

THESIS

EMPLOYEE ENGAGEMENT: UNDERSTANDING THE CONSTRUCT'S STABILITY

Submitted by

James W. Weston

Department of Psychology

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2016

Master's Committee:

Advisor: Zinta Byrne

Gwenith Fisher

Daniel Ganster

Jennifer Harman

Copyright by James Warren Weston 2016

All Rights Reserved

ABSTRACT

EMPLOYEE ENGAGEMENT: UNDERSTANDING THE CONSTRUCT'S STABILITY

Research has been contradictory in the definition and measurement of employee engagement. Despite being predominantly measured as a trait-like stable construct, engagement was originally introduced as a moment-to-moment fluctuating concept. The focus of the current research is on the conceptualization of the stability of the employee engagement construct. Specifically, I examined fluctuations in engagement as they related to varying levels of three theoretical antecedents (psychological safety, psychological availability, and psychological meaningfulness). Using experience sampling, forty nine participants were asked to complete surveys on cellular devices or workplace computers, for five weeks, twice per day at random moments, for a total of 30 data-points per participant. These daily surveys assessed fluctuations in engagement levels in relation to the other contextual variables, while their accumulation over the five-week period provided insight into the relative stability of the construct. Results showed momentary job engagement was positively related to momentary stress, affect, and the quality of coworker interactions. Additionally, between-person differences in engagement were positively related to job satisfaction, general positive affect, and general job engagement. The current study provides a glimpse into within-person fluctuations in engagement. Findings suggest that although engagement may vary within-employees, between-person differences are still present and represent valuable information.

ACKNOWLEDGEMENTS

This thesis began with an idea in my first semester of graduate school. Over the subsequent several years it developed into this, its final form. I would be remiss not to thank the multitude of people who helped make this possible.

First, I could not have completed this process without my advisor Dr. Zinta Byrne. Arriving in graduate school, I learned that your research interests had switched from justice to engagement, and this emphasis on motivation drew me in. You put up with countless drafts of the thesis proposal, and helped guide me through a data-collection process that we both look back on and say “never again” (for graduate school at least). In the last three years, you have helped me navigate classes, publications, assistantships, and research grants. Your guidance has been truly invaluable thus far, and I know the next few years that you endure with me as a student will be equally as beneficial.

Second, thank you to my committee members, Drs Jennifer Harman, Daniel Ganster, and Gwenith Fisher for your continued effort and input throughout this process. You all helped me develop this project into what it is today. Gwen, your statistics and measurement expertise proved invaluable when designing and evaluating this process, and your friendship outside of the classroom helped me push through when data-collection was frustrating. Jen, throughout our meetings you helped ensure that we always saw the bigger picture for psychology and our study as a whole. Frequently, I would get bogged down in I/O specific considerations and you were always able to bring us back to implications for the psychology field as a whole. Dan, as we went through this process I always kept one question in mind, and will always keep it in mind in my future work: “So, in the end, why does this matter?” Focusing on this question above others

helped me grow as a researcher. Research was no longer a project to check a box, but it became an endeavor designed to further our science and humanity.

Third, I have had amazing colleagues throughout my time at CSU. One of the big draws to CSU as a school was the tight-knit community among the graduate students and faculty. Thank you to all of you for being my social support system in graduate school and life. My cohort, Kevin, Megan, Steve R., Sam, and Charlie, I am honored simply to know you and truly privileged to call each of you good friends. To my lab-mates, Steve M., Kyla, and Kyle, you have helped me develop as a student beyond anything you could imagine. You set incredible standards to live up to, and I see each of you – cohort and lab-mates – as role models.

Finally, and most importantly, thank you to my family for supporting me through this thesis process and for the unconditional love you have given throughout my life. Mom, Dad, and Lee-Ann, you have always been incredibly supportive of me even when I wasn't... academically inclined. You have always taught a deep respect for education – although, I was slow to catch on – and because of that I am several years through graduate school, and now a thesis down. To my siblings, Jess, Amanda, and Robert, you have all been amazing to look up to as a kid, and now as an adult. Thanks for showing me how to be the best person I can be at all times. Alison, my incredible fiancée, you have been my rock over the past seven years, but especially in the past three as we have lived in Colorado. You continually put up with my graduate school jargon, and listen to my thesis ideas even when they are tiring. You learned an invaluable skill: ignoring most of what I say as soon as the words “thesis” or “psychology” are uttered. You will undoubtedly need this for years to come. Jokes aside, you truly are the reason I am in this position today, and I could not thank you enough, or be more in love with you.

DEDICATION

This thesis is dedicated to my family: Alison, Mom, Dad, Lee-Ann, Debbie, Bryan, Robert, Jessica, Norman, Riley, Amanda, Curtis, Brad, and Cameron. You have each been, and continue to be inspirations.

TABLE OF CONTENTS

ABSTRACT.....	ii
ACKNOWLEDGEMENTS.....	iii
DEDICATION.....	v
INTRODUCTION.....	1
BACKGROUND: EMPLOYEE ENGAGEMENT.....	4
THEORETICAL INCONSISTENCIES REGARDING STABILITY.....	7
RECENT STABILITY RESEARCH.....	10
THEORETICAL UNDERPINNINGS OF THE CURRENT STUDY.....	13
FLUCTUATIONS IN ENGAGEMENT.....	13
STABILITY IN ENGAGEMENT.....	15
RELATED VARIABLES.....	16
PSYCHOLOGICAL MEANINGFULNESS.....	16
PSYCHOLOGICAL AVAILABILITY.....	17
PSYCHOLOGICAL SAFETY.....	19
STATE-AFFECT.....	20
TRAIT-AFFECT.....	21
JOB SATISFACTION.....	22
SUMMARY.....	24
METHOD.....	25
PARTICIPANTS.....	25
PROCEDURES.....	26
EXPERIENCE SAMPLING METHODOLOGY.....	26
PILOT TEST.....	28
PILOT TEST PARTICIPANTS.....	29
PILOT TEST PROCEDURE.....	29
FINAL STUDY PROCEDURES.....	31
MEASURES.....	33
DEMOGRAPHICS.....	33
PERCEIVED WORK STRESS.....	33
AFFECTIVITY.....	34

ENGAGEMENT.....	35
JOB SATISFACTION.....	36
QUALITY OF COWORKER RELATIONS.....	36
WORK TASKS.....	37
OPEN-ENDED QUESTIONS.....	38
DATA ANALYSIS.....	39
RESULTS.....	41
MULTILEVEL ANALYSIS.....	43
DISCUSSION.....	48
WITHIN-DAY DATA.....	48
PERSON-LEVEL DATA.....	50
IMPLICATIONS AND FUTURE DIRECTIONS.....	51
LIMITATIONS.....	55
CONCLUSION.....	58
TABLE 1.....	59
TABLE 2.....	60
TABLE 3.....	61
TABLE 4.....	62
TABLE 5.....	63
TABLE 6.....	64
<i>FIGURE 1</i>	65
<i>FIGURE 2</i>	66
<i>FIGURE 3</i>	67
<i>FIGURE 4</i>	68
<i>FIGURE 5</i>	69
<i>FIGURE 6</i>	70
<i>FIGURE 7</i>	71
<i>FIGURE 8</i>	72
REFERENCES.....	73
APPENDIX A.....	84

Introduction

Engagement has become increasingly recognized as a key research topic in the organizational sciences (Sonnentag, 2011). For example, engagement is positively related to productivity (Rich et al., 2010), organizational commitment (Chalofsky & Krishna, 2009), and organizational citizenship behaviors (Moliner, Martinez-Tur, Ramos, Perio, & Cropanzano, 2008; Rich, 2006), and negatively related to outcomes such as turnover intentions, and burnout (Schaufeli, Bakker, & Van Rhenen, 2009).

Employee engagement refers to a motivational state (Meyer & Gagné, 2008; Rich, LePine, & Crawford, 2010), characterized by exerting one's full self in a work role (Kahn, 1990). Since Kahn's (1990) seminal paper on engagement, researchers have treated employee engagement as both a trait-like and momentary construct, resulting in confusion over the stability of the construct. For example, conceptualizations range from Kahn's (1990) supposition that engaged states have a highly dynamic temporal quality, to Schaufeli, Salanova, González-Romá, and Bakker's (2002) treatment of engagement as a lasting affective state. Other researchers have proposed engagement as a process inferring trait-like and momentary qualities. For example, Macey and Schneider (2008) suggested that engagement is a process with behavioral, moment-to-moment, and trait-like aspects.

Yet, none of these conceptualizations clearly articulates the stability of the construct; that is, how long does an engaged state actually last? Are employees engaged at a consistent level, or does their engagement naturally vary within months, weeks, or even single workdays? What affects an employee's level of engagement? Moreover, what affects *fluctuations* in levels of engagement? Recent theoretical review papers concerning fluctuations in engagement draw

attention to a need for additional research in this area (Bakker, 2014; Xanthopoulou, Bakker, & Ilies, 2012). These review papers provide a comprehensive description of current studies looking at what may be considered fluctuations in engagement. Yet, even in these reviews, authors arrive at a critical conclusion: the need for “capturing fluctuations [in engagement] over shorter periods of time [as opposed to across days and weeks]” (Bakker, 2014, p. 8). Furthermore, researchers purport that assessing engagement and other constructs in a moment-to-moment framework could “contribute to a better explanation of employee functioning at work on a day-to-day basis” (Xanthopoulou et al., 2012, p. 1052). Yet, despite these researchers clearly noting the need for additional research, no one to date has taken up the charge. Moreover, Bakker, Albrecht, and Leiter (2011) posited that among the key questions needing to be resolved regarding engagement, fluctuations in engagement was crucial for future research on the construct.

Therefore, the goal of the current study is to address the issue of the stability of the construct of employee engagement. By ‘stability’ I am referring to the duration of an engaged state. Fluctuations in engagement may indicate that engagement is not a trait-like construct, and that researchers and practitioners should embrace the moment-to-moment conceptualization of engagement in both their theories and applications.

In this study, I used an intensive longitudinal experience sampling data collection process to parse the issue of stability versus variability in employee engagement. In addition, consistent with recent trends in engagement research, I examined specific correlates of engagement that potentially influence its fluctuations. Researchers have begun to look at daily fluctuations of engagement (Bakker, 2014), and have been assessing what covaries with or causes these fluctuations in engagement (e.g., work-life-balance perceptions; Culbertson, Mills, & Fullager, 2012). Yet, none of these studies on daily fluctuations to date have looked at the specific

antecedents to engagement as proposed by Kahn (1990) - meaningfulness, safety, and availability. In this study, I use momentary stress (Kahn, 1992) and the quality of coworker interactions (Kahn, 1990) as substitutes for psychological availability and safety, respectively. Stress and coworker interactions represent more concrete variables than their theoretical counterparts (availability and safety), thus adding to the practicality of the results. In addition, the longevity of the engaged state is examined in the current study, adding to the overall contribution of the study to this line of inquiry and advance the theoretical understanding of the engagement construct.

The stability, or variability, of a construct can directly affect the relationships between that construct and organizational antecedents and outcomes (George, 1991). For example, when looking at trait vs. state mood, George found that *state* mood was positively related to pro-social behaviors. Contrarily, she found that *trait* mood was not related to pro-social behaviors. The distinction between state and trait conceptualizations in George's research made the difference between a significant prediction and a lack of relationship altogether. The same distinction may hold for employee engagement. Given employee engagement's frequently cited relationship with positive outcomes like satisfaction (Harter, Schmidt & Hayes, 2002), and performance (Christian, Garza, & Slaughter, 2011), accurately conceptualizing the stability of engagement may be critical for future research on outcomes associated with an engaged workforce.

The stability of engagement, and consequently its measurement, has serious implications for practice. To illustrate some of the practical implications of accurately conceptualizing employee engagement's stability, I begin with an example: In an annual review of a company's workforce, the organization assesses the engagement of its employees. To do so, the organization uses a tool that measures engagement with the assumption that it is a stable construct; that

employees' engagement levels have been steady all year. When assessing engagement, the organization may use a measure like the Utrecht Work Engagement Scale (UWES; Schaufeli et al., 2002), or the Gallup Organization's Q12 (Harter et al., 2002), both of which operationalize engagement as a stable, lasting state. Upon using this cross-sectional approach to gather information, let us suppose the organization finds their employees' engagement levels are less than desired. Consequently, they implement expensive engagement interventions to boost engagement levels. However, without understanding whether engagement is actually a stable construct or not, the organization may be attempting to "fix" a problem that does not exist. Namely, engagement levels may have been low *at the time* of the survey due to natural fluctuations in engagement levels, perhaps based when the survey was offered (end of the workday) or season (e.g., holidays). If this is the case, the organization just spent an unnecessary amount of money and time fixing a non-problem.

To adequately conceptualize the stability or variability of engagement, in this study fluctuations of engagement over time are examined. This paper begins with a discussion of current engagement theories and transitions into reasons for expansion. I then discuss concerns raised by previous research about fluctuations in engagement, introduce the intensive data collection process, analytic techniques, present results of hypothesis testing, and conclude with a discussion on the practical and theoretical implications of the findings, offering suggestions for practitioners and future research.

Background: Employee Engagement

Grounded in role theory (Goffman, 1959), Kahn (1990) focused on fluctuations in engagement, discussing the psychological and behavioral mechanisms associated with asserting oneself into one's work-role. Inherent in Kahn's conceptualization was the idea that engagement

was unlike other concepts of motivation in that employees could experience ebbs or flows in motivational state based on conscious or unconscious decisions (Kahn & Fellows, 2014).

Fittingly, Kahn (1990) defined engagement as the simultaneous moment-to-moment allocation of emotional, physical, and cognitive resources to one's work role.

To achieve this engaged state, Kahn (1990) proposed three antecedents: psychological availability, psychological safety, and psychological meaningfulness. He purported that the best way to conceptualize these antecedents was by an employee asking him or herself three questions: "How meaningful is it for me to bring myself into this performance? How safe is it to do so? How available am I to do so?"(p.703). An employee's answers to these questions may vary at times. For example, certain work tasks may be more meaningful to the employee than other tasks, or a certain group at work may be more supportive of the employee than other work groups are, thus changing that employee's perception of psychological safety. Rich, LePine, and Crawford (2010) investigated these antecedents to engagement using proxies including core self-evaluations to represent availability, value congruence to represent meaningfulness, and perceived organizational support as psychological safety. The authors found empirical support for each of these as antecedents to engagement. Additionally, May, Gilson, and Harter (2004) provided empirical support for the notion that psychological meaningfulness, safety, and availability are related to engagement (at the static level).

Although Kahn's (1990) definition and framework launched the engagement revolution, the most frequently used definition of engagement in the literature is that of Schaufeli and colleagues (2002). Their model of engagement stems from the burnout literature. According to Maslach, Schaufeli, and Leiter (2001), burnout is defined as a lasting state of disengagement in which workers experience emotional exhaustion, depersonalization, and decreased personal

accomplishment. These authors asserted that engagement, therefore, is the direct opposite of burnout. Schaufeli and colleagues (2002) further refined the definition of engagement as having three components: vigor, dedication, and absorption. Vigor refers to the energies expended at work, and the persistence to continue on tasks. Dedication is likened to involvement in that it entails a pride and enthusiasm in one's work. Finally, absorption refers to total concentration and being fully encapsulated in one's work. Given the theoretical juxtaposition between engagement and the stable construct of burnout, it follows that like burnout, work engagement was conceptualized as a stable trait-like construct by Maslach and colleagues, and Schaufeli et al.

In contrast to presenting engagement as either momentary or trait-like, Macey and Schneider (2008) proposed that engagement has both state- and trait-like characteristics. They asserted that engagement is a process that flows through three distinct stages – trait engagement, state engagement, and behavioral engagement. Trait engagement referred to a lasting engagement characterized by proactivity and dedication, mirroring personality, specifically conscientiousness. State engagement follows trait engagement and is described as an affective construct, mirroring absorption (Schaufeli et al., 2002) and energy (Kahn, 1990). Lastly, Macey and Schneider proposed behavioral engagement included observable behaviors most closely associated with extra-role behaviors, such as personal initiative and organizational citizenship. Despite criticisms of this tripartite distinction (see Newman & Harrison, 2008 for an example), Macey and Schneider's process model, nevertheless, has sparked conversation over the stability of engagement.

I frame the current research within Kahn's (1990) theoretical framework of engagement for several reasons. First and foremost, Kahn's definition was the first proposed, identifying engagement as a moment-to-moment state; a fluctuation that has yet to be supported or rejected.

Second, and perhaps more importantly, concerns surrounding other conceptualizations of engagement as confounded with other variables such as burnout, satisfaction, and intrinsic motivation respectively (Cole et al., 2012; Newman & Harrison, 2008) suggest a lack of clarity as to what engagement is and is not. Third, incorporating stable constructs into the definition of a possibly fluctuating construct prevents a clear assessment of whether the construct is or is not stable or fluctuating. Recent evidence shows that Kahn's model of engagement is theoretically and statistically distinct from satisfaction, intrinsic motivation, and job involvement (Rich et al., 2010), whereas recent work suggests Schaufeli et al.'s (2002) definition shares significant overlap with burnout (Cole et al., 2012), a rather steady state. Fourth, although interest in Macey and Schneider's (2008) process model has sparked recent attention (Culbertson et al., 2012), each component of their model is also defined with considerable construct confounding. For instance, state engagement includes involvement, commitment, empowerment, and satisfaction. Hence, Kahn's perspective is most relevant to the current study purpose, namely because of its distinctiveness from other constructs.

