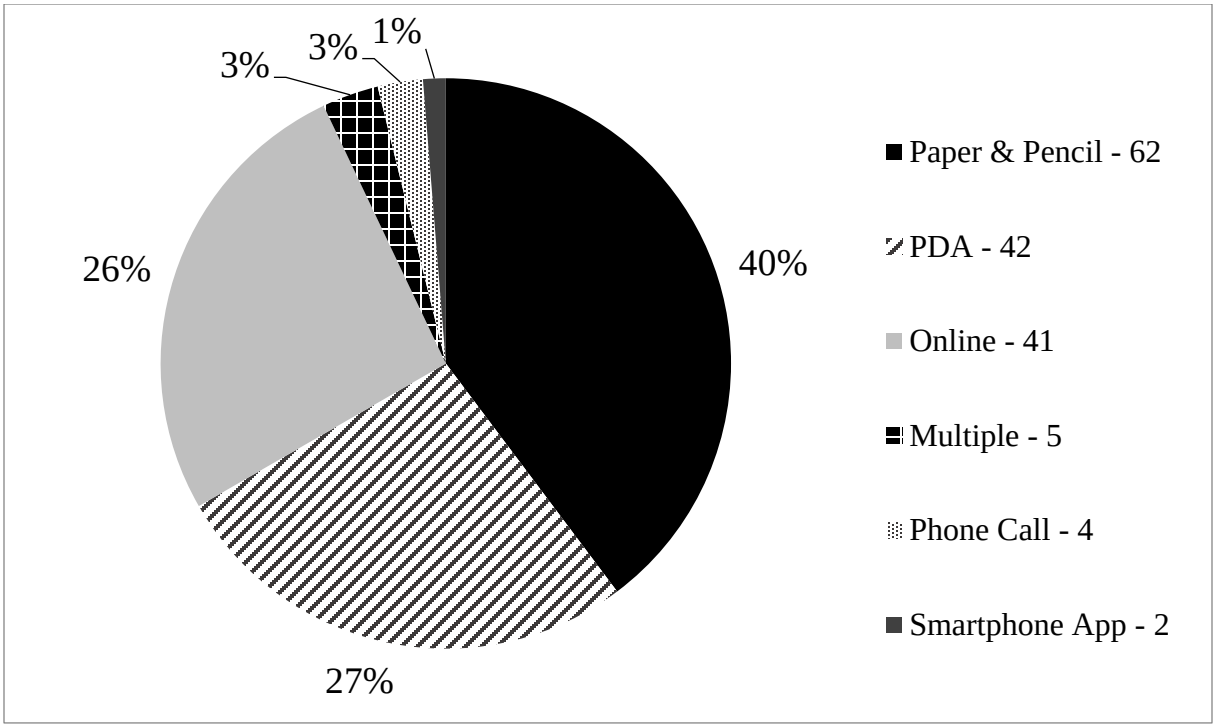

Figure 9. Survey media: Frequencies and percentages. (k = 156 studies reported this information).

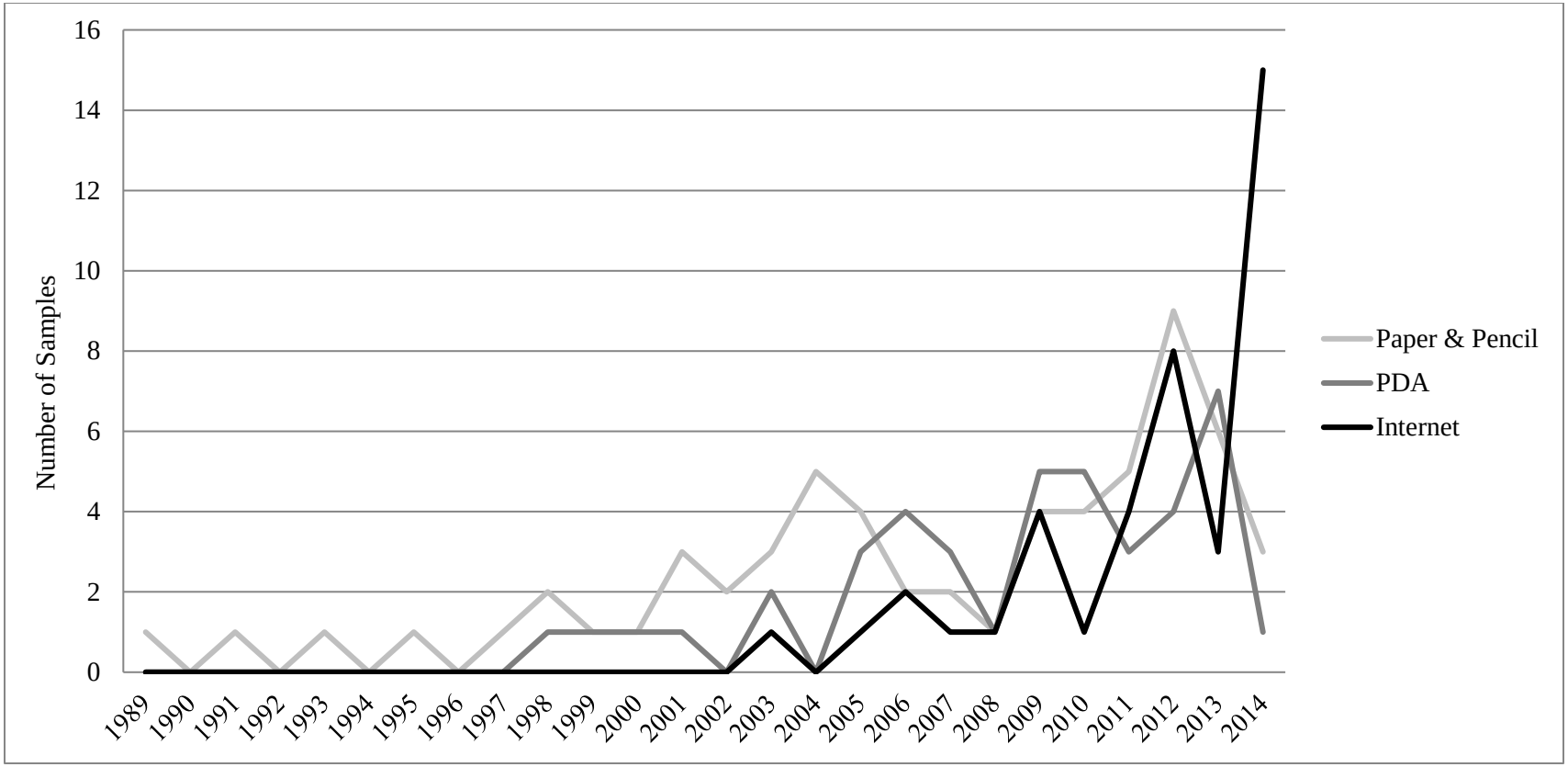

Figure 10. Frequencies of three major types of survey media in ESM studies from 1989-2014.

