PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Desiree Joy Zielke

Entitled
Ecological Momentary Assessment versus Traditional Retrospective Self-Reports as Predictors
of Health-Relevant Outcomes

For the degree of   Doctor of Philosophy

Is approved by the final examining committee:

Jesse C. Stewart, Ph..D.   Kevin L. Rand, Ph.D.
                    Chair
Jane Williams, Ph.D.   Melissa Cyders, Ph.D.

To the best of my knowledge and as understood by the student in the *Research Integrity and Copyright Disclaimer (Graduate School Form 20)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Approved by Major Professor(s): Jesse C. Stewart, Ph.D.

Approved by: John Hazer, Ph.D.   05/09/2012
                    Head of the Graduate Program   Date

PURDUE UNIVERSITY
GRADUATE SCHOOL

Research Integrity and Copyright Disclaimer

Title of Thesis/Dissertation:

Ecological Momentary Assessment versus Traditional Retrospective Self-Reports as Predictors of Health-Relevant Outcomes.

For the degree of ___Doctor of Philosophy_____

I certify that in the preparation of this thesis, I have observed the provisions of *Purdue University Executive Memorandum No. C-22,* September 6, 1991, *Policy on Integrity in Research.**

Further, I certify that this work is free of plagiarism and all materials appearing in this thesis/dissertation have been properly quoted and attributed.

I certify that all copyrighted material incorporated into this thesis/dissertation is in compliance with the United States' copyright law and that I have received written permission from the copyright owners for my use of their work, which is beyond the scope of the law. I agree to indemnify and save harmless Purdue University from any and all claims that may be asserted or that may arise from any copyright violation.

Desiree Joy Zielke
_____
Printed Name and Signature of Candidate

05/09/2012
_____
Date (month/day/year)

ECOLOGICAL MOMENTARY ASSESSMENT VERSUS TRADITIONAL

RETROSPECTIVE SELF-REPORTS AS PREDICTORS OF HEALTH-RELEVANT

OUTCOMES

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Desiree Joy Zielke

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

August 2012

Purdue University

Indianapolis, Indiana

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Jesse C. Stewart for his guidance, patience, and support with this research project. I would also like to thank my other committee members, Dr. Melissa Cyders, Dr. Kevin L. Rand, and Dr. Jane Williams, for their time and assistance. I would also like to thank M. E. Lee for providing her awarded funding to this research project and her assistance in programming the equipment. In addition, I would like to thank both M. E. Lee and Tabitha Ratcliffe for their hard work in assisting with data collection and data cleaning.

TABLE OF CONTENTS

LIST OF TABLES

ABSTRACT

Zielke, Desiree Joy. Ph.D., Purdue University, August 2012. Ecological Momentary Assessment versus Traditional Retrospective Self-Reports as Predictors of Health-Relevant Outcomes. Major Professor: Jesse C. Stewart.

Ecological momentary assessment (EMA) has been asserted by proponents of the technique as being superior to standard paper-and-pencil measurements in terms of the reliability and validity of the information obtained; however, this claim has not yet been fully evaluated in the literature. Accordingly, the purpose of this study was to evaluate one aspect of this assertion by comparing the utility of EMA and retrospective measures of depressive symptoms in predicting health-relevant biological and behavioral outcomes. It was hypothesized that (1) the EMA measure will have better predictive utility when examining objective sleep quality (a biological outcome), and that (2) the retrospective measure will have better predictive utility when examining blood donation intention (a behavioral outcome). Ninety-six undergraduate females participated in this 2-week study. Depressive symptoms were measured momentarily and retrospectively using the Center for Epidemiological Studies-Depression Scale (CES-D). The biological outcome was assessed by actigraphy, whereas the behavioral outcome was measured via a self-report questionnaire. Unfortunately, it was not possible to fully test these hypotheses due to the

failure to observe relationships between the predictor variables and the outcomes. The reported results, although limited, did not provide support for the hypotheses. Supplemental analyses revealed a moderate to high amount of shared variance between the EMA and retrospective measures, a similar extent of random error in both measures, and potentially a greater degree of systematic error in the retrospective measure. Due to the paucity of literature examining the claim of superior reliability and validity of EMA versus retrospective measures, as well as the failure of the current study to evaluate this assertion sufficiently, it appears that this claim remains unfounded. Therefore, suggestions for future research are provided.

INTRODUCTION

Ecological momentary assessment (EMA), also referred to as experience sampling, is a broadly defined assessment method in which repeated measures of variables of interest are obtained in real-time from individuals in their natural environment (Shiffman, Stone, & Hufford, 2008). EMA has been asserted by proponents of the technique as being superior to standard paper-and-pencil measurements in terms of the reliability and validity of the information obtained (Stone, Shiffman, Atienza, & Nebeling, 2007); however, this claim has not yet been fully evaluated. Accordingly, the overall objective of this project is to examine whether the predictive utility of self-reports obtained by EMA is superior to that of self-reports obtained by traditional retrospective measures for health-relevant outcomes.

This manuscript begins with a discussion of predictive utility and measurement error. Predictive utility is the ability of a measure to predict an outcome of interest (McDonald, 1999). The aspects of measurement error that are discussed in relation to predictive utility are random error and systematic error. Information will be presented to illustrate that measures that contain less measurement error should have greater predictive utility. Next, descriptions and a review of the strengths and weaknesses of traditional and EMA measures are provided. Following these sections, the

available empirical evidence is reviewed. After this evidence is presented, it is evaluated to determine whether EMA or traditional retrospective measures have superior predictive utility. Major limitations of this literature are also identified.

The current project, which sought to address the identified limitations, is then discussed. This cross-sectional, observational study investigated the utility of both an EMA measure and a traditional retrospective measure of depressive symptom severity in predicting health-relevant biological and behavioral outcomes that have previously been associated with depression. The biological outcome was objective sleep quality and the behavioral outcome was intention to donate blood.

Overview of Predictive Utility, Retrospective Self-Reports, and EMA Self-Reports

Predictive Utility and Measurement Error

Predictive utility is the ability of a measure to predict an outcome of interest (McDonald, 1999). Classical test theory states that the observed score on a measure is equal to the sum of the true score and the measurement error (T. Kline, 2005; Shultz & Whitney, 2005). Measurement error consists of both systematic error and random error (R. Kline, 2005); although it should be noted that systematic error is not consistently presented in the classical test theory literature as part of measurement error. Measures that contain less measurement error should have greater predictive utility when the measures are equally related to outcomes being predicted. To reduce the discrepancy between the true score and the observed score, it is necessary to decrease systematic error, random error, or both.

Systematic error is error that has a consistent biasing effect on the observed scores (Carmines & Zeller, 1979). Specifically, systematic error influences the observed scores in such a way that they are all increased or decreased relative to the true score (Kerlinger, 1986). Systematic error can occur when an individual responds in a socially desirable manner. In this situation, scores would be consistently overestimated for all desirable items and consistently underestimated for all undesirable items (Niemi, 1993). When an individual retrieves information from memory, systematic error can also occur. For example, when an individual provides a pain rating following a medical procedure, their rating could be higher than the actual overall pain experienced due to the tendency to remember the most painful part of the experience. Finally, changes in the environment during administration of a measure can also introduce systematic error. For example, loud traffic outside a classroom where individuals are completing the measure could systematically decrease the observed scores due to a reduced ability to concentrate. As illustrated by these examples, systematic error can impede a measure's ability to accurately assess the target construct. Therefore, systematic error can decrease the construct validity of the measure (Carmines & Zeller, 1979).

Construct validity can have an effect on predictive utility. If it has been established in the literature that a construct is predictive of a particular outcome, a measure that has high construct validity should have a stronger relationship to that outcome than a measure that has low construct validity. In other words, the measure with higher construct validity should have greater predictive utility when the construct is related to the outcome of interest. Importantly, the construct validity of a measure can be increased by decreasing its systematic error (Schwab, 1999). Of note, decreasing

systematic error will improve predictive utility only when systematic error varies across individuals. When systematic error is constant, it will have no effect on predictive utility. Furthermore, if the systematic error of the measure is predictive of the outcome of interest, predictive utility will increase as systematic error increases.

A measure must also be reliable in order for it to be valid (Nunnally, 1978). Reliability is the extent to which scores on a measure are repeatable and stable over a wide range of conditions; therefore, measures that are more consistent have a higher reliability (Carmines & Zeller, 1979; Nunnally, 1978). For a measure to be reliable, the random error associated with the measure must be low (Carmines & Zeller, 1979). Random error is unsystematic error that consists of all the influences that occur by chance in a testing situation and that interfere with the measurement of the true score. Potential sources of random error include: (a) fatigue, (b) guessing, (c) ambiguously written items, and (d) fluctuations in mood or memory (Kerlinger, 1986). If a measure with high reliability is administered many times, the random errors form a normal distribution around the true score. Therefore, random error does not bias the observed score in a particular direction and the expected value of the mean of the distribution of random errors is zero (T. Kline, 2005).

The reliability of a measure can be improved by reducing its random error. Two ways to reduce random error are standardization and aggregation (Strube, 2000). Standardization refers to the careful control of the measurement conditions so that unrelated sources of error do not influence the observed scores (Strube, 2000). Giving instructions for a measure in the same way to every participant is an illustration of standardization. Aggregation refers to the process of calculating the average score across

multiple administrations of a measure or the average score across multiple items of a measure (Strube, 2000). Aggregation improves reliability by suppressing random error through replication (Strube, 2000). That is, when more items are included in a measure or when more administrations are given, the aggregation of the assessments allows for the random sources of error that contaminate each observed score to cancel each other out, thereby leaving a better estimate of the true score (Strube, 2000).

Reliability coefficients help to estimate both the true score variance and the error variance associated with the observed score (Shultz & Whitney, 2005). Because it is not possible to calculate the true score variance directly, the reliability coefficient for a predictor (r_{xx}) is equal to one minus the proportion of variance due to random error (R. Kline, 2005). A large value for r_{xx} indicates that the predictor is reliable, as the proportion of variance due to random error is small. As r_{xx} approaches zero, the observed scores begin to represent random numbers. Therefore, a small value for r_{xx} indicates that the predictor is unreliable.

The reliability coefficient can be used to determine the maximal relationship (that can be measured) between a predictor (x) and an outcome (y). There is perfect relationship between the predictor and the outcome when the value of the observed validity coefficient of the predictor and the outcome (r_{xy}) is equal $\pm$ 1.00. Such a relationship can occur only if the scores of both the predictor and the outcome are perfectly reliable, given that the theoretical absolute value of $r_{xy} = \sqrt{r_{xx} \times r_{yy}}$ (R. Kline, 2005). This formula illustrates that when the scores of either the predictor or the outcome are unreliable, the correlation between them is attenuated. Stated another way, the predictive utility of a measure will decrease as its random error increases.

Retrospective Self-Report Measures

Traditional retrospective self-report measures require individuals to reflect on past events and report on their emotions, cognitions, behaviors, and/or experiences from those events (Kazdin, 2003). These reports are usually obtained using paper-and-pencil questionnaires and the typical length of recall is the past week or month. Reports can involve providing specific information for a particular amount of time, such as reporting emotions experienced during a recent argument with a friend. Reports can also involve providing global assessments of typical affective states and behaviors based on an accumulation of previous experiences and knowledge. Information obtained from retrospective measures can vary in the degree of subjectivity that is requested. Some reports are intentionally subjective, such as the amount of love you feel in your relationship with your spouse or the intensity of a cigarette craving. Other retrospective data are more objective, such as the number of cigarettes smoked in a day. As may be evidenced by the description of retrospective self-report, the data obtained from these measures typically represent the endpoints of change rather than the process that occurred (Metts, Sprecher, & Cupach, 1991).

Retrospective self-reports are widely used in clinical and research settings due to the advantages associated with them. Specifically, retrospective measures: (a) can be developed quickly, (b) are convenient and efficient, (c) are easy to administer and complete, (d) are a direct assessment of the individual's perception of their cognitions, emotions, and experiences, and (e) provide a comprehensive representation of the individual's everyday functioning (Gorin & Stone, 2001; Kazdin, 2003).

As with any assessment method, there are important limitations of retrospective measures. One of the most widely made criticisms of retrospective measures is related to memory bias. When individuals are asked to make retrospective reports, they rely on memories they have of their personal experiences to summarize information regarding past events (Tourangeau, 2000). These memories are constructed at the time of retrospection. These memories not only consist of the original memory but also of (a) logical inferences that are used to fill in missing details, (b) memories that are associated with the original memory, and (c) other relevant information (Plous, 1993). Because memory recall is often more of a process of reconstruction than of retrieval, it is subject to biases at the time of recall.

Recall biases are primarily due to the processes used to retrieve information and the cognitive heuristics used to estimate and summarize information. These biases are most likely to occur when individuals are requested to retrospect over longer periods of time, such as a week or a month (Smyth & Stone, 2003). Memory retrieval can be heavily influenced by the individual's situation and mental state at the time of recall. Additional recall biases include recency and saliency. Recency refers to the process in which events that have occurred most recently in the individual's life are more accessible to memory and are more likely to be retrieved (Yoshiuchi, Yamamoto, & Akabayashi, 2008). Saliency is when experiences that are the most memorable are more likely to be recalled (Yoshiuchi et al., 2008). These cognitive heuristics are very convenient when interacting and recalling in the everyday world. However, when being asked to accurately retrospect for research purposes, cognitive heuristics are less desirable and more problematic (Stone et al., 2007).

An additional limitation of retrospective measures is that researchers using these measures do not typically request information regarding the variability of moods or behaviors over time and across situations (Shiffman et al., 2008). Retrospective measures typically involve only one assessment at one point in time, thus missing information from real-world situations and circumstances that play out in day-to-day life.

EMA Self-Report Measures

EMA is an assessment method in which repeated measures of variables of interest are obtained in real-time from individuals in their natural environment (Stone et al., 2007). Although the nature of data collected using this technique varies from study to study, EMA frequently includes momentary self-reports of emotions, cognitions, behaviors, and/or experiences. In EMA studies, individuals are instructed to respond to items presented on a hand-held computer (e.g., a personal digital assistant or a smartphone) after being signaled by the device at various times during a monitoring period.

EMA was initially developed as an alternative to retrospective self-report with the hope that recall bias would be minimized, given that the assessments occur through immediate reports of current experiences (Smyth & Stone, 2003; Stone & Shiffman, 2002). Additional goals of EMA are maximization of ecological validity and the study of microprocesses that may influence behavior in the natural environment (Stone et al., 2007). EMA also contributes to research because it allows data to be aggregated across time to provide information regarding the individual's typical condition. Furthermore, it can be used in cross-sectional research designs to analyze contextual associations (e.g.,

emotional responses associated with stress) or different aspects of experience between two or more events that co-occur in time. EMA can also be used to determine the order in which events typically occur (Shiffman et al., 2008).

Studies using EMA methodology have numerous advantages over investigations using traditional retrospective measures. In EMA studies, behaviors and experiences are reported in the real-world and, therefore, are considered to be ecologically valid. These real-world assessments may generate data that is more generalizable to an individual's daily life, whereas reports acquired in a laboratory setting may be artificial due to being taken in an environment that is far removed from the real-world of the individual (Shiffman et al., 2008; Smyth & Stone, 2003). Another key feature of EMA is that it attempts to minimize recall biases that occur through retrospection and summary processes by assessing behaviors and experiences of interest in the moment, or close to the moment, during which they occurred (Stone & Shiffman, 2002). The use of EMA methodology allows for multiple assessments across a wide variety of situations and time periods. Multiple assessments allow for the aggregation of momentary reports to characterize an experience rather than relying on the individual's summary of the experience (Stone & Shiffman, 2002). Another important contribution is the fact that EMA allows more detailed information to be obtained regarding dynamic processes and situational influences on individual's behaviors, cognitions, and emotions (Shiffman et al., 2008). This detailed information permits researchers to examine acute or short-lived effects that may not be detected when relying on summary data (Shiffman & Stone, 1998).

In addition to EMA's advantages, there are also some important limitations of this assessment method. One potential limitation that has prompted a fair amount of research is reactivity to the assessment. Reactivity refers to the notion that the frequent recording of an individual's behavior or experience may lead to changes in that behavior or experience. To date, the research examining whether reactivity occurs when using EMA is mixed (Shiffman et al., 2008). It has been shown that reactivity can occur when individuals are trying to change the behavior that is the object of study. Additionally, when individuals report the behavior before they perform it, it may provide them an opportunity to control the behavior (Scollon, Kim-Prieto, & Diener, 2003). On the other hand, several studies have found little or no evidence of reactivity when reporting on pain intensity (Shiffman et al., 2008; Stone, Broderick, Schwartz, Shiffman, Litcher-Kelly, & Calvanese, 2003).

Other limitations associated with EMA include (a) participant burden, (b) sample bias, (c) poor compliance, (d) difficulty in the assessment of rare experiences, and (e) special populations' ability to use the technology. Completing EMA measures multiple times a day, especially during a busy day, can be a burden on participants. In addition, being signaled multiple times a day can prove to be irritating. Due to EMA being burdensome, studies using this methodology may unintentionally have sample bias because certain types of people may be more willing to bear the burden of multiple assessments. There could be additional bias in the data due to noncompliance if participants do not respond to every prompt and only complete reports when prompted at convenient times (Shiffman & Stone, 1998). Experiences or behaviors that occur rarely are difficult to capture when using EMA. Finally, certain groups of people (e.g.,

individuals with visual or hearing impairments) may have a more difficult time using the technology associated with EMA (Stone et al., 2007).

When comparing retrospective measures with EMA measures, differences between the two methodologies become even more apparent. It has been shown that retrospective measures have a tendency to generate higher scores for certain constructs than EMA ratings. For example, symptoms are usually described as more intense, longer lasting, and occurring more frequently on retrospective measures (Shiffman et al., 2008). In addition, there are inconsistent findings regarding the correspondence between EMA and retrospective measures. To illustrate this inconsistency, the correlation between aggregated EMA measures and retrospective measures ranges from 0.20 for headache intensity, frequency, and duration to 0.70 for pain intensity (Shiffman et al., 2008).

EMA and Retrospective Self-Reports, Measurement Error, and Predictive Utility

The information presented thus far allows the following hypotheses to be made regarding the relationships among EMA and retrospective self-reports, measurement error, and predictive utility. EMA measures might have superior predictive utility when compared to retrospective measures due to increased reliability because of less random error. Random error might be decreased because there are more assessments being administered, which are subsequently aggregated. In addition, EMA measures might have superior predictive utility because of improved construct validity due to less systematic error. Systematic error might be decreased because there is less recall bias, a lower likelihood of cognitive heuristic use, and less memory decay.

Conversely, retrospective measures might have superior predictive utility when compared to EMA measures due to increased reliability because of less random error. Random error might be decreased because measures are typically administered under controlled conditions and also might contain more items. Furthermore, retrospective measures might have superior predictive utility because of improved construct validity due to less systematic error. Systematic error might be decreased if a measure contains more items that are representative of the construct, which allows for a more complete assessment of the construct. Another reason why retrospective measures might have superior predictive utility is because the systematic error might be predictive of the outcome. A qualitative literature review was undertaken in order to determine whether EMA or retrospective measures have superior predictive utility for health-relevant outcomes. The methods and findings of this review are discussed in the following section.

Qualitative Literature Review

Review of Studies

A comprehensive literature search for articles written in English was conducted at two different times, February 22, 2009 and August 14, 2009 using the PsychInfo and MedLine databases. The search terms were the keywords, *ecological momentary assessment* and *experience sampling*. A total of 479 articles were identified on February 22, and a total of 548 articles were identified on August 14. Additional articles were

found by reviewing the reference sections of the relevant articles that were identified and through monitoring the publication release alerts of relevant journals.

Studies were examined in which both an EMA and a retrospective measure of the same construct were administered to the same participants. Additionally, these studies needed a health-relevant outcome being predicted by both an EMA and a retrospective measure. Studies that had only one momentary assessment per day were included due to the small number of investigations identified. There were seven studies that met these inclusion criteria. Appendix A contains diagrams summarizing the primary results of each study presented in the next section.

Evidence of the Superior Predictive Utility of EMA Self-Report Measures

Four studies provide evidence that EMA measures have superior predictive utility. Two studies conducted by Kamarck and colleagues (Kamarck, Muldoon, Shiffman, & Sutton-Tyrell, 2007; Kamarck, Muldoon, Shiffman, Sutton-Tyrrell, Gwaltney, & Janicki, 2004) examined whether demand and control in the workplace are associated with subclinical cardiovascular disease and its progression. Participants on whom the following analyses are based were 152 employed older adults enrolled in the Pittsburgh Healthy Heart Project. Participants completed two EMA periods, during which demand and control in the workplace were assessed. EMA measures of demand and control were obtained using different items than the retrospective measures of demand and control, which followed the EMA periods. Ultrasound assessments were performed to determine the mean carotid intima-medial thickness (IMT), which is a measure of subclinical cardiovascular disease (Mancini, Dahlof, & Diez, 2004). The

2004 study found that the EMA measure of demand in the workplace was more strongly associated with carotid IMT than the retrospective measure. The 2007 study's findings were for the 88 men who were employed, as the associations between momentary demand and control and carotid IMT progression were not significant among employed women. The results of the 2004 study were replicated in the 2007 study in that the EMA measure of demand in the workplace predicted progression of carotid IMT, whereas the retrospective measure did not (see Appendix A).