Theoretical inconsistencies regarding stability. Although Kahn's (1990) framework provides the best place to start, his model is not completely indicative of a moment-to-moment state. Specifically, his assertion of engagement being moment-to-moment relies on both theories of fluctuating states and also those suggesting stability. First, Kahn used Alderfer's existence-relatedness-growth (ERG) theory (Alderfer, 1969). Engagement antecedents - safety, meaningfulness and availability - roughly relate to Alderfer's growth, relatedness, and existence needs, respectively. Alderfer additionally included a principle called frustration-regression, which essentially states that individuals regress to previous positions in the model when their current needs are unsatisfied. Thus, they regress from attempting to satisfy their relatedness

needs back to attempting to satisfy their existence needs. This principle implies that individuals are constantly changing their efforts to fit the scenario. Kahn used ERG theory when discussing the antecedents to engagement. ERG is a needs satisfaction theory proposing that people have the ability to progress and digress to meet their needs, indicating that as needs go fulfilled or unfulfilled, motivation fluctuates. Similarly, Kahn posited that the presence or absence of antecedent conditions could lead to engaged or disengaged employees. Specifically, Kahn stated that individuals need to feel psychologically safe, that their work is meaningful, and that they are psychologically available, all at the same time in order to become engaged. This also means that employees' motivational states are subject to change depending on the scenario – or more specifically the fulfillment or lack thereof these critical antecedents. Hence, ERG formed the fluctuating bases of Kahn's engagement.

Second, Kahn (1990) relied on the Job Characteristics Theory (JCT; Hackman & Oldham, 1980) to develop his framework. In his seminal study on engagement, one of Kahn's goals was to evaluate the conditions that were necessary for engagement. While this could be interpersonal relations or needs satisfaction, he also recognized the importance of contextual job factors. The JCT is a model that proposes specific characteristics of the workplace lead to motivation. Thus, Kahn used the JCT to hypothesize that certain contextual aspects of the workplace lead to critical psychological states, which in turn result in affective and behavioral outcomes. The contextual aspects of work defined in the JCT are referred to as core job dimensions, which include skill variety (i.e., the necessity for the use of multiple skills in one's job), task significance (i.e., the value of the task to the organizational goal), task identity (i.e., an employee's felt internal identification with the task), autonomy (i.e., the ability to control one's work activities, plans, or goals), and feedback (i.e., timely and constructive feedback on job

activities and performance). The critical psychological states of the JCT refer to an employee's experienced meaningfulness, experienced responsibility, and their knowledge of results. Note the similarities between Kahn's (1990) meaningfulness, availability, and safety as psychological conditions antecedent to engagement, a motivational state. Importantly, the JCT is a job redesign strategy, which emphasizes lasting changes to the workplace to promote long-term changes in motivational level, and ultimately satisfaction and performance. As with most job redesign strategies (Morgeson & Campion, 2003), it is reasonable to expect that these core job characteristics do not change from day to day. Hence, the JCT makes up the more stable derivative of Kahn's (1990) engagement.

Based on the theoretical framework he derived from ERG, role theory (Goffman, 1959), and the JCT, Kahn used qualitative methods to understand people's investment of themselves at work and found variability in engagement levels based on participants' recollections of instances when engagement was high or low. Kahn (1992) outlined critical aspects of what should be looked at in engagement research, one of which was to assess "particular moments of role performances and people's immediate experiences and behaviors in those moments" (p. 343). Based on his results, Kahn (1990) defined engagement as a moment-to-moment phenomenon, even though his theoretical backing was based on the fluctuations inherent in ERG theory and stability of the JCT. Additionally, the distinction between engagement being a moment-to-moment state, and a lasting trait is ambiguous in Kahn's own recent accounts of engagement. Kahn and Fellows (2014) suggest that engagement can be "fleeting", or it can be a "steady state, punctuated by interludes of relative disengagement" (p.106). Thus, within Kahn's conceptualization there are discrepancies with how engagement is temporally described. I note

though, that the majority of Kahn's work conceptualizes engagement as a moment-to-moment state (Kahn, 1990, 1992; Kahn & Fellows, 2014).

Kahn's (1990) framework is not the only one that seems to combine foundations of variability with stability. As previously noted, the most researched framework for assessing engagement is Schaufeli and colleagues' (2002), in which they likened engagement to a lasting affective state. Schaufeli and colleagues (2002) created similar inconsistencies in their theoretical suppositions as Kahn. In particular, Schaufeli et al. conceptualized engagement as the opposite end of a spectrum shared with burnout, which is a *lasting* state of fatigue, cynicism, and reduced efficacy (Maslach et al. 2001). It seems that by placing engagement as the trait-like theoretical opposite of burnout, but including dynamic (fluctuating) aspects of the construct in the definition (e.g., vigor), Schaufeli et al.'s conceptualization of engagement also suffers from contradictory shortcomings.

Regardless of the model in which a researcher chooses to frame engagement, a lack of clarity around the conceptualization of engagement's stability exists and needs resolution. Of late, researchers have begun focusing on the issue of stability, by conducting longitudinal and experience sampling studies to look at how engagement changes over time.

Recent stability research. Researchers have begun using daily diary studies – sampling employees daily – to gain perspective on how an employee has changed from day to day (Ohly, Sonnentag, Niessen, & Zapf, 2010). This daily diary research has led to findings showing engagement is positively related to optimism, transformational leadership (Tims, Bakker, & Xanthopoulou, 2011), and daily coaching on the job (Xanthopoulou, Bakker, Demerouti, & Schaufeli, 2009). One of the issues with these daily diary studies, however, is that they are

conducted only once per workday. As Bakker (2014) notes, this provides valuable information across days, but fails to address fluctuations within days.

Although the results from these studies provide interesting first insights into fluctuations in engagement levels, there are multiple ways of looking at the within-person variations of employee engagement that may assess the within-day fluctuations. For instance, one could assess fluctuations using only a few time points over long spans of time (e.g., Brauchli, Schaufeli, Jenny, Füllemann, & Bauer, 2013 only used four data points per participant). The variability could be assessed using a single time-point in one day, which requires participants to reflect on their experiences during the day as a whole (Culbertson et al., 2012). Researchers can also use experience sampling, tracking employee attitudes, behaviors or characteristics across a multitude of time points (Bolger & Laurenceau, 2013), essentially capturing several data points per day for each participant. Statistically, fluctuations can be evaluated using any of these research designs.

Some of these aforementioned research designs are more efficacious than others in producing meaningful results. For example, gathering data over only a few time-points can allow for the assessment of the within- and between-person engagement. However, as researchers we must ask ourselves whether it makes sense to consider a fluctuation in engagement levels over a three year time period (e.g., Brauchli et al., 2013) a meaningful within-person difference. We must also ask ourselves how accurate an employee's recollection is if there is one daily time collection only, which typically falls after the employee has left work (e.g., Culbertson et al., 2012). Robinson and Clore (2002) purported that on self-report measures, timing is key because memories can become skewed as time passes between the actual event and the measurement of the event. This is particularly salient for emotionally driven situations (Robinson & Clore, 2002). Given the previously established relationship between engagement and affect (both positive and

negative; Bledow, Schmitt, Frese, & Kühnel, 2011; Steger, Littman-Ovadia, Miller, Menger, & Rothmann, 2013), it is logical to conclude that reducing this fallacious recall in any way possible is beneficial. Experience sampling at multiple times within each day attempts to do just that (Bolger & Laurenceau, 2013).

Recently researchers have begun to look at multiple data points within a single day to assess engagement levels. For example, Ouweneel, LeBlanc, Schaufeli, and van Wijhe (2012), examined engagement twice per day and found that fluctuations in engagement levels were positively related to positive emotions and hope from the previous day. Ouweneel et al. also found that all three dimensions of engagement as defined by Schaufeli et al. (2002) varied substantially from day to day. Another example of experience sampling is Bledow, Schmitt, Frese, and Kühnel's (2011) examination of engagement and affect. The researchers sampled participants twice per day, finding a positive shift in affectivity was accompanied by an increase in engagement.

Thus, examining these daily fluctuations in employee engagement reveals they exist (Bakker, 2014), and that within-person variation is meaningful and related to various aspects of an employee's well-being (Culbertson et al., 2012). Hence, the use of experience sampling methodologies that sample participants several times within the same day provides important information on fluctuations within days. In contrast, the use of one survey per day cannot provide this within-day variability, nor can the use of a few spread-out time-points over a longer period (i.e., Brauchli et al., 2013).

Xanthopoulou et al. (2012) and Bakker (2014) both provide comprehensive qualitative reviews of the literature on fluctuations in engagement, concluding that more research in this

area is needed. The current study answers this call for additional research clarifying the stability of engagement, specifically testing Kahn's (1990) conceptualization.

Theoretical Underpinnings of the Current Study

Based on my review of the literature on engagement, I propose a refined conceptualization of engagement that suggests it is a within-person moment-to-moment, fluctuating motivational state, that also possesses a degree of between-person stability. That is, although employees may experience fluctuations in their engagement levels throughout the day (and potentially across a day or two) indicating within-person variability, over time we can see enough relative stability that between-person comparisons appear meaningful. Most current research findings, aside from recent studies using diary and experience sampling methods, are based on engagement as a stable construct and these findings are compelling (e.g., Chalofsky & Krishna, 2009; Moliner et al., 2008; Rich et al., 2010). Although Kahn (1990) and Schaufeli et al. (2002) framed their definitions on both theories of stability and fluctuation, each took a stand as to what engagement is – either moment-to-moment or stable state. Perhaps what comes closest to reality is Kahn and Fellows (2014) recent supposition that engagement may be transitory, and yet still be fairly stable over time. I build my hypotheses in two stages: (1) to test the fluctuations of engagement and (2) to test its stability, to determine how fleeting or steady engagement is.

Fluctuations in engagement. Previous research has supported the notion of substantial variations in engagement levels within employees (Bakker, 2014; Bledow et al., 2011; Brauchli et al., 2013; Culbertson et al., 2012; Xanthopoulou et al., 2012). Consequently, researchers should not only look at these fluctuations across days, weeks, and months, but also *within* days,

to gain a rich perspective on whether engagement may be primarily a moment-to-moment phenomenon, or a more stable construct changing slowly over time.

Given the emotional component of employee engagement, it is salient to look at the literature on affect for models of mapping these fluctuations. The influx of experience sampling methodology, introduced with research by Larson and Csikszentmihalyi (1983), opened the door for research into the moment-to-moment fluctuations in affect. Specifically, Weiss and Cropanzano (1996) used this concept of momentary fluctuations measureable within an experience sampling methodology to propose that emotions vary as affect-driven events occur. Due to situations that arise throughout the day, these event-driven emotional fluctuations embody the assumption that our emotions vary widely during any given workday. Consequently, research on affect and mood using experience sampling has relied on Weiss and Cropanzano's Affective Events Theory (AET), as a conceptual basis for within-day variability in affect (Miner, Glomb, & Hulin, 2005). For example, Bledow and colleagues (2011) framed engagement in terms of the affective shift model. This model is similar to AET in that it looks at fluctuations of affect within the day. The affective shift model posits that as negative affect is experienced, people attempt to alleviate it and consequently switch their state-affect to a positive one (Bledow et al., 2011). Bledow et al. found the shift to positive affect was also a predictor of engagement. As various affectively charged events may occur throughout the day, it is efficacious to survey multiple times during one day to fully capture these fluctuations. Given the affective component of engagement, it makes sense that a shift in affect also signals a shift in engagement. This supposition is consistent with Macey and Schneider's (2008) proposition that state engagement may mirror fluctuations in affect. Therefore, I hypothesize fluctuations in engagement that represent significant within-person variability.

Hypothesis 1: Engagement levels show significant within-person variability.

Stability in engagement. Kahn (1990) drew inspiration from theories of motivation when developing his construct definition and theory of employee engagement. Specifically, he focused part of this development on JCT. Job characteristics can be seen as relatively stable from day to day. For instance, skill variety and role autonomy probably do not vary greatly within each work week. Additionally, several of the work and affective outcomes in JCT have been highly correlated with engagement, such as satisfaction, productivity, and absenteeism (negatively for the latter; Kamalanabhan, Prakashsai, & Mayuri, 2009; Rich et al., 2010; Schaufeli et al., 2009). Consequently, one could draw the conclusion that different jobs foster different levels of engagement; meaning that engagement can depend on the job or task.

Engagement has an affective aspect to it (*emotional*, cognitive, and physical). Affect can be seen several different ways, but trait affect is a stable construct. Thus, it is natural to assume there will be at least some dispositional factors affecting engagement, leading to the conclusion that individuals may differ from each other in their overall engagement levels. Specifically, though there may be within-person fluctuations in engagement, a significant portion of the variation seen in engagement levels may be due to between-person differences, based on differing dispositional and stable contextual factors for each person. Furthermore, Kahn (1990) discussed individual differences in terms of personal perceptions. He stated that since the antecedents to engagement (safety, meaningfulness, and availability) are primarily subjective, it would make sense that an individual's dispositions and personal life experiences would dictate how he or she perceives these antecedents. These individual differences in perceptions suggest there exist between-person variability in engagement levels, representing some construct stability. Therefore, I hypothesize:

Hypothesis 2: Engagement levels show significant between-person differences.

Related variables. Understanding the fluctuations in engagement may be further informed by understanding how the antecedents of engagement relate to engagement – does engagement fluctuate with its antecedents? Kahn (1990) proposed three antecedents to engagement: psychological meaningfulness, psychological availability, and psychological safety. In this study, I use work tasks that are considered meaningful as a proxy for psychological meaningfulness, strain as a proxy for psychological availability, and quality of coworker interactions as a proxy for psychological safety.

Psychological meaningfulness. Kahn (1990) discussed the critical part that the work tasks play in creating meaning in one's role, suggesting that certain work tasks, and characteristics of those work tasks, lead to felt meaningfulness (Kahn, 1990, 1992). For instance, Kahn and Fellows (2014) suggested that tasks linking employees' roles to the values and goals of the organization are more likely to lead to meaningfulness and then engagement. The authors also suggested that challenging job tasks foster engagement. Job redesign strategies aim to give workers more responsibility, but also roles that require greater skill variety. An increased need for differential skills indicates the worker is switching through a variety of tasks. While a worker may derive great meaning from certain tasks due to their challenging nature, responsibility, and direct ties organizational goals, not all tasks in a worker's day may be structured this way. In an academic example, a professor may enjoy researching, mentoring, and teaching, however the tasks of applying for departmental funds, or filling out applications for classroom textbooks may not be as meaningful. Job tasks can also differ throughout the day. Though employees may take meaning from one aspect of their job, they may have to switch to another task that is tedious and time consuming.

Hypothesis 3: Fluctuations in engagement within employees are significantly positively related to work tasks that are considered meaningful, such that high levels of employee engagement positively relate to work tasks that are considered meaningful.

Psychological availability. Based on AET, Fuller, Stanton, Fisher, Spitzmüller, Russell, and Smith (2003) used an experience sampling methodology to map out daily fluctuations in stressors and strain. Stressors represent the cause of stress, and strain represents the reaction to these stressors (Lazarus, 1993). Although job stress has been correlated with affect (Spector, Zapf, Chen, & Frese, 2000), scholars have debated the value of this correlation. Specifically, some purport that one should statistically control for affect in the relationship between stressors and strains (Payne, 2000), whereas others argue that in doing so valuable information is lost (Judge, Erez, & Thoreson, 2000; Spector et al.). The Job Demands-Resources (JD-R; Demerouti, Bakker, Nachreiner, & Schaufeli, 2001) model explains the variability in strains due to stressors that is unrelated to affect and that directly applies to this study on engagement.

The JD-R model proposes that employees apply resources to avoid the strain caused by stressors, which are otherwise known as demands. As these stressors vary from day to day, so do the resulting strains (the process of stressors leading to strains is collectively called stress; Lazarus, 1993). In support, daily fluctuations in stressors have been shown to result in daily fluctuations in strains (e.g., Fuller et al., 2003; Zohar, 1999). When applying the JD-R to engagement, adequate resources result in engagement, whereas too many demands take away from the ability to become engaged (Schaufeli & Bakker, 2004). Thus, as employees expend resources to handle demands and avoid the resulting strains, they lose resources that would otherwise lead to or maintain engagement levels. Thus, fluctuations in job stress could lead to fluctuations in engagement levels, due to the stress process usurping resources that would

otherwise be allocated to engagement. Such fluctuating resource attainment and depletion result in fluctuating levels of psychological availability.

In addition, social psychology literature sheds light on behavioral and volitional constraints that may play a role in perceived psychological availability. Specifically, people have a limited amount of cognitive resources, including conscious control of thought (Baumeister, Bratslavsky, Muraven, & Tice, 1998). The notion that employees can be continuously cognitively stimulated without rest does not fit with the empirical evidence of people's actual cognitive capacities. Given these limited attentional resources (Baumeister et al., 1998), it would be fallacious to think that employees have a never-ending supply of cognitive capital to have the psychological availability necessary for constantly sustained high levels of engagement. Consequently, engagement levels must fluctuate during the day along with natural fluctuating cognitive capital. This also means that because people have a limited cognitive capacity, increasing demands could deplete an employee's psychological availability. Seen from the perspective of the JD-R model, if employees are expending resources to deal with stressors (demands), they will have fewer resources to become engaged at work. For instance, Schaufeli and Bakker (2004) found that engagement was directly related to the presence of resources. Should these resources be depleted, the employee may experience less psychological availability, and then according to Kahn (1990, 1992) will be less engaged.