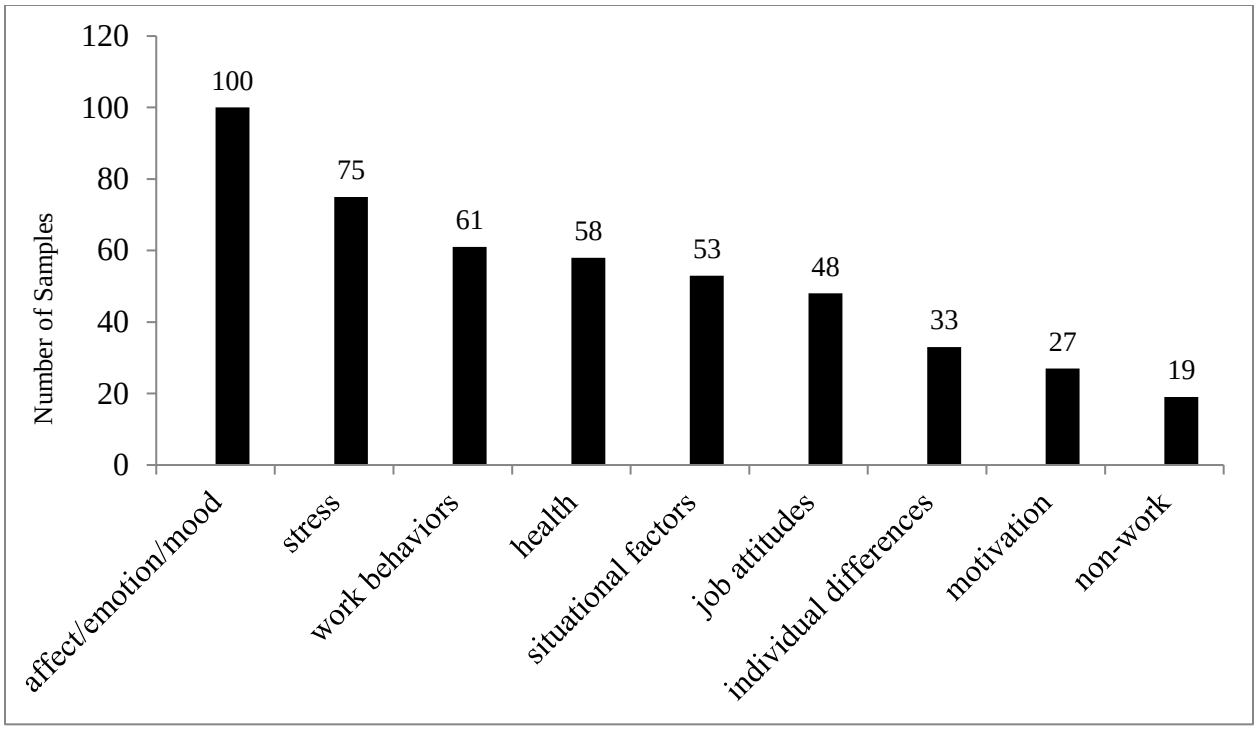

Figure 11. Broad construct category measured at the within-persons level.

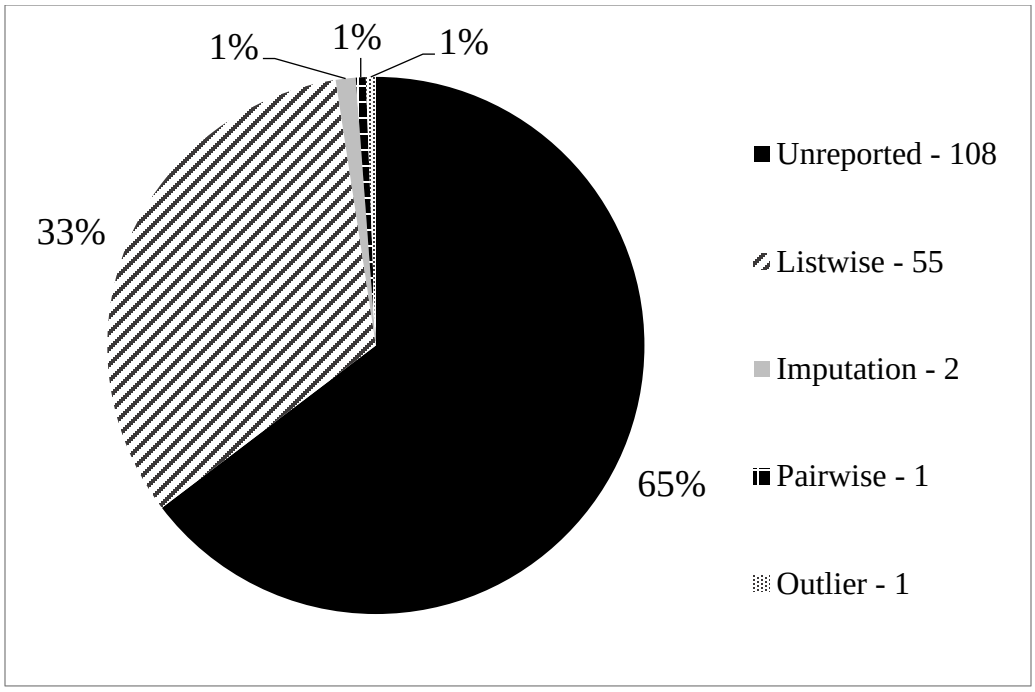

Figure 12. Missing data treatments: Frequencies and percentages.

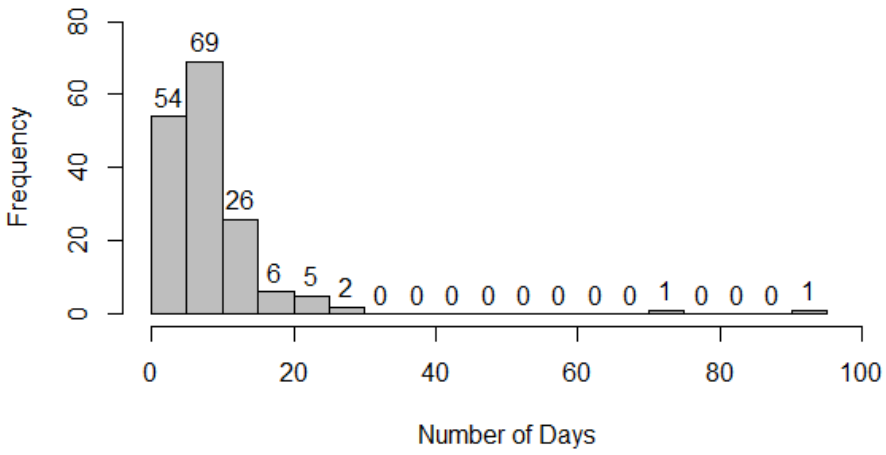

Figure 13. Distribution of studies by study duration (i.e., number of days).

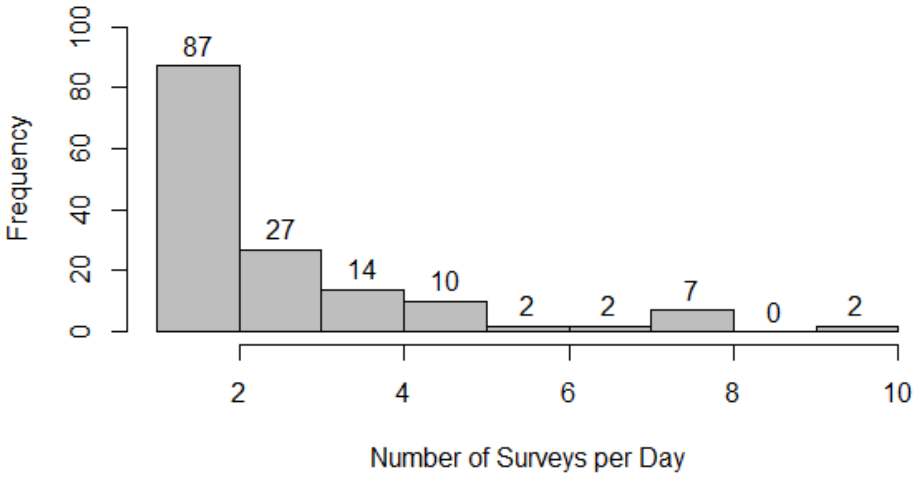

Figure 14. Distribution of studies by signal frequency (i.e., number of surveys per day).

REFERENCES

*Alliger, G. M., & Williams, K. J. (1993). Using signal-contingent experience sampling methodology to study work in the field: A discussion and illustration examining task perceptions and mood. *Personnel Psychology, 46*, 525-549. doi:10.1111/j.1744-6570.1993.tb00883.x

*Amabile, T. M., Barsade, S. G., Mueller, J. S., & Staw, B. M. (2005). Affect and creativity at work. *Administrative science quarterly, 50*, 367-403. doi:10.2189/asqu.2005.50.3.367

Anseel, F., Lievens, F., Schollaert, E., & Choragwicka, B. (2010). Response rates in organizational science, 1995–2008: A meta-analytic review and guidelines for survey researchers. *Journal of Business and Psychology, 25*, 335-349. http://dx.doi.org/10.1007/s10869-010-9157-6

Austin, J.T., Humphreys, L.G., & Hulin, C.L. (1989). Another view of dynamic criteria: A critical reanalysis of Barrett, Caldwell, and Alexander. *Personnel Psychology, 42*, 583-596. http://dx.doi.org/10.1111/j.1744-6570.1989.tb00670.x

*Baethge, A., & Rigotti, T. (2013). Interruptions to workflow: Their relationship with irritation and satisfaction with performance, and the mediating roles of time pressure and mental demands. *Work & Stress, 27*, 43-63. http://dx.doi.org/10.1080/02678373.2013.761783