Sonnenschein and colleagues (2007) examined the relationship between hypothalamic-pituitary-adrenal (HPA) axis function and clinical burnout symptoms among 42 individuals. The retrospective questionnaire was the Dutch version of the Maslach burnout inventory-general survey (MBI-GS), which has an exhaustion subscale. The EMA self-report exhaustion item was based on the emotional exhaustion item from the MBI-GS. HPA axis function was assessed by the dexamethosone suppression test (DST) using saliva samples collected during the two weeks of EMA monitoring. The EMA measure of exhaustion predicted the increase in cortisol levels after dexamethasone intake in a way that indicated a hypoactive HPA axis, whereas the retrospective measure did not predict any of the endocrine measures (see Appendix A).

In a study on smoking, Shiffman (2009) compared three types of self-report regarding cigarette consumption and then examined their ability to predict two biochemical indices of smoke exposure: cotinine levels and carbon monoxide levels. Participants in this study were 232 smokers enrolled in a smoking cessation study. The retrospective assessments included (a) a global retrospective report on the average number of cigarettes smoked per day and (b) time-line follow-back (TLFB) measures at

the end of each week, which required participants to report the number of cigarettes they smoked each day during the past week. The EMA measure consisted of participants pushing a button on the PDA each time they smoked a cigarette. Participants also reported any cigarettes they missed reporting during the day at the end of each day. At each clinic visit, participants provided a breath sample to measure carbon monoxide, and at the end of the first week, participants provided a saliva sample to assess cotinine concentration levels. Results revealed that the EMA measure predicted cotinine and carbon monoxide levels, whereas the retrospective measures did not (see Appendix A).

Evidence of the Superior Predictive Utility of Retrospective Self-Report Measures

In contrast to the above studies, two studies provide evidence that retrospective measures have superior predictive utility. In the first study, Helgeson, Lopez, and Kamarck (2009) investigated the consequences of the positive and negative aspects of friend relationships on psychological well-being and diabetes-related outcomes. A total of 76 adolescents with diabetes completed EMA measures at two different times which were separated by one month, as well as retrospective self-report measures at baseline. The predictor variables were interaction enjoyment and interaction upset, which were aggregated across the four days of the EMA period, and friend support and friend conflict, which were assessed retrospectively. The outcome of interest, for the purposes of this review, is metabolic control (most recent hemoglobin A1c), as it was the only outcome that was not retrospectively reported on. Aggregate interaction enjoyment (EMA), aggregate interaction upset (EMA), and friend support (retrospective) were not

predictors of metabolic control; however, friend conflict (retrospective) was a predictor of poorer metabolic control (see Appendix A).

In another study, Redelmeier, Katz, and Kahneman (2003) examined whether influencing the memory of patients undergoing a painful medical procedure predicted whether or not they returned for follow-up procedures. Participants were 682 patients undergoing a colonoscopy who were randomly assigned to a control or modified care group. The modified care intervention was designed to minimize pain during the final minutes of the procedure so that patients would have a more positive memory of the experience. Throughout the procedure, patients provided momentary reports regarding their *current level of pain* at 1-minute intervals. Following the procedure, patients were asked to rate the *total discomfort* of the procedure and how they would rank that procedure relative to eight other unpleasant personal events. Patients were followed for almost six years to determine whether or not they presented for a follow-up colonoscopy. Retrospective ratings of pain were associated with decreased return rates, whereas momentary pain ratings were not (see Appendix A). This study suggests that a person's memory for the pain experienced, rather than the actual pain experienced, determines their willingness to undergo the procedure again.

Mixed Findings

Steptoe, Gibson, Hamer, and Wardle (2007) examined whether positive affect predicts cortisol and cardiovascular responses to stressful laboratory tasks. In this study, 73 non-smoking employed men attended two stress testing sessions, during which blood pressure and heart rate data were obtained. At both sessions, participants also completed

the Positive and Negative Affect Schedule (PANAS), which is a retrospective measure. The EMA measure of positive affect asked participants to report their current level of happiness four times a day for two days. During the EMA period, participants collected saliva samples four times during one working day. The EMA measure of happiness, but not the retrospective measure of positive affect, predicted cortisol responses to stress during the one working day. However, there were mixed findings for cardiovascular responses to stress (see Appendix A). Overall, there were four findings in which retrospective positive affect predicted the outcomes of interest and nine findings in which EMA positive affect predicted the outcomes of interest.

Conclusions and Future Directions

To summarize, there were five studies that found EMA measures to be predictive of the outcome (Kamarck et al., 2004; Kamarck et al., 2007; Shiffman, 2009; Sonnenschein et al., 2007; Steptoe et al., 2007), two studies that found retrospective measures to be predictive of the outcome (Helgeson et al., 2009; Redelmeier et al., 2003), and one study in which both EMA and retrospective measures predicted the outcome (Steptoe et al., 2007). Interestingly, the five studies with a biological outcome found that EMA measures were predictive of the outcome, whereas retrospective measures were not. These outcomes involved physiological processes over which individuals have no conscious control. The outcomes included presence and progression of subclinical cardiovascular disease, cortisol levels after dexamethasone administration, biochemical indices of smoke exposure (i.e., cotinine levels and carbon monoxide levels), and cortisol responses to stress. In the two cases where retrospective measures were found to be

predictive of the outcome, the outcomes involved the behavior of the participants. One outcome was colonoscopy return rate. The other outcome was adolescents' ability to control their diabetes by engaging in self-care behaviors, such as taking insulin and monitoring food consumption.

This pattern of findings suggests that retrospective measures may be stronger predictors of behavioral outcomes because summary perceptions may be more important determinants of behavioral outcomes than reality. An example would be the decision to return for a colonoscopy. The individual may use the biased information from memory to make their decision because the experience itself has passed. Because information from memory may be more negative (i.e., the colonoscopy was painful and uncomfortable), the individual may choose not to return for a colonoscopy. For biological outcomes, it is likely that reality is more important than summary perceptions. An example would be job stress and carotid IMT progression. Actual exposure to job stress may be a more important determinant of carotid IMT progression than the perception of job stress. It is reasonable to tentatively conclude that EMA measures may be stronger predictors of biological outcomes, whereas retrospective measures may be stronger predictors of behavioral outcomes.

Only a small number of studies were examined due to the paucity of investigations in which EMA and retrospective measures were compared head-to-head as predictors of the same outcome. The lack of literature on this topic makes it difficult to develop any firm conclusions. A second limitation of the reviewed studies is that the difference between effect sizes was not examined. The studies compared each effect size to zero instead of comparing the EMA and retrospective self-report effect sizes to each

other. If the effect size of an EMA measure was statistically larger than the effect size of a retrospective measure, this would provide more definitive evidence that EMA measures are stronger predictors than retrospective measures. A third limitation is that in four studies the EMA and retrospective measures of the construct were not identical. Specifically, the wording of items and/or the response options were different for the EMA and retrospective measures. It is difficult to determine whether the differences in item wording and response options influenced the results. A fourth limitation is that it is not clear from the existing literature and this review if EMA measures contain less random error. No study reviewed reported any type of reliability coefficient. Reliability coefficients, particularly test-retest and internal consistency, would be helpful in determining the amount of random error present in the EMA and retrospective measures.

An overview of the methodological features of the ideal study illustrates many of the future directions that research in this area should take. The first component of the ideal study is the use of EMA and retrospective measures with the exact same items and response options. There would be multiple outcomes being predicted by the EMA and retrospective measures and the effect sizes would be compared with each other, not just with zero, to directly answer the predictive utility question. A correlation would be conducted between the two measures to ensure that they are not redundant. Mean differences for the two measures would also be examined in order to evaluate systematic error associated with each assessment method. The retrospective measure would be administered multiple times in order to obtain test-retest reliability. Test-retest reliability and internal consistency coefficients would be computed for the EMA and retrospective measures. These values would allow for an evaluation of random error associated with

each approach. To date, no study has contained all of these methodological features. This illustration of the ideal future study is the foundation for the present study.

Present Study

The studies presented above make it apparent that more research is needed to determine whether EMA or retrospective measures have superior predictive utility for biological and behavioral outcomes. The current study examined the utility of both an EMA measure and a traditional retrospective measure of depressive symptom severity in predicting health-relevant biological and behavioral outcomes that have previously been found to be related to depression. The biological outcome was objective sleep quality and the behavioral outcome was intention to donate blood.

Depression

Depression is the leading cause of disability in the United States in individuals ages 15-44 (National Institute of Mental Health, 2008). In any given year, approximately 6.7% of American adults suffer from depression (National Institute of Mental Health, 2008). Individuals with major depressive disorder (MDD) typically experience the following symptoms: (a) depressed mood, (b) loss of interest or pleasure in activities, (c) difficulties with either sleeping too much or not being able to sleep, (d) fatigue or loss of energy, (e) feelings of worthlessness, (f) psychomotor agitation or retardation, (g) diminished ability to think or concentrate, and/or (i) recurrent thoughts of death, recurrent suicidal ideation, or a suicide attempt (American Psychiatric Association, 2000). Symptoms of MDD can be present in individuals without a clinical diagnosis. These

depressive symptoms are important to study because they may be pre-cursors to MDD. In addition, subclinical depressive symptoms have been shown to increase risk of health conditions, including cardiovascular disease (Suls & Bunde, 2005).

Depression is also the number one cause of disability among college-aged individuals (National Institute of Mental Health, 2008). It has been demonstrated that, in a typical sample of undergraduate students, there is good variability in the presence and amount of depressive symptoms (Hawkins, Stewart, & Fitzgerald, 2010; Radloff, 1991; Stewart & Stines, 2008). In a study conducted at Indiana University-Purdue University Indianapolis (IUPUI), the mean Beck Depression Inventory-Second Edition (BDI-II) score was 9.6 (SD = 7.8) (Stewart & Stines, 2008). This mean and standard deviation indicate slightly elevated depressive symptoms with acceptable variability in the scores. In a study of 214 college students, the average Center for Epidemiologic Studies-Depression Scale (CES-D) score was 15.5 (SD = 9.7), again suggesting slightly elevated depressive symptoms with acceptable variability (Radloff, 1991). Therefore, it is reasonable to examine this construct in an undergraduate sample as a predictor of a biological and behavioral outcome related to depression.

Depression and Sleep Quality

Approximately 50% to 60% of young adults aged 21 to 30 with depression report difficulties with sleep (Nutt, Wilson, & Paterson, 2008). Individuals with MDD typically report difficulties falling asleep, staying asleep, early morning awakenings, nonrestorative sleep, decreased sleep duration, disturbing dreams, and daytime fatigue (Benca, 2005). Depressed individuals have been shown to have prolonged sleep latency,

increased wakefulness during sleep, and early morning awakenings as measured by polysomnography (Benca, 2005). Poor sleep quality is a health-relevant outcome because various indicators of poor sleep quality have been associated with a greater risk of hypertension, stroke, diabetes, and obesity (Ayas et al., 2003; Gangwisch, Malaspina, Boden-Albala, & Heymsfield, 2005; Kawakami, Takatsuka, & Shimizu, 2004; Schwartz, Anderson, Cole, Cornoni-Huntley, Hays, & Blazer, 1999; Walsh, Dement, & Dinges, 2005). In addition, difficulties with sleep have been associated with an increased risk of mortality (Hublin, Partinen, Koskenvuo, & Kaprio, 2007).

Sleep difficulties have traditionally been measured using subjective reports, which assess perceived sleep quality (Rotenberg, Indursky, Kayumov, Sirota, & Melamed, 2000). A widely used subjective measure is the Pittsburgh Sleep Quality Index, which asks individuals to report on their sleep quality, sleep duration, sleep latency, sleep habits, sleep disturbances, daytime dysfunction, and use of sleep medication for the past month (Buysse, Reynolds III, Monk, Berman, & Kupfer, 1989). Sleep duration is defined as the total time an individual is asleep, whereas sleep latency is the length of time it takes an individual to fall asleep once they are in bed (Caldwell & Redeker, 2009). On subjective measures, individuals with sleep difficulties tend to under-report sleep duration and overestimate the time they are awake during the night as compared to individuals without sleep difficulties (Means, Edinger, Glenn, & Fins, 2003). In addition, Tsuchiyama and colleagues (2003) found that (a) individuals with depression do not estimate their sleep duration accurately and (b) their estimation worsens as the objective sleep disturbances become more severe. Therefore, subjective measures of sleep provide inconsistent, unreliable reports of sleep quality.

Objective measures of sleep quality can be obtained by using polysomnography or actigraphy. Polysomnography involves the application of electrodes to an individual's face and scalp in order to measure brain activity, eye movement, and muscle activity (Edinger, Means, Stechuchak, & Olsen, 2004). Actigraphy involves using a device that can be attached to a belt around the waist, wrist, or ankle. This device collects data on truncal movement activity over extended periods of time, from which various indicators of sleep quality can be derived. In a study comparing polysomnography (the gold standard in objective sleep assessment), actigraphy, subjective sleep logs, and a device that measures eye movement, it was found that the actigraphic measures of time in bed and sleep latency were most strongly correlated with polysomnographic measures of the same dimensions (rs = 0.997 and 0.870, respectively) (Edinger et al., 2004).

Studies that have examined the relationship between depression and objective sleep quality, as measured by actigraphy, have found that individuals with depression symptoms or clinical depression have poorer objective sleep than non-depressed individuals. For instance, Kawada, Katsumata, Suzuki, and Shimizu (2007) examined actigraphic sleep duration in college students with and without depressive symptoms. They found that college students who reported more severe depressive symptoms had significantly shorter sleep duration than college students who reported fewer depressive symptoms.

In another study examining objective sleep quality and depression, Joffe and colleagues (2009) used actigraphy to measure objective sleep quality in depressed and non-depressed women with the vasomotor symptoms of hot flashes and night sweats. Depression was diagnosed using the Structured Clinical Interview for the DSM-IV Axis I

Disorders (SCID-I). Women diagnosed with depression had shorter sleep duration by 47.7 minutes and longer sleep latency by 13.8 minutes as compared to non-depressed women. Korszun and colleagues (2002) also examined actigraphic measures of sleep quality, although in a different sample which included healthy controls, adults with fibromyalgia only, adults with depression only, and adults with fibromyalgia and depression. A diagnosis of depression was determined by using the SCID-I. There were significant differences among groups for nighttime activity level and sleep efficiency (sleep duration divided by total time in bed) such that participants with depression had more nighttime activity and poorer sleep efficiency than the healthy controls and adults with fibromyalgia.

Mendlowicz and colleagues (1999) examined the relationship between objective sleep quality (measured by actigraphy) and depressed mood (measured by the Alzheimer's Disease Assessment Scale) in 32 individuals from the community. Although daytime activity level was found to be the best predictor of depressed mood; sleep latency, awake after sleep onset, and total time in bed also significantly predicted depressed mood. In a study with older men, Paudel and colleagues (2008) examined objective sleep quality and depressive symptoms in 3,051 men aged 65 and older. Objective sleep quality was measured by actigraphy and depressive symptoms were measured by the Geriatric Depression Scale (GDS). Men with depressive symptoms ($3 <$ GDS total score < 5) and men classified as having depression (GDS total score ≥ 6) had significantly longer sleep latency than men with normal levels of depressive symptoms (GDS total score ≤ 2).

There are other studies that have failed to find associations between actigraphic measures of objective sleep quality and depressive symptoms. Caldwell and Redeker (2009) examined objective sleep quality, using actigraphy, and depressive symptoms, using the BDI-II, in 115 inner-city women. No significant relationships between the BDI-II and the actigraphic measures of sleep efficiency, sleep duration, number of awakenings, time in bed, and awake after sleep onset were found. In another study, Dorheim and colleagues (2009) examined objective sleep quality (measured by actigraphy) in 21 depressed and 21 non-depressed women who had recently given birth. Depression was defined as a score of ≥ 10 on the Edinburgh Postnatal Depression Scale. No significant relationships were found between depression and the actigraphic measures of sleep latency, awake after sleep onset, sleep duration, sleep efficiency, early morning awakening, total wake time, and time in bed.

Given that both clinical depression and elevated depressive symptoms have been associated with poor subjective and objective sleep quality, it is reasonable to expect that college students with depressive symptoms will have prolonged sleep latency and shorter sleep duration as measured by actigraphy. Therefore, these two indicators of objective sleep quality were chosen as the health-relevant biological outcomes for this study.

Depression and Blood Donation

Blood donation intention was chosen as a health-relevant behavioral outcome for three reasons. First, it is typical for there to be at least one blood drive on the IUPUI campus each academic year. Therefore, local college students are likely to have been exposed to the opportunity to donate blood. Second, typical blood donors are between

the ages of 20 and 50 and are more likely to have at least some college education (Giles, McClenahan, Cairns, & Mallet, 2004). Lemmens and colleagues (2005) found that approximately 48% of their sample of 284 non-donor young adults had strong intentions to donate blood. Thus, it is reasonable to assume that there will be good variability in blood donation intention in college student sample.

Third, although the evidence is limited, depression may be associated with a decreased likelihood of blood donation. To my knowledge, there are no studies that have examined the relationship between depression and blood donation. Therefore, I conducted a secondary analysis of 2,209 individuals enrolled in the National Health and Nutrition Examination Survey (NHANES). A slight trend was found in the data suggesting that individuals who have a diagnosis of depression are less likely to donate blood than those free of depression. The relationship between depression and self-reported blood donation was examined in the NHANES data from three data cycles: 1999-2000, 2001-2002, and 2003-2004. Major depression was diagnosed by the Composite International Diagnostic Interview. Blood donation was assessed by the question "During the past 12 months, have you donated blood?" A crosstabs analysis revealed that 5.3% of individuals without depression reported having donated blood in the past year compared to 4.1% of individuals with depression. When the analysis was stratified by race (white versus non-white), 7.3% of white individuals without depression reported donating blood in the past year, whereas only 2.3% of white individuals with depression donated. In contrast, among the non-white participants, 3.8% of the non-depressed individuals, as compared to 6.6% of the depressed persons, reported donating blood. These findings provide preliminary support for the notion that elevated depressive

symptoms may be associated with a reduced intention to donate blood. Furthermore, these findings suggest that there may be differences in the association between depression and blood donation based on the race of the individual.

<u>Hypotheses</u>

Consistent with the findings of the literature reviewed above, I proposed that EMA measures would have better predictive utility than traditional retrospective measures when examining health-relevant biological outcomes. Conversely, I proposed that traditional retrospective measures would have better predictive utility than EMA measures when examining health-relevant behavioral outcomes. My specific hypotheses were that (1) an EMA measure of depressive symptoms would have better predictive utility than a retrospective measure when examining objective sleep quality, and that (2) a retrospective measure of depressive symptoms would have better predictive utility than an EMA measure when examining blood donation intention. Hypothesis 1 would be supported if it was found that the EMA measure of depressive symptoms was more strongly and positively associated with objective sleep latency and more strongly and negatively associated with objective sleep duration. Hypothesis 2 would be supported if it was found that the retrospective measure of depressive symptoms was more strongly and negatively associated with intention to donate blood.

In addition to testing these two hypotheses, the present study has three secondary objectives. Achieving these objectives will help me to understand the factors that may be responsible for the observed pattern of results. The first objective is to examine the bivariate correlation between the EMA and retrospective measures of depressive

symptoms. The bivariate correlation will assist in determining the amount of overlap between the two measures. The second objective is to evaluate whether there are differences between the Cronbach's alpha and test-retest reliability of the two measures. The Cronbach's alphas and test-retest reliabilities will help to determine the amount of random error present in each measure. The third objective is to compare the mean level of depressive symptoms of the EMA measure to that of the retrospective measure. Comparing the mean levels will provide information regarding whether there is greater systematic error (e.g., due to mood-congruent or peak-end memory heuristic use) present in the retrospective measure.

METHOD

<u>Participants</u>

Participants were 127 female undergraduate students from psychology classes at Indiana University-Purdue University Indianapolis. Students were recruited through the psychology web-based experiment scheduling system. Exclusion criteria were: male, less than 18 years of age, current use of psychotropic medication, a history of a sleep disorder other than insomnia (e.g., sleep apnea) or current use of prescription sleep medication, and a previous history of blood donation. Men were excluded from this study, as there are traditionally fewer male students in the IUPUI Psychology courses that participate in research studies and men, as a group, tend to have lower mean depression scores than women (Nolen-Hoeksema, Larson, & Grayson, 1999). Individuals who were currently using psychotropic medications were excluded because use of these medications could influence both depressive symptoms and sleep quality. In addition, persons who had been diagnosed with a sleep disorder or were currently taking prescription sleep medication were excluded, as they are likely to experience abnormal sleep patterns that are unrelated to depressive symptoms. Those with insomnia were included because excluding individuals with this common difficulty could considerably restrict the range of sleep quality. Persons with a history of blood donation were excluded because their previous blood donation experiences may have a strong influence on their intention to

donate blood in the future, which could mask the influence of depressive symptoms on this outcome.

