HOW PHYSICAL ACTIVITY AND TIME OF DAY INFLUENCE JUDGMENTS OF LEARNING AND FREE RECALL

by

Joseph J. Bassili, B.A.S.

A thesis submitted to the Graduate Council of
Texas State University in partial fulfillment
of the requirements for the degree of
Master of Arts
with a Major in Psychological Research
August 2015

Committee Members:

William Kelemen, Chair

Reiko Graham

Yueqin Hu

ACKNOWLEDGEMENTS

I would first like to recognize the contributions of my thesis committee. Thank you Dr. Hu for your instruction and many conversations over statistical concepts. Your influence extends the bounds of this work, as you have helped spark my passion for statistical analysis. Thank you Dr. Graham for your expectations and words of encouragement. Your high expectations pushed me, but I always left your office with the confidence that I could meet them. Thank you Dr. Kelemen for your mentorship, guidance, and the latitude to grow independently. Your open door and open ear are truly appreciated.

I would not have been able to accomplish this work without the love and support of my wife. Words cannot express the gratitude that I have for your faith in me. However you did more than provide support. You were and always will be an inspiration. Thank you, Kristy.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table **Page**

LIST OF ABBREVIATIONS

Abbreviation	Description
ADACL	Activation-Deactivation Adjective Checklist
ANOVA	Analysis of Variance
CDC	Centers for Disease Control
JOL	Judgment of Learning
MEQ	Morningness-Eveningness Questionnaire
MET	Metabolic Equivalent
TOD	Time of Day

ABSTRACT

Previous research suggests the existence of a time of day (TOD) effect in cognitive functioning such that people have an optimal time each day during which their cognitive functions peak: typically, older adults peak in the morning and younger adults peak in the evening (May, Hasher, & Stolzfus, 1993). Explicit memory is one cognitive function that has shown a TOD effect (e.g., Schmidt, Collette, Cajochen, & Peigneux, 2007). In addition, exercise has been shown to improve explicit memory (e.g., Labban & Etnier, 2011). Thus it may be possible to offset the decline in explicit memory that has been observed during non-peak times of the day. The main purposes of this study were to determine if the TOD influences metacognitive judgments, and to test if an acute bout of exercise would offset the TOD effect in a young adult, evening-type population. The study utilized a 2 (type of activity: exercise vs. sedentary) X 2 (TOD: morning vs. afternoon) repeated-measures design. A TOD effect (i.e., better performance in the afternoon than in the morning) emerged for memory and judgment of learning (JOL) magnitude. However, exercise did not offset the morning decrement in performance. These results suggest that young, evening-type adults have better recall in the afternoon and that they are aware of this TOD effect, as evidenced by the magnitude of their JOLs. Additional research is needed to determine if exercise or other interventions can offset this TOD effect.

CHAPTER I

Introduction

Much research over the past three decades has focused on factors that influence metacognition, which refers to the cognitive processes pertaining to one's own cognition (Dunlosky & Metcalfe, 2009; Larson, 2009; Metcalfe & Shimamura, 1994, Schwarz, 2015). Flavell (1979) introduced the terms "metacognitive knowledge" and "metacognitive experiences." He described metacognitive knowledge as that which people already have about their own cognitive processes and factors that can influence them. These factors may include the characteristics of the people themselves, the cognitive task, and strategies people use to complete the task. Flavell also defined metacognitive experiences as mental processes that occur along with and about "intellectual enterprise" (p. 906). For example, listening to a speaker at a conference would be an intellectual enterprise and the sudden realization that you are not paying attention to the speaker would be a metacognitive experience. In this case, the monitoring of attention is one aspect of metacognition.

Nelson and Narens subsequently (1990) described three key principles of metacognition, which have sparked much research in cognitive psychology. The first principle is that cognitive processes can be divided into at least two levels: the meta-level and the object-level. The object-level consists of cognitive processes (e.g., perception, attention, and memory) and the meta-level interacts with them. The second principle is that the meta-level contains a model of the object-level. This model, which may or may not be accurate at any given time, contains the metacognitive knowledge of the cognitive task being performed (Dunlosky & Metcalfe, 2009). The third principle is that metacognition involves two key processes known as monitoring and control. Monitoring refers to the assessment of object-level processes, such that information flows from the cognitive process "up" to the metacognitive process. The information from the object-level is used to modify the metacognitive model contained in the meta-level. In other words, monitoring refers to a form of feedback from the object-level to the meta-level. At the meta-level, the information collected via monitoring is assessed, and new information may be sent to object-level to alter the cognitive process itself. This informational flow from the meta-level back "down" to the object level is referred to as control. This model of metacognition has guided much research over the past 30 years (for reviews, see Larson, 2009; Metcalfe & Shimamura, 1994, Schwarz, 2015).

To illustrate these principles, consider the example of listening to the speaker at a conference. In this case, one cognitive process would be attention: at the object level, you are paying attention to the speaker. At the meta-level, you are monitoring the process of attention to maintain focus. At the heart of the meta-level is a model of attention. Within this model is the metacognitive knowledge of what you know about your own attention. One aspect of this knowledge may be that you know your attention is poor in the morning and that caffeine can improve it. When you find that your mind is beginning to wander (monitoring), you might decide to drink a cup of coffee (control). Thus, information flows to the meta-level from the object-level and vice-versa for optimal cognitive processing.

Memory Monitoring

Explicit memory is an important cognitive process that can be monitored and controlled by the meta-level. This process is referred to as "metamemory" and represents a large portion of metacognitive research (Dunlosky & Metcalfe, 2009). In fact, metacognition and metamemory are now largely used synonymously. When examining metacognition, most researchers have focused on the monitoring process only. Researchers typically ask participants to make a judgment of learning (JOL), during or soon after study; which represents the participant's assessment of how well he or she will remember the studied information. Generally, participants are usually presented with multiple stimuli and asked to make a JOL on each one.