*Bakker, A. B., & Xanthopoulou, D. (2009). The crossover of daily work engagement: Test of an actor–partner interdependence model. *Journal of Applied Psychology, 94*, 1562. http://dx.doi.org/10.1037/a0017525

*Beal, D. J., Trougakos, J. P., Weiss, H. M., & Dalal, R. S. (2013). Affect spin and the emotion regulation process at work. *Journal of Applied Psychology, 98*, 593. http://dx.doi.org/10.1037/a0032559

*Beal, D. J., Trougakos, J. P., Weiss, H. M., & Green, S. G. (2006). Episodic processes in emotional labor: Perceptions of affective delivery and regulation strategies. *Journal of Applied Psychology, 91*, 1053. http://dx.doi.org/10.1037/0021-9010.91.5.1053

Beal, D. J., & Weiss, H. M. (2003). Methods of ecological momentary assessment in organizational research. *Organizational Research Methods, 6*, 440-464. http://dx.doi.org/10.1177/1094428103257361

*Beattie, L., & Griffin, B. (2014). Accounting for within-person differences in how people respond to daily incivility at work. *Journal of Occupational and Organizational Psychology, 87*, 625-644. http://dx.doi.org/10.1111/joop.12067

*Binnewies, C., & Wörnlein, S. C. (2011). What makes a creative day? A diary study on the interplay between affect, job stressors, and job control. *Journal of Organizational Behavior, 32*, 589-607. http://dx.doi.org/10.1002/job.731

*Biron, M., & van Veldhoven, M. (2012). Emotional labour in service work: Psychological flexibility and emotion regulation. *Human Relations, 65*, 1259-1282. http://dx.doi.org/10.1177/0018726712447832

*Bledow, R., Schmitt, A., Frese, M., & Kühnel, J. (2011). The affective shift model of work engagement. *Journal of Applied Psychology, 96*, 1246-1257. http://dx.doi.org/10.1037/a0024532

Bliese, P. D., Chan, D., & Ployhart, R. E. (2007). Multilevel methods: Future directions in measurement, longitudinal analyses, and nonnormal outcomes. *Organizational Research Methods, 10*, 551-563. http://dx.doi.org/10.1177/1094428107301102

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual review of psychology, 54*, 579-616. http://dx.doi.org/10.1146/annurev.psych.54.101601.145030

*Bono, J. E., Foldes, H. J., Vinson, G., & Muros, J. P. (2007). Workplace emotions: The role of supervision and leadership. *Journal of Applied Psychology, 92*, 1357. doi:10.1037/0021-9010.92.5.1357

*Bono, J., Glomb, T., Shen, W., Kim, E., & Koch, A. (2012). Building positive resources: Effects of positive events and positive reflection on work-stress and health. *Academy of Management Journal, 56*, 1601-1627. http://dx.doi.org/10.5465/amj.2011.0272

*Brans, K., Koval, P., Verduyn, P., Lim, Y. L., & Kuppens, P. (2013). The regulation of negative and positive affect in daily life. *Emotion, 13*, 926. http://dx.doi.org/10.1037/a0032400

*Breevaart, K., Bakker, A. B., Demerouti, E., & Hetland, J. (2012). The measurement of state work engagement: A multilevel factor analytic study. *European Journal of Psychological Assessment, 28*, 305. http://dx.doi.org/10.1027/1015-5759/a000111

*Butler, A., Grzywacz, J., Bass, B., & Linney, K. (2005). Extending the demands-control model: A daily diary study of job characteristics, work-family conflict and work-family facilitation. *Journal of Occupational and Organizational Psychology, 78*, 155-169. http://dx.doi.org/10.1348/096317905x40097

*Carson, R. L. (2006). *Exploring the episodic nature of teachers' emotions as it relates to teacher burnout*. (Doctoral dissertation). Available from PsycINFO. (622007906; 2007-99005-023)

*Chandler, M. M. (2012). *The antecedents and consequences of core affect variability at work.* (Doctoral dissertation). Available from PsycINFO. (1437971806; 2013-99161-127)

*Christensen, A. L. (2014). *Feedback, affect, and creative behavior a multi-level model linking feedback to performance.* (Doctoral dissertation). Available from PsycINFO. (1653143757; 2015-99030-309)

*Claessens, B. J., Van Eerde, W., Rutte, C. G., & Roe, R. A. (2010). Things to do today...: A daily diary study on task completion at work. *Applied Psychology, 59*, 273-295. http://dx.doi.org/10.1111/j.1464-0597.2009.00390.x

*Clark, M. A. (2010). *Why do employees behave badly? An examination of the effects of mood, personality, and job demands on counterproductive work behavior*. (Doctoral dissertation). Available from PsycINFO. (750856893)

Cohen, J. (1992). A power primer. *Psychological bulletin*, *112*, 155.

*Cohen, T. R., Panter, A. T., Turan, N., Morse, L., & Kim, Y. (2014). Moral character in the workplace. *Journal of Personality and Social Psychology*. *107*, 943. doi:10.1037/a0037245

*Conway, N., & Briner, R. B. (2002). A daily diary study of affective responses to psychological contract breach and exceeded promises. *Journal of Organizational Behavior, 23*, 287-302. http://dx.doi.org/10.1002/job.139

Csikszentmihalyi, M. & Larson, R. (1987). Validity and reliability of the Experience-Sampling Method. *The Journal of Nervous and Mental Disease, 175*, 526-536. http://dx.doi.org/10.1007/978-94-017-9088-8_3

*Csikszentmihalyi, M., & LeFevre, J. (1989). Optimal experience in work and leisure. *Journal of personality and social psychology, 56*, 815. http://dx.doi.org/10.1037/0022-3514.56.5.815

Dalal, R. S., Bhave, D. P., & Fiset, J. (2014). Within-person variability in job performance A theoretical review and research agenda. *Journal of Management, 40*, 1396-1436. http://dx.doi.org/10.1177/0149206314532691

*Dalal, R.S., Lam, H., Weiss, H. M., Welch, E. R., & Hulin, C. L. (2009). A within-person approach to work behavior and performance: Concurrent and lagged citizenship-counterproductivity associations, and dynamic relationships with affect and overall job performance. *Academy of Management Journal, 52*, 1051-1066. http://dx.doi.org/10.5465/amj.2009.44636148

*Daniels, K., Boocock, G., Glover, J., Hartley, R., & Holland, J. (2009). An experience sampling study of learning, affect, and the demands control support model. *Journal of Applied Psychology, 94*, 1003. http://dx.doi.org/10.1037/a0015517

*Daniels, K., & Harris, C. (2005). A daily diary study of coping in the context of the job demands–control–support model. *Journal of Vocational Behavior, 66*, 219-237. http://dx.doi.org/10.1016/j.jvb.2004.10.004

*Daniels, K., Hartley, R., & Travers, C. J. (2006). Beliefs about stressors alter stressors' impact: Evidence from two experience-sampling studies. *Human Relations, 59*, 1261-1285. http://dx.doi.org/10.1177/0018726706069768

*Davis, K. D. (2008). *Daily positive and negative work-family spillover and crossover between mothers and children*. (Doctoral dissertation). Available from PsycINFO. (849723506)

*Debus, M. E., Sonnentag, S., Deutsch, W., & Nussbeck, F. W. (2014). Making flow happen: The effects of being recovered on work-related flow between and within days. *Journal of Applied Psychology, 99,* 713-722. http://dx.doi.org/10.1037/a0035881

*Demerouti, E., Bakker, A. B., Sonnentag, S., & Fullagar, C. J. (2012). Work-related flow and energy at work and at home: A study on the role of daily recovery. *Journal of Organizational Behavior, 33*, 276-295. http://dx.doi.org/10.1002/job.760

*Derks, D., & Bakker, A. B. (2012). Smartphone use, work–home interference, and burnout: A diary study on the role of recovery. *Applied Psychology, 63*, 411-440. http://dx.doi.org/10.1111/j.1464-0597.2012.00530.x

*Derks, D., van Mierlo, H., & Schmitz, E. B. (2014). A diary study on work-related smartphone use, psychological detachment and exhaustion: Examining the role of the perceived segmentation norm. *Journal of occupational health psychology, 19*, 74. http://dx.doi.org/10.1037/a0035076