Of the original sample of 127 female participants, six participants were not eligible to participate due to not meeting inclusion criteria and 21 participants failed to complete all study requirements. Of the 100 participants who completed the entire study, four were excluded because they did not complete at least 16 EMA questionnaires (75% compliance) and/or have three nights of sleep data. Thus, the final sample was comprised of 96 female undergraduate students with a mean age of 21.6 years ($SD = 5.2$, range = 18-47), of whom approximately 33% identified themselves as non-white (17% African American, 7% Asian, 6% Hispanic/Latino, 2% Other) .

Participants received four course credits and either a $15 gift card ($n = 70$) or entry into a prize raffle for a $100 gift card ($n = 30$) for completing the study. Complete participation consisted of four laboratory visits, one 3-day EMA period, and one 3-day actigraphy period (see Table 1 for the timing of procedures).

<div align="center">Measures</div>

<div align="center">Depressive Symptom Severity</div>

The Center for Epidemiologic Studies-Depression Scale (CES-D) was designed to assess depressive symptoms in the general population (Radloff, 1977). It specifically measures current level of depressive symptoms with an emphasis on the affective component. This scale consists of 20 items on which individuals are asked to report how they have felt over the past 7 days using a scale ranging from 0 (rarely or none of the

time) to 3 (most or all of the time) (see Appendix B). Total scores range from 0 to 60,

with higher scores indicating greater depressive symptom severity. A score $\geq$ 16 is

traditionally used to suggest clinically significant depressive symptoms (Radloff, 1977;

Radloff, 1991). Internal consistency for this scale ranges from 0.85 in community

samples to 0.90 in psychiatric samples (Radloff, 1977), and the coefficient alpha for

college samples has been found to be 0.87 (Radloff, 1991). The test-retest reliability

ranges from 0.51 at two weeks to 0.59 at eight weeks in patient populations (Radloff,

1977). The CES-D has been shown to have good construct validity, as well as good

concurrent validity with clinical and self-report criteria (Radloff, 1977). This measure

has also been shown to distinguish between non-depressed community members and

acutely depressed individuals (Radloff, 1991). In a study of 214 college students, the

average CES-D score was 15.46 with a standard deviation of 9.67 (Radloff, 1991), which

indicates acceptable variability. Of this sample of college students, 41% had a total score

$\geq$ 16, indicating clinically significant depressive symptoms (Radloff, 1991).

The CES-D was administered both retrospectively and momentarily to

participants. The retrospective measure of the CES-D was administered on a computer

using SurveyMonkey [TM] (Menlo Park, CA) during laboratory visit 1 and then again at

laboratory visit 2 (see Table 1 for timing of procedures). The instructions for the

retrospective version of the CES-D read "For each statement, please mark the response

option that best describes how you have been feeling *in the past week*." Participants had

the opportunity to complete 21 momentary versions of CES-D (7 per day) during the 3-

day EMA monitoring period. Previous studies have used EMA sampling rates ranging

from 4-15 assessments per day (Kamarck et al., 2004; Shiffman, 2009; Sonnenschein et

al., 2007; Steptoe et al., 2009). The momentary version of the CES-D was identical to the retrospective version except that it was administered on the smartphone and the instructions were modified to read "Over the *last 2 hours*, indicate the response option for each statement which best describes how often you have felt or behaved this way." Participants were allowed to select one of four response options for each question using the smartphone stylus. The momentary assessments were obtained using Palm® (United States) Centro smartphones.

Using interval contingent recording, in which the questionnaire is tied to a pre-determined period, participants were signaled every 2 hours between the hours of 8:55 a.m. and 8:55 p.m. to complete the CES-D. Participants were given 10 minutes following the first signal to complete the questionnaire. If participants had not completed the questionnaire within 10 minutes it was coded as a missed questionnaire. Participants were given a letter to present to their course instructors to inform them that the student was participating in a research study which involved completing a brief questionnaire on their smartphone every two hours. To receive full compensation for the EMA portion of the study, participants needed to complete at least 16 (75%) of the momentary CES-D administrations. If participants did not have 75% compliance, they were required to undergo an additional day of monitoring to receive full compensation ($n = 13$).

The software program that was used to format the customized EMA questionnaires is Satellite Forms 7.2^{TM} (Thacker Network Technologies Inc.; Lacombe, Alberta, Canada). The Satellite Forms software allows for the creation of a database that has tailored menus and data entry screens. The data collected with Satellite Forms was uploaded into PASW® Statistics 17.0. All relevant variables were calculated and

checked for outliers and computation errors. A total score for each EMA assessment of the CES-D was computed and then the total scores were aggregated to provide an average momentary score. It is important to note that the CES-D contains one item which assesses restless sleep. This item was not included in the computation of the total CES-D score for the EMA or retrospective measures. The inclusion of this item in the calculation of the total score could result in criterion contamination because sleep is being measured objectively by two of the outcomes and the self-report of restless sleep would be a subjective report that may be highly correlated with the objective measures.

Objective Sleep Quality

Objective sleep quality was measured by the Actigraph^{TM} GT3X (ActiGraph, LLC; Pensacola, FL). The GT3X is a small red box which has an outer plastic case and measures 1.5 in. x 1.44 in. x 0.70 in. It has a battery lasts up to 21 days. This small device is threaded onto an elastic belt that is worn around the waist against the body. The device is able to identify small and large movements of the individual that occur in all three dimensions of movement (e.g., vertical, horizontal, forward/backward). The device also identifies periods of minimal activity that resemble sleep.

The data obtained was extracted from the GT3X using ActiLife® software. This software utilizes the Sadeh Algorithm to determine sleep onset, sleep duration, sleep latency, number of awakenings, wake minutes, sleep efficiency, and average wake time (Sadeh, Sharkey, & Carskadon, 1994). For the purposes of this study, sleep duration and sleep latency were the sleep variables of interest. Sleep duration is the total number of minutes between sleep onset and the final awakening time. In order to assist in

determining sleep latency, which is the amount of time it takes a person to fall asleep once in bed, participants also completed a sleep diary to record the time they got into bed each night during the 3-day actigraphy period (Ancoli-Israel, Cole, Alessi, Chambers, Moorcroft, & Pollack, 2003). Specifically, participants were required to write down the time they got into bed to sleep each night using the time on a watch which was attached to the sleep diary. Each morning, participants also were required to write down the time they got out of bed or the time at which they were awake for the day. The time the participant got into bed to sleep was used to calculate sleep latency. Participants received an e-mail each evening to remind them to record their bedtime and awakening time (Appendix C).

To receive full compensation for the actigraphy portion of the study, participants needed to have data for 80% of the three days and nights of actigraphy monitoring. If participants did not have data for 80% of the monitoring period, they were required to undergo an additional day of monitoring to receive full compensation ($n = 6$).

In a study comparing actigraphy to polysomnography (the gold standard in objective sleep assessment), it was found that actigraphy has 97% sensitivity in identifying sleep episodes (de Souza, Benedito-Silva, Pires, Poyares, Tufik, & Calil, 2003). Polysomnographic and actigraphic measures of total time in bed and sleep latency have been found to correlate 0.99 and 0.87, respectively (Edinger et al., 2004). Additionally, it has been reported that the correlation between polysomnographic and actigraphic measure of sleep duration is 0.97 (Jean-Louis, von Gizycki, Zizi, Fookson, Spielman, Nunes, Fullilove et al., 1996). Based on these findings, it is reasonable to expect that actigraphy will provide accurate estimates of sleep duration and sleep latency.

Blood Donation Intention

Participants were administered the Blood Donation Intention Scale (France, France, Kowalsky, & Cornett, 2010) at laboratory visit 4 on a computer using SurveyMonkey [TM] (see Appendix D). This scale, which is designed to measure an individual's intention to donate blood, is composed of three items. Each item is rated on a 1 to 7 scale (1 = likely, agree, probable; 7 = unlikely, disagree, improbable). The total score of this measure ranges from 3 to 21. The test-retest reliability of this scale ranges from 0.89 for same day to 0.82 for 5 to 7 days later (France et al., 2010). Internal consistency was found to range from 0.95-0.98 (France et al., 2010).

Secondary Outcomes

To increase the number of opportunities to compare the predictive utility of EMA and traditional retrospective measures, several secondary outcomes were assessed. The secondary biological outcome was average energy expenditure, as measured by actigraphy. The secondary behavioral outcomes were: (a) handgrip duration, (b) cumulative grade point average (GPA), (c) semester GPA, (d) number of missed classes in a typical week, and (e) number of work shifts missed in a typical week.

It was expected that participants who reported greater depressive symptom severity would have: (a) lower energy expenditure due to less physical activity which may be contributed to experiencing a depressed mood (Azar, Ball, Salmon, & Cleland, 2011), (b) shorter handgrip duration (Tangney, Baumeister, & Boone, 2004) and more missed classes (Heiligenstein & Guenther, 1996) and work shifts (Lerner et al., 2004) in a typical week, possibly due to the depressive symptom of avolition; and (c) lower GPAs

because these participants are less likely to be able to put forth the effort required to obtain higher grades in their courses or difficulties in concentration (Andrews & Wilding, 2004; Haines, Norris, & Kashy, 1996; Hysenbegasi, Hass, & Rowland, 2005).

Other Factors

During laboratory visit 1, participants were asked to report: (a) their date of birth; (b) sex; (c) race/ethnicity; (d) tobacco use; and (e) alcohol use on a computer using SurveyMonkey™ (Appendix E). Responses to the race/ethnicity item were combined into two groups, white and non-white. Individuals who identified their race/ethnicity as other than White/Caucasian were placed in the non-white category. Tobacco use responses were also combined into two groups, current users and former/never users. Daily alcohol intake (grams/day) was computed using the quantity-frequency method (Garg, Wagener, & Madans, 1993). Height and weight, which was used to calculate body mass index (BMI), was obtained using a standard medical scale during laboratory visit 2. Previous studies have shown that age, sex, race/ethnicity, tobacco use, alcohol use, and BMI are associated with sleep quality (Buysse, Reynolds III, Monk, Berman, & Kupfer, 1989; Foster & Peters, 1999; Hall et al., 2009; Jennings, Muldoon, Hall, Buysse, & Manuck, 2007; Phillips & Danner, 1995). Thus, these factors were measured and examined as potential confounding variables if a relationship was detected between depressive symptoms and objective sleep quality. In addition, previous studies have shown that age, sex, and race/ethnicity are associated with blood donation, such that younger individuals, men, and whites are more likely to donate blood (Boulware, Ratner, Ness, Cooper, Campbell-Lee, LaVeist, & Powe, 2002; Tscheulin & Lindenmeier, 2005;

Wu, Glynn, Schreiber, Wright, Lo, Murphy, Kleinman et al., 2001). Therefore, these factors were examined as potential confounding variables if a relationship was detected between depressive symptoms and intention to donate blood.

<p style="text-align:center">Procedure</p>

Table 1 outlines the timing of procedures for the proposed study. Before enrolling in the study, potential participants were presented with a list of the exclusion criteria on the psychology web-based experiment scheduling system. To start laboratory visit 1, which took place in the IUPUI Psychology Department computer laboratory, participants answered a series of brief questions to ensure that none of exclusion criteria were met. Participants were then asked to review and sign the informed consent. Participants were also asked to sign a form to grant the study investigators permission to obtain their GPA for their entire academic career at IUPUI as well as for the current semester. After written consent was obtained, participants were told that the purpose of the study was to examine how multiple health and academic behaviors as well as emotional factors may influence sleep quality. Next, they were asked to complete the retrospective CES-D, the demographic items, the smoking items, the alcohol items, and the missed class and work items on the computer. The momentary assessment aspect of the study was then explained. Specifically, participants were given a demonstration of how to complete a momentary questionnaire, completed one "practice" administration of the momentary CES-D, and were informed as to what steps they should take if they encounter equipment problems. Laboratory visit 1 was conducted in groups of 5-12 participants. Participants completed this visit on a Monday and returned for the second

visit on Thursday. Participants scheduled the times for the remaining three laboratory visits at the end of the first visit.

During the 3-day EMA monitoring period, participants were signaled every 2 hours to complete the momentary CES-D. Participants were given 10 minutes following the signal to complete the questionnaire. If the questionnaire has not been completed within 10 minutes it was coded as a missed questionnaire.

At laboratory visit 2, three days later, participants returned the smartphone and completed the retrospective version of the CES-D on a computer. Each participant's height, weight, and performance on the handgrip task were also measured during this visit. Laboratory visit 2 took place in the research laboratory where there was a computer and a standard medical scale.

During laboratory visit 3, which occurred on the Monday of the next week, participants were instructed on how to wear the actigraph belt and under which conditions to remove the belt. Specifically, participants were shown how to secure the actigraph belt around their waist under their clothes. Participants were also instructed to remove the actigraph belt only when they bathed or swam. Laboratory visit 3 took place in the research laboratory. In addition to wearing the actigraph belt, participants also completed the previously described brief sleep diary. In an attempt to increase compliance with the sleep diary, participants were informed that the experimenter would be able to tell if they were not accurate in their recording of their sleep and wake time because this information was also being measured by the actigraph.

During the 3-day actigraphy monitoring period, participants wore the actigraph belt all day every day except for when they were bathing or swimming. Participants

received an e-mail each evening reminding them to record their bedtime and an e-mail each morning reminding them to record their awakening time.

Laboratory visit 4 occurred on Thursday of the second week and lasted approximately 10 minutes. This visit took place in the research laboratory. During this visit, participants returned the actigraph belt and completed the Blood Donation Intention Scale on a computer. Debriefing occurred at the end of study participation before participants left the laboratory. Lastly, participants were compensated for their participation.

Data Analysis

Primary Analyses

To test Hypothesis 1 (i.e., the EMA measure of depressive symptoms will have better predictive utility than the retrospective measure when examining objective sleep quality), multiple linear regression analyses and a hand calculation were performed. In the first regression model, aggregate momentary CES-D was entered as the predictor variable and sleep duration was entered as the criterion variable. The second regression model had retrospective CES-D as the predictor variable and sleep duration as the criterion variable. A hand calculation was performed to determine whether the unstandardized regression coefficients were significantly different from each other. The formula that was used to compare the coefficients is as follows (Cohen & Cohen, 1983):

$$Z = (B_{i1} - B_{i2}) / (SE^2_{Bi1} + SE^2_{Bi2})^{1/2}$$

B_{i1} is the coefficient for the EMA measure of depressive symptoms, whereas B_{i2} is the coefficient for the retrospective measure. *SE* is the standard error for each coefficient. Because *Z* is a difference score that is normally distributed, it allows for the use of the standard normal distribution to determine significance. The same set of regression analyses and hand calculations was used to examine which form of self-report is a stronger predictor of objective sleep latency.

To test Hypothesis 2 (i.e., the retrospective measure of depressive symptom will have better predictive utility than the EMA measure when examining blood donation intention), the same set of regression analyses and hand calculations was performed. In these models, the same predictors were used but the criterion variable was the Blood Donation Intention Scale total score.

The same set of analyses was repeated for each of the secondary outcomes (i.e., average energy expenditure, handgrip duration, cumulative GPA, and semester GPA). Average energy expenditure was log-transformed prior to analysis to reduce positive skew (skewness = 3.2 before transformation, skewness = 0.5 after transformation).

Because the initial analyses did not include any covariates, the regression analyses and hand calculations were repeated for the following combinations of outcomes and covariates: (a) average sleep duration and average energy expenditure analyses were adjusted for age, race/ethnicity, BMI, alcohol use, and tobacco use; (b) blood donation intention, cumulative GPA, and semester GPA analyses were adjusted for age and race/ethnicity; and (c) the handgrip duration analysis was adjusted for age, race/ethnicity, and handgrip strength.

Secondary Analyses

Three secondary objectives were examined to understand the factors that may be responsible for the pattern of results. One secondary objective was to calculate the bivariate correlation between the total scores of the EMA and retrospective measures of depressive symptoms.

Another secondary objective was to examine the difference between the Cronbach's alphas and the test-retest reliabilities of the EMA and the retrospective measures. Cronbach's alphas were obtained by conducting reliability analyses in PASW® Statistics 17.0. Cronbach's alpha for the EMA measure was computed by aggregating each EMA CES-D item and using these aggregated item values in the analysis. Test-retest reliability for the EMA measure was obtained by performing a Pearson's correlation between the aggregate scores for Day 1 and 2, Day 1 and 3, and Day 2 and 3. Test-retest reliability for the retrospective measure was obtained by performing a Pearson's correlation between the total score for laboratory visit 1 and visit 2. To determine whether there was a significant difference between the Cronbach's alphas, the following formula (the Feldt method) was used (Charter & Feldt, 1996):

$$t_{N-2} = \left| \alpha_1 - \alpha_2 \right| \sqrt{N-2} / \sqrt{4(1-\alpha_1)(1-\alpha_2)(1-r_{12}^2)}$$

With this method, α_1 is the Cronbach's alpha value for one test, α_2 is the Cronbach's alpha value for the other test, and r_{12}^2 is the correlation between the two tests squared. The t distribution table was used to determine significance.

To determine whether there was a significant difference between the test-retest reliabilities, a Fisher's z' Transformation was performed for each Pearson correlation

value (*R*) in order to produce standardized values that are normally distributed. The following formula was used to transform the correlations:

$$Z_f = \tfrac{1}{2} * \ln ((1 + R) / (1 - R)).$$

Once the Pearson correlations were transformed, a difference score was calculated between the Fisher's *z'* values using the following formula (Cohen & Cohen, 1983):

$$Z = (z'_1 - z'_2) / ((1 / n_1 - 3) + (1 / (n_2 - 3))^{\frac{1}{2}}$$

n_1 and n_2 are the respective sample sizes for each measure. Because *Z* is a difference score that is normally distributed, it allows for the use of the standard normal distribution to determine significance.

The final secondary objective was to conduct a dependent samples *t* test between the aggregate momentary CES-D total score and the retrospective CES-D total score.

RESULTS

Study Variables

Average sleep latency, class attendance, and work attendance were not examined as outcomes. Average sleep latency was not examined due to obtaining exceptionally small values ($M = 4.34$, $SD = 3.44$, range = 0-16 minutes). Based on these results, most participants fell asleep within 5 minutes of turning off their light to go to sleep. These findings are inconsistent with the literature assessing sleep latency using actigraphy, which report average sleep latencies between 15 and 32 minutes among adults without sleep difficulties (Edinger et al., 2004; Means et al., 2003). Our values suggest that the participants may not have been accurate in recording their bed time. The class and work attendance variables were not examined due to severe restriction of range. The number of reported missed class sessions in a typical week ranged from 0 (85% of participants) to 4 ($M = 0.23$, $SD = 0.64$, skewness = 3.56). Fifty-seven participants reported that they were employed, and the number of reported missed 8-hour work shifts in a typical week ranged from 0 (89% of participants) to 2.5 ($M = 0.17$, $SD = 0.56$, skewness = 3.53). It is possible that the restriction of range in the class and work attendance variables was due to selection bias, as the individuals who choose to participate in a multi-session research study may be higher in conscientiousness than the typical student.

Descriptive Statistics

Descriptive statistics for the covariates, predictors, and outcomes are shown in Tables 2 and 3. The average number of EMA observations ranged from 16 to 25 ($M =$ 19.4, $SD = 2.1$) and corresponded to a compliance rate of 92%, which may be an inflated value because 10 participants had to complete an extra day of monitoring. Both the mean momentary CES-D total score and the retrospective total score fell below the cut point indicative of clinically significant depressive symptoms (≥ 16; Radloff, 1991); however, 32 (33%) individuals did have scores above this cut point. Participants, on average, obtained fewer than eight hours of sleep each night, and the average energy expenditure was approximately two times higher than the recommended 150 kcal/day suggested by the United States Surgeon General (U.S. Department of Health and Human Services, 2012). Thirty-five (37%) participants obtained the maximum total score of 21 on the Blood Donation Intention Scale. Participants' demonstrated effort in the handgrip task, as only two participants persisted in the task for less than 10 seconds. The mean for both cumulative GPA and semester GPA was slightly lower than a B average. Not surprisingly, approximately 77% of the variance was shared between the two GPA variables ($r(88) = .88, p < .001$).

Independent samples t tests and Pearson chi-square analyses were conducted to examine if there were differences between the participants who did not complete the study versus those who did complete the study. These analyses revealed no differences between these groups on age, race/ethnicity, BMI, tobacco use, alcohol use, or the predictor and outcomes variables (see Table 4).

Primary Analyses

Biological Outcomes

Two sets of multiple regression analyses were conducted: one with no covariates and another adjusting for age, race/ethnicity, BMI, tobacco use, and alcohol use. In the unadjusted analyses (see Table 5), the EMA measure of depressive symptoms was not a predictor of average sleep duration; however, it fell just short of significance for average energy expenditure ($p = .06$). The retrospective measure was not related to either of the biological outcomes. When the regression coefficients were compared, the EMA measure was not a stronger predictor than the retrospective measure of either of the biological outcomes (ps $= .48$ and $.22$, respectively).

As shown in Table 6, neither the EMA measure nor the retrospective measure predicted average sleep duration or average energy expenditure in the fully adjusted models. Once again, the EMA measure was not a stronger predictor than the retrospective measure when the coefficients were compared (ps $= .68$ and $.46$, respectively).