Based on the JD-R model, literature on Baumeister et al.'s (1998) findings that we have a limited amount of consciously asserted energy, and previous findings of availability predicting engagement (Rich et al., 2010), I expect that psychological availability and engagement will be highly related at the moment-to-moment level.

Hypothesis 4: Fluctuations in engagement within employees are significantly inversely related to fluctuations in stressors, such that moments of high strain perceptions correspond with low levels of engagement.

Psychological safety. Kahn (1990, 1992) referred to psychological safety as employees' experiencing a climate of trust that allowed them to be themselves and invest in their work roles without fear of negative criticism. Employees feeling psychologically safe could speak up or challenge policies or actions, with little fear that such challenges would result in career sabotage. When relations with coworkers are deemed positive (i.e., supportive or trusting), psychological safety is felt by employees (Kahn, 1992). Rich et al. (2010) found that supervisor support, serving as a proxy for safety, was positively related to engagement. Depending on the work role though, individuals may have just as much, if not more interpersonal interactions with fellow job incumbents than they do with their supervisor. Additionally, not all of these coworker relations are equally positive. For example, in the Q12 measure of engagement (Harter et al., 2002) there are questions indicating that we think of our interactions with employees as singular, and these singular interactions may make a large difference for engagement. One question states "I have a best friend at work" and another states "My supervisor, or *someone* at work, seems to care about me as a person" (Harter et al., 2002, p. 269). Thus, individual employee interactions may have a large effect on engagement. Importantly, the notion of individual employee interactions reflects the idea that employees interact with various coworkers throughout the day. Moreover, these interactions may vary in terms of their quality. For example, an employee could interact with their 'best friend at work', and come away energized and supported. Later, that same employee may interact with a peer in a negative manner and walk away feeling drained, unjustly treated, and unsafe to fully express themselves in their work-role. Positive coworker interactions may

increase employee engagement as a function of employees feeling it is safe to assert their full selves in their work role at that time. Indeed, May et al. (2004) found that rewarding coworker relations were positively related to perceptions of psychological safety. Therefore, I hypothesize that engagement is positively related to the quality of coworker interactions, which vary from day to day indicating a fluctuating psychologically safe space.

Hypothesis 5: Fluctuations in engagement within employees are significantly positively related to the quality of coworker interactions, such that high levels of employee engagement are positively related to high quality coworker interactions.

State-affect. Engagement is the simultaneous exertion of *emotional*, cognitive, and physical resources into one's work role; thus, a portion of experiencing engagement is affective. However, research relating engagement to affectivity shows mixed findings (Bledow et al., 2011; Steger et al., 2013). For instance, contrary to what one might expect, Bledow and colleagues found that negative affect may be a precursor to engagement; negative mood and affect in the morning led to *higher* engagement levels in the afternoon. Bledow et al. used an experience sampling longitudinal study where employees were surveyed each day in the morning and afternoon for nine days. Framed using the affective shift model, Bledow et al. concluded that the presence of negative affect could signal a need for a shift in disposition, and the subsequent shift to a higher positive affect in the afternoon triggered higher levels of engagement later in the day. It is important to note that when looking at affect and engagement simultaneously, Bledow et al. found that positive affect was related to higher engagement, and that negative affect acted as an antecedent to this relationship. In contrast, in a cross-sectional study framed using the JD-R model, Steger et al. (2013) found that work engagement was only correlated with high levels of affectivity (i.e., positive affect). One reason for the discrepancy in findings between these two

studies could be differences between cross-sectional vs. longitudinal methodology. Bledow et al. were able to model changes in affect and engagement over multiple time points, whereas Steger and colleagues' findings mirrored a single point in time such as the latter findings from Bledow et al.

Based on the AET, affective shift model, and Kahn's (1990) conceptualization, I expect engagement to be positively related to momentary positive affect, and negatively related to momentary negative affect.

Hypothesis 6a: High levels of engagement are positively related to high levels of momentary positive affect.

Hypothesis 6b: High levels of engagement are negatively related to high levels of momentary negative affect.

Trait-affect. Although affect has been examined as a moment-to-moment state, represented by fleeting emotions (Miner et al., 2005; Weiss & Cropanzano, 1996), it has also been looked at as a stable, enduring trait that characterizes a person's lasting disposition (Watson, Clark, & Tellegen, 1988). I previously hypothesized a relationship between engagement and moment-to-moment affect (Hypothesis 6), but it is efficacious to also discuss the relationship between dispositional affect and engagement. The intra-class correlation coefficients (ICCs) seen in research examining fluctuations in engagement have shown that engagement is highly variable within people (Bledow et al., 2011; Culbertson et al., 2012; Tims et al., 2011). The ICC, expressed as a ratio in decimal form, is a measure of the within-person differences to between-person differences effects on the variability of a construct (e.g., an ICC of .3 indicates that 30% of the variability seen in the construct is due to between-person differences, and that 70% of the variability is due to within-person differences). Yet, the ICCs in previous

studies are over .2, which indicates that at least 20% of the variability seen in employee engagement is due to between-person differences. This implies that individuals have somewhat differing average levels of engagement, despite within-person fluctuations. Given that affect can be trait-like, it may play a role in previously observed between-person differences. For example, Dalal, Baysinger, Brummel, and LeBreton (2012) found that engagement was highly correlated with positive trait-affect ($r = .66$). The relationship between engagement and trait-affect could be explaining some of the observed variance between people. Consequently, I hypothesize that engagement is a somewhat stable construct such that meaningful between-person variability exists and correlates with a stable construct like trait-affectivity.

Hypothesis 7a: High levels of engagement are positively related to high levels of trait positive affect.

Hypothesis 7b: High levels of engagement are negatively related to high levels of trait negative affect.

Job satisfaction. Previous studies have observed a moderate to high positive relationship between engagement and job satisfaction (Harter et al., 2002; Rich et al., 2010). The moderate correlations between engagement and job attitudes, such as commitment, satisfaction, and involvement (e.g., Harter et al., 2002) have prompted some scholars to argue that engagement is not a new construct (Newman & Harrison, 2008). Instead, they have argued that engagement is a higher-order latent construct that represents a combination of several attitudes. Although researchers have responded to this claim with empirical studies demonstrating the distinctiveness of engagement from these attitudes (Christian et al., 2011; Rich et al., 2010; Shuck, Ghosh, Zigarmi, & Nimon, 2013), there is no denying that engagement is positively related to job attitudes like satisfaction, commitment, and involvement. These moderate correlations suggest

potential common underlying constructs or conceptual overlaps and could be expected given that one dimension of engagement is affect, also shown to be positively correlated with the same work attitudes.

Job satisfaction is considered a relatively stable attitude that is not prone to daily fluctuations (Shuck et al., 2013; Staw & Ross, 1985). Previous work looking at job satisfaction and engagement as stable trait-like variables has resulted in a moderate correlation between the two constructs ($r = .53$; Christian et al., 2011). Since the two are related on a stable temporal level, it follows that they will be related at a more dynamic temporal level (i.e., moment-to-moment). However, although engagement may fluctuate significantly within employees, there still may be between-person differences in mean engagement levels. The positive correlation between job satisfaction and engagement suggests that employees who have high job satisfaction should also have a higher mean engagement than employees with low job satisfaction.

I posit that mean engagement over time will be related to job satisfaction. The previously supported relationship provides initial justification for this assumption. Also in support, engagement is partially framed in the JCT and a confirmed outcome of the JCT is job satisfaction. It turns out that job satisfaction is one of the most empirically supported outcomes of the JCT, a motivational model (Loher et al., 1985). Job satisfaction is positively correlated with critical psychological states (i.e., meaningfulness, and responsibility; Loher et al., 1985), and therefore, when looking at a lasting stable attitude such as job satisfaction (Staw & Ross, 1985), it follows that consistent levels of motivation are closely associated. Thus, the stable between-person aspects of engagement (in this study the mean level of engagement) are hypothesized as the motivational force that is associated with positive stable levels of job satisfaction.

Hypothesis 8: Mean engagement levels within employees positively correlate with job satisfaction.

Summary

Theoretically, engagement has been conceptualized as a moment-to-moment state (Kahn, 1990), a lasting affective state (Schaufeli et al., 2002), or some combination of the two (Macey & Schneider, 2008). The current study posits that while there are significant between-person differences in engagement, the majority of the variability is due to within-person fluctuations. To test these hypotheses, I used an intensive longitudinal data collection in which employees responded to momentary surveys on the variables of interest.

Method

I used experience sampling methodology over five weeks to assess engagement at various times during the day. Additionally, at baseline data collection points, I assessed several other constructs to see how they vary over time along with engagement. Collecting data on these other constructs will allow me to show how other variables may be affecting or not affecting engagement levels (i.e., affect, stressors, and strain). Having a personalized look into employees' well-being could be important for explaining fluctuations (or lack-there-of) in their personal engagement levels.

Participants

Forty nine (final) participants were employees at two different organizations in two separate industries. Twelve participants came from a technology focused organization based on the East Coast. Thirty seven participants came from a large health care provider in the Pacific Northwest. Four hundred and eighty nine employees were initially contacted to complete the baseline surveys, of which 146 completed them, representing a 29.9% response rate for the baseline. Of these 146 participants, 134 responded to the moment-to-moment surveys. Fifty total employees completed 15 or more of the daily surveys, making them eligible for inclusion in the statistical analysis (one participant was then removed due to invariant responding on all of the moment-to-moment response variables) for a retention rate of 36.6% on the daily surveys.

Participants were 72.1% female, had a mean age of 41.02, with an average job tenure of 3.67 years, worked full-time (93%; 7% worked between 20 and 40 hours per week), and were primarily Caucasian (83.7%; 11.6% were multi-racial). A variety of job types and levels were represented by the participants. At the technology firm organizational level was assessed and

30% were mid-level managers, 60% were in a non-management role, and 10% were upper-level management, at the health care firm job type was collected and medical assistants made up 39.5% of the sample, patient services representatives were 21.1%, administrative assistants, disease management support specialists, practice managers, and RN's each made up 5.3% of the sample, and coding specialists, executive assistants, health information specialists, care management managers, physician practice managers, nursing clinic supervisors, and surgery schedulers accounted for an additional 2.6% of the sample each. As employees came from different job types, organizations, and industries, demographic characteristics were assessed as potential covariates in the analyses to ensure their use as one sample was appropriate.

Procedures

Experience sampling methodology. Researchers have used daily diaries to assess engagement in multiple studies (Ouweneel et al., 2012; Bakker & Xanthopoulou, 2009). Ohly, Sonnentag, Niessen, and Zapf (2010) suggest the use of diary studies to look at fluctuations in individual variables because this methodology allows employees to fill out surveys after work hours, and provide a daily view of the focal study variables. The authors also noted that daily diary studies belong to a family of methodologies looking at multiple within-person data points. Based on the review and recommendations by Ohly et al., this family of methodologies served as the basis for conducting the current research.

Daily diaries are a form of intensive longitudinal methods (Bolger & Laurenceau, 2013). Within this family, experience sampling has grown in use and acceptance. Experience sampling methodology entails surveying employees on a daily basis or more. When studying engagement, the goal is to see exactly how people are feeling while at work, so sampling multiple times per day in the workplace is the logical approach. Sampling “in situ” (Hormouth, 1986, p. 1)

heightens the external validity of the results of the study, but it also allows for an in-the-moment view of the work variables being studying.

Within experience sampling there are several different types of research designs. Researchers can choose between event- or signal-contingent reporting (Bolger & Laurenceau, 2013). Event-contingent reporting asks participants to respond to a survey, each time a behavior, mood, or event occurs. For example, regarding engagement, an event-contingent design would ask participants to actively recognize their engaged state and *then* report. Although this approach to data collection is valuable, it can lead to a reliance on the participant's own knowledge of exactly when these states are occurring and subsequently interrupt that state to take the measure.

Signal-contingent reporting relies on signaling employees at random times via some mechanism such as a smartphone, alarm, or email, to let them know it is time to report their *current* state (be it engaged or otherwise) at that moment. This approach provides a wider representation of states reported over time, thus limiting range restriction in responses.

Another advantage of using signal-contingent experience sampling can be seen when comparing that methodology to the use of daily diary studies that rely on data collected once per day after work hours. The problem with this single point post-work daily diary study is that it requires employees to accurately recall events from their day, potentially hours after they occurred. Tversky and Kahneman (1982) demonstrated some of the problems with conscious recollection, and how easily memories can be affected. In particular, the authors' work focused on eye-witnesses of car crashes, and how easily the recall of objective facts (e.g., car colors were incorrectly stated, and even crash culpability was incorrectly recalled) can be skewed by retrospective recollection. However, a signal-contingent methodology would reduce any recollective biases due to the simultaneous nature of recognition of the construct level within

one's self, and response to the alarm thereby allowing these attitudes to be sampled as they are actually happening (Tversky & Kahneman, 1982).

The current study's procedure is based on previously utilized experience sampling methodologies. For example, Miner, Glomb, and Hulin (2005) provided cellphones that alerted participants to complete surveys at various times of the day. A similar methodology was used by Fuller et al. (2003), where they prompted participants to respond within an hour to a signal to complete online surveys. Both groups of researchers used signal-contingent methodologies and required multiple data points per day for participants. Miner et al. (2005) introduced the efficacy of using randomized time points to account for any differences in time of day, or day of the week. The rationale was that by randomizing time of day, measurements of engagement levels would not be based on the mornings or afternoons, rather it represents fluctuations in engagement regardless of time of day. The same could be said for days of the week. The current research also randomized times that participants were surveyed within a day, as well randomizing the three days within a work-week in which participants completed multiple surveys.

Pilot test. I conducted a pilot test to assess the efficacy of my methodology prior to completing the reported study. Due to the intensive nature of the sampling, as well as the cost to the researcher and company, correctly tailoring the methodology before the start of the official study was crucial. Given that the goal of the pilot test was to assess the methodology and not collect actual participant data, research regulatory board approval was not obtained.

My two major questions in the pilot test were: What is the most effective medium to use to survey people multiple times per day over multiple days? And, what will motivate participants to complete at least 80% of the surveys they are sent? I created experimental conditions to

answer these questions. I assigned participants to take the surveys on either their phone, or their computer. They were instructed that they could take the surveys on their assigned device only. Additionally, I wanted to see if a reinforcement schedule for rewards would make a difference in participation rates, a ploy designed to obtain an 80% or more response rate. In my “fixed” condition, participants were told they were able to make \$10 if they completed 80% or more of their surveys. In my “variable” condition, participants were told that they were able to make up to \$10 by completing the study, and would be rewarded after some of the surveys. The fixed condition participants saw a ‘thank you for completing the survey’ message after each survey. The variable condition saw a ‘thank you for completing the survey, you have just earned \$1’ after the surveys, or just simply a ‘thank you for completing the survey message’. By changing the message after a certain number of completion times, I created a variable reward system. These two sets of conditions were combined to make four different conditions. That is, participants were grouped into phone-fixed, phone-variable, computer-fixed, or computer-variable conditions.

Pilot test participants. I recruited 22 graduate students from Colorado State University’s department of psychology. One graduate student was recruited from Florida International University’s law school. Although I personally knew each participant, invitations to participate were sent out via a graduate student list-serve to mimic the final study. I created randomized four digit identification numbers for each participant so that I could track their data collection over time without tying identifying information to the data. Based on a random number generator, the participants were placed into one of the four conditions.

Pilot test procedure. Participants completed within-day surveys over two weeks. Thus, participants were surveyed twice a day, for three days of the workweek, for two consecutive

weeks. These daily surveys had a time limit for completion of 1hr for the first week, and 1.5hrs for the second week. At the end, participants were given a final open-ended survey assessing issues with the methodology. I wanted participants to complete 80% of the surveys to mimic my desired completion percentage in the actual study. So, if participants completed 10 out of 13 of the surveys, I considered their completion rate at 80% (the actual percentage is lower, but given the small number of surveys, this seemed to be a better option than going with 11 out of 13).

Overall, I had an attrition rate of 48%, meaning that of the original 23 participants, 12 completed 80% or more of the surveys. There were no differences in completion rates between the variable and fixed groups, or the computer and phone groups. One participant of the 23 contacted me partway through the study to tell me that issues external to the pilot study prevented participation.

I received 20 final open-ended surveys, and gained valuable insight into participant perceptions of the methodology. The biggest issues were: a lack of checking email for the reminders of the survey, the inability to switch between taking surveys on the phone or on the computer, and a misunderstanding of some of the academic language. To address these issues participants in the final study were allowed to use both their computers and phones, increasing flexibility. I limited participants to one or the other in the pilot test to ensure participation on both media. Finally, through the use of explanatory emails from me and the HR managers at both organizations, I was in closer touch with the actual participants of the study to describe what the measures mean, and changed some of the wording in the surveys to be more clear and concise.

The last issue I faced from the pilot study was that the majority of participants said money was not a motivating factor for completing the study. The problem here is that if \$10 is not motivating for graduate students working on a limited yearly stipend, then it will probably

not be motivating for full-time working professionals who typically make more than students. To address this challenge I incentivized participation through the use of a lottery system. If the participants in the final study completed the necessary number of surveys, he or she was entered into a lottery to win one of several cash prizes. Given the financial constraints of the study, this technique allowed participants to potentially make a meaningful amount of money for the completion of these surveys, yet kept the total research costs within a \$600 grant fund for the study and prevented perceptions of coercion with higher award amounts.