Diener, E., Larsen, R. J., Levine, S., & Emmons, R. A. (1985). Intensity and frequency: Dimensions underlying positive and negative affect. *Journal of Personality and Social Psychology, 48*, 1253-1265. http://dx.doi.org/10.1037/0022-3514.48.5.1253

*Dimotakis, N. (2010). *The relationship of work demands and resources to subjective well-being: The role of self-efficacy and coping*. (Doctoral dissertation). Available from PsycINFO. (881737079)

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

*Duckett, E. (1997). *Parents' work in a family context: Linkages between paid work, family work, and young adolescents' emotional well-being*. (Doctoral dissertation). Available from PsycINFO. (304356454)

Dyer, N. G., Hanges, P. J., & Hall, R. J. (2005). Applying multilevel confirmatory factor analysis techniques to the study of leadership. *The Leadership Quarterly, 16*, 149-167. http://dx.doi.org/10.1016/j.leaqua.2004.09.009

Eatough Cooley, E. (2013). Illegitimate tasks and employee well-being: A daily diary study.* (Doctoral dissertation). Available from PsycINFO. (1346690042)

*Edmondson, D., Shaffer, J. A., Chaplin, W. F., Burg, M. M., Stone, A. A., & Schwartz, J. E. (2013). Trait anxiety and trait anger measured by ecological momentary assessment and their correspondence with traditional trait questionnaires. *Journal of research in personality, 47*, 843-852. http://dx.doi.org/10.1016/j.jrp.2013.08.005

Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press.

*Fay, D., & Sonnentag, S. (2012). Within-person fluctuations of proactive behavior: How affect and experienced competence regulate work behavior. *Human Performance, 25*, 72-93. http://dx.doi.org/10.1080/08959285.2011.631647

*Feuerhahn, N., Sonnentag, S., & Woll, A. (2014). Exercise after work, psychological mediators, and affect: A day-level study. *European Journal of Work and Organizational Psychology, 23*, 62-79. http://dx.doi.org/10.1080/1359432X.2012.709965

*Fifield, J., McQuillan, J., Armeli, S., Tennen, H., Reisine, S., & Affleck, G. (2004). Chronic strain, daily work stress and pain among workers with rheumatoid arthritis: Does job stress make a bad day worse? *Work & Stress, 18*, 275-291. http://dx.doi.org/10.1080/02678370412331324996

*Fisher, C. D. (2000). Mood and emotions while working-missing pieces of job satisfaction. *School of Business Discussion Papers, 64*.

*Fisher, C. D., Minbashian, A., Beckmann, N., & Wood, R. E. (2013). Task appraisals, emotions, and performance goal orientation. *Journal of Applied Psychology, 98,* 364. http://dx.doi.org/10.1037/a0031260

*Foo, M. D., Uy, M. A., & Baron, R. A. (2009). How do feelings influence effort? An empirical study of entrepreneurs' affect and venture effort. *Journal of Applied Psychology, 94,* 1086. http://dx.doi.org/10.1037/a0015599

*Foster, J. B. (2003). *Child care disruptions and working mothers: An experience sampling method approach*. (Doctoral dissertation). Available from PsycINFO. (305309132)

*Fuller, J. A., Stanton, J. M., Fisher, G. G., Spitzmüller, C., Russell, S. S., & Smith, P. C. (2003). A lengthy look at the daily grind: Time series analysis of events, mood, stress, and satisfaction. *Journal of Applied Psychology, 88,* 1019. http://dx.doi.org/10.1037/0021-9010.88.6.1019

Francis, G. (2012). The psychology of replication and replication in psychology. *Perspectives on Psychological Science, 7,* 585-594.

*Gabriel, A. S., Diefendorff, J. M., Chandler, M. M., Moran, C. M., & Greguras, G. J. (2014). The dynamic relationships of work affect and job satisfaction with perceptions of fit. *Personnel Psychology, 67,* 389-420. http://dx.doi.org/10.1111/peps.12042

*Gabriel, A. S., Diefendorff, J. M., & Erickson, R. J. (2011). The relations of daily task accomplishment satisfaction with changes in affect: A multilevel study in nurses. *Journal of Applied Psychology, 96,* 1095. http://dx.doi.org/10.1037/a0023937

*Gallo, L. C., Bogart, L. M., Vranceanu, A. M., & Matthews, K. A. (2005). Socioeconomic status, resources, psychological experiences, and emotional responses: A test of the reserve capacity model. *Journal of Personality and Social Psychology, 88,* 386. http://dx.doi.org/10.1037/0022-3514.88.2.386

*Garrosa-Hernández, E., Carmona-Cobo, I., Ladstätter, F., Blanco, L. M., & Cooper-Thomas, H. D. (2013). The relationships between family-work interaction, job-related exhaustion, detachment, and meaning in life: A day-level study of emotional well-being. *Revista de Psicología del Trabajo y de las Organizaciones, 29*, 169-177. http://dx.doi.org/10.5093/tr2013a23

Ghiselli, E. E. (1956). Dimensional problems of criteria. *Journal of Applied Psychology, 40,* 1-4.

*Glomb, T. M., Bhave, D. P., Miner, A. G., & Wall, M. (2011). Doing good, feeling good: Examining the role of organizational citizenship behaviors in changing mood. *Personnel Psychology, 64*, 191-223. http://dx.doi.org/10.1111/j.1744-6570.2010.01206.x

*Gooty, J., Gavin, M. B., Ashkanasy, N. M., & Thomas, J. S. (2014). The wisdom of letting go and performance: The moderating role of emotional intelligence and discrete emotions. *Journal of occupational and organizational psychology, 87*, 392-413. http://dx.doi.org/10.1111/joop.12053

*Grandey, A. A., Tam, A. P., & Brauburger, A. L. (2002). Affective states and traits in the workplace: Diary and survey data from young workers. *Motivation and emotion, 26*, 31-55.

*Grebner, S., Elfering, A., Semmer, N. K., Kaiser-Probst, C., & Schlapbach, M. L. (2004). Stressful situations at work and in private life among young workers: An event sampling approach. *Social Indicators Research, 67*, 11-49. http://dx.doi.org/10.1023/b:soci.0000007333.96382.3a

*Grech, M. R., Neal, A., Yeo, G., Humphreys, M., & Smith, S. (2009). An examination of the relationship between workload and fatigue within and across consecutive days of work: Is the relationship static or dynamic? *Journal of occupational health psychology, 14*, 231. http://dx.doi.org/10.1037/a0014952

*Gross, S., Semmer, N. K., Meier, L. L., Kälin, W., Jacobshagen, N., & Tschan, F. (2011). The effect of positive events at work on after-work fatigue: They matter most in face of adversity. *Journal of Applied Psychology, 96*, 654. http://dx.doi.org/10.1037/a0022992

*Guenter, H., Schreurs, B., Van Emmerik, I. J., Gijsbers, W., & Van Iterson, A. (2013). How adaptive and maladaptive humor influence well-being at work: A diary study. *Humor, 26*, 573-594. http://dx.doi.org/10.1515/humor-2013-0032

*Hahn, S. E. (2000). The effects of locus of control on daily exposure, coping and reactivity to work interpersonal stressors: A diary study. *Personality and Individual Differences, 29*, 729-748. http://dx.doi.org/10.1016/s0191-8869(99)00228-7

*Harris, C., & Daniels, K. (2005). Daily affect and daily beliefs. *Journal of occupational health psychology, 10*, 415. http://dx.doi.org/10.1037/1076-8998.10.4.415

*Harris, C., Daniels, K., & Briner, R. B. (2003). A daily diary study of goals and affective well-being at work. *Journal of Occupational and Organizational Psychology, 76*, 401-410. http://dx.doi.org/10.1348/096317903769647256

*Harrison, M. M. (2008). *Finding meaning in the every day: An in-depth investigation of meaningful work experiences.* (Doctoral dissertation). Available from PsycINFO. (622062130; 2009-99100-061)

*Heller, D. (2003). *The dynamic process of life satisfaction: the role of job and marital satisfaction, life events, personality, and mood*. Unpublished doctoral dissertation, University of Iowa.

*Hill, C. (2004). *The phenomenon of flow among community college faculty: A study of the relationship between optimal experience and variables in the work environment.* (Doctoral dissertation). Available from PsycINFO. (305161479)

Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. Beverly Hills, CA: Sage.