Behavioral Outcomes

In the unadjusted analyses (see Table 5), the EMA measure of depressive symptoms was not a predictor of blood donation intention or handgrip duration; however, it fell just short of significance for cumulative GPA ($p = .05$) and semester GPA ($p = .06$). The retrospective measure was not a predictor of any of the four behavioral outcomes. The EMA measure was not a stronger predictor than the retrospective

measure of any of the behavioral outcomes when the coefficients were compared (ps = .50, .64, .10, and .20, respectively).

As shown in Table 6, neither the EMA nor the retrospective measure was a predictor of blood donation intention, cumulative GPA, or semester GPA when the analyses were adjusted for age and race. When the analyses predicting handgrip duration were adjusted for handgrip strength, age, and race/ethnicity, neither the EMA nor the retrospective measure predicted handgrip duration (see Table 6). Comparisons of the coefficients again revealed the EMA measure was not a stronger predictor than the retrospective measure of any of the behavioral outcomes (ps = .64, .80, .12, and .20, respectively; see Table 6).

Secondary Analyses

Correlation between EMA and Retrospective Self-Report

The bivariate correlation between the EMA and the retrospective measures indicates that these two measures are strongly and positively correlated ($r(94)$ = .75, $p <$ .001).

Random Error

To evaluate the extent of random error in the two measures, the Cronbach's alphas and test-retest reliabilities were compared. The Cronbach's alpha was 0.91 for the EMA measure and was 0.93 for the retrospective measure. The difference between these two values was not significant ($t(94)$ = 1.57, $p >$ .05). Because aggregation can inflate

reliability, one EMA time point was used to calculate Cronbach's alpha to provide a more parallel comparison to the retrospective measure. The EMA time point used was the 6:55 p.m. administration on Day 1, given that this time point had the most data. The Cronbach's alpha for this EMA time point was 0.86. The difference between this single EMA time point and the retrospective measure was significant, such that the retrospective measure had better internal consistency ($t(88) = 4.21, p < .05$).

Test-retest reliability for the retrospective measure was calculated using the Session 1 and Session 2 CES-D total scores ($r(93) = .77, p < .001$). Test-retest reliability for the EMA measure was calculated between the aggregate values of each day of data collection: Day 1 and Day 2 ($r(94) = .81, p < .001$); Day 1 and Day 3 ($r(94) = .68, p < .001$); and Day 2 and Day 3 ($r(94) = .75, p < .001$). Comparisons revealed no difference between the test-retest reliability of EMA and retrospective measures of depressive symptoms: (a) Day 1-Day 2 versus retrospective ($Z = .63, p = .52$) (b) Day 1-Day 3 versus retrospective ($Z = -.87, p = .38$); and (c) Day 2-Day 3 versus retrospective ($Z = -.26, p = .80$).

Systematic Error

A dependent samples t test revealed a difference in the mean levels of depressive symptoms ($t(95) = 4.80, p < .001$) such that participants reported greater depressive symptom severity on the retrospective measure ($M = 13.1$) than on the EMA measure ($M = 9.6$).

Exploratory Analyses

Exploratory Analysis 1

Participants in the Fall 2010 semester were provided different compensation for participation (entry into a $100 raffle) than participants in the other two semesters ($15 gift card). In addition, it is possible that students who participated in the Fall 2010 semester may have significantly different performance on academic outcomes, particularly if it was their first semester of college. For these reasons, one-way ANOVAs and Pearson chi-square analyses were performed to examine mean differences between the three semesters of participation (Fall 2010, Spring 2011, Summer 2011) on the covariates, predictors, and outcomes (see Table 7). These analyses revealed mean differences for semester GPA, such that the Summer 2011 participants ($n = 10$, $M = 3.60$, $SD = .47$) had a higher semester GPA than the Spring 2011 participants ($n = 76$, $M = 2.84$, $SD = .88$). No other mean differences were detected between semesters of participation.

Two sets of multiple regression analyses were conducted for the individuals who participated during the Spring 2011 semester: one with no covariates and another adjusting for age, race/ethnicity, BMI, tobacco use, alcohol use, and handgrip strength. As shown in Table 8 and Table 9, neither the EMA measure nor the retrospective measure predicted any of the outcomes in either the unadjusted or fully adjusted models. When the regression coefficients were compared, there was no difference between the measures in their predictive ability. Due to the small number of participants in the Fall 2010 semester ($n = 5$) and the Summer 2011 semester ($n = 12$), separate regressions were

not conducted for these semesters. These analyses suggest that the difference in compensation between the semesters did not have an effect on the findings of the study. Mean differences were found between semesters of participation, such that Summer 2011 participants had a higher GPA than the Spring 2011 participants. This finding may be due to taking fewer classes in the summer semester than the spring semester.

Exploratory Analysis 2

Participants who completed additional EMA and/or actigraphy monitoring due to poor compliance may have influenced the findings of this study. To examine whether this influence was present, two sets of multiple regression analyses were conducted for the individuals who were *not* sent out for an additional day of monitoring ($n = 81$): one with no covariates and another adjusting for age, race/ethnicity, BMI, tobacco use, alcohol use, and handgrip strength.

In the unadjusted analyses (see Table 10), the EMA measure of depressive symptoms was a predictor of average energy expenditure ($p = .04$); whereas the retrospective measure was not. Neither the EMA measure nor the retrospective measure was a predictor of average sleep duration, blood donation intention, handgrip duration, cumulative GPA, or semester GPA. When the regression coefficients were compared, the EMA measure was not a stronger predictor than the retrospective measure.

As shown in Table 11, neither the EMA measure nor the retrospective measure predicted the outcomes in the fully adjusted models. Once again, when the regression coefficients were compared, there was no difference between the measures in their

predictive ability. Therefore, participants who completed an additional day of monitoring did not influence the original findings of the study.

Exploratory Analysis 3

In order to determine if there were time of day effects for the EMA questionnaires, the seven daily EMA administrations were divided into three variables: morning EMA (8:55 a.m. and 10:55 a.m. administrations), afternoon EMA (12:55 p.m., 2:55 p.m., and 4:55 p.m.), and evening EMA (6:55 p.m. and 8:55 p.m.), and the average EMA CES-D total score was calculated for these three time points. Two sets of multiple regression analyses were conducted for each of the three time points: one with no covariates and another adjusting for age, race/ethnicity, BMI, tobacco use, and alcohol use.

In the unadjusted analyses (see Tables 12, 13, and 14), the afternoon time point for the EMA measure of depressive symptoms was a predictor of average energy expenditure ($p = .03$); whereas the morning and evening EMA time points and the retrospective measure was not. It is likely that this finding is due to Type I error given the number of analyses conducted. Neither the three EMA time points nor the retrospective measure was a predictor of average sleep duration, blood donation intention, handgrip duration, cumulative GPA, or semester GPA. When the regression coefficients were compared, none of the EMA measures were a stronger predictor than the retrospective measure.

As shown in Tables 15, 16, and 17, none of the three EMA time points or the retrospective measure predicted the outcomes in the fully adjusted models. Once again,

when the regression coefficients were compared, there was no difference between the measures in their predictive ability. Overall, these findings suggest that there were no time of day effects for the EMA questionnaires. The one exception is that the afternoon assessments were predictive of average energy expenditure in the unadjusted analyses, which is likely explained by Type I error.

DISCUSSION

The failure to observe associations between the depressive symptom measures and the biological and behavioral outcomes in this study may be due to the following possibilities: (1) the findings reflect the actual state of nature, (2) insufficient power, (3) limited variability in the predictor variables, (4) limited variability in the outcomes, and/or (5) imperfect measurement of the primary behavioral outcome.

First, it is possible that depressive symptom severity is truly not related to objective sleep duration, average energy expenditure, blood donation intention, handgrip duration, cumulative GPA, and semester GPA. However, this seems unlikely, given that these outcomes were chosen because the literature supports their relationship with depression, with the exception of blood donation intention (see Introduction for a review of these literatures).

Second, this study may not have had sufficient power to detect the relationships of interest. Consistent with this idea, nonsignificant trends were observed between the EMA measure and several of the outcomes (i.e., average energy expenditure, blood donation intention, cumulative GPA, and semester GPA; see Tables 5 and 6). If the sample size had been larger, it is possible that these relationships would have become statistically significant. The low power of this study made it particularly difficult to

detect differences between the regression coefficients, given that more participants are needed to have sufficient power for these calculations.

No study has examined the independent and dependent variables in the way this study has, which made it difficult to obtain an accurate number of participants needed to obtain power equal to or greater than .80. The study on which my power analysis was based examined the relationship between depressive symptoms (as measured by the CES-D) and objective sleep duration (as measured by actigraphy) in a sample of undergraduate students without psychiatric disorders (Kawada et al., 2007). The observed effect size of the CES-D total score was Adjusted $R^2 = 0.084$ for sleep duration. A power analysis for a two-tailed, point biserial correlation indicated that, for $r = 0.28$, power $= 0.80$, and $\alpha = .05$, a sample of 89 participants would be required. The current study has 96 participants. Therefore, based on this power analysis, it should have adequate power. However, it should be noted that this power analysis was conducted for a different statistical method than the one used in this study, and the sample was 75% male, whereas the current study has only female participants. In addition, in order to have sufficient power to compare regression coefficients against each other instead of against zero, even more participants would be needed.

Third, descriptive statistics indicated that only 33% of participants reported clinically significant depressive symptoms. Furthermore, the observed mean level of depressive symptoms (CES-D Total = 13.1) is lower than that reported by Radloff (1991) in a sample of 214 college students (CES-D Total = 15.5). At the same time, the larger standard deviations and range for both the EMA and retrospective measures suggests

ample variability (see Table 2). Therefore, it is unlikely that limited variability in the predictor variables explains the null results.

Fourth, the outcomes also did not show a pattern of limited variability. The observed mean (7.5 hours) and standard deviation (1.3 hours) for sleep duration were similar to values found in previous research. One study obtained a mean of seven hours with a standard deviation of one hour for women (Joffe et al., 2009), while another study reported a mean of 7.8 hours with a standard deviation of 1.2 hours (Mendlowicz et al., 1999). The considerable standard deviations and ranges for the other outcomes are indicative of adequate variability (see Table 2). Thus, limited variability in the outcomes does not appear to account for the null results.

Fifth, blood donation intention was conceptualized as a behavioral measure; however, the assessment approach (i.e., a self-report questionnaire) did not measure the actual behavior of interest. Intention to participate in an activity may be prone to social desirability, which is consistent with the high scores observed on this measure. It is possible that participant reports of their intention to engage in a behavior are artificially elevated regardless of their current depression level, which in turn could mask an association between depressive symptom severity and blood donation intention.

It seems that insufficient power and the lower percentage of participants with clinically significant depressive symptoms are the most plausible reasons for the null results. Previous research has found relationships between depressive symptoms and sleep duration (Joffe et al., 2009; Kawada et al. 2007; Korszun et al. 2002; Mendlowicz et al., 1999; Paudel et al., 2008), physical activity (Azar et al., 2008), handgrip duration (Tangney et al., 2004), and cumulative GPA and semester GPA (Andrews & Wilding,

2004; Haines, Norris, & Kashy, 1996; Hysenbegasi, Hass, & Rowland, 2005). To fully test the hypotheses, it was necessary to observe relationships between the predictor variables and outcomes. Unfortunately, this was not the case in the current study.

<div align="center">Primary Objective</div>

The primary objective of this study was to examine the utility of EMA and retrospective measures of depressive symptom severity in predicting objective sleep duration (a biological outcome) and blood donation intention (a behavioral outcome). The following additional outcomes, which have been shown to be related to depression (see Introduction for a review of these literatures), were also examined to increase the number of opportunities to compare the predictive utility of the measures: average energy expenditure (biological), handgrip duration (behavioral), cumulative GPA (behavioral), and semester GPA (behavioral). As was noted above, the failure to observe relationships between the predictor variables and the outcomes did not allow for a full test of the hypotheses.

Given that this study does not provide an ideal context for the hypotheses to be tested, it is not surprising that the first hypothesis – i.e., the EMA measure of depressive symptom severity would have greater predictive utility than the retrospective measure for the biological outcomes – was not supported by the results. Neither the EMA nor the retrospective measure predicted objective sleep duration. There was a trend in the data that suggested the EMA measure might predict average energy expenditure if the sample size was increased, whereas the retrospective measure did not predict this outcome.

When the coefficients of the two different assessment methods were compared, a difference was not detected.

The findings of no relationship between depression and objective sleep duration contrast with past investigations. The majority of the studies that have examined depressive symptom severity and objective sleep duration (as measured by actigraphy) have reported an association between these variables (Kawada et al., 2007; Mendlowicz et al., 1999; Paudel et al., 2008). Of particular relevance, Kawada and colleagues (2007) detected a negative relationship between these factors in a sample of college students. In studies involving clinical samples, it has been found that patients with major depressive disorder have shorter sleep duration than nondepressed individuals (Benca, 2005; Joffe et al., 2009). At least two previous studies have also not observed a relationship between depression and objective sleep duration. In one study, no relationship was found between the BDI-II and the actigraphic measure of sleep duration in inner-city women (Caldwell & Redeker, 2009). In the other study, there was no relationship between the Edinburgh Postnatal Depression Scale and the actigraphic measure of sleep duration in 21 depressed women who had recently given birth (Dorheim et al., 2009).

As was mentioned above, this study does not provide an ideal context in which to fully test the hypotheses. Therefore, once again it was not surprising that the second hypothesis – i.e., the retrospective measure of depressive symptoms would have greater predictive utility than the EMA measure for behavioral outcomes – was not supported by the results. The EMA measure of depressive symptoms did not predict blood donation intention or handgrip duration. However, the EMA measure fell just short of significance for cumulative GPA and semester GPA in the unadjusted analyses and for cumulative

GPA in the adjusted analyses. The retrospective measure was not a predictor of any of the four behavioral outcomes. No difference was detected when the regression coefficients of the two assessment methods were compared.

There is no published evidence supporting a relationship between depression and blood donation intention; however, a secondary analysis of the NHANES data found a slight trend suggesting white individuals with a diagnosis of depression may be less likely to donate blood (see p. 32). Such a trend was not found in the current study. Previous research has detected a relationship between depressive symptom severity and lower GPA that was partially replicated in this study (Andrews & Wilding, 2004; Haines, Norris, & Kashy, 1996; Hysenbegasi, Hass, & Rowland, 2005) in that there was a trend between the EMA measure and cumulative GPA. It is possible that this relationship would have been detected in a larger sample.

Overall, the observed pattern of results is not consistent with previous studies that have compared the predictive utility of EMA and retrospective measures. In the studies reviewed on pp. 19-23, EMA measures were generally stronger predictors of biological outcomes (e.g., carotid IMT), and retrospective measures were stronger predictors of behavioral outcomes (e.g., colonoscopy return rate). An interesting, although nonsignficant, pattern in the current data is that the regression coefficients for the EMA measure are larger (and in the expected direction) than those for the retrospective measure for the majority of outcomes. This observation raises the possibility that, in a larger sample, the EMA measure may have been a better predictor of all outcomes, not only the biological ones.

Secondary Objectives

Three secondary objectives were pursued to understand the factors that may be responsible for the pattern of results. Specifically, the amount of overlap between the EMA and retrospective measures of depressive symptom severity, as well as the extent of random and systematic error in these measures, were examined. The amount of overlap between the measures was moderate to high, as they shared 56% of the variance ($r = .75$). Although no study was found that reported a correlation between EMA and retrospective measures of depressive symptoms, one study reported a correlation of 0.63 between EMA and retrospective measures of negative emotions (Feldman Barrett, 1997). The current study's correlation may be higher than that observed by Feldman Barrett (1997) due to using the exact same items and response set for both measures. Previous research has used different questionnaires for EMA and retrospective measures or has substantially altered the items in order to fit on the mobile device screen or to make the questions more relevant to a shorter time frame of assessment. The strong correlation between the EMA and retrospective measures in this study indicates that they both are likely assessing the same or a similar construct (i.e., depressive symptom severity). An additional explanation for this strong correlation could be common method variance, given that the same questions and response set were used for both the EMA and retrospective measures.

Random error was assessed by comparing the Cronbach's alphas and the test-retest reliabilities of the EMA and retrospective measures. Both measures had high internal consistency ($\alpha = .91$ for the EMA measure and $\alpha = .93$ for the retrospective measure), which suggests that they both consist of items that are measuring a single construct. There was no difference in internal consistency between the aggregate EMA

and retrospective measures. When only one EMA time point (α = .86) was examined, the retrospective measure had slightly (although significantly) better internal consistency, possibly because it was administered under controlled conditions. Nevertheless, a Cronbach's alpha of 0.70 or greater is considered to be acceptable (Shultz & Whitney, 2005). Therefore, regardless of whether the internal consistency for EMA is evaluated at one time point or across time points, it is still at an acceptable level. There was no difference in the test-retest reliabilities between the EMA and retrospective measures. Taken together, these results suggest that there is likely no meaningful difference in the amount of random error present in the two measures.

Systematic error was examined by comparing the mean levels of the EMA and retrospective measures. Participants reported greater depressive symptom severity on the retrospective measure (M = 13.1) than on the EMA measure (M = 9.6). The higher mean for the retrospective measure is consistent with the idea that participants may be using cognitive heuristics when reporting retrospectively. For instance, participants may be reporting on the day when they remembered feeling the most depressed and/or for the most recent day instead of their typical level of depression across two weeks (peak-end heuristic; Fredrickson, 2000). It is also possible that individuals who were feeling more depressed at the time of their laboratory sessions may have reported on days that were congruent with their current mood (mood-congruent heuristic; Smyth & Stone, 2003). It is also possible that the lower level of mean depressive symptoms on the EMA measure are due to practice effects, given that participants may have automatically completed the questionnaire without considering each item as they became more familiar with the measure.

Limitations and Recommendations for Future Research

Limitations of this study include the smaller sample size, the cross-sectional design, the relatively brief monitoring periods for EMA and actigraphy, and the inadequate behavioral outcome. As was discussed earlier, this study may not have had sufficient power due to its smaller sample size, which may have resulted in the failure to detect relationships that do exist. In addition, this study was designed to compare the *predictive* utility of EMA and retrospective measures. Due to restricted time and resources, however, cross-sectional relationships, and not longitudinal relationships, were examined. Another consequence of restricted time and resources was that the EMA and actigraphy monitoring periods were only three days each. Longer monitoring periods would have allowed for aggregation over more observations and, therefore, would have provided more reliable and representative estimates of typical depressive symptom severity and sleep duration. A final limitation was that there is little empirical support for a relationship between depression and the primary behavioral outcome of blood donation intention.

Future studies should incorporate the following changes: (a) include a larger sample, (b) recruit more participants with clinically significant depressive symptoms, (c) utilize a longitudinal design, (d) conduct EMA and actigraphy for a longer period of time, and (e) include behavioral outcomes known to be strongly related to depressive symptoms. Researchers should ensure that their sample is large enough to detect relationships between the EMA and retrospective measures and the selected outcomes, as well as differences between the regression coefficients for the measures. In future studies, approximately 50% of the participants should be experiencing clinically

significant depressive symptoms to increase the likelihood of detecting these relationships. A longitudinal research design would allow for a better evaluation of predictive utility, given that that directionality of the observed associations could be established. Collecting EMA data and measuring objective sleep duration for a longer period, perhaps over a 2-week period (M. Okun, personal communication, April 11, 2011), would provide a better assessment of participants' typical depressive symptoms and sleep patterns. In addition, it would be preferable to have participants record their getting into bed time using a smartphone to ensure more accurate time stamping, as they may complete the paper-and-pencil diaries the following day. Finally, future studies should include behavioral outcomes known to have strong and consistent relationships with depressive symptoms. Physical fitness behaviors may be one behavioral outcome worth examining, given that physical activity has been demonstrated to have a negative relationship with depressive symptoms (Strohle, 2009).

Conclusions

The primary objective of this study was to examine the utility of both an EMA measure and a retrospective measure of depressive symptom severity in predicting health-relevant biological and behavioral outcomes. Unfortunately, it was not possible to fully test the hypotheses due to the failure to observe relationships between the predictor variables and the outcomes. The reported results, although limited, did not provide support for Hypothesis 1 (i.e., the EMA measure of depressive symptom severity would have greater predictive utility than the retrospective measure for the biological outcomes)

or Hypothesis 2 (i.e., the retrospective measure of depressive symptoms would have greater predictive utility than the EMA measure for behavioral outcomes).

Supplemental analyses revealed that the EMA and retrospective measures of depressive symptoms shared a moderate to high amount of the variance. In addition, there was no difference in internal consistency or test-rest reliability between the EMA and retrospective measures, indicating that the extent of random error present in the two measures was similar. The retrospective measure's mean level was higher than that of the EMA measure, which is consistent with the idea that participants are using cognitive heuristics when reporting retrospectively. Therefore, the retrospective measure may contain greater systematic error than the EMA measure.