Final Study Procedures. Participants were recruited from two organizations. I collaborated with the HR manager at the technology firm, and an organizational development (OD) specialist at the health care firm. The OD specialist recruited business unit leaders to reach out to their employees for an employee engagement effort. The HR manager and business unit leaders contacted employees in their departments about a survey effort to aide in their organization's employee engagement assessments. They informed employees – briefly – of the nature of the study, and to expect an email from me. I reached out to the potential participants in an initial email, introducing both myself and the project, letting them know that at the start of the next week I would be sending a baseline survey to them that would take around 20 minutes to complete. Employees were informed that full participation in these surveys would make them eligible for a lottery-style cash drawing. Due to organizational concerns, the monetary reward varied between the two organizations. At the technology firm, employees were entered into a drawing to win one of two \$100 cash prizes. At the healthcare firm, employees were entered into a drawing to win one of twenty \$30 cash prizes.

A graphical depiction of the data collection is displayed in Figure 1. All 489 potential participants received the baseline survey. This survey contained measures of trait affect (positive

and negative), general engagement levels, the meaningfulness of their work tasks, and overall job satisfaction. The survey concluded with demographic questions on age, gender, ethnicity, tenure with organization, supervisor, and job, job level, and employment status (i.e., part-time vs. full-time). Employees then created a unique identifier that was used to link their moment-to-moment surveys to each other.

Employees who completed this baseline effort were emailed within-day surveys. These surveys occurred twice per day (one in the morning, the other in the afternoon), three days per week, for five total weeks. To boost participation, an hour and a half window to complete these surveys was given. Surveys were never be closer than two hours apart (from finish of the first, to start of the next).

Employees at the technology organization all worked Monday to Friday from 9:00am to 5:00pm. Each of these employees were sent the surveys at the same time, on the same days as each other. However, at the healthcare organization, employees worked variable hours, and variable days as many were shift workers. Thus, in the baseline these participants completed a scheduling survey, indicating whether they had a fixed or variable work schedule. For employees with a variable work schedule, they were asked to complete a scheduling survey at the start of each week to indicate their new work schedule. In addition, these employees received surveys at different times and days as each other to reflect their varying schedules and work hours. In the daily surveys, participants completed momentary stress, engagement, affect (positive and negative), quality of coworker interactions, and task meaningfulness inventories, along with two optional open-ended questions. Fuller and colleagues (2003) found these open-ended responses to be valuable sources of alternative explanations for responses. These open-ended responses also provide an intriguing window into possible affective events throughout the day.

One week after the conclusion of the daily surveys, employees completed a follow up longer survey, measuring the same variables as the baseline effort. Participants were given contact information to gain access to results of the study or express concerns over the study. One reason for using these baseline and follow-up data collection points is that it provides data that can be related to both trait and state engagement, similar to the methodology from George's (1991) study on mood. George found that state mood and trait mood had different relationships with pro-social behaviors. Once data were compiled, and recipients of the cash prizes determined, data were de-identified and linked by the unique identifiers employees provided (for clarity, in the discussion of individual results in this manuscript, participant IDs were re-numbered to be 1-50).

Measures

Reliabilities for baseline scales are represented by alphas (α) from the baseline measurement occasion. To assess the reliability of the moment-to-moment scales, I used a technique developed by Cranford, Shrout, Iida, Rafaeli, Yip, and Bolger (2006), that decomposes variance in the scales to within- and between-person levels and gives an estimate of internal consistency. Thus, it will be referred to as within- and between-person α in this section.

Demographics. Demographic information was collected in the first data collection session. Participants were asked to voluntarily report their sex, age, job tenure, organizational tenure, tenure with supervisor, job level, job type, and ethnicity.

Perceived work stress. Stress was assessed on the daily surveys using a 6-item Stress in General scale (SIG; Stanton, Balzer, Smith, Parra, & Ironson, 2001). This original scale was shortened (Fuller et al., 2003), and studies have reported adequate reliability of scores for both the 3 item-subcales ($\alpha = .70$, $\alpha = .71$). However, in this study I used an alpha that reflects the

full scale as the scale was represented by one factor. The sample reliability estimates in this study were $\alpha=.67$ at the within-person level, and $\alpha=.85$ at the between-person level. The SIG scale asks participants to respond to “Indicate how much each of the following words/phrases describes your job situation today in the past 3 hours”, and rate each word or phrase from 1 = *not at all* to 5 = *definitely*. Example words (see Appendix A for the full survey) are “relaxed”, “calm”, and “pushed.” Accumulated validity evidence for the SIG is shown through correlations between SIG scores, and scores on stressor measures, physiological stress indicators, and general job attitude measures (Stanton et al., 2001).

Affectivity. Affectivity was assessed using the positive and negative affect scale (PANAS; Watson et al., 1988). Reported reliability of scores has previously ranged from .84 to .90 (Watson et al., 1988). It should be noted that discrepancies in observed reliability occurred due to changes in the instructions for participants (i.e., consider the past month, week, or day etc.). At the baseline level in the current survey, sample reliabilities were $\alpha=.87$ and $\alpha=.74$ for positive and negative affect respectively. Sample reliability estimates for within-day data were $\alpha=.58$ for within-person positive affectivity, and $\alpha=.86$ for between person positive affectivity. Sample reliability estimates for within-day data were $\alpha=.45$ for within-person negative affectivity, and $\alpha=.50$ for between person negative affectivity. The PANAS lists affective states (e.g., excited, upset), and asks participants to select a number, representing an intensity rating for the specified time period (e.g., you have felt this way today). The intensities range from 1 = *very slightly or not at all* to 5 = *extremely*. During the two long surveys at the start and end of the study, the PANAS reflected affectivity in “general.” During the daily surveys, the PANAS reflected affectivity in that “moment.” Validity evidence for the PANAS has been obtained in the

form of correlations between scores on the PANAS, and state anxiety scores, depression scores and psychological dysfunction (Watson et al., 1988).

Engagement. The Job Engagement Scale (JES; Rich et al., 2010) closely mirrors Kahn's (1990) seminal work on engagement as a simultaneous expression of emotional, cognitive, physical presence at work. Thus, I measured engagement using the 18-item Job Engagement Scale on the baseline and follow-up surveys. The 9-item shortened version of the JES (Crawford, LePine, & Buckman, 2013) was used in the moment-to-moment surveys to accommodate length concerns. As employees needed to complete these surveys in real-time, 30 times, brevity was of the utmost concern. The item stems were altered during the within-day surveys to reflect a moment-to-moment view of engagement. Previous analysis on the full length JES has shown high internal consistencies ($\alpha = .95$; Rich et al., 2010). In the current sample, the reliability estimate for the JES was $\alpha = .95$ on the baseline measurement. Reliability estimates remained acceptable on the shortened JES during within-day measurements (i.e., within-person $\alpha = .76$, and between-person $\alpha = .91$). Items on the measure assess three dimensions of engagement: cognitive (e.g., "At work my mind is focused on my job"), emotional (e.g., "I am enthusiastic in my job"), and physical (e.g., "I devote a lot of energy to my job"). Items are answered on a 5-point Likert scale (1 = *strongly disagree*, 5 = *strongly agree*). Validity evidence for the full length JES is seen through its distinction from other constructs such as job satisfaction, job involvement, and intrinsic motivation (Rich et al., 2010). The measure was created to be scored as an aggregate of the three dimensions of engagement. This is in accordance with the supposition (Kahn, 1990) that engagement is the simultaneous exertion of all three types of energy on the job. Moreover, when creating the scale, Rich et al. noted that dimension correlations were higher than .6,

indicating support for their combination of all three dimensions to represent a single engagement construct.

Job satisfaction. Job satisfaction was assessed using the 8-item Abridged Job in General scale (AJIG; Russell et al., 2004). Scores on the AJIG have previously shown internal consistency reliabilities of .87 (Russell et al., 2004). In the current study, the observed reliability estimate for this sample was $\alpha=.83$. Items are listed as descriptors to an employee's job (e.g., "Excellent") and the employee is asked to respond to them as either 0 = no, 1 = ?, or 3 = yes. The AJIG measures an overall job satisfaction. Russell and colleagues accumulated criterion-related validity evidence through the relationship between the AJIG and the five dimensions of the job descriptive index, an intention to quit survey, and an organizational commitment measure.

Quality of coworker relations. Quality of coworker relations was assessed using May et al.'s (2004) measure of rewarding coworker relations. This is a 10-item scale measuring agreement to statements on a 5-point Likert scale (1 = *Strongly Disagree*, 5 = *Strongly Agree*). Items refer to an employee's perceived relationship quality with his or her coworkers (e.g., "My interactions with my coworkers are rewarding"). May et al. reported a Cronbach's alpha of .93 for the scale in their sample. Quality of coworker relations was measured during the within-day measurement periods, and the reliability estimate for this sample was $\alpha=.75$ at the within-person level, and $\alpha=.76$ at the between-person level. Criterion-related validity evidence was accumulated through the expected convergent relationship between coworker interactions, and psychological safety and meaningfulness.

At the within-day level, instructions asked participants to rate their level of agreement to several statements regarding their coworker relationships *in the last three hours*. I elected to use five of the questions from the May et al. (2004) coworker relations scale for the proposed study.

Several of the questions referred only to lasting perceptions of coworker relationships (e.g., “I believe that my coworkers appreciate who I am”), and my study requires the participant to make judgments of moment-to-moment coworker interactions. Additionally, I modified question stems to reflect this moment-to-moment context (e.g., “In the past three hours...”, “My interactions with coworkers were rewarding”). The use of this moment-to-moment stem, and the five coworker relation statements allows for employees to report their recent coworker relations since the previous survey.

Work Tasks. Assessing task meaningfulness at the moment-to-moment level posed a challenge. Other measures of meaningfulness (May et al., 2004; Steger, Dik, & Duffy, 2012) assess work meaningfulness as a whole, and at a static level, wherein most questions do not relate to real-time shifts in work tasks. Therefore, I adapted three items from May et al.’s meaningfulness survey, and assessed each item as it pertained to tasks on the baseline survey. For example “These tasks are personally meaningful to me.” From that data, I created a meaningfulness score for each of eight tasks for each employee. The eight work tasks were compiled through the use of O*Net’s work activities data. O*Net is essentially an online database for job analyses. The database has information on a multitude of jobs, categorized by groupings such as industry, level, occupation, or job. Additionally, O*Net provides a listing of work activities common across all jobs. O*Net groups these work activities into four categorizations: Information input, mental processes, work output, and interaction with others. I chose eight specific work activities from O*Net, that span three of the four categories of work activities. I chose not to include work activities from the information input category, because they appeared less generalizable across occupations than the other categories. Additionally, I chose work tasks representing a variety of different interaction levels, autonomy levels, and skill

levels. Based on feedback from the pilot test indicating that activity titles could be ambiguous at times, to clearly indicate the meaning of work activities I listed O*Net's definition of the activity next to it in the survey.

On the baseline survey, participants were asked to respond to three meaningfulness items relating to each of the eight activities. On the daily surveys, participants were asked to check off which of the eight activities they had participated in during the last three hours. Based on this, I created a composite meaningfulness score for each data collection point for each employee, consisting of an average of the meaningfulness of the work tasks they had completed in the last three hours (meaningfulness as indicated by their baseline ratings of these tasks). For example, if a participant indicated that "Making decisions and solving problems" was 3.9 (out of 5) in terms of meaningfulness (on his or her baseline), and participated in only that task during the three hours, his or her momentary task meaningfulness score for that data collection point would have been 3.9. If that same participant had indicated that "Performing general physical activities" was a 2.1 in terms of meaningfulness (on his or her baseline survey) least meaningful work activity and he/she participated in that task as well as "Making decisions...", then his or her momentary task meaningfulness score would be a 3 (the average of 3.9 and 2.1). The checklist nature of this daily scale also served the purpose of easing a participant's burden of question responding at the daily level. A checklist of eight activities is far more manageable than answering twenty four questions about meaningfulness every three hours. The overall estimate of reliability for all 24 items in this sample was $\alpha=.92$. Individual three-item scales for each work task had reliability estimates ranging from $\alpha=.92$ to $\alpha=.96$.

Open-ended questions. Each survey ended with two open-ended questions asking participants "Is there anything out of the ordinary that happened since the last survey," and

“Have any unexpected work scenarios happened since the last survey.” Fuller et al. (2013) found that at times important information could be gained from a qualitative response. By giving open ended questions, participants were able to indicate anything that may have affected data collection that was not assessed in the rest of the survey.

Data Analysis

To test my hypotheses, I analyzed the experience sampling data using multi-level modeling (MLM). Hierarchical linear modeling (referred to as MLM for distinction from the statistical software package called HLM) for repeated-measures designs has several advantages over more traditional longitudinal analyses such as a repeated measures ANOVA. First, repeated-measures MLM data inherently violates the assumption of an independence of errors (sphericity); however, allowing data to be nested within individuals alleviates this concern (Tabachnick & Fidell, 2013). Second, repeated-measures ANOVAs are particularly sensitive to missing data, and require equal intervals of measurement for each participant, whereas MLM can model change even when data are incomplete (Luke, 2004; Tabachnick & Fidell). Alternatively, a time-series analysis could have addressed similar research questions by showing which variables affect fluctuations in engagement and controlling for autocorrelation of errors. However, time-series analyses require at least 50 data points per participant to determine a definite pattern of responses (McDowall, McCleary, Meidinger, & Hay, 1980).

As such, MLM represents the most efficacious technique for answering my research questions using the data collected. Through the use of MLM, I created a two level model for my data, with the first level indicating within-day predictors and the second level indicating person-level data. Hypotheses 1 and 2 were tested using the intra-class correlation coefficient (ICC), produced in the null model, and were verified through time-lagged engagement beta-weights and

level two general engagement level relationships. Hypotheses 3-6b were tested using separate level one models, and verified in an overall model including all level one predictors – a standard procedure in MLM is to only include significant predictors in the final model (Singer & Willett, 2003), thus predictors were first tested independently, and then significant predictors were added to an overall model. Hypotheses 7-8 were tested at the second level of the MLM.

Results

All data analyses were conducted using several different statistical programs. Baseline measure reliabilities were analyzed using SPSS Statistics, version 23 (IBM Corp., 2015). Moment-to-moment measure reliabilities were assessed using SAS software. Confirmatory factor analyses and measurement invariance calculations were run in MPlus statistical software, version 6.12 (Muthén & Muthén, 1998-2011). Finally, hypotheses were tested using HLM, version 6 (Raudenbush, Bryk, & Congdon, 2004).

To ensure that the two samples did not differ in terms of their engagement, I first assessed differences in engagement by organization, and no significant differences appeared. Thus, I combined to two organizations (healthcare and technology firm) for all analyses. Prior to assessing relationships, I performed confirmatory factor analyses on all variables in the study using MPlus. All scales had been previously developed and tested in other published studies. The intent here was to verify scale structure, rather than to refine and construct scales. To show good fit, scales should demonstrate a non-significant chi-square (χ^2) statistic, comparative fit (CFI) and Tucker-Lewis indices (TLI) above .95, a root mean square error of approximation (RMSEA) below .08, and a standardized root-mean-square residual (SRMR) below .10 (Kline, 2016).

All scales demonstrated some deviation from recommended cutoffs for certain statistics (Table 1). However, the relatively small sample size for the baseline measures ($n = 134$) may have hindered these statistics. The three factor JES came close to acceptability on all statistical cutoffs (Table 1). Importantly, each fit statistic in the three factor model (Table 1) was significantly better than a one factor model (RMSEA = .239; $\chi^2[135] = 1205.522$, $p < .001$; CFI = .587; TLI = .522; SRMR = .049). For example CFI of .931 in the three factor model is more than .02 higher than the CFI of .587 in the one factor model (Cheung & Rensvold, 2001). Since

the dimensions of the JES were all correlated moderately highly ($r < .55$), and theory indicates (Kahn, 1990; Rich et al., 2010) that engagement is the simultaneous exertion of all three dimensions, I used an aggregate score for the three dimensions of engagement in subsequent analyses. This decision is also consistent with Rich et al.'s construction of the measure as a single scale score. Positive affectivity also initially performed poorly ($\chi^2[35] = 92.572$, $p < .001$; CFI = .887; TLI = .855; RMSEA = .107; SRMR = .063), but also showed increased model fit once items 6 and 10 were dropped (Table 1). Job satisfaction performed poorly as a one factor model ($\chi^2[20] = 90.773$, $p < .001$; CFI = .831; TLI = .763; RMSEA = .161; SRMR = .078), however, when items were separated into two factors by positive and negative wording (see Appendix) model fit improved (Table 1). Since job satisfaction, as measured by the AJIG, is theoretically unidimensional, and the two factors were highly correlated ($r = .80$), I used an aggregate score for the two dimensions in subsequent analyses.

Negative affectivity did not perform well as a one factor model ($\chi^2[35] = 133.804$, $p < .001$; CFI = .690; TLI = .601; RMSEA = .141; SRMR = .093), however as a two factor model with question 5 dropped, all indices indicated a better model fit (Table 1). MacKinnon, Jorm, Christensen, Korten, Jacomb, and Rodgers (1999) noted that the negative affect portion of the PANAS does not always perform the way it should, and used this to modify the existing scale as well. I separated the negative affectivity scale into two subscales and dropped one item that did not load on either factor (.179 and .186). The first factor represented words in the scale that indicated an angered or negative mood (i.e., distressed, upset, irritable), and the second factor represented a more anxious side of negative affectivity (i.e., guilty, scared, ashamed, nervous, jittery, afraid). Despite the poor performance of a one factor scale, theoretically negative affectivity represents a singular construct. Additionally, the two factors were moderately

correlated ($r = .38$), and thus I used an aggregate score of the two dimensions in subsequent analyses.