Hofstede, G. & Hofstede, G. J. (2010). *Cultures and organizations: Software of the mind* (Rev. 3rd ed.). New York, NY: McGraw-Hill.

Hormuth, S. E. (1986). The sampling of experiences in situ. *Journal of Personality, 54*, 262-293. http://dx.doi.org/10.1111/j.1467-6494.1986.tb00395.x

*Huang, J. L., & Ryan, A. (2011). Beyond personality traits: A study of personality states and situational contingencies in customer service jobs. *Personnel Psychology, 64*, 451-488. http://dx.doi.org/10.1111/j.1744-6570.2011.01216.x

*Hülsheger, U. R., Alberts, H. J., Feinholdt, A., & Lang, J. W. (2013). Benefits of mindfulness at work: The role of mindfulness in emotion regulation, emotional exhaustion, and job satisfaction. *Journal of Applied Psychology, 98*, 310. http://dx.doi.org/10.1037/a0031313

*Hunton, J. E. (2005). Behavioral self-regulation of telework locations: Interrupting interruptions! *Journal of Information Systems, 19*, 111-140. http://dx.doi.org/10.2308/jis.2005.19.2.111

*Igic, I., Ryser, S., & Elfering, A. (2013). Does work stress make you shorter? An ambulatory field study of daily work stressors, job control, and spinal shrinkage. *Journal of occupational health psychology, 18*, 469-480. http://dx.doi.org/10.1037/a0034256

*Ilies, R., Schwind, K. M., Wagner, D. T., Johnson, M. D., DeRue, D. S., & Ilgen, D. R. (2007). When can employees have a family life? The effects of daily workload and affect on work-family conflict and social behaviors at home. *Journal of Applied Psychology, 92*, 1368-1379. http://dx.doi.org/10.1037/0021-9010.92.5.1368

*Ilies, R., Scott, B. A., & Judge, T. A. (2006). The interactive effects of personal traits and experienced states on intraindividual patterns of citizenship behavior. *Academy of Management Journal, 49*, 561-575. http://dx.doi.org/10.5465/AMJ.2006.21794672

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Med, 2*, e124. http://dx.doi.org/10.1371/journal.pmed.0020124

*Jones, F., O'Connor, D. B., Conner, M., McMillan, B., & Ferguson, E. (2007). Impact of daily mood, work hours, and iso-strain variables on self-reported health behaviors. *Journal of Applied Psychology, 92*, 1731-1740. http://dx.doi.org/10.1037/0021-9010.92.6.1731

*Judge, T. A., Scott, B. A., & Ilies, R. (2006). Hostility, job attitudes, and workplace deviance: Test of a multilevel model. *Journal of Applied Psychology, 91*, 126-138. http://dx.doi.org/10.1037/0021-9010.91.1.126

*Judge, T. A., Woolf, E. F., & Hurst, C. (2009). Is emotional labor more difficult for some than for others? A multilevel, experience-sampling study. *Personnel Psychology, 62*, 57-88. http://dx.doi.org/10.1111/j.1744-6570.2008.01129.x

*Karoly, P., Okun, M. A., Enders, C., & Tennen, H. (2014). Effects of pain intensity on goal schemas and goal pursuit: A daily diary study. *Health Psychology, 33*, 968-976. http://dx.doi.org/10.1037/hea0000093

*Kiffin-Petersen, S., Murphy, S. A., & Soutar, G. (2012). The problem-solving service worker: Appraisal mechanisms and positive affective experiences during customer interactions. *Human Relations, 65*, 1179-1206. http://dx.doi.org/10.1177/0018726712451762

*Kim, J.M. (2012). *A longitudinal study on the relationship between workplace harassment and cyberloafing*. (Doctoral dissertation). Available from PsycINFO. (1319305586)

*King, E. B., Mohr, J. J., Peddie, C. I., Jones, K. P., & Kendra, M. (2014). Predictors of identity management: An exploratory experience-sampling study of lesbian, gay, and bisexual workers. *Journal of Management*. http://dx.doi.org/10.1177/0149206314539350

*Korunka, C., Kubicek, B., Prem, R., & Cvitan, A. (2012). Recovery and detachment between shifts, and fatigue during a twelve-hour shift. *Work: A Journal of Prevention, Assessment and Rehabilitation, 41*, 3227-3233.

*Kotrba, L.M. (2007). *The antecedents and consequences of the variability of job satisfaction*. Unpublished doctoral dissertation, Wayne State University.

Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling.* Thousand Oaks, CA: Sage.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*,159-174. http://dx.doi.org/10.2307/2529310

*Larson, R., Verma, S., & Dworkin, J. (2001). Men's work and family lives in India: The daily organization of time and emotion. *Journal of Family Psychology, 15*, 206-224. http://dx.doi.org/10.1037/0893-3200.15.2.206

*Loi, R., Yang, J., & Diefendorff, J. M. (2009). Four-factor justice and daily job satisfaction: A multilevel investigation. *Journal of Applied Psychology, 94*, 770-781. http://dx.doi.org/10.1037/a0015714

*Low, S.Z. (2012). *Workplace harassment and counterproductive work behaviors: a daily diary investigation*. (Doctoral dissertation). Available from PsycINFO. (1319304677)

*Luong, A., & Rogelberg, S. G. (2005). Meetings and more meetings: The relationship between meeting load and the daily well-being of employees. *Group Dynamics: Theory, Research, and Practice, 9*, 58-67. http://dx.doi.org/10.1037/1089-2699.9.1.58

*Madrid, H. P., Patterson, M. G., Birdi, K. S., Leiva, P. I., & Kausel, E. E. (2014). The role of weekly high-activated positive mood, context, and personality in innovative work

behavior: A multilevel and interactional model. *Journal of Organizational Behavior, 35*, 234-256. http://dx.doi.org/10.1002/job.1867

Mathieu, J. E., Aguinis, H., Culpepper, S. A., & Chen, G. (2012). Understanding and estimating the power to detect cross-level interaction effects in multilevel modeling. *Journal of Applied Psychology, 97*, 951. http://dx.doi.org/10.1037/a0028380

*Matjasko, J. L., & Feldman, A. F. (2006). Bringing work home: The emotional experiences of mothers and fathers. *Journal of Family Psychology, 20*, 47-55. http://dx.doi.org/10.1037/0893-3200.20.1.47

*Matta, F. K., Erol-Korkmaz, H. T., Johnson, R. E., & Biçaksiz, P. (2014). Significant work events and counterproductive work behavior: The role of fairness, emotions, and emotion regulation. *Journal of Organizational Behavior, 35*, 920-944. http://dx.doi.org/10.1002/job.1934

*McCance, A.S. (2010). *Emotional labor in intercultural service encounters: An experience sampling study.* (Doctoral dissertation). Available from PsycINFO. (868856535)

*Meier, L. L., Semmer, N. K., & Hupfeld, J. (2009). The impact of unfair treatment on depressive mood: The moderating role of self-esteem level and self-esteem instability. *Personality and Social Psychology Bulletin, 35*, 643-655. http://dx.doi.org/10.1177/0146167208331337

*Miner, A. G. (2001). *Experience sampling events, moods, behaviors, and performance at work*. (Doctoral dissertation). Available from PsycINFO. (619962546; 2002-95004-072)

*Miner, A. G., & Glomb, T. M. (2010). State mood, task performance, and behavior at work: A within-persons approach. *Organizational Behavior and Human Decision Processes, 112*, 43-57. http://dx.doi.org/10.1016/j.obhdp.2009.11.009

*Miner, A. G., Glomb, T. M., & Hulin, C. (2005). Experience sampling mood and its correlates at work. *Journal of Occupational and Organizational Psychology, 78*, 171-193. http://dx.doi.org/10.1348/096317905X40105

*Mojza, E. J., Lorenz, C., Sonnentag, S., & Binnewies, C. (2010). Daily recovery experiences: The role of volunteer work during leisure time. *Journal of Occupational Health Psychology, 15*, 60-74. http://dx.doi.org/10.1037/a0017983

*Montgomery, A. J., Panagopoulou, E., Peeters, M. C. W., & Schaufeli, W. B. (2009). Exploring types of interference between work and non-work: Using a diary study approach. *Community, Work & Family, 12*, 455-471. http://dx.doi.org/10.1080/13668800903192101