Proponents of EMA assert that EMA measures are superior to traditional retrospective measures in terms of reliability and validity. However, as evidenced by the paucity of literature examining this claim, as well as the failure of the current study to evaluate this assertion sufficiently, it seems that this claim is currently unfounded. Future research is needed in order to fully evaluate this contention as the use of EMA methodology can be demanding for both the researcher and the participants. Therefore, if EMA is not shown to have superior psychometric properties (including predictive utility), it may be difficult to justify the extensive amount of time and resources required to use this assessment method for the purpose of assessing psychosocial traits. On the other hand, EMA may still be of use in assessing moment-to-moment in psychosocial states.

LIST OF REFERENCES

LIST OF REFERENCES

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorder* (Revised 4th ed.). Washington, DC: Author.

Ancoli-Israel, S., Cole, R., Alessi, C., Chambers, M., Moorcroft, W., & Pollak, C. P. (2003). The role of actigraphy in the study of sleep and circadian rhythms. American Academy of Sleep Medicine Review Paper. *Sleep, 26,* 342-392.

Andrews, B., & Wilding, J. M. (2004). The relation of depression and anxiety to life-stress and achievement in students. *British Journal of Psychology, 95,* 509-521.

Azar, D., Ball, K., Salmon, J., & Cleland, V. (2011). Individual, social, and physical environmental correlates of physical activity among young women at risk of depression. *Journal of Physical Activity and Health, 8,* 133-140.

Ayas, N. T., White, D. P., Al-Delaimy, W. K., Manson, J. E., Stampfer, M. J., Speizer, F. E., Patel, S., et al. (2003). A prospective study of self-reported sleep duration and incident diabetes in women. *Diabetes Care, 26,* 380-384.

Benca, R. M. (2005). Mood disorders. In: M. H. Kryger, T. Roth, & W. C. Dement (Eds.), *Principles and Practice of Sleep Medicine* (4th ed.; pp. 1311 – 1326). Philadelphia, PA: Elsevier Saunders.

Boulware, L. E., Ratner, L. E., Ness, P. M., Cooper, L. A., Campbell-Lee, S., LaVeist, T. A., & Powe, N. R. (2002). The contribution of sociodemographic, medical, and attitudinal factors to blood donation among the general public. *Transfusion, 42,* 669-678.

Bureau of Labor Statistics. (January 26, 2010). *Volunteering in the United States—2009.* Retrieved from http://www.bls.gov/news.release/volun.nr0.htm

Buysse, D. J., Reynolds III, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research, 28,* 193-213.

Caldwell, B. A., & Redeker, N. S. (2009). Sleep patterns and psychological distress in women living in an inner city. *Research in Nursing & Health, 32,* 177-190.

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment.* Beverly Hills, CA: Sage Publications Inc.

Charter, R. A., & Feldt, L. S. (1996). Testing the equality of two alpha coefficients. *Perceptual and Motor Skills, 82,* 763-768.

de Souza, L., Benedito-Silva, A. A., Pires, M. L. N., Poyares, D., Tufik, S., & Calil, H. M. (2003). Further validation of actigraphy for sleep studies. *Sleep, 26*, 81-85.

Dorheim, S. K., Bondevik, G. T., Eberhard-Gran, M., & Bjorvatn, B. (2009). Subjective and objective sleep among depressed and non-depressed postnatal women. *Acta Psychiatrica Scandinavica, 119,* 128-136.

Edinger, J. D., Means, M. K., Stechuchak, K. M., & Olsen, M. K. (2004). A pilot study of inexpensive sleep-assessment devices. *Behavioral Sleep Medicine, 2,* 41-49.

Feldman Barrett, L. (1997). The relationships among momentary emotion experiences, personality descriptions, and retrospective ratings of emotion. *Personality and Social Psychology Bulletin, 23,* 1100-1114.

Foster, J. H., & Peters, T. J. (1999). Impaired sleep in alcohol misusers and dependent alcoholics and the impact upon outcome. *Alcoholism: Clinical and Experimental Research, 23,* 1044-1051.

France, C. R., France, J. L., Kowalsky, J. M., & Cornett, T. L. (2010). Education in donation coping strategies encourages individuals to give blood: Further evaluation of a donor recruitment brochure. *Transfusion, 50,* 85-91.

Fredrickson, B. L. (2000). Extracting meaning from past affective experiences: The importance of peaks, ends, and specific emotions. *Cognition and Emotion, 14,* 577-606.

Gangwisch, J. E., Malaspina, D., Boden-Albala, B., & Heymsfield, S. B. (2005). Inadequate sleep as a risk factor for obesity: Analyses of the NHANES I. *Sleep, 28,* 1289-1296.

Garg, R., Wagener, D. K., & Madans, J. H. (1993). Alcohol consumption and risk of ischemic heart disease in women. *Archives of Internal Medicine, 153,* 1211-1216.

Giles, M., McClenahan, C., Cairns, E., & Mallet, J. (2004). An application of the theory of planned behaviour to blood donation: The importance of self-efficacy. *Health Education Research, 19,* 380-391.

Gorin, A. A., & Stone, A. A. (2001). Recall biases and cognitive errors in retrospective self-reports: A call for momentary assessments. In: A. Baum, T. Revenson, & J. Singer (Ed.). *Handbook of health psychology* (pp. 405-413). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Haines, M. E., Noris, M. P., & Kashy, D. A. (1996). The effects of depressed mood on academic performance in college students. *Journal of College Student Development, 37,* 519-526.

Hall, M. H., Matthews, K. A., Kravitz, H. M., Gold, E. B., Buysse, D. J., Bromberger, J. T., et al. (2009). Race and financial strain are independent correlates of sleep in midlife women: The SWAN sleep study. *Sleep, 32,* 73-82.

Hawkins, M. A., Stewart, J. C., & Fitzgerald, G. J. (March, 2010). *The combined effect of depressive symptoms and hostility on autonomic nervous system function.* Poster presented at the meeting of The American Psychosomatic Society, Portland, OR.

Heiligenstein, E., & Guenther, G. (1996). Depression and academic impairment in college students. *Journal of American College Health, 45,* 59-65.

Helgeson, V. S., Lopez, L. C., & Kamarck, T. (2009). Peer relationships and diabetes: Retrospective and ecological momentary assessment approaches. *Health Psychology, 28,* 273-282.

Hublin, C., Partinen, M., Koskenvuo, M., & Kaprio, J. (2007). Sleep and mortality: A population-based 22-year follow-up study. *Sleep, 30,* 1245-1253.

Hysenbegasi, A., Hass, S. L., & Rowland, C. R. (2005). The impact of depression on the academic productivity of university students. *The Journal of Mental Health Policy and Economics, 8,* 145-151.

Jean-Louis, G., von Gizycki, H., Zizi, F., Fookson, J., Spielman, A., Nunes, J., Fullilove, R., et al. (1996). Determination of sleep and wakefulness with the actigraph data analysis software (ADAS). *Sleep, 19,* 739-743.

Jennings, J. R., Muldoon, M. F., Hall, M., Buysse, D. J., & Manuck, S. B. (2007). Self-reported sleep quality is associated with the metabolic syndrome. *Sleep, 30,* 219-223.

Joffe, H., Soares, C. N., Thurston, R. C., White, D. P., Cohen, L. S., & Hall, J. E. (2009). Depression is associated with worse objectively and subjectively measured sleep, but not more frequent awakenings, in women with vasomotor symptoms. *Menopause: The Journal of the North American Menopause Society, 16,* 671-679.

Kamarck, T. W., Muldoon, M. F., Shiffman, S., Sutton-Tyrell, K. (2007). Experiences of demand and control during daily life are predictors of carotid atherosclerotic progression among healthy men. *Health Psychology, 26,* 324-332.

Kamarck, T. W., Muldoon, M. F., Shiffman, S., Sutton-Tyrell, K., Gwaltney, C., & Janicki, D. L. (2004). Experiences of demand and control in daily life as correlates of subclinical carotid atherosclerosis in a healthy older sample. *Health Psychology, 23,* 24-32.

Kawada, T., Kastsumata, M., Suzuki, H., & Shimizu, T. (2007). Actigraphic predictors of the depressive state in students with no psychiatric disorders. *Journal of Affective Disorders, 98,* 117-120.

Kawakami, N., Takatsuka, N., & Shimizu, H. (2004). Sleep disturbance and onset of type 2 diabetes. *Diabetes Care, 27,* 282-283.

Kazdin, A. E. (2003). *Research design in clinical psychology* (4th ed.). Boston, MA: Allyn and Bacon.

Kerlinger, F. N. (1986). *Foundations of behavioral research* (3rd ed.). New York: Holt, Rinehart, & Winston.

Kim, J., & Pai, M. (2010). Volunteering and trajectories of depression. *Journal of Aging and Health, 22,* 84-105.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: The Guilford Press.

Kline, T. J. B. (2005). *Psychological testing: A practical approach to design and evaluation.* Thousand Oaks, CA: Sage Publications.

Korszun, A., Young, E. A., Engleberg, N. C., Brucksch, C. B., Greden, J. F., & Crofford, L. A. (2002). Use of actigraphy for monitoring sleep and activity levels in patients with fibromyalgia and depression. *Journal of Psychosomatic Research, 52,* 439-443.

Lemmens, K. P. H., Abraham, C., Hoekstra, T., Ruiter, R. A. C., De Kort, W. L. A. M., Brug, J., & Schaalma, H. P. (2005). Why don't young people volunteer to give blood? An investigation of the correlates of donation intentions among young nondonors. *Transfusion, 45,* 945-955.

Lerner, D., Adler, D. A., Chang, H., Lapitsky, L., Hood, M. Y., Perissinotto, C., et al. (2004). Unemployment, job retention, and productivity loss among employees with depression. *Psychiatric Services, 55,* 1371-1378.

Mancini, G. B., Dahlof, B., & Diez, J. (2004). Surrogate markers for cardiovascular disease: structural markers. *Circulation, 109,* IV22-IV30.

McDonald, R. P. (1999). *Test theory: A unified treatment.* Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Means, M. K., Edinger, J. D., Glenn, D. M., & Fins, A. I. (2003). Accuracy of sleep perceptions among insomnia sufferers and normal sleepers. *Sleep Medicine, 4,* 285-296.

Mendlowicz, M. V., Jean-Louis, G., von Gizycki, H., Zizi, F., & Nunes, J. (1999). Actigraphic predictors of depressed mood in a cohort of non-psychiatric adults. *Australian and New Zealand Journal of Psychiatry, 33,* 553-558.

Metts, S., Sprecher, S., & Cupach, W. R. (1991). Retrospective self-reports. In: B. M. Montgomery & S. Duck (Ed.). *Studying Interpersonal Interaction* (pp. 162-178). New York, NY: The Guilford Press.

Musick, M. A., & Wilson, J. (2003). Volunteering and depression: The role of psychological and social resources in different age groups. *Social Science & Medicine, 56,* 259-269.

National Institute of Mental Health. (2008). *The numbers count: Mental disorders in America.* Retrieved from http://www.nimh.nih.gov/health/publications/the-numbers-count-mental-disorders-in-america/index.shtml#MajorDepressive.

Nolen-Hoeksema, S. Larson, J., & Grayson, C. (1999). Explaining the gender difference in depression. *Journal of Personality and Social Psychology, 77,* 1061-1072.

Nunnally, J. (1978). *Psychometric Theory* (2nd ed.). New York, NY: McGraw-Hill.

Nutt, D., Wilson, S., & Paterson, L. (2008). Sleep disorders as core symptoms of

    depression. *Dialogues in Clinical Neuroscience, 10,* 329-336.

Paudel, M. L., Taylor, B. C., Diem, S. J., Stone, K. L., Ancoli-Israel, S., Redline, S.,

    Ensrud, K. E., et al. (2008). Association between depressive symptoms and sleep

    in community-dwelling older men. *Journal of the American Geriatrics Society,*

    56, 1228-1235.

Phillips, B. A., & Danner, F. J. (1995). Cigarette smoking and sleep disturbance.

    Archives of Internal Medicine, 155, 734-737.

Pierce, J. P. (2009). Electronic recording, self-report, and bias in measuring cigarette

    consumption [Commentary]. *Health Psychology, 28,* 527-528.

Plous, S. (1993). *The psychology of judgment and decision making.* New York, NY:

    McGraw-Hill.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in

    the general population. *Applied Psychological Measurement, 1,* 385-401.

Radloff, L. S. (1991). The use of the Center for Epidemiologic Studies Depression Scale

    in adolescents and young adults. *Journal of Youth and Adolescence, 20,* 149-166.

Redelmeier, D. A., Katz, J., & Kahneman, D. (2003). Memories of colonoscopy: A

    randomized trial. *Pain, 104,* 187-194.

Rotenberg, V. S., Indursky, P., Kayumov, L., Sirota, P., & Melamed, Y. (2000). The

    relationship between subjective sleep estimation and objective sleep variables in

    depressed patients. *International Journal of Psychophysiology, 37,* 291-297.

Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct validity: The principle of aggregation. *Psychological Bulletin, 94,* 18-38.

Sadeh, A., Sharkey, K. M., & Carskadon, M. A. (1994). Activity-based sleep-wake identification: An empirical test of methodological issues. *Sleep, 17,* 201-207.

Schwab, D. P. (1999). *Research methods for organizational studies.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Schwartz, S., Anderson, W. M., Cole, S. R., Cornoni-Huntley, J., Hays, J. C., & Blazer, D. (1999). Insomnia and heart disease: A review of epidemiologic studies. *Journal of Psychosomatic Research, 47,* 313-333.

Scollon, C. N., Kim-Prieto, C., & Diener, E. (2003). Experience sampling: Promises and pitfalls, strengths and weaknesses. *Journal of Happiness Studies, 4,* 5-34.

Shiffman, S. (2009). How many cigarettes did you smoke? Assessing cigarette consumption by global report, time-line follow-back, and ecological momentary assessment. *Health Psychology, 28,* 519-526.

Shiffman, S., & Stone, A. A. (1998). Introduction to the special section: Ecological momentary assessment in health psychology. *Health Psychology, 17,* 3-5.

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology, 4,* 1-32.

Shultz, K. S., & Whitney, D. J. (2005). *Measurement theory in action.* Thousand Oaks, CA: Sage Publications.

Smyth, J. M., & Stone, A. A. (2003). Ecological momentary assessment research in behavioral medicine. *Journal of Happiness Studies, 4,* 35-52.

Sonnenschein, M., Mommersteeg, P. M. C., Houtveen, J. H., Sorbi, M. J., Schaufeli, W. B., van Doornen, L. J. P. (2007). Exhaustion and endocrine functioning in clinical burnout: An in-depth study using the experience sampling method. *Biological Psychology, 75,* 176-184.

Steptoe, A., Gibson, E. L., Hamer, M., & Wardle, J. (2007). Neuroendocrine and cardiovascular correlates of positive affect measured by ecological momentary assessment and by questionnaire. *Psychoneuroendocrinology, 32,* 56-64.

Stewart, J. C., & Stines, J. A. (March, 2008). *Shared and unique features of overlapping negative emotions as predictors of sleep quality.* Poster presented at the meeting of The Society of Behavioral Medicine, San Diego, CA.

Stone, A. A., Broderick, J. E., Schwartz, J. E., Shiffman, S., Litcher-Kelly, L., & Calvanese, P. (2003). Intensive momentary reporting of pain with an electronic diary: Reactivity, compliance, and patient satisfaction. *Pain, 104,* 343-351.

Stone, A. A., & Shiffman, S. (2002). Capturing momentary, self-report data: A proposal for reporting guidelines. *Annals of Behavioral Medicine, 24,* 236-243.

Stone, A. A., Shiffman, S., Atienza, A. A., & Nebeling, L. (2007). *The science of real-time data capture: Self-reports in health research.* Oxford, NY: Oxford University Press.

Strohle, A. (2009). Physical activity, exercise, depression and anxiety disorders. *Journal of Neural Transmission, 116,* 777-784.

Strube, M. J. (2000). Reliability and generalizability theory. In L. G. Grimm & P. R. Yarnold (Ed.), *Reading and understanding more multivariate statistics* (pp. 23-66). Washington, DC: APA.

Suls, J., & Bunde, J. (2005). Anger, anxiety, and depression as risk factors for

cardiovascular disease: The problems and implications of overlapping affective

dispositions. *Psychological Bulletin, 131,* 260-300.

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts

good adjustment, less pathology, better grades, and interpersonal success. *Journal

of Personality, 72,* 271-322.

Tourangeau, R. (2000). Remembering what happened: Memory errors and survey reports.

In A. A. Stone, J. S. Turkkan, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman, & V.

S. Cain (Eds.). *The science of self-report: Implications for research and practice*

(pp. 29-48). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Tscheulin, D. K., & Lindenmeier, J. (2005). The willingness to donate blood: An

empirical analysis of socio-demographic and motivation-related determinants.

Health Services Management Research, 18, 165-174.

Tsuchiyama, K., Nagayama, H., Kudo, K., Kojima, K., & Yamada, K. (2003).

Discrepancy between subjective and objective sleep in patients with depression.

Psychiatry and Clinical Neurosciences, 57, 259-264.

U.S. Department of Health and Human Services. (2012). Overweight and obesity: What

you can do. Retrieved from

http://www.surgeongeneral.gov/topics/obesity/calltoaction/fact_whatcanyoudo.ht

ml

Walsh, J. K., Dement, W. C., & Dinges, D. F. (2005). Sleep Medicine, public policy, and public health. In: M. H. Kryger, T. Roth, & W. C. Dement (Eds.), *Principles and Practice of Sleep Medicine* (4th ed.; pp. 648 – 656). Philadelphia, PA: Elsevier Saunders.

Wu, Y., Glynn, S. A., Schreiber, G. B., Wright, D. J., Lo, A., Murphy, E. L., Kleinman, S. H. et al. (2001). First-time blood donors: Demographic trends. *Transfusion, 41,* 360-364.

Yoshiuchi, K., Yamamoto, Y., & Akabayashi, A. (2008). Application of ecological momentary assessment in stress-related diseases. *BioPsychoSocial Medicine, 2,* 13-19.

TABLES

Table 1.
Timing of Procedures

Visit	Duration	Assessments
1 Monday or Tuesday, Week 1	60 minutes	Obtain Informed Consent Assess Exclusion Criteria Complete Retrospective CES-D Complete Demographic, Smoking, and Alcohol Items Complete EMA Training Schedule Visits 2-4
EMA Monitoring Period Week 1	3 days	Complete Momentary CES-D Assessments
2 Thursday or Friday, Week 1	10 minutes	Return EMA Equipment Complete Retrospective CES-D Obtain Height and Weight
3 Monday or Tuesday, Week 2	10 minutes	Complete Actigraphy Training
Actigraphy Monitoring Period, Week 2	3 days	Wear Actigraphy Belt Except While Bathing Complete Sleep Diary Twice Each Day
4 Thursday or Friday,	10 minutes	Return Actigraphy Equipment Complete Blood Donation Intention

Week 2	Scale
	Debriefing and Compensation

Note. EMA = Ecological Momentary Assessment. CES-D = Center for Epidemiologic Studies-Depression Scale.

Table 2.
Descriptive Statistics for the Covariates, Predictors, and Outcomes

Variable	*M*	*SD*	Range	Cronbach's Alpha
Covariate Variables				
Age, years	21.6	5.2	18-47	
Race/ethnicity, % white	67			
BMI, kg/m^2	25.8	7.1	17.5-57.6	
Tobacco Use, % yes	18.8			
Alcohol Use (g/day)	5.4	8.3	0-43.7	
Handgrip Strength (kg)	23.5	5.6	7-38	
Predictor Variables				
EMA CES-D Total Mean	9.6	6.4	0.4-39.8	.91
Session 1 CES-D Total	13.1	10.6	0-54	.93
Biological Outcomes				
Sleep Duration (hrs)	7.5	1.3	4.6-11.1	
Energy Expenditure (kcal)	294	188.1	55.7-1321.5	
Behavioral Outcomes				
Blood Donation Total Score	14.7	6.8	3-21	.94
Handgrip duration (seconds)	30.2	14.6	0-70	
Cumulative GPA	2.91	0.75	1.43-4.00	
Semester GPA	2.94	0.88	0.54-4.00	

Note. $N = 96$. BMI = Body Mass Index. EMA = Ecological Momentary Assessment. CES-D = Center for Epidemiologic Studies-Depression Scale. GPA = Grade Point Average.

Table 3.