To assess whether scales had equivalent structure across multiple measurements, the moment-to-moment surveys were tested for measurement invariance (Vandenberg & Lance, 2000). Acceptable fit statistics, as previously introduced, indicate measurement equivalence across occasions. I chose two surveys at random, one from the first half of the measurement periods, and one from the second half of the periods (survey #6 and survey #21). Importantly, these statistics should be interpreted with care as sample sizes were only 39 and 37 respectively. The moment-to-moment stress, coworker relations, engagement, and positive affect scales showed poor fit in some, but not all indices (Table 2). Given the small sample sizes, and that most of the indices approached good fit, I interpreted these to indicate measurement invariance for these scales. Negative affectivity did not show measurement invariance at the moment-to-moment level (Table 1), indicating that psychometric properties may not be equivalent for that scale across measurement occasions. As the negative affectivity portion of the PANAS also did not perform well at the baseline measurement, it follows that a smaller sample size at the moment-to-moment level may produce even worse and more unstable psychometrics. I proceeded to still test negative affectivity in an MLM model, but found no significant results (discussed in the following section), which may be due to these psychometric issues, or to a true lack of relationship between engagement and negative affectivity.

Multilevel Analysis

Descriptive statistics and correlations for all variables are listed in Table 3. Between- and within-person correlations are calculated and listed for all variables, as are means and standard deviations.

The first step I took in analyzing my data was to graph the dependent variable, as it occurred over time. The initial step is the creation of a spaghetti plot (Figure 2; Luke, 2004), which in this case shows the trajectory of engagement over measurement occasions. The plots indicated no specific pattern of responses, so I plotted each participant individually (Figures 3-8). These individual plots allowed me to check for the quality of my data, and obtain an initial impression of the fluctuations in engagement. Participant 31 is omitted from this Figure as I removed this participant from the data-set due to invariant responding. Although several other participants show stagnant engagement levels, they were left in the analysis due to responses on the other variables that indicated they were indeed taking the survey seriously (hence, their lack of variation may represent an important stable component of engagement).

To test hypotheses concerning daily level data, MLM was used. First, a null model (Formula 1) with no predictors was tested to check the independence of errors, and obtain the ICC. The ICC represents the proportion of variance that is due to between-group (i.e., participant) differences. It is attained by dividing the between-group variation by the total variation.

$$Engagement_{ii} = \beta_{00} + r_{0i} + e_{ii} \quad \text{[Formula 1]}$$

In this case, the null model indicates that this equation is $.722/ (.443 + .722)$, which produces an ICC of .620, indicating that 62% of the variation in engagement levels is due to between-person differences. An ICC of .620 represents a majority of the variance, which was not expected. Importantly, for hypothesis 1, this ICC indicates that 38% of the variation in engagement is due to within-person differences, and for hypothesis 2, the rest of the variation is due to between-person differences. Thus, both hypotheses 1 and 2 were supported.

MLM typically takes place as a series of steps or model building (Nezlek, 2012; Tabachnick & Fidell, 2013). The first step is checking the null model to ensure MLM is appropriate, as was previously done to check for the ICC. Second, level one predictors are added one at a time (Singer & Willett, 2003). Thus, the model for daily engagement switches from the null model (Formula 1), to a model with level one predictors (Formula 2).

$$Engagement_{ii} = \beta_{00} + \beta_{10}*(Predictor_{ii}) + r_{0i} + e_{ii} \quad [Formula 2]$$

Notice, the addition of level one predictors added error terms at the second level, accounting for systematic within-person variation. Formula 2 indicates the mixed model formula for MLM. This mixed model combines the level one and level two equations in MLM into a single formula (it represents all of the same information as separate formulas would, just simplified down). For brevity's sake, I will only present the mixed model formulae from here on out (the non-mixed model becomes multiple different equations). All level one predictors were added using group-mean centering. Essentially, this involves centering the predictor on the mean for that person (as opposed to the sample as a whole). This is advantageous in MLM at the first level as it aids the interpretability of results.

I first added each predictor independently into the model (Nezlek, 2012). If the predictor was significant, I used it in the final level one model. Results from these models indicated that negative affectivity ($\beta_{10} = .070, p = .134$), and meaningfulness ($\beta_{10} = -.060, p = .360$) were not significantly related to momentary fluctuations in engagement (Table 3), indicating that hypotheses 3 and 6b were not upheld. However, positive affectivity ($\beta_{10} = .469, p < .001$), the quality of coworker interactions ($\beta_{10} = .188, p < .001$), and time lagged engagement ($\beta_{10} = .142, p < .001$) were all significant predictors of daily engagement (Table 3), thus supporting hypotheses 5 and 6a. Although the relationship between stress and engagement was significant

($\beta_{10} = .066, p = .016$), it was in the opposite direction as hypothesized, rejecting hypothesis 4. Moving forward with these results, I created a model in which momentary engagement was regressed on stress, positive affectivity, and quality of coworker relations, at level 1 (Model 8 in Table 4). All predictors remained significant at the .05 level. Interestingly, the simultaneous regression of these variables on engagement heightened the relationship between stress and engagement, and weakened (although still significant) the relationship between quality of coworker relations and engagement with respect to their previously tested individual models (Table 3). To further address hypotheses 1 and 2 about variations in engagement levels, I added time lagged engagement to this model (Model 9 in Table 4). All predictors again remained significant, indicating that even controlling for the previous measurement period's engagement levels, these were significant predictors of engagement at that moment. This lends support to both hypothesis 1, and 2. That is, the beta-weight for time lagged engagement was smaller than any of the other moment-to-moment predictors, demonstrating that previous engagement levels made the least difference here, thus indicating that engagement levels do fluctuate substantially. However, as this was still a significant predictor of engagement, it can be concluded that prior engagement levels do have a modest effect on current engagement levels, thereby signifying that some portion of engagement remains relatively stable.

Next, I added level two predictors to the MLM. As with creating the first level, I added predictors one at a time, keeping only significant variables. Six participants had no data for affectivity and job satisfaction at this level, and thus were omitted from this stage of the analysis. Since my hypotheses only dealt with the relationship between daily engagement and level two predictors (as opposed to some type of moderated relationship), these predictors were only added to interact with the intercept (person-level engagement) as opposed to being added to interact

with level one predictors (this would create a moderation effect). As was the case with the moment-to-moment data, general level NA was not significantly related to daily engagement ($\beta_{10} = .004, p = .989$), thereby rejecting hypothesis 7b. Contrarily, hypotheses 7a and 8 were supported as general-level PA ($\beta_{10} = .478, p = .006$), and job satisfaction ($\beta_{10} = .425, p = .031$), were significantly related to engagement. Finally, as a last bit of testing for hypotheses 1 and 2, general-engagement was tested as a level two predictor. The high beta-weight in this model ($\beta_{10} = .803, p < .001$) indicates support for hypotheses 2, such that a person's average across their moment-to-moment engagement levels is highly related to their engagement score at the baseline. However, upon entering all three simultaneously in the model, general engagement levels remained the only significant predictor of engagement (Table 5), indicating that much of the variance accounted for by job satisfaction and general positive affect was actually represented by general engagement levels.

Discussion

The current study sought to investigate the relationship between engagement and several hypothesized antecedents and outcomes at a within-day level. Moreover, it attempted to serve as a rigorous test of Kahn's model of engagement beyond existing studies. Kahn's conceptualization suggested that depending on certain contextual and interpersonal factors, individuals fluctuate in terms of how much they choose to invest themselves in their work-roles, yet they also experience some stability in engagement over time. My study is a first to assess both the fluctuations and stability of engagement.

Within-day Data

The support and rejection of my hypotheses indicates several interesting and important revelations about the nature of engagement, and how employees approach their work roles. As expected, the quality of coworker relations, and positive affect were significant predictors of momentary engagement levels. This suggests that as employees feel their recent interactions have been positive, they tend to be more engaged in what they are doing. Similarly, as certain events dictate an individual's affect, her or his engagement levels may follow suit (Bledow et al., 2011; Weiss & Cropanzano, 1996). That is, an event may occur that causes an employee to have a positive outlook for the afternoon, which would in turn signal an uptick in her or his engagement levels as well. However, the relationship between engagement and stress was actually in an opposite direction than hypothesized, as engagement was positively related to stress levels. One possible explanation for this modest positive relationship comes from the human performance literature (e.g., Keeley, Zayac, & Correia, 2008). As individuals experience low to moderate levels of anxiety, their performance is actually heightened. They become more

aware and mindful of the situation. It could be that a similar reaction is happening when employees experience a small amount of stress; they may become more absorbed with their work as their senses are heightened, indicating a readiness for a physical and cognitive exertion into their work role (i.e. ready to engage).

Negative affectivity and meaningfulness were not related to employee engagement at the within-day level, contrary to my hypotheses. While affectivity in general may theoretically have ties with the emotional dimension of engagement, research has been mixed on the relationship between negative affectivity and engagement. For instance, Halbesleben, Harvey, and Bolino (2009) found negative affectivity correlated $-.55$, $-.2$, and $.01$ with engagement, respectively, across three samples. Additionally, job engagement, as conceptualized by Kahn (1990), is a motivational construct that deals with the thoughtful exertion of one's energy. Although positive affect inherently deals with feeling energetic (Watson et al., 1988), and overlaps quite well with the active nature of engagement, negative affect is more concerned with feelings of fear, distress, and pessimism.

Of note is the lack of a relationship found between task meaningfulness and job engagement at the momentary level. Here though, I offer a few explanations. First, task completion had a small within day variance, indicating that *typically* individuals complete the same or similar work tasks each day. This may be a function of the sample, or it may be telling of workers in general. Although it is generally assumed that work tasks vary across days, it may be the case that the majority of workers end up completing relatively similar tasks quite frequently. For instance, a hospital administrative worker may help check patients in, complete paperwork, and record information for the majority of the day, every day. Due to the method for collecting meaningfulness data, this lack of variance would constrain the correlation between

engagement and task meaningfulness. Second, although the collection strategy for meaningfulness data intuitively makes sense, it may be that meaningfulness at work is a stable between-person predictor – something that an assessment at the momentary level would have missed. As the measure of task meaningfulness did not take into account the valence of these tasks at a between-person level, it could not be added to the second level of the MLM, thereby it is missed in the current data analysis.

Finally, engagement was significantly predicted by a time lagged engagement level; meaning an employee's afternoon engagement was significantly related to her or his morning engagement. However, this relationship was relatively modest, and compared to other predictors, such as positive affect, stress, and quality of coworker relations, engagement at the previous assessment period had a minor effect on current engagement. This result reveals the inherently fluctuating nature of engagement. If Tuesday afternoon's engagement levels were best predicted by Tuesday morning's engagement levels (as opposed to other predictors) it would be safe to conclude that engagement may not actually vary as much as theorized. Similarly, even in its own model, the coefficient representing that relationship did not exceed .15, indicating that relatively little of the variance in current engagement levels could be explained by directly preceding engagement levels.

Person-level Data

At the second level of analysis, I found similar support (or lack-there-of) for my hypotheses. Again, positive affect was significantly related to engagement, and negative affect was not related to engagement (both at the general level). In this case, both general engagement levels, and job satisfaction levels were related to engagement. However, when added into the same model, general positive affect and job satisfaction were no longer significantly related to

engagement, only general engagement levels were. Importantly, between-person differences in daily engagement were significantly related to job satisfaction. Individuals who reported being more satisfied were likely the same individuals who reported a tendency to be more engaged.

That general engagement levels are related to between-person differences in engagement should come as no surprise. Yet, this represents an important bit of practical information. Experience sampling studies that involve the intensive surveying of employees allow for a nuanced look into direct drivers of engagement, but if the purpose of a survey effort is merely to assess levels and figure out where individual differences lie, more general single time-point engagement surveys may serve as perfectly acceptable and more efficient techniques. Collecting moment-to-moment experience sampling data is a more involved and more intensive procedure than collecting static single time-point measurements of the same constructs. Inherently, these strategies serve different purposes; where yearly general assessment of a workforce's engagement may be efficacious for addressing gaps in engagement, and employees in need of help, a moment-to-moment assessment of engagement can provide employers with a fine-grained and direct look at real-time drivers and consequences of engagement. The high positive relationship between these two survey methodologies for assessing engagement indicates that although the two are not totally interchangeable for answering research questions, both may be substitutable for workforce assessment.

Implications and Future Directions

The intent of this study was to provide knowledge about the stability of engagement to advance the literature and practice of employee engagement. The current findings fill a void in the literature concerning the stability of engagement as a construct. The study serves as a

theoretical base for a multitude of future research considerations, and provides valuable information to employers regarding the impetus behind engagement interventions.

The findings of this study provide a solid base for the growing need for a universally accepted conceptualization of engagement as either a stable or dynamic construct. Bakker and colleagues (2011) noted that an important issue in engagement research is creating measures that adequately capture fluctuations and stability in engagement. My findings support the notion that engagement does fluctuate significantly within-people, and that a portion of these fluctuations are due to contextual variables that coincide with engagement levels at that moment.

One of the interesting avenues left unexplored by this research is the direct indication of engagement spikes (positive and negative). Examining the open-ended comments allows for anecdotal rather than statistical inferences about waxes and wanes in engagement levels. However, because I opted to have open ended responses as optional questions, employees did not provide data in this manner regularly. Future research could expand upon these anecdotal findings in two ways. First, additional qualitative data collection could provide rich data that details the nature of fluctuations in engagement, and employees' opinions as to why they were more or less involved in their work role at a given time. Second, the open ended responses I did receive echoed Culbertson et al. (2012), such that work-life interference issues appeared to play a role in lapses in engagement. For instance, in Figure 4, Participant 14 experiences a large dip in their engagement scores at measurement occasion 10 and on this particular occasion he or she provided an open ended response indicating a family emergency. Similarly, in Figure 3, Participant 1 experiences a drastic drop in engagement during measurement occasion 9, and indicated through an open-ended response that they had an abnormally poor interaction with a coworker in the last few hours. An exploration of these work-life and coworker interaction issues

from an occurrence specific standpoint would provide valuable information for both science and practice. As my data collection was signal-contingent (employees provided responses based on pre-determined signaled times), event-contingent ESM methodologies (employees would respond when a given scenario or situation happened at work; Bolger & Laurenceau, 2013) are still unrepresented in the literature and could further the understanding of engagement as a construct.

This research was unable to measure important momentary outcomes of engagement, and as such it would behoove future research to consider assessing these variables. For instance, in a similar sample to the present study, researchers could look at patient satisfaction scores and tie that to the state-engagement levels at the time the patient was seen. As noted by George (1991), trait-like versus moment-to-moment measurements of variables can relate differentially to organizational outcomes. Researchers should consider this when designing future studies. Similarly, employers must consider this when choosing how to measure engagement levels within their workforce. Measurement of trait-like engagement, as opposed to momentary engagement, may be measuring engagement that is differently related to the desired outcomes.

Concerning businesses, a variable view on engagement levels has several important implications. First, although results indicate a significant variance within-individuals in their engagement, a trait-like stable engagement remains. These results provide support for the view that even though individuals may vary in their moment-to-moment engagement, they may have stable overall levels of engagement over an extended period. Research has shown that engagement levels can vary according to job characteristics like autonomy (Menguc, Auh, Fisher, & Haddad, 2013), as well as supervisor support (Saks, 2006). That engagement levels can fluctuate depending on contextual and individual issues, and also that they may stay stable on

average over time is positive given the relationships between engagement and administrative outcomes such as performance (see Rich et al., 2010). For instance, if an employer can increase an employee's average level of engagement, it follows that the employee may also be able to achieve a stable increase in performance as well.

Additionally, the measurement of engagement has been a point of contention (Cole, Walter, Bedeian, & O'Boyle, 2012). The findings in this proposed study inform best practices for measurement. In particular, engagement can be measured at the general, trait-like, level. Yet, practitioners should not overlook the fluctuating nature of engaged states. As daily mood and stress levels vary, an individual's exertion into her or his work role will simultaneously wax or wane. A quality supervisor will recognize indications of these temporary fluctuations in engagement and both proactively and prescriptively manage employees' moment-to-moment engagement.

Finally, engagement theory was furthered by this study in several intriguing ways. First, Kahn's (1990) model was empirically tested and generally supported. However, stress may play a more complex role in engaging employees than previously thought. A small positive relationship with engagement may indicate a reconsideration of the burnout model of engagement (Schaufeli et al., 2002) that relies on the JD-R to explain a negative relationship between engagement and stress. Job engagement as operationalized by the JES and conceptualized by Kahn, may be a slightly different construct than work engagement as operationalized by the UWES and conceptualized by Schaufeli et al. (Byrne, Peters, & Weston, in press). Second, a non-significant relationship with negative affectivity at both the static (baseline), and moment-to-moment levels indicates that although engagement may have theoretical and empirical ties to positive affectivity, the same is not true for negative affectivity.

The active and energetic portion of positive affectivity aligns with the motivational nature of employee engagement, however a negative mood operationalized by feelings of distress and anxiety may not represent a meaningful antecedent to engagement. In the JD-R, resources (including personal ones) lead to high levels of engagement, yet demands do not affect engagement levels (Bakker & Demerouti, 2007). Similarly, personal energy and activity appears to lead to engagement in this study, while conditions that produce anxiety, distress, and unrest do not relate to engagement. Third, the variables used in this study represent practical and malleable constructs in the workplace. I specifically chose certain proxies for antecedents to engagement because of their value to managers and supervisors hoping to affect change. Although safety and availability represent important and meaningful constructs, providing a practitioner with a recommendation to reduce stress, increase the quality of coworker interactions in the workplace, and provide more meaningful tasks (that are directly identified through the use of a survey like mine) represents a concrete recommendation that can be easily enacted. This study provides evidence that stress and the quality of coworker interactions in the preceding few hours do directly affect engagement levels during that time.