*Navarro, J., & Arrieta, C. (2010). Chaos in human behavior: The case of work motivation. *The Spanish Journal of Psychology, 13*, 244-256. http://dx.doi.org/10.1017/S1138741600003826

Newman, D. A. (2014). Missing data five practical guidelines. *Organizational Research Methods, 17*, 372-411. http://dx.doi.org/10.1177/1094428114548590

Newman, D. A., & Sin, H. P. (2007). How do missing data bias estimates of within-group agreement? Sensitivity of SDWG, CVWG, rWG (J), rWG (J)*, and ICC to systematic nonresponse. *Organizational Research Methods, 12,* 113-147. http://dx.doi.org/10.1177/1094428106298969

*Nielsen, K., & Cleal, B. (2010). Predicting flow at work: Investigating the activities and job characteristics that predict flow states at work. *Journal of Occupational Health Psychology, 15*, 180-190. http://dx.doi.org/10.1037/a0018893

*Niessen, C., Sonnentag, S., & Sach, F. (2012). Thriving at work—A diary study. *Journal of Organizational Behavior, 33*, 468-487. http://dx.doi.org/10.1002/job.763

*Nohe, C., Michel, A., & Sonntag, K. (2014). Family–work conflict and job performance: A diary study of boundary conditions and mechanisms. *Journal of Organizational Behavior, 35*, 339-357. http://dx.doi.org/10.1002/job.1878

*Norton, T.A., Hammoud, K., Zacher, H. & Ashkanasy, N.M. (2014, April). *A daily diary study of environmental workplace behavior*. Poster presented at the annual conference of the Society for Industrial and Organizational Psychology, Honolulu, HI.

*Ohly, S., & Fritz, C. (2010). Work characteristics, challenge appraisal, creativity, and proactive behavior: A multi-level study. *Journal of Organizational Behavior, 31*, 543-565. http://dx.doi.org/10.1002/job.633

Ohly, S., Sonnentag, S., Niessen, C., & Zapf, D. (2010). Diary studies in organizational research. *Journal of Personnel Psychology, 9,* 79-93. doi:10.1027/1866-5888/a000009

*Oliver, P.H. (2001). *Parents psychological distress and children's temperament as moderators of the effects of daily work stress on family interactions*. (Doctoral dissertation). Available from PsycINFO. (250981376)

Ostroff, C. (1993). Comparing correlations based on individual-level and aggregated data. *Journal of Applied Psychology, 78,* 569-582. http://dx.doi.org/10.1037/0021-9010.78.4.569

*Ouweneel, E., Le Blanc, P. M., Schaufeli, W. B., & van Wijhe, C. I. (2012). Good morning, good day: A diary study on positive emotions, hope, and work engagement. *Human Relations, 65*, 1129-1154. http://dx.doi.org/10.1177/0018726711429382

*Park, Y., & Sprung, J. M. (2015). Weekly work–school conflict, sleep quality, and fatigue: Recovery self-efficacy as a cross-level moderator. *Journal of Organizational Behavior, 36*, 112-127. http://dx.doi.org/10.1002/job.1953

*Peeters, M. C. W., Buunk, B. P., & Schaufeli, W. B. (1995). Social interactions and feelings of inferiority among correctional officers: A daily event-recording approach. *Journal of Applied Social Psychology, 25*, 1073-1089. http://dx.doi.org/10.1111/j.1559-1816.1995.tb00618.x

*Pomerance, M.H. (2011). *Investigating how self-schema characteristics influence the motivational orientation and performance of individuals within a work setting: A within-person approach*. (Doctoral dissertation). Available from PsycINFO. (1582291303)

*Pride, R. E. (2006). *Understanding the experience of work overload: A computerized experience-sampling study*. (Doctoral dissertation). Available from PsycINFO. (621559757; 2006-99004-096)

*Reicherts, M., & Pihet, S. (2000). Job newcomers coping with stressful situations: A micro-analysis of adequate coping and well-being. *Swiss Journal of Psychology, 59*, 303-316. http://dx.doi.org/10.1024//1421-0185.59.4.303

*Roberts, N.A. & Levenson, R. W. (2001). The remains of the workday: Impact of job stress and exhaustion on marital interaction in police couples. *Journal of Marriage and Family, 63*, 1052-1067. http://dx.doi.org/10.1111/j.1741-3737.2001.01052.x

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

*Rodell, J. B., & Judge, T. A. (2009). Can "good" stressors spark "bad" behaviors? The mediating role of emotions in links of challenge and hindrance stressors with citizenship and counterproductive behaviors. *Journal of Applied Psychology, 94*, 1438-1451. http://dx.doi.org/10.1037/a0016752

Roth, P. L., & BeVier, C. A. (1998). Response rates in HRM/OB survey research: Norms and correlates, 1990-1994. *Journal of Management, 24*, 97-117. http://dx.doi.org/10.1177/014920639802400107

*Sanz-Vergel, A., Rodríguez-Muñoz, A., Bakker, A. B., & Demerouti, E. (2012). The daily spillover and crossover of emotional labor: Faking emotions at work and at home. *Journal of Vocational Behavior, 81*, 209-217. http://dx.doi.org/10.1016/j.jvb.2012.07.003

*Schmitt, A., Zacher, H., & Frese, M. (2012). The buffering effect of selection, optimization, and compensation strategy use on the relationship between problem solving demands and occupational well-being: A daily diary study. *Journal of Occupational Health Psychology, 17*, 139-149. http://dx.doi.org/10.1037/a0027054

*Schreurs, B., Guenter, H., Hülsheger, U., & van Emmerik, H. (2014). The role of punishment and reward sensitivity in the emotional labor process: A within-person perspective. *Journal of Occupational Health Psychology, 19*, 108-121. http://dx.doi.org/10.1037/a0035067

*Schreurs, B. H. J., Hetty, v. E., Günter, H., & Germeys, F. (2012). A weekly diary study on the buffering role of social support in the relationship between job insecurity and employee performance. *Human Resource Management, 51*, 259-280. http://dx.doi.org/10.1002/hrm.21465

Schwartz, J. E., & Stone, A. A. (1998). Strategies for analyzing ecological momentary assessment data. *Health Psychology, 17*, 6-16. http://dx.doi.org/10.1037/0278-6133.17.1.6

*Seiger, C. P., & Wiese, B. S. (2009). Social support from work and family domains as an antecedent or moderator of work-family conflicts? *Journal of Vocational Behavior, 75*, 26-37. http://dx.doi.org/10.1016/j.jvb.2009.03.001

*Shiu, A. T. (1998). The significance of sense of coherence for the perceptions of task characteristics and stress during interruptions amongst a sample of public health nurses in Hong Kong: Implications for nursing management. *Public Health Nursing, 15*, 273-280. http://dx.doi.org/10.1111/j.1525-1446.1998.tb00350.x

*Simbula, S. (2010). Daily fluctuations in teachers' well-being: A diary study using the job demands–resources model. *Anxiety, Stress & Coping: An International Journal, 23*, 563-584. http://dx.doi.org/10.1080/10615801003728273

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological science, 22*, 1359-1366. http://dx.doi.org/10.1037/e519702015-014

*Song, Z., Foo, M., & Uy, M. A. (2008). Mood spillover and crossover among dual-earner couples: A cell phone event sampling study. *Journal of Applied Psychology, 93*, 443-452. http://dx.doi.org/10.1037/0021-9010.93.2.443

*Song, Z., Foo, M. D., Uy, M. A., & Sun, S. (2011). Unraveling the daily stress crossover between unemployed individuals and their employed spouses. *Journal of Applied Psychology, 96*, 151. http://dx.doi.org/10.1037/a0021035

*Sonnenschein, M., Sorbi, M. J., van Doornen, Lorenz J. P., Schaufeli, W. B., & Maas, C. J. M. (2007). Electronic diary evidence on energy erosion in clinical burnout. *Journal of Occupational Health Psychology, 12*, 402-413. http://dx.doi.org/10.1037/1076-8998.12.4.402

*Sonnentag, S., & Bayer, U. (2005). Switching off mentally: Predictors and consequences of psychological detachment from work during off-job time. *Journal of Occupational Health Psychology, 10*, 393-414. http://dx.doi.org/10.1037/1076-8998.10.4.393