Correlation Table for the Covariates, Predictors, and Outcomes

Pearson Correlation Coefficient

Variable	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Age	.25*	.24*	.22*	.14	-.07	-.07	.06	-.08	.12	.22*	.13	-.12	-.06
2. Race/Ethn	--	.26**	.07	-.09	-.22*	.18	.13	-.29**	-.01	-.01	-.02	-.31**	-.32**
3. BMI	--	--	-.11	-.09	.26*	-.09	.02	-.19	.48**	-.14	.04	-.24*	-.20
4. Tobacco	--	--	--	.36**	-.07	.35**	.29**	-.11	-.17	.08	.13	-.18	-.22*
5. Alcohol	--	--	--	--	-.01	.02	.02	.05	-.08	.05	-.04	-.10	-.05
6. Handgrip S	--	--	--	--	--	-.06	.01	.05	.30**	-.11	.17	.06	.10
7. EMA CES-D	--	--	--	--	--	--	.75**	-.08	-.19	-.15	-.04	-.21*	-.20
8. S1 CES-D	--	--	--	--	--	--	--	.01	-.11	-.12	.03	-.01	-.05
9. Sleep Dur	--	--	--	--	--	--	--	--	-.18	.03	-.22*	.05	.06
10. Energy Exp	--	--	--	--	--	--	--	--	--	-.21*	.06	.05	.10
11. BD Score	--	--	--	--	--	--	--	--	--	--	-.05	-.00	.05

12. Handgrip D	--	--	--	--	--	--	--	--	--	--	--	.05	.03
13. Cum. GPA	--	--	--	--	--	--	--	--	--	--	--	--	.88**
14. Sem. GPA	--	--	--	--	--	--	--	--	--	--	--	--	--

Note. Race/Ethn = Race/Ethnicity. BMI = Body Mass Index. Tobacco = currently a tobacco user. Alcohol = grams per day of alcohol. Handgrip S = Handgrip Strength in kg. EMA = Ecological Momentary Assessment. CES-D = Center for Epidemiologic Studies-Depression Scale. S1 = Session 1. Sleep Dur = Average Sleep Duration. Energy Exp = Average Energy Expenditure in kcal. BD Score = Blood Donation Intention Total Score. Handgrip D = Handgrip Duration in seconds. Cum. = Cumulative. GPA = Grade Point Average. Sem. = Semester.

*p < .05.

**p < .01.

Table 4.
Mean Differences between Completers and Non-completers

Variable	Completers	Non-completers	t/χ^2	p
	M (SD)	*M (SD)*		
Demographic and Covariate Variables				
Age	21.57 (5.25)	22.20 (6.29)	.53	.60
Race/Ethnicity, % white	67.7	72.0	.17	.68
BMI, kg/m^2	25.82 (7.13)	24.67 (3.88)	-.70	.49
Tobacco Use, % yes	17.7	12.0	.76	.68
Daily Alcohol Use (g/day)	10.05 (10.43)	7.15 (5.28)	-.85	.40
Handgrip Strength (kg)	23.45 (5.57)	24.03 (6.57)	.41	.68
Predictor Variables				
EMA CES-D Total Mean	9.75 (6.37)	9.65 (5.11)	.0001	.99
Session 1 CES-D Total	13.19 (10.57)	10.40 (6.48)	-1.22	.23
Biological Outcomes				
Sleep Duration (hrs)	7,.49 (1.28)	7.41 (0.64)	-.09	.93
Energy Expenditure	2.41 (0.22)	2.16 (0.13)	-1.95	.05
Behavioral Outcomes				
Blood Donation Total Score	14.64 (6.75)	11.20 (9.07)	-1.11	.27
Handgrip duration (seconds)	29.90 (14.38)	31.51 (16.65)	.35	.72
Cumulative GPA	2.91 (0.75)	3.17 (0.55)	1.54	.13
Semester GPA	2.95 (0.88)	3.12 (0.89)	.87	.39

ote. BMI = Body Mass Index. EMA = Ecological Momentary Assessment. CES-D = Center for Epidemiologic Studies-Depression Scale. GPA = Grade Point Average.

Table 5.
Unadjusted Regression Analyses

Variable	N	EMA CES-D Total Mean				Session 1 CES-D Total				*B* Comparison	
		B	*SE_B*	*β*	*p*	*B*	*SE_B*	*β*	*p*	*Z*	*p*
Biological Outcomes											
Sleep Duration	96	-.94	1.23	-.08	.45	.06	.75	.01	.93	-.70	.48
Energy Expenditure	94	-.01	.00	-.19	.06	.00	.00	-.11	.30	-1.25	.22
Behavioral Outcomes											
Blood Donation Total Score	95	-.16	.11	-.15	.14	-.07	.07	-.12	.27	-.69	.50
Handgrip duration	96	-.09	.24	-.04	.70	.04	.14	.03	.78	-.48	.64
Cumulative GPA	90	-.03	.01	-.21	.05	.00	.01	-.01	.89	-1.69	.10
Semester GPA	91	-.03	.02	-.20	.06	-.01	.01	-.05	.63	-1.30	.20

Note. EMA = Ecological Momentary Assessment. CES-D = Center for Epidemiologic Studies-Depression Scale. GPA = Grade Point Average

Table 6.
Fully Adjusted Regression Analyses

Variable	N	EMA CES-D Total Mean				Session 1 CES-D Total				B Comparison	
		B	SE_B	β	p	B	SE_B	β	p	Z	p
Biological Outcomes											
Sleep Duration[a]	96	.05	1.32	.00	.97	.67	.76	.09	.38	-.41	.68
Energy Expenditure[a]	94	.00	.00	-.12	.23	.00	.00	-.06	.54	-.75	.46
Behavioral Outcomes											
Blood Donation Total Score[b]	95	-.14	.11	-.13	.21	-.08	.07	-.12	.23	-.47	.64
Handgrip duration[c]	96	-.04	.24	-.02	.86	.03	.14	.02	.84	-.26	.80
Cumulative GPA[b]	90	-.02	.01	-.18	.08	.00	.01	.01	.89	-1.56	.12
Semester GPA[b]	91	-.03	.02	-.16	.11	.00	.01	-.03	.77	-1.28	.20

Note. EMA = Ecological Momentary Assessment. CES-D = Center for Epidemiologic Studies-Depression Scale. GPA = Grade Point Average.
[a]Adjusted for age, race/ethnicity, BMI, tobacco use, and alcohol use.
[b]Adjusted for age and race/ethnicity.
[c]Adjusted for age, race/ethnicity, and handgrip strength.

Table 7.
Mean Differences between Semesters

Variable	Fall 2010 M (SD)	Spring 2011 M (SD)	Summer 2011 M (SD)	F/χ^2	p
Demographic and Covariate Variables					
Age	19.00 (1.23)	21.63 (5.64)	22.08 (2.58)	.66	.52
Race/Ethnicity, % white	80.0	64.6	83.3	2.04	.36
BMI, kg/m^2	23.85 (4.63)	26.04 (7.52)	25.19 (5.25)	.27	.76
Tobacco Use, % yes	0	21.5	8.3	2.41	.30
Daily Alcohol Use (g/day)	0.27 (0.54)	5.06 (7.36)	10.10 (13.34)	3.05	.05
Handgrip Strength (kg)	22.67 (6.44)	23.72 (5.86)	22.92 (5.71)	.18	.84
Predictor Variables					
EMA CES-D Total Mean	7.58 (2.78)	9.88 (6.60)	8.70 (6.34)	.44	.65
Session 1 CES-D Total	5.00 (1.87)	14.19 (10.82)	9.25 (8.72)	2.80	.07
Biological Outcomes					
Sleep Duration (hrs)	8.12 (1.44)	7.37 (1.27)	7.92 (1.19)	1.67	.19

Energy Expenditure	2.34 (0.26)	2.41 (0.21)	2.44 (0.23)	.33	.72
Behavioral Outcomes					
Blood Donation Total Score	14.80 (8.32)	14.41 (6.77)	16.58 (6.23)	.54	.59
Handgrip duration (seconds)	39.00 (21.10)	28.74 (13.96)	36.24 (14.58)	2.40	.07
Cumulative GPA	3.09 (1.05)	2.84 (0.74)	3.34 (0.53)	2.14	.12
Semester GPA	3.09 (1.05)	2.84 (0.88)	3.60 (0.47)	3.56*	.03

Note. BMI = Body Mass Index. EMA = Ecological Momentary Assessment. CES-D = Center for Epidemiologic Studies-Depression Scale. GPA = Grade Point Average

$*p < .05$

Table 8.

Unadjusted Regression Analyses for Spring 2011 Semester Participants

Variable	N	EMA CES-D Total Mean				Session 1 CES-D Total				B Comparison	
		B	SE_B	β	p	B	SE_B	β	p	Z	p
Biological Outcomes											
Sleep Duration	79	-.47	1.31	-.04	.72	.57	.80	.08	.48	.68	.50
Energy Expenditure	77	-.01	.00	-.15	.20	-.00	.00	-.11	.33	.67	.50
Behavioral Outcomes											
Blood Donation Total Score	78	-.14	.12	-.13	.24	-.04	.07	-.07	.56	-.70	.48
Handgrip duration	79	.01	.24	.01	.96	.14	.15	.11	.36	-.43	.67
Cumulative GPA	75	-.02	.02	-.17	.15	.01	.01	.08	.51	-1.56	.12
Semester GPA	76	-.03	.02	-.19	.10	.00	.01	.03	.79	-1.60	.11

Note. EMA = Ecological Momentary Assessment. CES-D = Center for Epidemiologic Studies-Depression Scale. GPA = Grade Point Average.

Table 9.

Adjusted Regression Analyses for Spring 2011 Semester Participants

Variable	N	EMA CES-D Total Mean				Session 1 CES-D Total				B Comparison	
		B	SE$_B$	β	p	B	SE$_B$	β	p	Z	p
Biological Outcomes											
Sleep Duration	79	.47	1.43	.04	.74	1.17	.81	.17	.16	-.42	.67
Energy Expenditure	77	-.00	.00	-.03	.77	-.00	.00	-.07	.49	.00	1.00
Behavioral Outcomes											
Blood Donation Total Score	78	-.13	.12	-.13	.27	-.05	.07	-.08	.48	-.58	.56
Handgrip duration	79	.04	.24	.02	.87	.12	.15	.09	.42	-.27	.79
Cumulative GPA	75	-.02	.02	-.16	.16	.01	.01	.09	.44	-1.65	.10
Semester GPA	76	-.03	.02	-.18	.12	.00	.01	.03	.77	-1.52	.13

Note. EMA = Ecological Momentary Assessment. CES-D = Center for Epidemiologic Studies-Depression Scale. GPA = Grade Point Average.
[a]Adjusted for age, race/ethnicity, BMI, tobacco use, and alcohol use.
[b]Adjusted for age and race/ethnicity.
[c]Adjusted for age, race/ethnicity, and handgrip strength.

Table 10.

Unadjusted Regression Analyses for Participants without Additional Day of Monitoring

Variable	N	EMA CES-D Total Mean				Session 1 CES-D Total				B Comparison	
		B	SE_B	β	p	B	SE_B	β	p	Z	p
Biological Outcomes											
Sleep Duration	81	-.98	1.45	-.08	.50	.06	.90	.01	.94	-.61	.54
Energy Expenditure	80	-.01	.00	-.23	.04*	-.00	.00	-.10	.38	-1.33	.18
Behavioral Outcomes											
Blood Donation Total Score	80	-.06	.12	-.06	.60	-.03	.08	-.04	.71	-.25	.80
Handgrip duration	81	-.13	.26	-.06	.61	-.04	.16	-.03	.82	-.31	.76
Cumulative GPA	76	-.03	.02	-.18	.12	.00	.01	.04	.74	-1.52	.13
Semester GPA	77	-.03	.02	-.15	.18	-.00	.01	-.01	.96	-1.09	.28

Note. EMA = Ecological Momentary Assessment. CES-D = Center for Epidemiologic Studies-Depression Scale. GPA = Grade Point Average.

Table 11.

Adjusted Regression Analyses for Participants with an Additional Day of Monitoring

Variable	N	EMA CES-D Total Mean				Session 1 CES-D Total				B Comparison	
		B	SE$_B$	β	p	B	SE$_B$	β	p	Z	P
Biological Outcomes											
Sleep Duration	81	-.10	1.52	-.01	.95	.66	.90	.08	.47	-.43	.67
Energy Expenditure	80	-.00	.00	-.11	.27	-.00	.00	-.04	.65	-.83	.41
Behavioral Outcomes											
Blood Donation Total Score	80	-.04	.12	-.04	.73	-.03	.08	-.05	.69	-.09	.93
Handgrip duration	81	-.16	.27	-.07	.56	-.07	.17	-.05	.67	-.28	.78
Cumulative GPA	76	-.02	.02	-.16	.16	.01	.01	.07	.55	-1.58	.11
Semester GPA	77	-.02	.02	-.13	.24	.00	.01	.02	.89	-1.12	.26

Note. EMA = ecological momentary assessment. CES-D = Center for Epidemiologic Studies-Depression Scale. GPA = grade point average.

[a]Adjusted for age, race/ethnicity, BMI, tobacco use, and alcohol use.

[b]Adjusted for age and race/ethnicity.

[c]Adjusted for age, race/ethnicity, and handgrip strength.

Table 12.

Unadjusted Regression Analyses—Morning EMA (8:55AM, 10:55AM)

Variable	N	Morning EMA CES-D Total Mean				Session 1 CES-D Total				*B* Comparison	
		B	*SE_B*	*β*	*p*	*B*	*SE_B*	*β*	*p*	*Z*	*p*
Biological Outcomes											
Sleep Duration	96	-1.10	1.25	-.09	.38	.06	.75	.01	.93	-.80	.42
Energy Expenditure	94	-.01	.00	-.15	.16	.00	.00	-.11	.30	-.67	.50
Behavioral Outcomes											
Blood Donation Total Score	95	-.16	.11	-.15	.15	-.07	.07	-.12	.27	-.67	.50
Handgrip duration	96	-.16	.24	-.07	.50	.04	.14	.03	.78	-.72	.47
Cumulative GPA	90	-.02	.01	-.19	.08	.00	.01	-.01	.89	-1.43	.15
Semester GPA	91	-.03	.02	-.17	.12	-.01	.01	-.05	.63	-1.09	.28

Note. EMA = Ecological Momentary Assessment. CES-D = Center for Epidemiologic Studies-Depression Scale. GPA = Grade Point Average.

Table 13.
Unadjusted Regression Analyses—Afternoon EMA (12:55PM, 2:55PM, 4:55PM)

Variable	N	Afternoon EMA CES-D Total Mean				Session 1 CES-D Total				*B* Comparison	
		B	SE$_B$	β	p	B	SE$_B$	β	p	Z	p
Biological Outcomes											
Sleep Duration	96	-1.20	1.18	-.10	.32	.06	.75	.01	.93	-.90	.37
Energy Expenditure	94	-.01	.00	-.22	.03*	.00	.00	-.11	.30	-1.39	.16
Behavioral Outcomes											
Blood Donation Total Score	95	-.09	.11	-.09	.40	-.07	.07	-.12	.27	-.14	.89
Handgrip duration	96	-.05	.23	-.02	.84	.04	.14	.03	.78	-.32	.75
Cumulative GPA	90	-.03	.01	-.19	.07	.00	.01	-.01	.89	-1.49	.14
Semester GPA	91	-.03	.02	-.20	.06	-.01	.01	-.05	.63	-1.42	.16

Note. EMA = Ecological Momentary Assessment. CES-D = Center for Epidemiologic Studies-Depression Scale. GPA = Grade Point Average.

Table 14.

Unadjusted Regression Analyses—Evening EMA (6:55PM, 8:55PM)

Variable	N	Evening EMA CES-D Total Mean				Session 1 CES-D Total				B Comparison	
		B	SE_B	β	p	B	SE_B	β	p	Z	p
Biological Outcomes											
Sleep Duration	96	-.72	1.13	-.07	.53	.06	.75	.01	.93	-.58	.56
Energy Expenditure	94	-.01	.00	-.19	.06	.00	.00	-.11	.30	-1.11	.27
Behavioral Outcomes											
Blood Donation Total Score	95	-.18	.10	-.19	.07	-.07	.07	-.12	.27	-.94	.35
Handgrip duration	96	-.03	.22	-.02	.88	.04	.14	.03	.78	-.28	.78
Cumulative GPA	90	-.03	.01	-.19	.07	.00	.01	-.01	.89	-1.57	.12
Semester GPA	91	-.03	.02	-.19	.07	-.01	.01	-.05	.63	-1.31	.19

Note. EMA = Ecological Momentary Assessment. CES-D = Center for Epidemiologic Studies-Depression Scale. GPA = Grade Point Average.

Table 15.
Fully Adjusted Regression Analyses—Morning EMA (8:55AM, 10:55AM)

Variable	N	Morning EMA CES-D Total Mean				Session 1 CES-D Total				B Comparison	
		B	SE_B	β	p	B	SE_B	β	p	Z	p
Biological Outcomes											
Sleep Duration[a]	96	-.38	1.30	-.03	.77	.67	.76	.09	.38	-.70	.48
Energy Expenditure[a]	94	-.00	.00	-.08	.44	.00	.00	-.06	.54	-.56	.58
Behavioral Outcomes											
Blood Donation Total Score[b]	95	-.13	.11	-.12	.23	-.08	.07	-.12	.23	-.43	.67
Handgrip duration[c]	96	-.13	.24	-.06	.59	.03	.14	.02	.84	-.57	.57
Cumulative GPA[b]	90	-.02	.01	-.18	.08	.00	.01	.01	.89	-1.57	.12
Semester GPA[b]	91	-.02	.02	-.15	.14	.00	.01	-.03	.77	-1.14	.25

Note. EMA = Ecological Momentary Assessment. CES-D = Center for Epidemiologic Studies-Depression Scale. GPA = Grade Point Average.
[a]Adjusted for age, race/ethnicity, BMI, tobacco use, and alcohol use.

[b]Adjusted for age and race/ethnicity.
[c]Adjusted for age, race/ethnicity, and handgrip strength.

Table 16.

Fully Adjusted Regression Analyses—Afternoon EMA (12:55PM, 2:55PM, 4:55PM)

Variable	N	Afternoon EMA CES-D Total Mean				Session 1 CES-D Total				*B* Comparison	
		B	*SE*$_B$	*β*	*p*	*B*	*SE*$_B$	*β*	*p*	*Z*	*p*
Biological Outcomes											
Sleep Duration[a]	96	-.29	1.29	-.03	.82	.67	.76	.09	.38	-.64	.52
Energy Expenditure[a]	94	-.01	.00	-.14	.17	.00	.00	-.06	.54	-1.11	.27
Behavioral Outcomes											
Blood Donation Total Score[b]	95	-.08	.11	-.08	.46	-.08	.07	-.12	.23	.00	1.00
Handgrip duration[c]	96	-.03	.23	-.02	.88	.03	.14	.02	.84	-.23	.82
Cumulative GPA[b]	90	-.02	.01	-.16	.12	.00	.01	.01	.89	-1.44	.15
Semester GPA[b]	91	-.03	.02	-.17	.10	.00	.01	-.03	.77	-1.25	.21

Note. EMA = Ecological Momentary Assessment. CES-D = Center for Epidemiologic Studies-Depression Scale. GPA = Grade Point Average.
[a]Adjusted for age, race/ethnicity, BMI, tobacco use, and alcohol use.

[b]Adjusted for age and race/ethnicity.
[c]Adjusted for age, race/ethnicity, and handgrip strength.

Table 17.
Fully Adjusted Regression Analyses—Evening EMA (6:55PM, 8:55PM)

Variable	N	Evening EMA CES-D Total Mean				Session 1 CES-D Total				B Comparison	
		B	SE_B	β	p	B	SE_B	β	p	Z	p
Biological Outcomes											
Sleep Duration[a]	96	.12	1.22	.01	.92	.67	.76	.09	.38	-.55	.58
Energy Expenditure[a]	94	-.00	.00	-.11	.26	.00	.00	-.06	.54	-.83	.41
Behavioral Outcomes											
Blood Donation Total Score[b]	95	-.16	.10	-.17	.10	-.08	.07	-.12	.23	-.72	.47
Handgrip duration[c]	96	.00	.22	.00	1.0	.03	.14	.02	.84	-.11	.91
Cumulative GPA[b]	90	-.02	.01	-.18	.08	.00	.01	.01	.89	-1.57	.12
Semester GPA[b]	91	-.03	.02	-.17	.09	.00	.01	-.03	.77	-1.31	.19

Note. EMA = Ecological Momentary Assessment. CES-D = Center for Epidemiologic Studies-Depression Scale. GPA = Grade Point Average.
[a]Adjusted for age, race/ethnicity, BMI, tobacco use, and alcohol use.

[b]Adjusted for age and race/ethnicity.
[c]Adjusted for age, race/ethnicity, and handgrip strength.