Limitations

Due to the intensive nature of the study, participants needed to be compensated for their time, yet the current research was conducted on a tight budget that constrained the ability to compensate for each completed survey. Because of budget limitations, attrition was higher than desired, and those who completed the entire study may vary significantly from those did not. Although participants in the pilot test indicated that money was not a motivating factor, and attrition rates did not vary in that case regardless of payment condition, participants in the current study may have felt that the monetary value of the incentive was too low to be motivating.

Additionally, as employees never interacted directly with the researcher, I relied on organizational contacts to fully support this data collection effort. Although participants were introduced to me via email, and interacted with me through this online medium, having a more trusting relationship and the opportunity to further explain the study in person could have also bolstered participation rates. For example, in the pilot test I had trusting relationships and face-to-face contact with all participants and attrition rates were much lower.

Furthermore, because of the many data collection points, surveys needed to be short for employees to (1) complete them often, and (2) to feel the compensation was worth their effort. The brevity of these surveys means that certain contextual information was lost that might otherwise have provided additional insight into engagement's stability. For instance, I chose shortened scales, and scales that relied on an endorsement of a word rather than a full statement, and subsequently those scales were not all psychometrically sound. The ability to use longer and more validate measures could have aided this issue. Additionally, although I measured five variables at a moment-to-moment level, there were clearly other forces at play, and capturing that information is invaluable to the furthering of engagement literature.

Another limitation is that the two samples were collected during different months of the year, and participants were from different industries. Thus, when aggregated, the relationships between variables may have been minimized due to inherent differences in the two employers. Although the participant's organization did not significantly affect engagement levels when added to a regression model, the small uneven sample size (one organization had 12 participants, and the other had 37) may have hindered finding an actual difference by organizational affiliation. In a study designed to address specific antecedents at a momentary level, these nuances may have been lost between vastly different job types. Nurse practitioners face far

different conditions on the job than do web designers, and motivational drivers may also vary widely between these two samples. The small sample size here prevents in confident statistical conclusions about the differences in these samples. However, the inclusion of two very different industries across a variety of different jobs and job levels indicates that the results can generalize across organizational samples, thus addressing the population of interest which was employees in general.

Additionally, the measure of task meaningfulness was applied in a different way than it was in its original development. Scores were aggregated and averaged at a daily level based on a task checklist rather than having employees rate meaningfulness at each time point. Task meaningfulness was not found to be significantly related to engagement levels, which is directly in contrast to theory and other empirical research. However, as previously stated this may be an issue of the temporal aspect of meaningfulness rather than a measurement deficiency. Meaningfulness may be a relatively stable characteristic of work, unaffected by varying work tasks. As such, this would be a significant between-person predictor, and not something that would explain within-person variations in engagement.

Lastly, the proposed study does not directly resolve several concerns with the measurement of engagement (Cole et al., 2012; Newman & Harrison, 2008). Although resolving measurement issues was beyond the scope of this study, the consequence is that the study was conducted using measures with aforementioned deficiencies, which could potentially bias the results. By using engagement scales with questions that inherently imply stability, I may have affected the way the participants define engagement in their responses (Allen & Potkay, 1981). However, since the use of an engagement scale with the moment-to-moment question stem is theoretically supported, it follows that using an empirically established scale with modified

question stems would be a best practice. Also, many of the current criticisms with measuring engagement involve Macey and Schneider's (2008) conceptualization of engagement, or Schaufeli et al.'s (2002) conceptualization and measurement (i.e., the UWES) of engagement. By looking at engagement through the context of Kahn's (1990) definition and using a scale that is directly based on this conceptualization (i.e., the JES), I attempted to avoid most of the current criticisms (e.g., confounding with other similar constructs; Newman & Harrison, 2008).

Conclusion

Employee engagement represents an evolving construct. As our understanding of the relationship between engagement with constructs in its nomological network advances at a static level, it is also necessary to address momentary ebbs and flows in engagement levels. Doing so allows for a rigorous test of Kahn's (1990) seminal conceptualization of engagement, and aides in making conclusions about direct antecedents and covariates of engagement. Capturing real-time data indicated that engagement and positive mood fluctuate together; the investment of energy in one's work role is directly tied to the momentary disposition, stress, and the interpersonal interactions that an employee experiences. Importantly, these fluctuations are also distinct enough between employees that they predict important organizational outcomes like job satisfaction. The further pursuit of experience sampling data in the employee engagement literature will permit a fine-grained understanding of the construct as a whole.

Table 1

Confirmatory Factor Analyses for Baseline Variables

Scale	Dimensions	χ^2	(df)	RMSEA	CI	TLI	CFI	SRMR
Engagement	3	311.86	132*	.099	.085-.113	.920	.931	.049
Pos. Affect	1	55.105	20*	.111	.076-.146	.889	.921	.053
Neg. Affect	2	61.774	26*	.098	.067-.130	.831	.878	.064
Job Sat.	2	71.22	19*	.142	.107-.177	.816	.875	.070
Task Mean.	8	655.84	224*	.121	.110-.132	.857	.884	.051

Note. * $p < .05$; Pos. Affect = Positive Affect, Neg. Affect = Negative Affect, Job Sat. = Job

Satisfaction, Task Mean. = Task Meaningfulness.

Table 2

Measurement Invariance across Survey Occasions

Scale	χ^2	(df)	RMSEA	CI	TLI	CFI	SRMR
Engagement	126.98	45*	.219	.175-.264	.887	.929	.096
Pos. Affect	73.08	55	.093	.000-.146	.947	.968	.110
Neg. Affect	185.20	37*	.325	.279-.372	.453	.639	.645
Stress	44.04	21*	.170	.099-.240	.892	.924	.132
Coworker Relations	53.31	15*	.259	.186-.336	.912	.934	.257

Note. * $p < .05$; Pos. Affect = Positive Affect, Neg. Affect = Negative Affect.

Table 3

Variable Correlations Within and Between Individuals^a, Means, and Standard Deviations

Variable	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9	10	11
1 General Engagement	4.13	0.63		.66*	-.12	.50*	.57*	.00	.01	.70*	.70*	.34*	.43*
2 General Positive Affect	3.38	0.62			-.25	.42*	.45*	-.12	-.25	.41*	.43*	.28	.19
3 General Negative Affect	1.57	0.40				-.25	-.18	-.39*	.42*	.00	.00	-.02	.09
4 Job Satisfaction	2.56	0.56					.39*	.09	-.23	.33*	.33*	.27	.28
5 State Positive Affect ^b	2.68	0.91						-.18	-.33*	.59*	.61*	.49*	.45*
6 State Negative Affect	1.25	0.40					-.12*		.67*	.16	.15	.00	.31*
7 Stress	2.59	0.89					-.29*	.56*		.12	.12	-.13	.26
8 State Engagement	3.56	0.84					.57*	.11*	.12*		.99*	.57*	.45*
9 Lagged Engagement	3.55	0.85					.46*	.10*	.10*	.76*		.57*	.45*
10 Coworker Relations	3.89	0.80					.39*	-.07*	-.14*	.44*	.24*		.15
11 Task Meaning	4.04	0.53					.36*	.20*	.20*	.36*	.33*	.12*	

Note. Gen. = General, State = State; * $p < .05$.

^aWithin-individual correlations are below the main diagonal and between-individual correlations are above.

^bWithin-individual variables were averaged to calculate the correlations between individuals.

Table 4

MLM Equations and Coefficients for Individual Level 1 Models.

Model ID	Fixed Effect	β	SE	T-ratio
Model 1	Intercept (Null Model)	3.585**	.103	34.816
Model 2	Intercept	3.567**	.109	32.672
	Task Meaningfulness	-.060	.066	-.916
Model 3	Intercept	3.584**	.104	34.467
	Momentary Stress	.066*	.027	2.411
Model 4	Intercept	3.584**	.105	34.265
	Quality of Coworker Relations	.188**	.024	7.925
Model 5	Intercept	3.587**	.104	34.425
	Momentary Positive Affect	.469**	.027	17.507
Model 6	Intercept	3.587**	.104	34.423
	Momentary Negative Affect	.070	.047	1.498
Model 7	Intercept	3.584**	.106	33.955
	Time Lagged Engagement	.142**	.029	4.898

Note. * $p < .05$; ** $p < .001$.

Table 5

MLM Equations and Coefficients for the Full Level 1 Models.

Model ID	Fixed Effect	Coef.	SE	T-ratio
Model 8	Intercept	3.585**	.105	34.816
	Momentary Stress	.180**	.027	17.370
	Quality of Coworker Relations	.124**	.024	7.464
	Momentary Positive Affect	.474**	.022	5.732
Model 9	Intercept	3.582**	.106	33.872
	Momentary Stress	.175**	.024	7.190
	Quality of Coworker Relations	.122**	.022	5.665
	Momentary Positive Affect	.466**	.028	16.827
	Time Lagged Engagement	.109**	.025	4.391

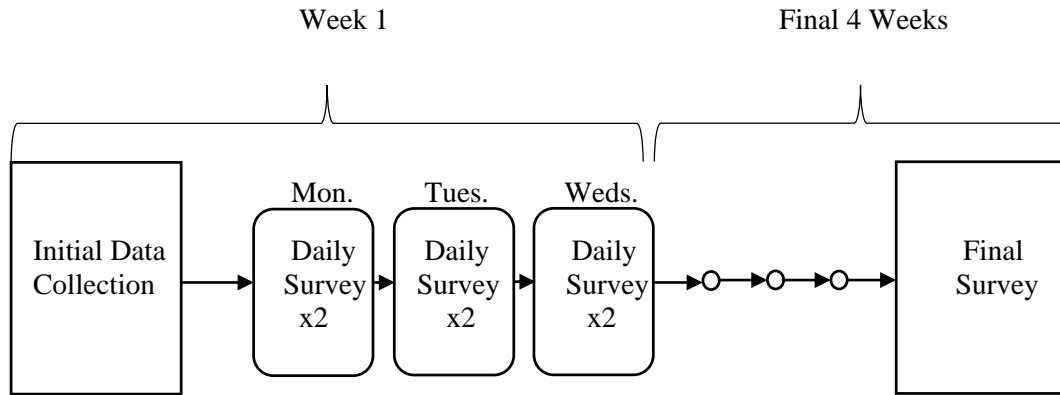
Note. *p < .05; **p < .001.

Table 6

MLM Equations and Coefficients for the Full Level 2 Models.

Model ID	Fixed Effect	Coef.	SE	T-ratio
Model 10	Intercept	3.561**	.100	35.600
	General Positive Affect	.478*	.165	2.906
Model 11	Intercept	3.561**	.112	31.883
	General Negative Affect	.004	.283	.014
Model 12	Intercept	3.561**	.105	33.776
	Job Satisfaction	.425*	.190	2.240
Model 13	Intercept	3.561**	.080	44.641
	General Engagement Levels	.803**	.128	6.280
Model 14	Intercept	3.561**	.081	43.733
	General Positive Affect	-.096	.177	-.545
	Job Satisfaction	-.026	.171	-.151
	General Engagement Levels	.878**	.184	4.765

Note. *p < .05; **p < .001.



This figure depicts beginning with the large data collection point at time 1, and then maps out an example of one week of data collection, followed by 4 more weeks, and then the final survey is taken. This creates a total of 5 weeks' worth of data collection, and 30 possible daily surveys per participant – 32 overall with the initial and final surveys included.

Figure 1. Data Collection Map.

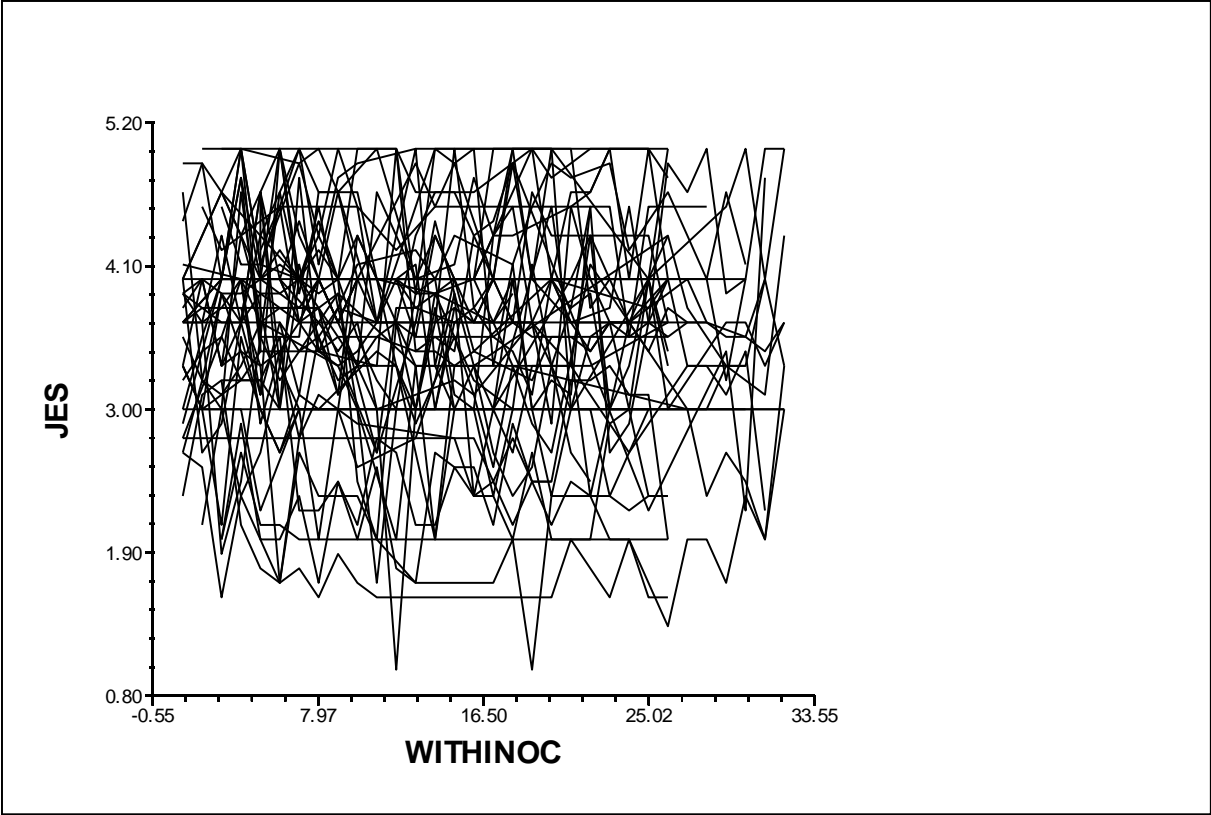


Figure 2. Fluctuations in moment-to-moment engagement levels (JES).

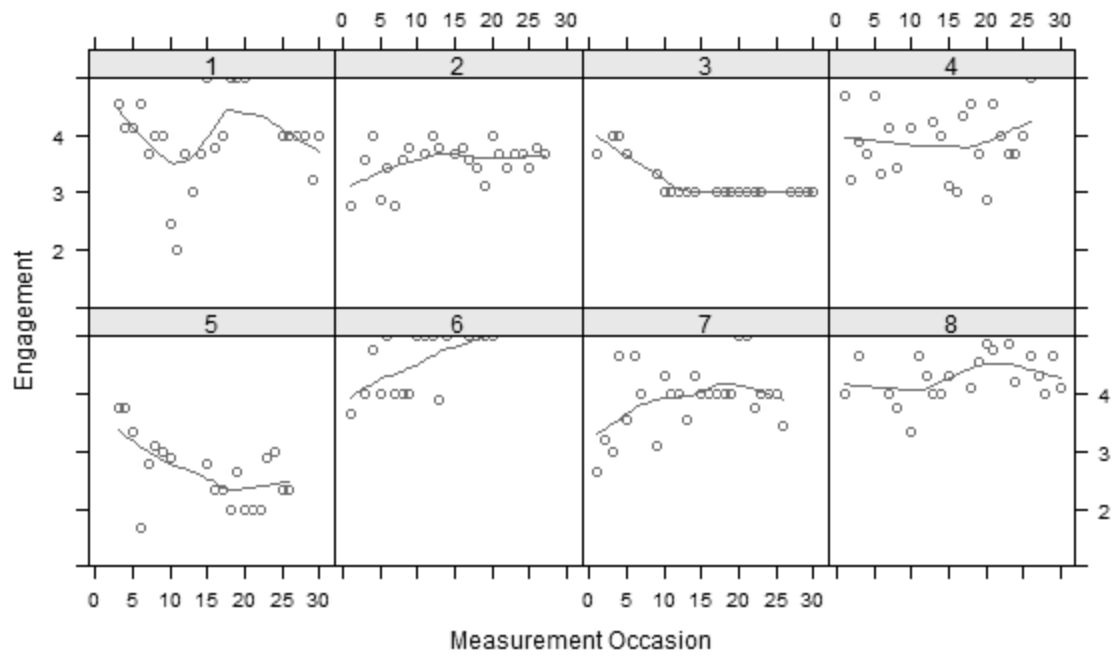


Figure 3. Line chart of moment to moment engagement for Participants 1-8.