*Sonnentag, S., Binnewies, C., & Mojza, E. J. (2008). "Did you have a nice evening?" A day-level study on recovery experiences, sleep, and affect. *Journal of Applied Psychology, 93*, 674-684. http://dx.doi.org/10.1037/0021-9010.93.3.674

*Sonnentag, S., & Grant, A. M. (2012). Doing good at work feels good at home, but not right away: When and why perceived prosocial impact predicts positive affect. *Personnel Psychology, 65*, 495-530. http://dx.doi.org/10.1111/j.1744-6570.2012.01251.x

*Sonnentag, S., Mojza, E. J., Demerouti, E., & Bakker, A. B. (2012). Reciprocal relations between recovery and work engagement: The moderating role of job stressors. *Journal of Applied Psychology, 97*, 842-853. http://dx.doi.org/10.1037/a0028292

*Spence, J. R., Brown, D. J., Keeping, L. M., & Lian, H. (2014). Helpful today, but not tomorrow? Feeling grateful as a predictor of daily organizational citizenship behaviors. *Personnel Psychology, 67*, 705-738. http://dx.doi.org/10.1111/peps.12051

*Stein, J. (2010). *Situational and trait influences on dynamic justice*. (Doctoral dissertation). Available from PsycINFO. (759472880)

*Stocker, D., Jacobshagen, N., Krings, R., Pfister, I. B., & Semmer, N. K. (2014). Appreciative leadership and employee well-being in everyday working life. *Zeitschrift Für Personalforschung, 28*, 73-95. http://dx.doi.org/10.1177/239700221402800105

*Story, L. B., & Repetti, R. (2006). Daily occupational stressors and marital behavior. *Journal of Family Psychology, 20*, 690-700. http://dx.doi.org/10.1037/0893-3200.20.4.690

*Stucky, E. R., Dresselhaus, T. R., Dollarhide, A., Shively, M., Maynard, G., Jain, S., . . . Rutledge, T. (2009). Intern to attending: Assessing stress among physicians. *Academic Medicine, 84*, 251-257. http://dx.doi.org/10.1097/ACM.0b013e3181938aad

*Tadić, M., Bakker, A. B., & Oerlemans, W. G. M. (2013). Work happiness among teachers: A day reconstruction study on the role of self-concordance. *Journal of School Psychology, 51*, 735-750. http://dx.doi.org/10.1016/j.jsp.2013.07.002

Tay, L., Woo, S. E., & Vermunt, J. K. (2014). A Conceptual and Methodological Framework for Psychometric Isomorphism Validation of Multilevel Construct Measures. *Organizational Research Methods, 17*, 77-106.

*Ten Brummelhuis, L. L., & Bakker, A. B. (2012). Staying engaged during the week: The effect of off-job activities on next day work engagement. *Journal of Occupational Health Psychology, 17*, 445-455. http://dx.doi.org/10.1037/a0029213

*Teuchmann, K., Totterdell, P., & Parker, S. K. (1999). Rushed, unhappy, and drained: An experience sampling study of relations between time pressure, perceived control, mood, and emotional exhaustion in a group of accountants. *Journal of Occupational Health Psychology, 4*, 37-54. http://dx.doi.org/10.1037/1076-8998.4.1.37

Thorndike, E. L. (1939). On the fallacy of imputing the correlations found for groups to the individuals or smaller groups composing them. *The American Journal of Psychology, 52*, 122-124. http://dx.doi.org/10.2307/1416673

*Tims, M., Bakker, A. B., & Derks, D. (2014). Daily job crafting and the self-efficacy—Performance relationship. *Journal of Managerial Psychology, 29*, 490-507. http://dx.doi.org/10.1108/JMP-05-2012-0148

*To, M. L., Fisher, C. D., Ashkanasy, N. M., & Rowe, P. A. (2012). Within-person relationships between mood and creativity. *Journal of Applied Psychology, 97*, 599-612. http://dx.doi.org/10.1037/a0026097

*Totterdell, P., Hershcovis, M. S., Niven, K., Reich, T. C., & Stride, C. (2012). Can employees be emotionally drained by witnessing unpleasant interactions between coworkers? A diary study of induced emotion regulation. *Work & Stress, 26*, 112-129. http://dx.doi.org/10.1080/02678373.2012.681153

*Totterdell, P., & Holman, D. (2003). Emotion regulation in customer service roles: Testing a model of emotional labor. *Journal of Occupational Health Psychology, 8*, 55-73. http://dx.doi.org/10.1037/1076-8998.8.1.55

*Totterdell, P., Kellett, S., Teuchmann, K., & Briner, R. B. (1998). Evidence of mood linkage in work groups. *Journal of Personality and Social Psychology, 74*, 1504-1515. http://dx.doi.org/10.1037/0022-3514.74.6.1504

*Totterdell, P., Wood, S., & Wall, T. (2006). An intra-individual test of the demands-control model: A weekly diary study of psychological strain in portfolio workers. *Journal of Occupational and Organizational Psychology, 79*, 63-84. http://dx.doi.org/10.1348/096317905x52616

*Trougakos, J. P., Beal, D. J., Cheng, B. H., Hideg, I., & Zweig, D. (2014). Too drained to help: A resource depletion perspective on daily interpersonal citizenship behaviors. *Journal of Applied Psychology, 100,*227-236. http://dx.doi.org/10.1037/a0038082

*Tschan, F., Rochat, S., & Zapf, D. (2005). It's not only clients: Studying emotion work with clients and co-workers with an event-sampling approach. *Journal of Occupational and Organizational Psychology, 78*, 195-220. http://dx.doi.org/10.1348/096317905x39666

*Tuckey, M. R., & Neall, A. M. (2014). Workplace bullying erodes job and personal resources: Between- and within-person perspectives. *Journal of Occupational Health Psychology, 19*, 413-424. http://dx.doi.org/10.1037/a0037728

*Unger, D., Niessen, C., Sonnentag, S., & Neff, A. (2014). A question of time: Daily time allocation between work and private life. *Journal of occupational and organizational psychology, 87*, 158-176. http://dx.doi.org/10.1111/joop.12045

Vancouver, J. B., Thompson, C. M., Tischner, E. C., & Putka, D. J. (2002). Two studies examining the negative effect of self-efficacy on performance. *Journal of Applied Psychology, 87*, 506-516. http://dx.doi.org/10.1037/0021-9010.87.3.506

*Van den Tooren, M., & de Jonge, J. (2011). Job resources and regulatory focus as moderators of short-term stressor-strain relations: A daily diary study. *Journal of Personnel Psychology, 10*, 97. http://dx.doi.org/10.1027/1866-5888/a000042

*Van Gelderen, B. R., Bakker, A. B., Konijn, E. A., & Demerouti, E. (2011). Daily suppression of discrete emotions during the work of police service workers and criminal investigation officers. *Anxiety, Stress & Coping, 24*, 515-537. http://dx.doi.org/10.1080/10615806.2011.560665

*Van Gelderen, B., Heuven, E., van Veldhoven, M., Zeelenberg, M., & Croon, M. (2007). Psychological strain and emotional labor among police-officers: A diary study. *Journal of Vocational Behavior, 71*, 446-459. http://dx.doi.org/10.1016/j.jvb.2007.09.001

*Van Hooff, Madelon L. M., Geurts, S. A. E., Kompier, M. A. J., & Taris, T. W. (2006). Work-home interference: How does it manifest itself from day to day? *Work & Stress, 20*, 145-162. http://dx.doi.org/10.1080/02678370600915940

*Volman, F. E., Bakker, A. B., & Xanthopoulou, D. (2013). Recovery at home and performance at work: A diary study on self–family facilitation. *European Journal of Work and Organizational Psychology, 22*, 218-234. http://dx.doi.org/10.1080/1359432x.2011.648375

*Volmer, J., Binnewies, C., Sonnentag, S., & Niessen, C. (2012). Do social conflicts with customers at work encroach upon our private lives? A diary study. *Journal of occupational health psychology, 17*, 304. http://dx.doi.org/10.1037/a0028454

*Wang, P. S., Beck, A. L., Berglund, P., McKenas, D. K., Pronk, N. P., Simon, G. E., & Kessler, R. C. (2004). Effects of major depression on moment-in-time work performance. *American Journal of Psychiatry, 161*, 1885-1891. http://dx.doi.org/10.1176/appi.ajp.161.10.1885