APPENDICES

Appendix A. Diagrams for Studies Included in the Literature Review

Kamarck, Muldoon, Shiffman, Sutton-Tyrrell, Gwaltney, & Janicki, 2004

Retrospective paper and pencil measure of Psychological Demands and Decision Latitude

$b_1 = .00, R^2_1 = .00;$
$b_2 = .00, R^2_2 = .00$

Presence of carotid IMT

Momentary assessment of work-related and non-work related task demand

$b_1 = .02, R^2_1 = .03*;$
$b_2 = .02, R^2_2 = .02*$

Kamarck, Muldoon, Shiffman, & Sutton-Tyrell, 2007

Retrospective paper and pencil measure of Psychological Demands and Decision Latitude

$R^2_1 = .02; R^2_2 = .01$

Progression of carotid IMT

Momentary assessment of work-related and non-work related task demand
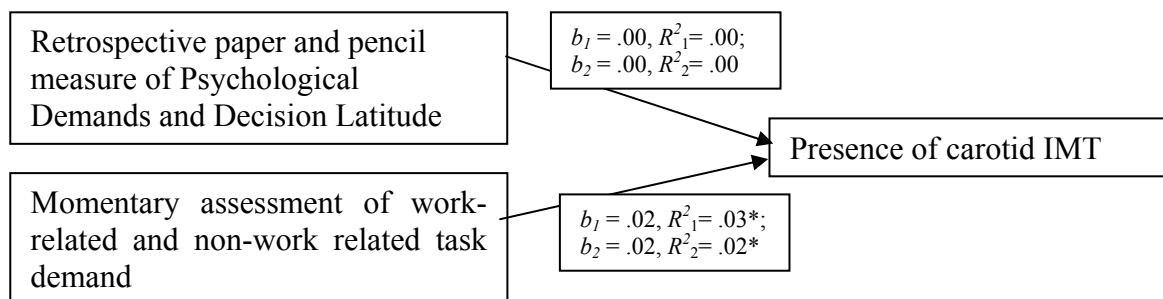
$R^2_1 = .08*; partial R^2_2 = .07*$
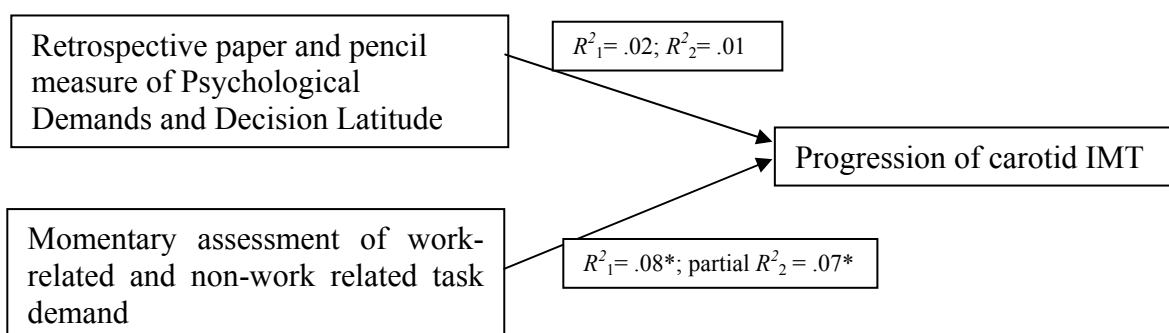
Sonnenschein, Mommersteeg, Houtveen, Sorbi, Schauefeli, & van Doornen, 2007

Retrospective paper and pencil measure of exhaustion

$\beta = .17$

Prediction of HPA axis functioning

Momentary measure of exhaustion

$\beta = -.69*$

Shiffman, 2009

Time-line follow-back (TLFB) measure of cigarettes smoked

$\beta_1 = .10; \beta_2 = .01$

Prediction of cotinine levels and carbon monoxide levels

Momentary measure of cigarettes smoked

$\beta_1 = .33*; \beta_2 = .34*$
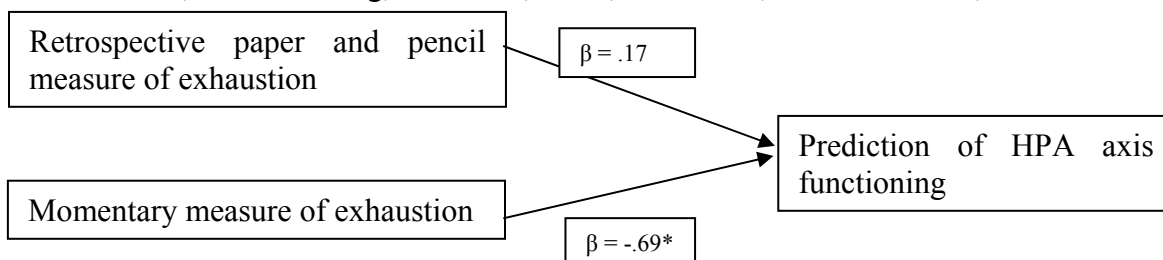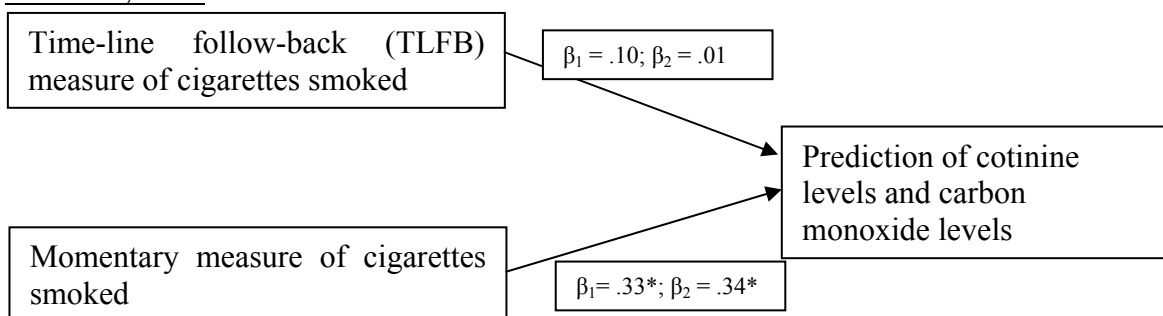
* = $p < .05$

Helgeson, Lopez, & Kamarck, 2009

| Retrospective paper and pencil measure of friend support and friend conflict | $\beta_1 = .22; \beta_2 = -.97*$ |
| Momentary measure of interaction enjoyment and interaction upset | $\beta_1 = .10, \beta_2 = .00$ |

Presence of metabolic control
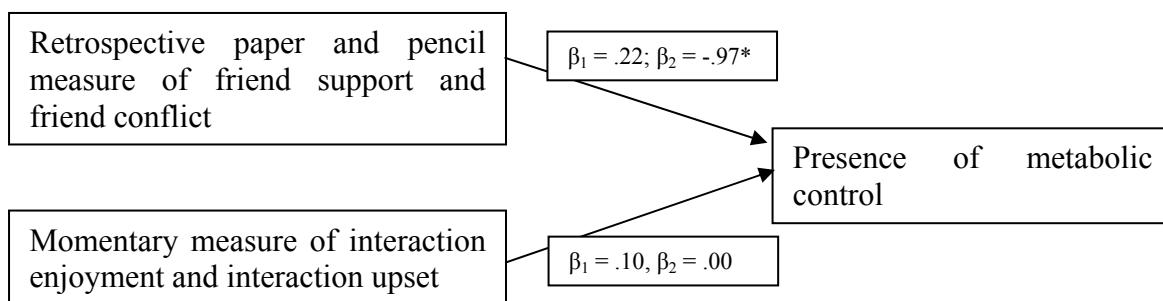
Redelmeier, Katz, & Kahneman, 2003

| Retrospective paper and pencil measure of total discomfort | $p < .002*$ |
| Momentary measure of current pain | $p > 0.10$ |

Colonoscopy return rate
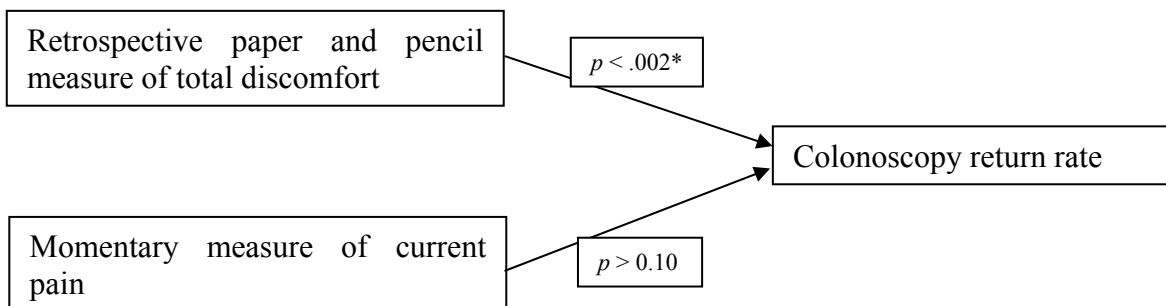
Steptoe, Gibson, Hamer, & Wardle, 2007 (non-significant βs were not consistently reported)

| Retrospective paper and pencil PANAS positive affect | $\beta s = n.s.$ |
| Momentary ratings of positive affect | $\beta s = -.23$ to $-.32*$ |

Prediction of cortisol responses

| Retrospective paper and pencil PANAS positive affect | $\beta s = -.32$ to $-.43*$ for SBP and $\beta s = n.s.$ for DBP |
| Momentary ratings of positive affect | $\beta s = -.24$ to $-.29*$ for SBP and $\beta s = -.23$ to $-.35*$ for DBP |

Prediction of cardiovascular responses
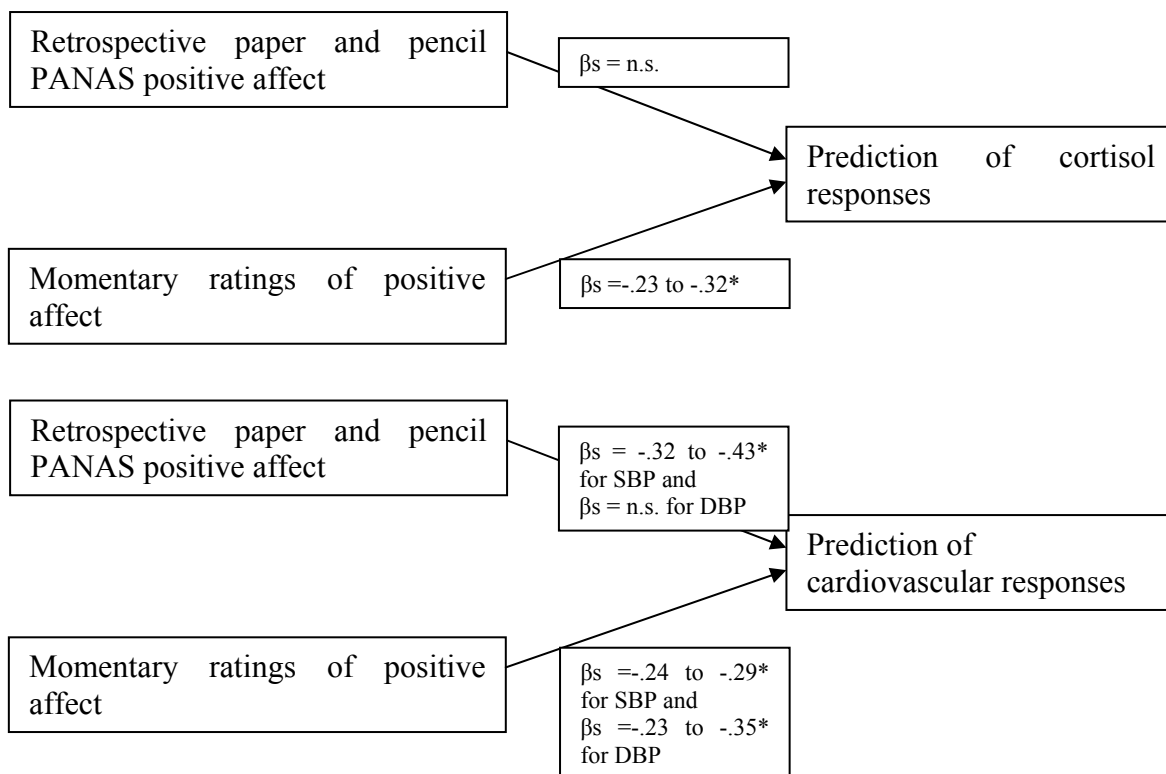
Appendix B. Retrospective CES-D

INSTRUCTIONS: For each statement, please mark the response option that best describes how you have been feeling *in the past week*.

		Rarely or none of the time (less than 1 day)	Some or a little of the time (1 – 2 days)	Occasionally or a moderate amount of the time (3 – 4 days)	Most or all of the time (5 – 7 days)
1.	I was bothered by things that usually don't bother me.				
2.	I did not feel like eating; my appetite was poor.				
3.	I felt that I could not shake off the blues, even with the help from family or friends.				
4.	I felt that I was just as good as other people.				
5.	I had trouble keeping my mind on what I was doing.				
6.	I felt depressed.				
7.	I felt that everything I did was an effort.				
8.	I felt hopeful about the future.				
9.	I thought my life had been a failure.				

		Rarely or none of the time (less than 1 day)	Some or a little of the time (1 – 2 days)	Occasionally or a moderate amount of the time (3 – 4 days)	Most or all of the time (5 – 7 days)
10.	I felt fearful.				
11.	My sleep was restless.				
12.	I was happy.				
13.	I talked less than usual.				
14.	I felt lonely.				
15.	People were unfriendly.				
16.	I enjoyed life.				
17.	I had crying spells.				
18.	I felt sad.				
19.	I felt that people dislike me.				
20.	I could not get "going".				

Appendix C. Sleep Diary Items

1. At what time did you shut off the light to go to bed (not reading or watching TV)?

   ___ ___:___ ___ A.M./P.M.

2. At what time did you wake up this morning?

   ___ ___:___ ___ A.M./P.M.

Appendix D. Blood Donation Intention Scale

1. I intend to give blood in the next 8 weeks.

    1   2   3   4   5   6   7

    Likely   Unlikely

2. I have decided to give blood in the next 8 weeks.

    1   2   3   4   5   6   7

    Agree   Disagree

3. I will try to give blood in the next 8 weeks.

    1   2   3   4   5   6   7

    Probable   Improbable

Appendix E. Demographic, Smoking, Alcohol, Class Attendance, and
Work Attendance Items

1. What is your date of birth (e.g., 06/28/1980)?

2. What is your sex?   (a) male (b) female

3. What race/ethnicity do you consider yourself?   (a) White/Caucasian; (b)
    Black/African American; (c) Hispanic/Latino; (d) Asian; (e) Native
    Hawaiian/Pacific Islander; (f) American Indian/Alaskan Native; (g) Other

4. Have you ever smoked cigarettes? (a) Yes, currently; (b) Yes, but not anymore;
    (c) No, never used

5. Have you ever used chewing tobacco? (a) Yes, currently; (b) Yes, but not
    anymore; (c) No, never used

6. Have you ever smoked a pipe or cigar? (a) Yes, currently; (b) Yes, but not
    anymore; (c) No, never used

7. During the past 12 months, how often did you drink wine? (a) Every day; (b)
    Nearly every day; (c) 3-4 times per week; (d) 1-2 times per week; (e) 2-3 times
    per month; (f) 1 time per month; (g) Less than 1 time per month but at least 1 time
    per year; (h) Less than 1 time per year; (i) Never drink wine

8. On the days you drank wine, on average, how many glasses of wine (4 oz) did
    you drink? (a) 0; (b) 1; (c) 2; (d) 3; (e) 4; (f) 5; (g) 6 or more

9. On the days that you drank wine, how often did you drink five or more glasses?
    (a) Always; (b) About three-quarters of the time; (c) About half of the time; (d)
    About one-quarter of the time; (e) Less than one-quarter of the time; (f) Never

10. During the past 12 months, how often did you drink beer? (a) Every day; (b) Nearly every day; (c) 3-4 times per week; (d) 1-2 times per week; (e) 2-3 times per month; (f) 1 time per month; (g) Less than 1 time per month but at least 1 time per year; (h) Less than 1 time per year; (i) Never drink beer

11. On the days you drank beer, on average, how many bottles or cans of beer (12 oz) did you drink? (a) 0; (b) 1; (c) 2; (d) 3; (e) 4; (f) 5; (g) 6 or more

12. On the days that you drank beer, how often did you drink five or more bottles or cans? (a) Always; (b) About three-quarters of the time; (c) About half of the time; (d) About one-quarter of the time; (e) Less than one-quarter of the time; (f) Never

13. During the past 12 months, how often did you drink liquor, including whiskey, rum, gin, vodka, bourbon, scotch, or liquers? (a) Every day; (b) Nearly every day; (c) 3-4 times per week; (d) 1-2 times per week; (e) 2-3 times per month; (f) 1 time per month; (g) Less than 1 time per month but at least 1 time per year; (h) Less than 1 time per year; (i) Never drink whiskey or liquor

14. On the days you drank liquor, on average, how many shots of beer (1.0 – 1.5 oz) did you drink? (a) 0; (b) 1; (c) 2; (d) 3; (e) 4; (f) 5; (g) 6 or more

15. On the days that you drank liquor, how often did you drink five or more shots? (a) Always; (b) About three-quarters of the time; (c) About half of the time; (d) About one-quarter of the time; (e) Less than one-quarter of the time; (f) Never

16. How often did you experience a hangover in the past 12 months? (a) Never; (b) 1 time per year; (c) 2-3 times per year; (d) 4-5 times per year; (e) 1 time every 2 months; (f) 1 time per month; (g) 2-3 times per month; (h) 1 time per week; (i) At least 2 times per week

17. Are you a full-time or part-time student? (a) full-time; (b) part-time

18. How many credit hours are you registered for this semester?

19. How many total class sessions do you have each week? (e.g., PSY-B 380 meets 2 times + BIOL-N 212 meets 2 times + ENG-W 131 meets 2 times + CSCI-N 100 meets 2 times + SOC-R 325 meets 2 times = 10 class sessions each week?

20. In a typical week, how many of these class sessions do you miss?

21. Are you currently employed? (a) No; (b) Yes

22. If you are currently employed, how many total work shifts (1 shift = 8 hours) do you have each week?

23. If you are currently employed, how many of these work shifts do you miss?

VITA

VITA

Desiree Joy Zielke

<u>Education</u>

August 2012	Doctor of Philosophy
	Clinical Psychology
	Degree Granting Institution: Purdue University
	Campus: Indiana University-Purdue University Indianapolis (IUPUI), Indianapolis, IN
	APA Accredited Program
	Dissertation Title: Ecological Momentary Assessment versus Traditional Retrospective Self-Reports as Predictors of Health-Relevant Outcomes
August 1, 2011 – July 27, 2012	Predoctoral Internship
	Salem Veterans Affairs Medical Center
	Major Rotations: Behavioral Medicine, Outpatient Psychological Services
	Minor Rotations: Psychiatric Emergency Room, Substance Abuse Liaison Team, Palliative Care
	Director of Training: Dana Holohan, Ph.D.
2007	Master of Science
	Clinical Psychology
	North Dakota State University, Fargo, ND
	Thesis Title: Providing post-task social support after an acute stress situation to enhance cardiovascular recovery.
2005	Bachelor of Science
	North Dakota State University, Fargo, ND
	Major: Psychology, Minors: Chemistry and History
2001	Associate of Arts
	Bismarck State College, Bismarck, ND

<u>Clinical Experience</u>

Position: Predoctoral Intern
Site: Outpatient Psychological Services; Salem VA Medical Center,
                             Salem, VA
Dates: February 2012 – Present; Major Rotation
Supervisors: Susan Duma, Psy.D., Sarah Voss Horrell, Ph.D., Theodore
                             Wright, Ph.D.

<u>Duties</u>: Conduct intake evaluations and individual therapy, using empirically supported treatments (e.g., CBT, ACT), for veterans with a variety of mental health problems (e.g., depression, obsessive-compulsive disorder, anger management, couples therapy) in a general mental health outpatient setting. I have endeavored to develop an emphasis in acceptance and commitment therapy (ACT) as part of this rotation by co-facilitating an ACT for Depression group and using ACT, when appropriate, for individual therapy. Conduct suicide risk assessments for veterans. Write intake evaluation reports, progress notes, treatment plans, integrated assessment reports, and discharge summaries. Conduct differential diagnosis for referring providers using thorough clinical interviews, the SCID-I, the SCID-II, the Clinician-Administered PTSD Scale (CAPS), and personality assessments. I have also co-led a Dialectical Behavior Therapy (DBT) Skills group for men.
<u>Tests given</u>: Minnesota Multiphasic Personality Inventory-Second Edition (MMPI-2), Millon Clinical Multiaxial Inventory-III (MCMI-III), Alcohol Use Disorder Test (AUDIT), Beck Depression Inventory-II (BDI-II), Beck Anxiety Inventory (BAI), PTSD Checklist-Civilian Version (PCL-C), and Repeatable Battery for the Assessment of Neuropsychological Status (RBANS).

Position: Predoctoral Intern
Site: Palliative Care; Salem VA Medical Center, Salem, VA
Dates: March 2012 – Present; Minor Rotation
Supervisors: Betty Gillespie, Ph.D.
<u>Duties</u>: Conduct dignity therapy and life review for veterans at the end-of-life. Provide support for family members of patients as well as staff members. Co-facilitate a grief support group for family members of veterans who were Palliative Care patients.

Position: Predoctoral Intern
Site: Long-Term Therapy; Salem VA Medical Center, Salem, VA
Dates: August 2011 – July 2012
Supervisors: MK Burton, Ph.D.
<u>Duties</u>: Conducted individual therapy, using empirically supported treatments (e.g., DBT, CBT for panic disorder), for two veterans that require long-term therapy. One veteran met criteria for post-traumatic stress disorder, panic disorder, and borderline personality disorder. I was his individual therapist for the DBT treatment protocol as borderline personality disorder was his presenting problem. The other veteran had difficulty reading so I adapted a CBT manualized treatment to treat his panic disorder

with agoraphobia. I also conducted prolonged exposure therapy with one veteran as part of my long-term therapy experience.
Tests given: BDI-II, BAI, and PCL-C.