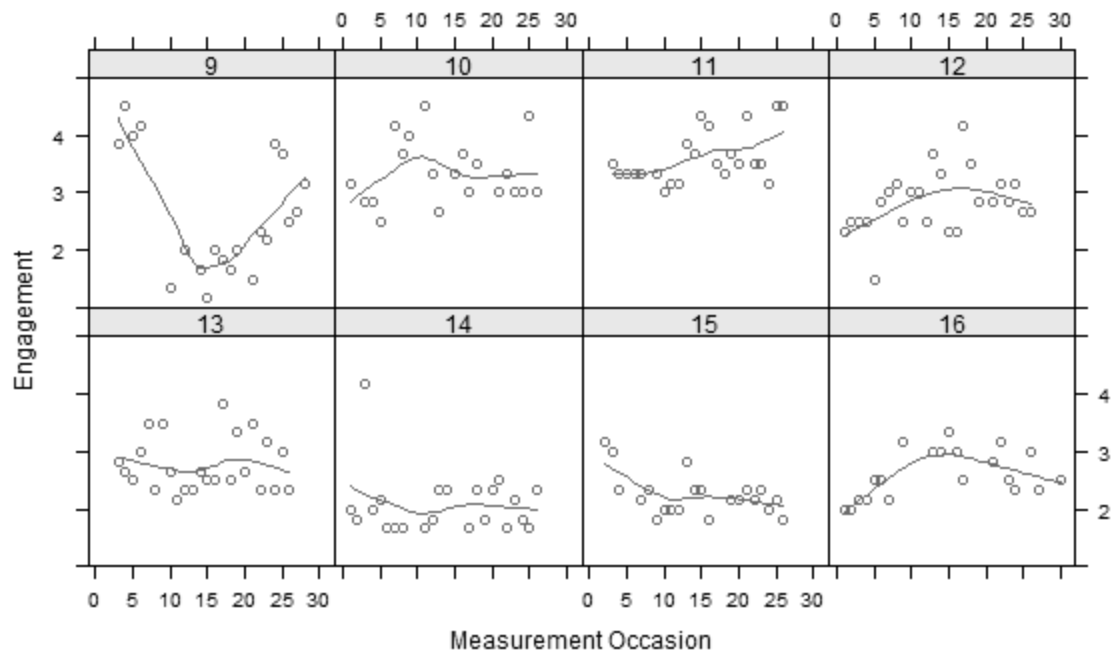


Figure 4. Line chart of moment to moment engagement for Participants 9-16.

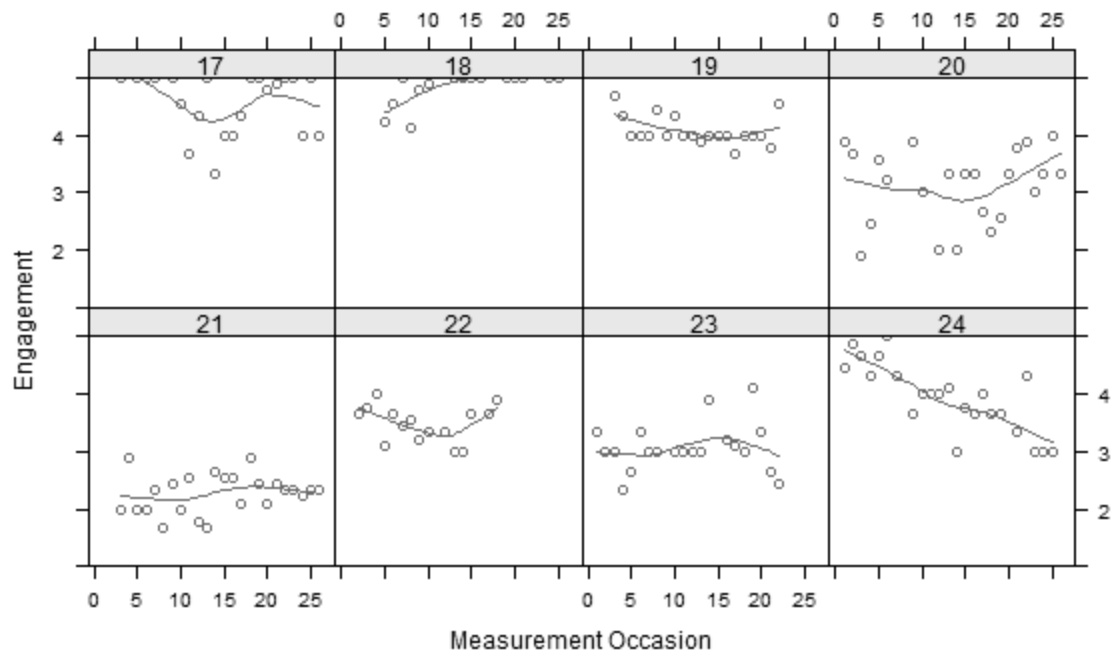


Figure 5. Line chart of moment to moment engagement for Participants 17-24.

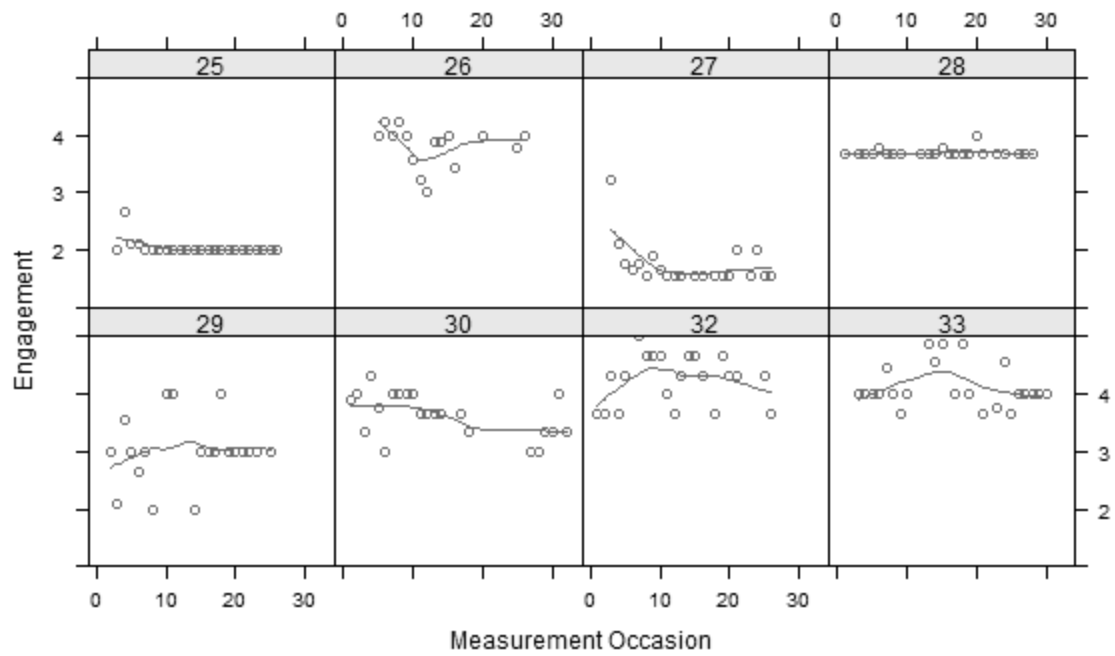


Figure 6. Line chart of moment to moment engagement for Participants 25-33.

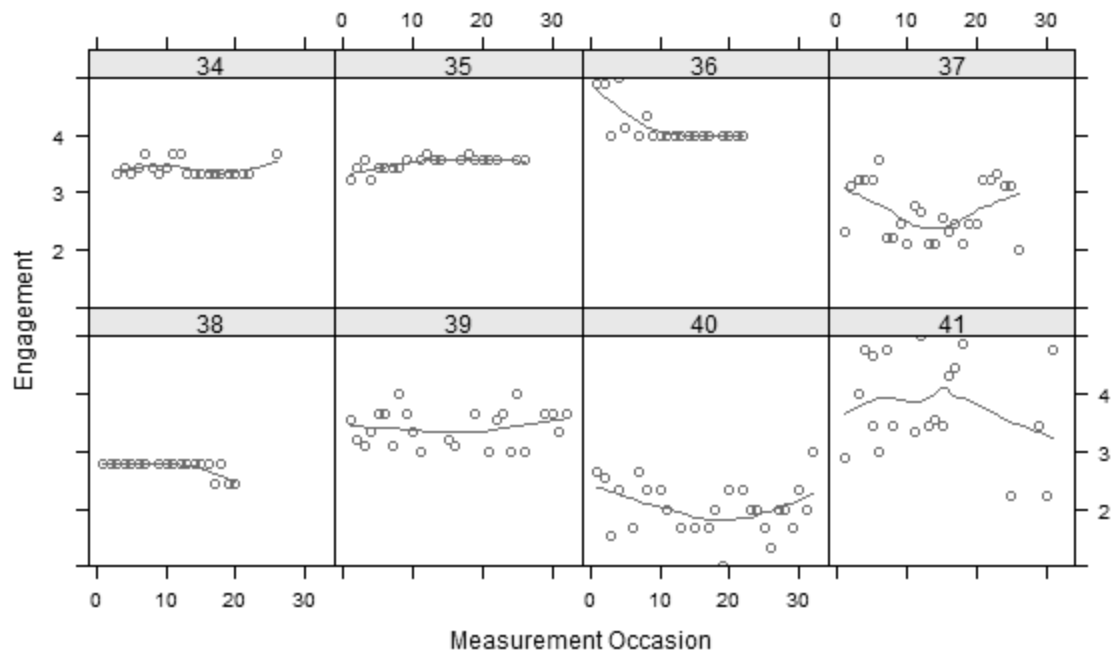


Figure 7. Line chart of moment to moment engagement for Participants 34-41.

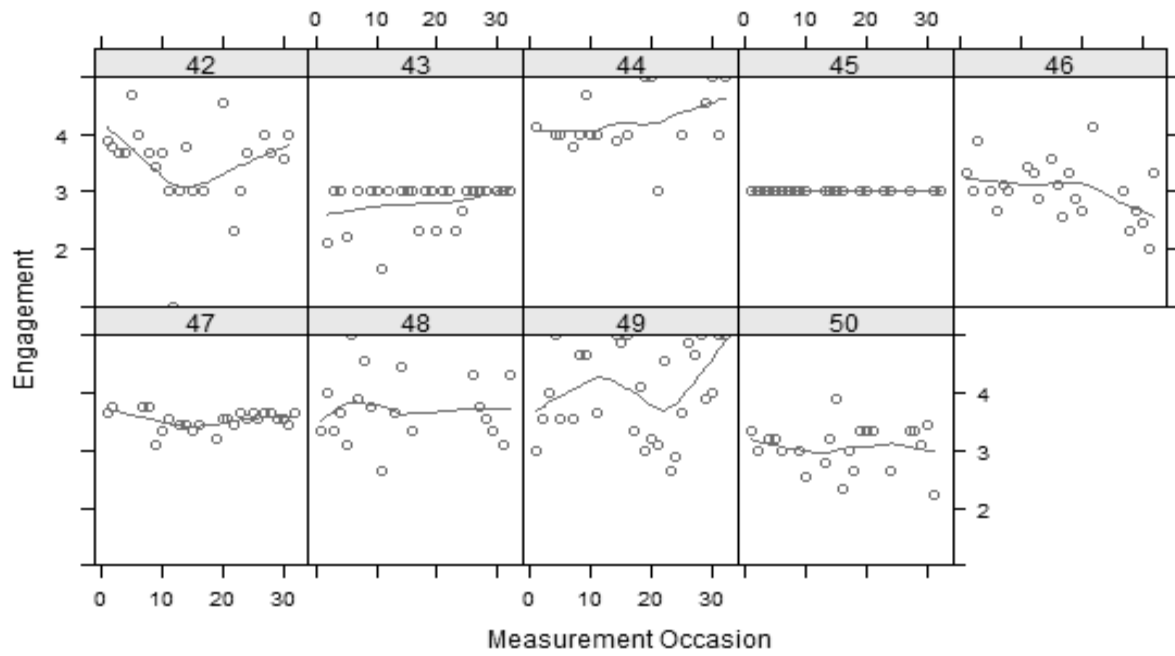


Figure 8. Line chart of moment to moment engagement for Participants 42-50.

References

- Alderfer, C. P. (1969). An empirical test of a new theory of human needs. *Organizational Behavior and Human Performance*, 4, 142-175.
- Allen, B. P., & Potkay, C. R. (1981). On the arbitrary distinction between states and traits. *Journal of Personality and Social Psychology*, 41(5), 916-928. doi: 10.1037/0022-3514.41.5.916
- Bakker, A. B. (2014). Daily fluctuations in work engagement: An overview and current directions. *European Psychologist*, 19(4), 1-10. doi: 10.1027/1016-9040/a000160
- Bakker, A. B., Albrecht, S. L., & Leiter, M. P. (2011). Key questions regarding work engagement. *European Journal of Work and Organizational Psychology*, 20(1), 4-28.
- Bakker, A. B., & Xanthopoulou, D. (2009). The crossover of daily work engagement: Test of an actor-partner interdependence model. *Journal of Applied Psychology*, 94(6), 1562-1571. doi: 10.1037/a0017525
- Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, 74(5), 1252-1265.
- Bledow, R., Schmitt, A., Frese, M., & Kühnel, J. (2011). The affective shift model of work engagement. *Journal of Applied Psychology*, 96(6), 1246-1257. doi: 10.1037/a0024532
- Bolger, N., & Laurenceau, J. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. New York, NY: The Guilford Press.

- Brauchli, R., Schaufeli, W. B., Jenny, G. J., Fülleman, D., & Bauer, G. F. (2013). Disentangling stability and change in job resources, job demands, and employee well-being – A three wave study on the job-demands resources model. *Journal of Vocational Behavior*, *83*, 117-129.
- Byrne, Z. S., Peters, J. M., & Weston, J. W. (in press). The struggle with employee engagement: Measures and construct clarification using five samples. *Journal of Applied Psychology*.
- Chalofsky, N., & Krishna, V. (2009). Meaningfulness, commitment and engagement: The intersection of a deeper level of intrinsic motivation. *Advances in Developing Human Resources*, *11*, 189-205. doi: 10.1177/1523422309333147
- Cheung, G. W., & Rensvold, R. B. (2001). The effects of parsimony and sampling error on the fit of structural equation models. *Organizational Research Methods*, *4*, 236-264. doi: 10.1177/109442810143004
- Christian, M. S., Garza, A. S., & Slaughter, J. E. (2011). Work engagement: a quantitative review and test of its relations with task and contextual performance. *Personnel Psychology*, *64*, 89-136. doi: 10.1111/peps.12070
- Cole, M. S., Walter, F., Bedeian, A. G., & O'Boyle, E. H. (2011). Job burnout and employee engagement: A meta-analytic examination of construct proliferation. *Journal of Management*, *38*(5), 1550-1581. doi: 10.1177/0149206311415252
- Culbertson, S. S., Mills, M. J., & Fullager, C. J. (2012). Work engagement and work-family facilitation: Making homes happier through positive affective spillover. *Human Relations*, *65*(9), 1155-1177. doi: 10.1177/0018726712440295
- Cranford, J. A., Shrout, P. E., Iida, M., Rafaeli, E., Yip, T., & Bolger, N. (2006). A procedure for evaluating sensitivity to within-person change: Can mood measures in diary studies

- detect change reliably? *Personality and Social Psychology Bulletin*, 32, 917-929. doi: 10.1177/0146167206287721
- Crawford, E. R., LePine, J. A., & Buckman, B. R. (2013). Job engagement scale short form items. *Unpublished*.
- Dalal, R. S., Baysinger, M., Brummel, B. J., & LeBreton, J. M. (2012). The relative importance of employee engagement, other attitudes, and trait affect as predictors of job performance. *Journal of Applied Social Psychology*, 42(1), 295-325.
- Demerouti, E., Bakker, A. B., Nachreiner, F., & Schaufeli, W. B. (2001). The job demands-resources model of burnout. *Journal of Applied Psychology*, 86, 499-512.
- 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(6), 1019-1033. doi: 10.1037/0021-9010.88.6.1019
- George, J. M. (1991). State or trait: Effects of positive mood on prosocial behaviors at work. *Journal of Applied Psychology*, 76(2), 299-307. doi: 10.1037/0021-9010.76.2.299
- Goffman, E. (1959). The presentation of self in everyday life. In E. Goffman (Ed.), *The Presentation of Self in Everyday Life* (pp. 17-25). Oxford, England. Doubleday.
- Hackman, J. R., & Oldham, G. R. (1980). *Work Redesign*. Reading, Mass.: Addison-Wesley.
- Halbesleben, J. R. B., Harvey, J., & Bolino, M. C. (2009). Too engaged? A conservation of resources view of the relationship between work engagement and work interference with family. *Journal of Applied Psychology*, 94, 1452-1465. doi: 10.1037/a0017595

- Harter, J. K., Schmidt, F. L., & Hayes, T. L. (2002). Business unit-level relationship between employee satisfaction, employee engagement, and business outcomes: A meta-analysis. *Journal of Applied Psychology, 87*(2), 268-279. doi: 10.1037//0021-9010.87.2.268
- Hormouth, S. E. (1986). The sampling of experiences in situ. *Journal of Personality, 54*(1), 262-293. doi: 10.1111/j.1467-6494.1986.tb00395.x
- Hurd, B. M. (2011). *Virtually engaged: The effect of a virtual work environment on task engagement, its antecedents, and consequences* (Unpublished Dissertation). Colorado State University, Colorado.
- IBM Corp. Released 2015. IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY: IBM Corp.
- Judge, T. A., Erez, A., & Thoreson, C. J. (2000). Why negative affectivity (and self-deception) should be included in job stress research: bathing the baby with the bath water. *Journal of Organizational Behavior, 21*, 101-111.
- Kahn, W. A. (1990). Psychological conditions of personal engagement and disengagement at work. *Academy of Management Journal, 33*(4): 692–724. doi: 10.2307/256287
- Kahn, W. A. (1992). To be fully there: Psychological presence at work. *Human Relations, 45*(4), 321-349.
- Kahn, W. A., & Fellows, S. (2014). Employee engagement and meaningful work. In B. J. Dik, Z. S. Byrne, & M. F. Steger (Eds.) *Purpose and meaning in the workplace*: 105-126. Washington D. C.: American Psychological Association.
- Kamalanabhan, T. J., Prakashsai, L., & Mayuri, D. (2009). Employee engagement and job satisfaction in the information technology industry. *Psychological Reports, 105*, 759-770. doi: 10.2466/pr0.105.3.759-770