*Wang, M., Liu, S., Liao, H., Gong, Y., Kammeyer-Mueller, J., & Shi, J. (2013). Can't get it out of my mind: Employee rumination after customer mistreatment and negative mood in the next morning. *Journal of Applied Psychology, 98*, 989-1004. http://dx.doi.org/10.1037/a0033656

*Weiss, H. M., Nicholas, J. P., & Daus, C. S. (1999). An examination of the joint effects of affective experiences and job beliefs on job satisfaction and variations in affective experiences over time. *Organizational Behavior and Human Decision Processes, 78*, 1-24. http://dx.doi.org/10.1006/obhd.1999.2824

*White, R. C. (2012). *A few feet to failure: Two essays on enhancing the service experience through customer contact employee performance*. (Doctoral dissertation). Available from PsycINFO. (875560302)

*Williams, K. J., Suls, J., Alliger, G. M., Learner, S. M., & Wan, C. K. (1991). Multiple role juggling and daily mood states in working mothers: An experience sampling study. *Journal of Applied Psychology, 76*, 664-674. http://dx.doi.org/10.1037/0021-9010.76.5.664

*Xanthopoulou, D., Baker, A. B., Heuven, E., Demerouti, E., & Schaufeli, W. B. (2008). Working in the sky: A diary study on work engagement among flight attendants. *Journal of Occupational Health Psychology, 13*, 345-356. http://dx.doi.org/10.1037/1076-8998.13.4.345

*Yang, J., & Diefendorff, J. M. (2009). The relations of daily counterproductive workplace behavior with emotions, situational antecedents, and personality moderators: A diary

study in hong kong. *Personnel Psychology, 62*, 259-295. http://dx.doi.org/10.1111/j.1744-6570.2009.01138.x

*Zacher, H., Brailsford, H. A., & Parker, S. L. (2014). Micro-breaks matter: A diary study on the effects of energy management strategies on occupational well-being. *Journal of Vocational Behavior, 85*, 287-297. http://dx.doi.org/10.1016/j.jvb.2014.08.005

*Zacher, H., & Wilden, R. G. (2014). A daily diary study on ambidextrous leadership and self-reported employee innovation. *Journal of Occupational and Organizational Psychology, 87*, 813-820. http://dx.doi.org/10.1111/joop.12070

*Zelenski, J. M., Murphy, S. A., & Jenkins, D. A. (2008). The happy-productive worker thesis revisited. *Journal of Happiness Studies, 9*, 521-537. http://dx.doi.org/10.1007/s10902-008-9087-4

Zevon, M. A., & Tellegen, A. (1982). The structure of mood change: An idiographic/nomothetic analysis. *Journal of Personality and Social Psychology, 43*, 111-122.

*Zhang, Y. (2014). *Leaders' daily work demands, recovery, and leadership behaviors*. (Doctoral dissertation). Available from PsycINFO. (1506432598; 2014-99030-380)

*Zohar, D., Tzischinski, O., & Epstein, R. (2003). Effects of energy availability on immediate and delayed emotional reactions to work events. *Journal of Applied Psychology, 88*, 1082-1093. http://dx.doi.org/10.1037/0021-9010.88.6.1082

APPENDIX A: SEARCH TERMS

Appendix A. List of search terms used in the literature search for ESM studies.

Search terms	
affective events theory	organizational citizenship behavior
contextual performance	organizational commitment
counterproductive work behavior	organizational justice
coworker	stress
deep acting	supervisor
diary study	surface acting
ecological momentary assessment	task performance
emotional labor	withdrawal
employee	within-person
experience sampling method	work
job performance	work-family conflict
job satisfaction	workplace deviance

APPENDIX B: COHEN'S D VALUES FOR 3 RESPONSE RATES

Appendix B. Cohen's *d* values for 3 response rates with incentives and reminders.

	Between Person Response Rate	Between Person Analysis Rate	Pooled Within-Persons Response Rate
No Incentives vs. Incentives	*d*= .05 / *k's*= 46, 39 [-0.39, 0.49]	*d*= -.18 / *k's*= 41, 42 [-0.63, 0.26]	*d*= -.34 / *k's*= 26, 25 [-.91, 0.24]
No Reminder vs. Reminder	*d*= -.09 / *k's*= 43, 42 [-0.53, 0.34]	*d*= -.31 / *k's*= 34, 49 [-0.76, 0.14]	*d*= .74 / *k's*= 5, 46 [-.24, 1.72]

Note: Each effect size *d* is followed by the sample sizes (*k*) of its two groups (starting with the No/Unreported Incentives or No Reminder); 95% CI's appear in brackets below each *d* value.

APPENDIX C: CODING DISCREPENCIES AND RESOLUTIONS

Appendix C. Summary of coding discrepancies and resolutions used to calculate corrected Cohen's kappa for within-persons construct subcategories.

Construct Categories	Item(s) Coded Discrepantly	Resolution
Activities Outside Work, Family	family/home activities	family/home activities coded as Activities Outside Work
Affect, Creativity, Job Satisfaction, Motivation,	Csikszentmihalyi, M., & LeFevre, J. (1989) – coder 1 coded the authors' general measurement category "quality of experience" whereas coder 2 coded the subfacets within the quality of experience category (motivation, concentration, relaxation, creativity, satisfaction, potency, and affect)	subfacets coded rather than general construct
Beliefs about Self, Motivation	mastery experiences	mastery experiences coded as Motivation
Beliefs about Self, Motivation	empowerment	empowerment coded as Beliefs about Self
Beliefs about Self, Perceived Support	sense of belonging	sense of belonging coded as Perceived Support
Beliefs about Self, Work Behavior	identity management	identity management coded as Work Behavior
Conflict/Constraint, Stress/Strain	hassles	hassles coded as Stress/Strain
Conflict/Constraint, Work Events,	work interruptions	work interruptions coded as Work Events
Conflict/Constraint, Work Family Conflict	marital conflict, school-work conflict	marital conflict, school-work conflict coded as Conflict/Constraint
CWB, Stress/Strain	sexual harassment	sexual harassment coded as CWB
Demands, Perceived Support, Stress/Strain	Gallo, L. C., Bogart, L. M., Vranceanu, A. M., & Matthews, K. A. (2005) – coder 1 coded the authors' general measurement category "psychosocial experience" whereas coder 2 coded the subfacets within the psychosocial experience category (perceived environmental demands, perceived control, social intimacy/support, social strain)	subfacets coded rather than general construct
Emotional Labor/Regulation, Family	surface acting at home	surface acting at home counts as only Emotional Labor/Regulation
Family, Life Satisfaction	marital satisfaction	marital satisfaction coded as Family
Family, Work Family Conflict	work-family-interaction, work-family-facilitation	work-family-interaction, work-family-facilitation coded as Work Family Conflict

Flow, Work Engagement	workflow interruptions, absorption	workflow interruptions and absorption were coded as Flow
Job Characteristics, Job Satisfaction, Work Behavior	job crafting	job crafting coded as Job Characteristics
Job Characteristics, Work Behavior	time use	time use coded as Work Behavior
Job Characteristics, Work Behavior	task juggling	task juggling coded as Job Characteristic
Job Satisfaction, Life Satisfaction, Positive Affect,	happiness at work	happiness at work coded as Job Satisfaction
Life Satisfaction	meaning in life	meaning in life coded as Life Satisfaction
Motivation	wants	wants coded as Motivation
Motivation, Work Behavior	energy management strategies	energy management strategies coded as Work Behavior
Motivation, Work Engagement	vigor, vitality	vigor, vitality coded as Work Engagement
Motivation, Work Engagement	concentration	concentration coded as Motivation
Performance, Work Behavior	self-rated quality of service, personal and work outcomes	self-rated quality of service, personal and work outcomes coded as Performance
Personality	person-job fit, person-organization fit	P-O fit, P-J fit coded as Personality
Recovery	relaxation	relaxation coded as Recovery
Stress/Strain	consequences	consequences coded as Stress/Strain
Work Behavior	activity, selection/optimization/compensation strategy use, smartphone use	activity, selection/optimization/compensation strategy use, smartphone use coded as Work Behavior
Work Behavior, Work Events	frequency of daily communication, humor expression, appreciate behavior/situation	frequency of daily communication, humor expression, appreciate behavior/situation coded as Work Behavior
Work Family Conflict	objective function	objective function coded as Work Family Conflict