Position: Predoctoral Intern
Site: Behavioral Medicine; Salem VA Medical Center, Salem, VA
Dates: August 2011 – February 2012; Major Rotation
Supervisors: Hani Shabana, Ph.D., Derek Bacchus, Ph.D., Sarah Hartley,
                   Ph.D.
Duties: Conducted intake evaluations and individual therapy, using empirically supported treatments, for veterans with a variety of medical problems (e.g., chronic pain, insomnia, diabetes management) in primary care, behavioral medicine, and a variety of specialty clinic settings (e.g., oncology, infectious disease).   In addition, gained approximately 30 hours of supervised experience in home-based primary care. Conducted suicide risk assessments for veterans.   Wrote intake evaluation reports, progress notes, and treatment plans. Lead or co-facilitated a chronic pain group, a weight management group, a shared medical appointment for female veterans with high cholesterol, and a shared medical appointment for veterans being evaluated for obstructive sleep apnea. Conducted and wrote integrated reports for bariatric surgery evaluations and organ transplant evaluations.
Tests given: Montreal Cognitive Assessment (MOCA), Millon Behavioral Medicine Diagnostic (MBMD), Patient Health Questionnaire-9 (PHQ-9), Geriatric Depression Scale, Diabetes Distress Scale, Alcohol Use Disorder Test-Core (AUDIT-C), Drug Abuse Screening Test (DAST), Multidimensional Health Locus of Control, Questionnaire on Weight and Eating Patterns (QWEP).

Position: Predoctoral Intern
Site: Substance Abuse Liaison Team; Salem VA Medical Center,
                   Salem, VA
Dates: October 2011 – Present; Minor Rotation
Supervisors: Josephine DeMarce, Ph.D.
Duties: Conducted motivational enhancement therapy utilizing personalized feedback reports for veterans with substance use concerns. The motivational enhancement therapy sessions were recorded and coded by Dr. DeMarce for adherence to motivational interviewing in order to assist in the development of motivational interviewing skills. Provided individual therapy for early intervention for one veteran with substance use concerns. Conducted personality inventories and clinical interviews and wrote integrated reports to aide in diagnostic clarification and differential diagnosis for veterans in the outpatient substance abuse program. Lead or co-facilitated a mental illness and recovery group and a Seeking Safety group for veterans with co-morbid psychiatric and substance abuse diagnoses.
Tests given: MMPI-2, MCMI-III, BDI-II, and PCL-C.

Position: Predoctoral Intern
Site: Psychiatric Emergency Room; Salem VA Medical Center,
                     Salem, VA
Dates: August 2011 – October 2011; Minor Rotation
Supervisors: Arven Bhandari, M.D.
Duties: Conducted brief intake evaluations for veterans presenting with psychiatric emergencies in order to provide crisis intervention. Wrote intake reports for veterans and collaborated with other mental health professionals involved in their care.

Position: Practicum Student
Site: Richard L. Roudebush VA Medical Center, Red Team
                     Primary Care Clinic, Indianapolis, IN
Dates: May 2010 – November 2010
Supervisor: Jennifer Lydon-Lam, Ph.D.
Duties: Conducted intake interviews for veterans referred by their primary care physician. Conducted brief neuropsychological and cognitive testing as well as personality and symptom inventories with veterans. Wrote nine integrated reports of psychological, neurological, and cognitive findings for referring physicians. Provided individual therapy sessions using empirically supported treatments to veterans on issues of chronic pain, alcohol and drug dependence, and mood/anxiety disorders. Provided couples therapy sessions to a veteran and spouse regarding marital difficulties. Led or co-facilitated weight loss management and chronic pain management groups with veterans. Developed and presented a case conceptualization to the Primary Care Psychologists. Provided assertive communication training to scheduling staff in order to enhance effective communication with veterans.
Tests given: Shipley Institute of Living Scale, Neurobehavioral Cognitive Status Examination (Cognistat/NCSE), Mini-Mental Status Evaluation, Beck Anxiety Inventory (BAI), MMPI-2, MOCA, BDI-II, PHQ-9, and PTSD Checklist-Military Version (PCL-M).

Position: Practicum Student
Site: Clarian Adult Endocrinology Clinic, Indianapolis, IN
Dates: March 2010 – September 2010
Supervisor: Mary de Groot, Ph.D.
Duties: Conducted intake interviews and individual therapy sessions using empirically supported treatments with individuals with diabetes in a multidisciplinary clinic which included a clinical psychologist, endocrinologists, endocrinology nurses, and diabetes educators. Focus of treatment included difficulties with self-care behaviors, medication adherence, depression, and/or anxiety. Completed didactic training in type 1 and type 2 diabetes, treatment of type 1 and type 2 diabetes, psychological issues associated with type 1 and type 2 diabetes, co-morbid conditions, and other general health psychology topics and interventions.

Position: Practicum Student
Site: Larue D. Carter Memorial Hospital, Borderline Personality Disorder Unit, Indianapolis, IN
Dates: August 2009 – February 2010
Supervisor: Joan Farrell, Ph.D.

Duties: Conducted individual therapy focused on distress management and coping using schema therapy. Led or co-facilitated Schema Therapy, Process, Interpersonal Skills, and Distress Management groups. Received training in effectively managing and defusing difficult patients and group situations. Observed and provided feedback to medical residents implementing schema therapy in session with individual clients. This placement was on a specialized inpatient unit for individuals with Borderline Personality Disorder. Worked within a multidisciplinary treatment approach which included a clinical psychologist, psychiatrist, a pastoral psychotherapist, and licensed social worker.

Position: Psychometrist
Site: Dr. Wayne Samuelson, Private Practice, Fargo, ND
Dates: November 2007 – July 2008
Supervisor: Wayne Samuelson Ph.D.

Duties: Administered intelligence tests, achievement tests, and assessments for attention problems to children and adolescents.

Tests given: Woodcock Johnson III NU Tests of Cognitive Abilities and Tests of Achievement, Wechsler Abbreviated Scale of Intelligence (WASI), IVA Continuous Performance Test

Position: Psychometrist
Site: Knowlton, O'Neill, and Associates, Fargo, ND
Dates: August 2007 – November 2007
Supervisor: Glenn Knowlton, Ph.D.

Duties: Administered intelligence tests, achievement tests, and assessments for attention problems to children, adolescents, and adults.

Tests given: Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV), Wechsler Adult Intelligence Scale-Third Edition (WAIS-III), Woodcock Johnson Test of Achievement-Revised, Wide Range Achievement Test-Fourth Edition (WRAT-4), IVA Continuous Performance Test

Position: Practicum Student
Site: Prairie St. John's Clinic, Fargo, ND
Dates: August 2006 – May 2007
Supervisor: Kevin Schumacher, Ph.D.

Duties: Administered, scored, and interpreted intelligence tests, achievement tests, and personality tests for children, adolescents, and adults. Conducted assessments for attention problems in children and adolescents. I have written 23 integrated reports of psychological, neurological, and cognitive findings for this placement. Conducted intake interviews with children and adolescents experiencing depression, anxiety, and behavioral difficulties. Conducted individual therapy sessions using empirically

supported treatments with children experiencing depression, anxiety, and oppositional defiant disorder.

Tests given: WISC-IV, WPPSI-III, KBIT-2, WAIS-III, WMS-III, WIAT-II, Woodcock Johnson Test of Achievement-Revised, WRAT-4, MMPI-II, MMPI-A, MCMI-III, MACI-II, Draw-A-Person, Conners-CPT, BASC, Conners' Parent Rating Scale, Conners' Teacher Rating Scale

Position: Practicum Student
Site: Fraser, Ltd., Fargo, ND
Dates: January 2006 – May 2006
Supervisor: Raymond Miltenberger, Ph.D.

Duties: Conducted behavioral observations and wrote behavior intervention plans for developmentally disabled adults living in a group home setting. Assisted group home staff in implementing behavior intervention plans.

Position: Psychiatric Technician
Site: Prairie St. John's Hospital, Fargo, ND
Dates: July 2004 – May 2006
Supervisor: Michelle Petrik, M.A.

Duties: Assisted in observations, planning, interventions, and evaluation of care to adult inpatient psychiatric clients with mental illness or chemical dependency concerns. Primary responsibilities were to interact with clients, lead groups focused on establishing goals for the day and whether those goals were achieved, ensure vital signs were within normal range, assist in crisis interventions, and help promote healthy distress management and coping behaviors.

Research Experience

May 2009 – May 2012   Dissertation Research
                       Department of Psychology, IUPUI
                       "Ecological Momentary Assessment versus Traditional
                       Retrospective Self-Reports as Predictors of Health-
                       Relevant Outcomes"
                       Chair: Jesse C. Stewart, Ph.D.

The purpose of this study was to examine which type of assessment method, ecological momentary assessment (EMA) or retrospective self-report, has greater predictive utility for health-relevant outcomes. More specifically, I examined whether EMA or retrospective measures of depressive symptoms are better predictors of objective sleep quality and intention to donate blood. Based on a qualitative literature review I conducted for my preliminary examination, I believed that EMA measures would have superior predictive utility for objective sleep quality whereas retrospective measures would have superior predictive utility for intention to donate blood. I defended my dissertation in May 2012.

August 2008 – May 2012   Research Assistant
                           Department of Psychology, IUPUI
                           Cardiovascular and Behavioral Medicine Laboratory
                           Supervisor/Principal Investigator: Jesse C. Stewart, Ph.D.
I prepared a comprehensive literature review for publication as well as helped prepare and write NIH and other grant applications. I was responsible for obtaining limited-use datasets which I have conducted analyses on for publication. I collected data using psychophysiological recording of cardiovascular and autonomic nervous system responses to stress, including EKG and impedance cardiography. I scored psychophysiological data using Mindware Technologies software. I was primarily responsible for learning to use momentary assessment and actigraphy technology for my dissertation as well as other research projects. I am able to use these technologies to collect and score data using Palm® Centro smartphone technology as well as ActiLife® and Satellite Forms® software. I am currently responsible for supervising a two bachelor's level research assistants; one student is conducting her honor's thesis research within my dissertation research.

August 2007-August 2008   Research Associate
                           Department of Psychology, North Dakota State University
                           Health Psychology Laboratory
                           Supervisor/Principal Investigator: Kevin D. McCaul, Ph.D.
I conducted telephone interviews with community members in order to determine cancer risk perceptions and health behaviors. Additional responsibilities included writing manuscripts, coding qualitative data to determine inter-rater reliability, and analyzing and presenting information regarding ACT scores and math placement exams for incoming freshmen at NDSU.

August 2006-August 2008   Master's Thesis Research
                           Department of Psychology, North Dakota State University
                           "Providing Post-task Social Support after an Acute Stress Situation to Enhance Cardiovascular Recovery"
                           Chair: Clayton J. Hilmert, Ph.D.
The purpose of this study was to investigate the effects of the timing of social support on cardiovascular recovery from and rumination about a stressful speech task. The results of this study provide preliminary evidence that social support facilitates cardiovascular recovery, particularly if support occurs after the stressor, and that the content of thoughts, especially positive thoughts, maybe more important in hastening cardiovascular recovery than the amount of rumination. Two manuscripts from this thesis have been published and one manuscript is in progress.

August 2006-August 2008   Research Assistant, Laboratory Coordinator
                           Department of Psychology, North Dakota State University
                           Social Psychophysiology and Health Laboratory
                           Supervisor/Principal Investigator: Clayton J. Hilmert, Ph.D.

I designed and conducted empirical stress and health studies. I collected data using psychophysiological recording of cardiovascular and autonomic responses to stress, using BIOPAC Systems, Inc. devices, including blood pressure, EKG, and impedance cardiography. I scored psychophysiological data using Mindware Technologies software. I supervised and trained two doctoral-level, three master's-level, and 13 undergraduate research assistants in the experimental protocol and in the use of psychophysiological recording devices. I was primarily responsible for data management and storage.

January 2004-August 2007   Research Assistant
                             Department of Psychology, North Dakota State University
                             Health Psychology Laboratory
                             Supervisor/Principal Investigator: Kevin D. McCaul, Ph.D.
I helped prepare an NIH grant application for research to examine implicit processes in smoking behavior. I assisted in participant recruitment and data collection for an ecological momentary assessment palm-pilot study designed to study smokers' worry about smoking consequences. I conducted telephone interviews with spouses of individuals with cancer in order to determine risk perceptions and health behaviors among the spouses. Additional responsibilities included creating questionnaires in MediaLab, conducting data analyses, and conducting follow-up phone calls for research purposes.

<u>Publications</u>

Stewart, J. C., Zielke, D. J., Hawkins, M. A., Williams, D. R., Carnethon, M. R., Knox,
    S. S., & Matthews, K. A., (in press). Depressive symptom clusters as predictors
    of 5-year incidence of coronary calcification: The CARDIA Study. *Circulation.*

Hilmert, C. J., Ode, S., Zielke, D. J., & Robinson, M. D. (2010). Blood pressure
    reactivity predicts somatic reactivity to stress in daily life. *Journal of Behavioral
    Medicine, 33,* 282-292.

Ode, S., Hilmert, C. J., Zielke, D. J., & Robinson, M. D. (2010). Neuroticism's
    importance in understanding the daily life correlates of heart rate variability.
    Emotion, 10, 536-543.

Magnan, R. E., Köblitz, A. R., Zielke, D. J., & McCaul, K. D. (2009). The effects of
    warning smokers on risk, worry, and motivation to quit. *Annals of Behavioral
    Medicine, 37,* 46-57.

<u>Manuscripts in Preparation</u>

Zielke, D. J., Lee, M. E., Ratcliffe, T., & Stewart, J. C. (in preparation). Associations
    between negative emotional factors and the discrepancy between subjective and
    objective reports of sleep quality.

Zielke, D. J. & Hilmert, C. J. (in preparation). Associations between cortisol reactivity and rumination in a social stress task.

Conference Presentations

Lee, M. E., Zielke, D. J., & Stewart, J. C. (2011, April). *Shared and unique features of emotional factors and their relationships to objective sleep quantity and quality.* Poster presented at the IUPUI Psychology Capstone Poster Session, Indianapolis, IN.

Zielke, D. J. & Stewart, J. C. (2011, April). *The predictive utility of ecological momentary assessment measures versus retrospective self-report measures: A qualitative literature review.* Poster presented at IUPUI Research Day, Indianapolis, IN.

Stewart, J. C., Zielke, D. J., Hawkins, M. A. W., Williams, D. R., Carnethon, M. R., Knox, S. S., & Matthews, K. A. (2011, March). *Depressive symptom clusters as predictors of 5-year incidence of coronary calcification: The CARDIA Study.* Paper presented at the American Psychosomatic Society's 69th Annual Scientific Meeting, San Antonio, TX.

Teoh, A. N., Hilmert, C. J., & Zielke, D. J. (2011, March). *Primary and secondary emotional effects of social support.* Paper presented at the American Psychosomatic Society's 69th Annual Scientific Meeting, San Antonio, TX.

Zielke, D. J. & Stewart, J. C. (2010, March). *The predictive utility of ecological momentary assessment measures versus retrospective self-report measures: A qualitative literature review.* Poster presented at the American Psychosomatic Society's 68th Annual Scientific Meeting, Portland, OR.

Stewart, J. C., Hawkins, M. A., & Zielke, D. J. (2009, March). *Associations of positive and negative psychological factors with indices of cardiac autonomic balance and regulatory capacity.* Poster presented at the American Psychosomatic Society's 67th Annual Scientific Meeting, Chicago, IL.

Zielke, D. J. & Hilmert, C. J. (2009, March). *Cortisol reactivity to stress predicts tendencies to ruminate.* Poster presented at the American Psychosomatic Society's 67th Annual Scientific Meeting, Chicago, IL.

Kvasnicka, L. R., Zielke, D. J., Hilmert, C. J. (2009, February). *Gender of social support provider and receiver during stress.* Poster presented at the Society for Personality and Social Psychology's 10th Annual Conference, Tampa Bay, FL.

Ode, S, Zielke, D. J., Robinson, M. D., & Hilmert, C. J. (2009, February). *Heart rate variability as a mediator of the neuroticism-negative affect relationship.* Poster presented at the Society for Personality and Social Psychology's 10th Annual Conference, Tampa Bay, FL.

Zielke, D. J. & Hilmert, C. J. (2008, February). *Timing of social support affects rumination and cardiovascular recovery.* Poster presented at the Society for Personality and Social Psychology's 9th Annual Conference, Albuquerque, NM.

Seamands, E. D., Zielke, D. J., & Hilmert, C. J. (2008, March). *Affect and its effect on post-task rumination.* Poster presented at the 23rd Annual Red River Psychology Conference, Moorhead, MN.

Zielke, D. J., Magnan, R. E., Koblitz, A. R., & McCaul, K. D. (2007, March). *"On-line" thoughts about smoking and motivation to quit.* Poster presented at Society of Behavioral Medicine's 28th Annual Meeting & Scientific Sessions, Washington, D. C.

Zielke, D., Koblitz, A. R., Dillard, A. J., & McCaul, K. D. (2005, March). *Self-monitoring does not increase reactivity.* Poster presented at 22nd Annual Red River Psychology Conference, Fargo, ND.

<u>Invited Talks</u>

Zielke, D. J. (2012, April). *Ecological momentary assessment versus traditional retrospective self-reports as predictors of health-relevant outcomes.* Salem Virginia Veterans Affairs Medical Center, Mental Health Service Line Staff Meeting.

Zielke, D. J. (2011, October). *Case presentation on the management of diabetes with co-morbid depression and PTSD.* Salem Virginia Veterans Affairs Medical Center, Mental Health Service Line Staff Meeting.

Zielke, D. J. (2010, December). *Acceptance and Commitment Therapy for a male with chronic pain.* Richard L. Roudebush Veterans Affairs Medical Center, Primary Care Psychology Staff Meeting.

Zielke, D. J. (2010, October). *Case presentation on somatoform disorder assessment.* Indiana University-Purdue University, Department of Psychology Pro-seminar Series.

Zielke, D. J. (2008, May). *Heart Rate Variability, Blood Pressure, and Rumination.* North Dakota State University, Department of Psychology Health/Social Brown-Bag Series.

Zielke, D. J. (2007, May). *Timing of social support and cardiovascular recovery.* North Dakota State University, Department of Psychology Colloquium Series.

Teaching Experience

June 2011-May 2012   Instructor
August 2010-December 2010   B104: Psychology as a Social Science; on
                               campus and online
                               Department of Psychology, IUPUI
                               Supervisor: Lisa Contino, Ph.D.
I developed and delivered lectures and administered class activities to an undergraduate class of approximately 50 students during the regular academic year of 2010. Managed an online course and conducted problem solving sessions with students doing poorly in the course for the 2011 summer semester and the 2011-2012 academic year.

May 2011-June 2011   Instructor
                               B305: Statistics
                               Department of Psychology, IUPUI
                               Supervisor: Jane Williams, Ph.D.
I developed and delivered lectures, conducted SPSS training sessions, and administered class activities to an undergraduate class of approximately 50 students.

June 2010-May 2011   Instructor
                               B380: Abnormal Psychology
                               Department of Psychology, IUPUI
                               Supervisor: John Guare, Ph.D.
I developed and delivered lectures and administered class activities to an undergraduate class of approximately 70 students during the regular academic year and approximately 20 students during the summer semester.

January 2009-August 2009   Teaching Assistant
                               B360: Child and Adolescent Psychology
                               Department of Psychology, IUPUI
                               Supervisor: Terri Tarr, Ph.D.
I was responsible for grading course assignments for a class of approximately 50 students.

January 2009-May 2009   Teaching Assistant
                               B310: Lifespan Development
                               Department of Psychology, IUPUI
                               Supervisor: Alex Khislavsky, M.S.
I was responsible for grading course assignments and compiling quizzes for a class of approximately 50 students.

August 2008-December 2008   Teaching Assistant
                               B481: Capstone Laboratory in Clinical
                               Rehabilitation Psychology
                               Department of Psychology, IUPUI

Supervisor: Jesse C. Stewart, Ph.D.

I was responsible for delivering lectures regarding the use of SPSS as well as grading course assignments for a class of approximately 30 students.

August 2005-May 2006   Teaching Assistant
                              111: Introduction to Psychology
                              Department of Psychology, North Dakota State
                              University
                              Supervisor: Cathy Waters, M.S.

I was responsible for administering and grading make-up essay exams for a class of approximately 200 students.

Honors and Awards

2010   Educational Enhancement Grant, Research Funding Award,
                   Graduate Student
              Organization, Indiana University-Purdue University Indianapolis
              Educational Enhancement Grant, Travel Award, Graduate Student
                   Organization,
              Indiana University-Purdue University Indianapolis

2009   Educational Enhancement Grant, Travel Award, Graduate Student
                   Organization
              Indiana University-Purdue University Indianapolis

2005-2007   Psi Chi Honor Society—National Honor Society
              Phi Kappa Phi Honor Society—National Honor Society

Professional Memberships

2010-present   American Psychological Association Division 38

2008-2009   Society for Personality and Social Psychology

2007-2009   American Psychological Association Division 12

2006-2007   Society of Behavioral Medicine