- Keeley, J., Zayac, R., & Correia, C. (2008). Curvilinear relationships between statistics anxiety and performance among undergraduate students: Evidence for optimal anxiety. *Statistics Education Research Journal*, 7, 4-15.
- Kline, R. B. (2016). *Principles and practice of structural equation modeling, fourth edition*. New York, NY: The Guilford Press.
- Larson, R., & Csikszentmihalyi, M. (1983). The experience sampling method. *New Directions for Methodology of Social and Behavioral Science*, 15, 41-56.
- Lazarus, R. S. (1993). From psychological stress to the emotions: A history of changing outlooks. *Annual Review of Psychology*, 44, 1-21.
- Loher, B. T., Noe, R. A., Moeller, N. L., & Fitzgerald, M. P. (1985). A meta-analysis of the relation of job characteristics to job satisfaction. *Journal of Applied Psychology*, 70(2), 280-289.
- Luke, D. A. (2004). Multilevel modeling. In M. S. Lewis-Beck (Ed.), *Series: Quantitative applications in the social sciences*. Thousand Oaks, CA: Sage Publications.
- Macey, W. H., & Schneider, B. (2008). The meaning of employee engagement. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 1, 3–30. doi: 10.1111/j.1754-9434.2007.0002.x
- Mackinnon, A., Jorm, A. F., Christensen, H., Korten, A. E., Jacomb, P. A., & Rodgers, B. (1999). A short form of the positive and negative affect schedule: evaluation of factorial validity and invariance across demographic variables in a community sample. *Personality and Individual Differences*, 27, 405-416.
- Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology*, 52, 397-422. doi: 10.1146/annurev.psych.52.1.397

- May, D. R., Gilson, R. L., & Harter, L. M. (2004). The psychological conditions of meaningfulness, safety, and availability and the engagement of the human spirit at work. *Journal of Occupational and Organizational Psychology, 77*, 11-37.
- McDowall, D., McCleary, R., Meidinger, E. E., & Hay Jr., R. A. (1980). Interrupted time series analysis. In R. E. Weber (Ed.), *Series: Quantitative applications in the social sciences*. Thousand Oaks, CA: Sage Publications.
- Menguc, B., Auh, S., Fisher, M., & Haddad, A. (2013). To be engaged or not to be engaged: The antecedents and consequences of service employee engagement. *Journal of Business Research, 66*(11), 2163-2170. doi: 10.1016/j.jbusres.2012.01.007F
- Meyer, J. P., & Gagné, M. (2008). Employee engagement from a self-determination theory perspective. *Industrial and Organizational Psychology, 1*, 60-62. doi: 10.1111/j.1754-9434.2007.00010.x
- 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. doi: 10.1348/096317905X40105
- Moliner, C., Martinez-Tur, V., Ramos, J., Perio, J., & Cropanzano, R. (2008). Organizational justice and extra role customer service: The mediating role of well-being at work. *European Journal of Work and Organizational Psychology, 17*(3), 327-348. doi: 10.1080/13594320701743616
- Morgeson, F., & Campion, M. (2003). Work design. In W. Borman, D. Ilgen, & R. Klimoski (Eds.), *Handbook of Psychology: Industrial and Organizational Psychology, Volume 12* (423-452). Hoboken, NJ: Wiley.

- Muthén, L. K., & Muthén, B. O. (2011). *Mplus User's Guide: Statistical Analysis with Latent Variables: User's Guide*. Muthén & Muthén.
- Newman, D. A., & Harrison, D. A. (2008). Been there, bottled that: Are state and behavioral work engagement new and useful construct "wines"? *Industrial and Organizational Psychology, 1*, 31-35. doi: 10.1111/j.1754-9434.2007.00003.x
- Nezlek, J. B. (2012). Multilevel modeling analyses of diary-style data. In M. R. Mehl, & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 357-383). New York, NY: The Guilford Press.
- Ohly, S., Sonnentag, S., Niessen, C., & Zapf, D. (2010). Daily studies in organizational research: An introduction and some practical recommendations. *Journal of Personnel Psychology, 9*(2), 79-93. doi: 10.1027/1866-5888/a000009
- Ouweneel, E., LeBlanc, 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*(9), 1129-1154. doi: 10.1177/0018726711429382
- Payne, R. L. (2000). Comments on 'Why negative affectivity should not be controlled in job stress research: Don't throw out the baby with the bath water'. *Journal of Organizational Behavior, 21*, 97-99.
- Raudenbush, S.W., Bryk, A.S., & Congdon, R. (2004). HLM 6 for Windows [Computer software]. Skokie, IL: Scientific Software International, Inc.
- Rich, B. L. (2006). *Job engagement: Construct validation and relationship with job satisfaction, job involvement, and intrinsic motivation* (Unpublished Dissertation). University of Florida, Florida.

- Rich, B. L., LePine, J. A., & Crawford, E. R. (2010). Job engagement: Antecedents and effects on job performance. *Academy of Management Journal*, *53*(3), 617-635. doi: 10.5465/AMJ.2010.51468988
- Robinson, M. D. & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, *128*, 934-960.
- Russell, S. S., Spitzmüller, C., Lin, L. F., Stanton, J. M., Smith, P. C., & Ironson, G. H. (2004). Shorter can also be better: The abridged job in general scale. *Educational and Psychological Measurement*, *64*(5), 878-893. doi: 10.1177/0013164404264841
- Saks, A. M. (2006). Antecedents and consequences of employee engagement. *Journal of Managerial Psychology*, *21*(7), 600-619. doi: 10.1108/02683940610690169
- Schaufeli, W. B. & Bakker, A. B. (2004). Job demands, job resources, and their relationship with burnout and engagement: a multi-sample study. *Journal of Organizational Behavior*, *25*, 293-315. doi: 10.1002/job.248
- Schaufeli, W. B., Bakker, A. B. & Van Rhenen, W. (2009). How changes in job demands predict burnout, work engagement and sickness absenteeism. *Journal of Organizational Behavior*, *30*, 893-917. doi: 10.1002/job.595
- Schaufeli, W. B., Salanova, M., González-Romá, V. & Bakker, A. B. (2002). The measurement of engagement and burnout: A two sample confirmatory factor analytic approach. *Journal of Happiness Studies*, *3*, 71-92. doi: 10.1023/A:1015630930326
- Shuck, B., Ghosh, R., Zigarmi, D., & Nimon, K. (2013). The jingle jangle of employee engagement: Further exploration of the emerging construct and implications for workplace learning and performance. *Human Resource Development Review*, *12*(1), 11-35. doi: 10.1177/1534484312463921

- Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press.
- Sonnentag, S. (2011). Commentary: Research on work engagement is alive and well. *European Journal of Work and Organizational Psychology, 20(1)*, 29-38. doi: 10.1080//1359432X.2010.510639
- Spector, P. E., Zapf, D., Chen, P. Y., & Frese, M. (2000). Why negative affectivity should not be controlled in job stress research: Don't throw out the baby with the bath water. *Journal of Organizational Behavior, 21*, 79-95.
- Stanton, J. M., Balzer, W. K., Smith, P. C., Parra, L. F., & Ironson, G. (2001). A global measure of work stress: The stress in general (SIG) scale. *Educational and Psychological Measurement, 61*, 866-888.
- Steger, M. F., Dik, B. J., & Duffy, R. D. (2012). Measuring meaningful work: The Work and Meaning Inventory (WAMI). *Journal of Career Assessment, 20*, 322-337. doi: 10.1177/1069072711436160
- Steger, M. F., Littman-Ovadia, H., Miller, M., Menger, L., & Rothmann, S. (2013). Engaging in work even when it is meaningless: Positive affective disposition and meaningful work interact in relation to work engagement. *Journal of Career Assessment, 21(2)*, 348-361. doi: 10.1177/1069072712471517
- Staw, B. M., & Ross, J. (1985). Stability in the midst of change: A dispositional approach to job attitudes. *Journal of Applied Psychology, 70(3)*, 469-480.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Upper Saddle Lake, NJ: Pearson Education Inc.

- Tims, M., Bakker, A. B., & Xanthopoulou, D. (2011). Do transformational leaders enhance their followers' daily work engagement? *The Leadership Quarterly*, 22, 121-131. doi: 10.1016/j.leaqua.2010.12.011
- Tversky, A., & Kahneman, D. (1982). Evidential impact of base rates. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment Under Uncertainty. Heuristics and Biases* (153-160). New York, NY: Cambridge University Press.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3, 4-70. doi: 10.1177/109442810031002
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063–1070. doi: 10.1037/0022-3514.54.6.1063
- Weiss, H. M., & Cropanzano, R. (1996). Affective events theory: A theoretical discussion of the structure, causes and consequences of affective experiences at work. In Staw, B. M., & Cummings, L. L. (Eds.), *Research in organizational behavior: An annual series of analytical essays and critical reviews* (pp. 1-74). San Diego, CA: Elsevier Science/JAI Press.
- Xanthopoulou, D., Bakker, A. B., Demerouti, E., & Schaufeli, W. B. (2009). Work engagement and financial returns: A diary study on the role of job and personal resources. *Journal of Occupational and Organizational Psychology*, 82, 183-200. doi: 10.1348/096317908X285633

Xanthopoulou, D., Bakker, A. B., & Ilies, R. (2012). Everyday working life: Explaining within-person fluctuations in employee well-being. *Human Relations*, *65*(9), 1051-1069. doi: 10.1177/0018726712451283

Zohar, D. (1999). When things go wrong: The effect of daily work hassles on effort, exertion, and negative mood. *Journal of Occupational and Organizational Psychology*, *72*, 265-283.

Appendix A

Baseline Surveys

PANAS (Watson, Clark, & Tellegen, 1988)

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Indicate to what extent you generally feel this way, that is, how you feel on average. Use the following scale to record your answers:

1	2	3	4	5
very slightly or not at all	a little	moderately	quite a bit	extremely
	<input type="checkbox"/> interested		<input type="checkbox"/> irritable	
	<input type="checkbox"/> distressed		<input type="checkbox"/> alert	
	<input type="checkbox"/> excited		<input type="checkbox"/> ashamed	
	<input type="checkbox"/> upset		<input type="checkbox"/> inspired	
	<input type="checkbox"/> strong		<input type="checkbox"/> nervous	
	<input type="checkbox"/> guilty		<input type="checkbox"/> determined	
	<input type="checkbox"/> scared		<input type="checkbox"/> attentive	
	<input type="checkbox"/> hostile		<input type="checkbox"/> jittery	
	<input type="checkbox"/> enthusiastic		<input type="checkbox"/> active	
	<input type="checkbox"/> proud		<input type="checkbox"/> afraid	

Daily Survey Contextual Issues (O*Net; May, Gilson & Harter, 2004)

Please indicate which of the following tasks or interactions you have been a part of in the last 3 hours. Mark all that apply:

Making Decisions and Solving Problems:

- Analyzing information and evaluating results to choose the best solution and solve problems
- Please rate the extent to which these statements are true or untrue (repeated for each task):

1	2	3	4	5
Absolutely	Untrue	Neither	True	Absolutely

Untrue	True					
<input type="radio"/>	These tasks are personally meaningful to me.	1	2	3	4	5
<input type="radio"/>	These tasks are worthwhile.	1	2	3	4	5
<input type="radio"/>	These tasks are significant to me.	1	2	3	4	5

__ Processing Information

- Compiling, coding, categorizing, calculating, tabulating, auditing, or verifying information or data

__ Documenting/Recording Information

- Entering, transcribing, recording, storing, or maintaining information in written or electronic/magnetic form

__ Performing General Physical Activities

- Performing physical activities that require considerable use of your arms and legs and moving your whole body, such as climbing, lifting, balancing, walking, stopping, and handling of materials.

__ Selling or Influencing Others

- Convincing others to buy merchandise/goods or to otherwise change their minds or actions.

__ Assisting and Caring for Others

- Providing personal assistance, medical attention, emotional support, or other personal care to others such as coworkers, customers, or patients.

__ Scheduling Work and Activities

- Scheduling events, programs, and activities as well as the work of others.

__ Performing Administrative Activities

- Performing day-to-day administrative tasks such as maintaining information files and processing paperwork.

The Abridged Job in General Scale (AJIG; Russell et al., 2004)

Please use the following indications to answer these questions:

Indication	0	1	3
Meaning	“No”	“?”	“Yes”

Please indicate whether the following words or phrases describe your job in general:

- __ Good
- __ Undesirable
- __ Better than most

- Disagreeable
- Makes me content
- Excellent
- Enjoyable
- Poor

Full Job Engagement Scale (Rich, LePine, & Crawford, 2010)

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

1. I work with intensity on my job.
2. I exert my full effort to my job.
3. I devote a lot of energy to my job.
4. I try my hardest to perform well on my job.
5. I strive as hard as I can to complete my job.
6. I exert a lot of energy on my job.
7. I am enthusiastic about my job.
8. I feel energetic about my job.
9. I am interested in my job.
10. I am proud of my job.
11. I feel positive about my job.
12. I am excited about my job.
13. At work, my mind is focused on my job.
14. At work, I pay a lot of attention to my job.
15. At work, I concentrate on my job.
16. At work, I focus a great deal of attention on my job.
17. At work, I am absorbed in my job.
18. At work, I devote a lot of attention to my job.

Within-Day Surveys

PANAS (Watson, Clark, & Tellegen, 1988)

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Indicate to what extent you feel this way right now, that is, at the present moment. Use the following scale to record your answers:

1	2	3	4	5
very slightly or not at all	a little	moderately	quite a bit	extremely

interested irritable

- | | |
|--|--------------------------------------|
| <input type="checkbox"/> _distressed | <input type="checkbox"/> _alert |
| <input type="checkbox"/> _excited | <input type="checkbox"/> _ashamed |
| <input type="checkbox"/> _upset | <input type="checkbox"/> _inspired |
| <input type="checkbox"/> _strong | <input type="checkbox"/> _nervous |
| <input type="checkbox"/> _guilty | <input type="checkbox"/> _determined |
| <input type="checkbox"/> _scared | <input type="checkbox"/> _attentive |
| <input type="checkbox"/> _hostile | <input type="checkbox"/> _jittery |
| <input type="checkbox"/> _enthusiastic | <input type="checkbox"/> _active |
| <input type="checkbox"/> _proud | <input type="checkbox"/> _afraid |

Job Engagement Scale, Short Form (Crawford, LePine & Buckman, 2013)

Directions: Following are a number of statements regarding how you invest your energies at work. Read each statement carefully. Then, fill in the bubble indicating your level of agreement with each statement:

- | | | | | |
|-------------------|----------|---------|-------|----------------|
| 1 | 2 | 3 | 4 | 5 |
| Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |

These statements describe how you might focus your *physical energy* **while at work**. Right now at work...

- | | | | | | |
|---------------------------------------|---|---|---|---|---|
| 1.1 I am working with high intensity. | 1 | 2 | 3 | 4 | 5 |
| 1.2 I am exerting my full effort. | 1 | 2 | 3 | 4 | 5 |
| 1.3 I am devoting a lot of my energy. | 1 | 2 | 3 | 4 | 5 |

These statements describe how you might focus your *emotional energy* **while at work**. Right now at work...

- | | | | | | |
|--|---|---|---|---|---|
| 1.4 I am putting my emotions into what I do. | 1 | 2 | 3 | 4 | 5 |
| 1.5 I am emotionally connected. | 1 | 2 | 3 | 4 | 5 |
| 1.6 I am putting my feelings into my work. | 1 | 2 | 3 | 4 | 5 |

These statements describe how you might focus your *cognitive energy* **while at work**. Right now at work...

- | | | | | | |
|---|---|---|---|---|---|
| 1.7 I am giving my full attention to my job. | 1 | 2 | 3 | 4 | 5 |
| 1.8 I am concentrating completely. | 1 | 2 | 3 | 4 | 5 |
| 1.9 My mind is focused on the work that I do. | 1 | 2 | 3 | 4 | 5 |

Stress in General Scale (Fuller, Stanton, Fisher, Spitzmüller, Russell, & Smith, 2003)

Indicate how much each of the following words/phrases describes your job situation today based on the following scale:

1	2	3	4	5
Not at all	A little	Moderately	Quite a bit	Extremely

- Relaxed
- Comfortable
- Pushed
- Calm
- Smooth-Running
- More Stressful than I'd Like

Rewarding Co-worker Relations (May, Gilson & Harter, 2004)

Please indicate the extent to which you agree or disagree with how the following statements describe your coworker interactions in the last three hours.

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

In the last three hours...

- 1. My interactions with coworkers have been rewarding. 1 2 3 4 5
- 2. My coworkers listened to what I had to say. 1 2 3 4 5
- 3. I trusted the coworkers I interacted with. 1 2 3 4 5
- 4. My coworkers valued my input. 1 2 3 4 5
- 5. I sensed a real connection with my coworkers. 1 2 3 4 5