The rating is often made in terms of a percentage from 0 (definitely will not recall) to 100 (definitely will recall). Participants are sometimes also asked to provide aggregate JOLs, in which they rate how many total words they will remember from the list. Aggregate JOLs can be made prior to encoding, after study, or both. The values of these ratings, either item-by-item or aggregate, are referred to as JOL magnitude.

Metacognitive accuracy is the relationship between JOL magnitude and subsequent recall. Researchers often assess two types of metacognitive accuracy: calibration and resolution. Calibration is an absolute measure of participants' ability to predict their overall level of performance and can be assessed using bias scores. To calculate bias scores, researchers compare the average JOL magnitude (item-by-item JOLS, pre-study aggregate JOLs, or post-study JOLs) to the percentage of words correctly recalled. A subject with perfect calibration would have a bias score of zero. A bias value greater than zero would reflect overconfidence and a bias value less than zero would reflect underconfidence.

In addition to calibration, researchers can also measure the relative accuracy of JOLs (i.e, resolution). Resolution is the accuracy of a participant's JOL for each stimulus relative to the JOL accuracy for all other stimuli. It can be assessed using a correlation that is computed between a participant's JOL magnitude and recall for each item in a list. In other words, resolution assesses whether or not stimuli with higher JOL magnitudes are in fact more likely to be recalled than stimuli with lower magnitudes. When relative accuracy is reported as a correlation coefficient, values of zero indicate a complete lack of relationship, values of -1.0 indicate a perfect inverse relation between JOLs and recall, and values of +1.0 indicate perfect metacognitive accuracy. The Goodman-Kruskal correlation is the most widely used measure of resolution, though other effective measures do exist (Nelson, 1984; Benjamin & Diaz, 2008).

Time of Day Effect in Cognition

The effects of time of day (TOD) on cognitive processes have been the subject of research for over a century. In 1885, Ebbinghaus wrote that not only does memory differ between people, but "… also when different phases of the existence of the same individual are compared: morning and evening, youth and old age, find him different in this respect" (1903, p. 3). O'Shea (1901) noted that in self-reports, a majority of students thought their minds worked best in the morning. Later, Kleitman (1939) identified two chronotypes, "morning type" and "evening type," according to a person's sleep patterns: morning types go to bed early and awaken early, whereas evening types go to bed late and awaken late. To formalize these ideas, Horne and Ostberg (1976) developed the Morningness-Eveningness Questionnaire (MEQ), which is a self-report measure that identified three chronotypes: morning, evening, and intermediate types. They found that these three chronotypes correspond to circadian patterns in body temperature, such that these patterns are phase-shifted by chronotype. Although the three chronotypes follow the same curve, they peak at different times. Morning types start the day with higher body temperatures that stay higher throughout the daytime and peak in the evening. Body temperatures of evening types continue to rise after morning types peak and remain higher during the nighttime; intermediate types have body temperature curves that generally fall between morning and evening types (Horne & Ostberg, 1976).

Additional research has identified individual differences in chronotypes (Schmidt, Collette, Cajochen, & Peigneux, 2007). Typically, there is a trend toward young adults being classified as evening types and older adults being morning types. However the majority of young adults tested often turn out to be classified as intermediate types (May et al., 1993). For example, May et al. administered the MEQ to 210 college students and 91 older adults to screen participants for a subsequent study. Older adults were classified as morning type (74 percent) or intermediate type (26 percent). The young adults tended to be intermediate (50 percent) and evening type (44 percent), with only 6 percent identified as morning type. More recently, Bassili and Kelemen (2014) found a similar pattern in a sample of 69 college students, with 58 percent classified as intermediate type, 40 percent as evening type, and only 2 percent as morning type.

Decades of research have shown that a person's chronotype can play a major role in determining that person's optimal period of cognitive function and that a person’s chronotype tends to correspond with his or her optimal TOD, such that morning types perform better in the morning and evening type perform better in the afternoon (Kraemer et al., 2000; May, Hasher, & Foong, 2005; May, Hasher, & Stolzfus, 1993; Schmidt, Collette, Cajochen, & Peigneux, 2007; West, Murphy, Armillo, Craik, & Stuss, 2002; Yoon, May, & Hasher, 2000). For example, Anderson and others (1991) tested semantic memory in a sample of 54 evening-type and 45 morning-type college students as identified by the MEQ. Participants were shown word pairs and had to answer whether the words had the same name or were from the same semantic category. Participants were tested at either 9 a.m., 2 p.m., or 8 p.m. and the dependent variable was reaction time. Reaction times increased across the TOD for the morning-type participants and decreased for evening-types as the day progressed. Thus, the optimal TOD in terms of semantic memory access corresponded to the participants' chronotype.

May and colleagues (1993) tested recognition in evening-type young adults and morning-type older adults. Half of each group was tested in the morning (8 a.m. or 9 a.m.) and the other half was tested in the afternoon (4 p.m. or 5 p.m.). Participants read 10 short stories and then received an immediate test for sentence recognition. Test stimuli consisted of 30 sentences (three from each story) and 30 novel sentences. The sentences were presented one at a time and participants responded whether or not they appeared in one of the stories. If participants correctly recognized a sentence, it was scored as a hit and if they incorrectly recognized a novel sentence, it was scored a false alarm. An accuracy score was computed by subtracting the number of false alarms from the number of hits. Accuracy was higher for the older adults tested in the morning compared to those tested in the afternoon. Inversely, accuracy was higher for the young adults tested in the afternoon compared to those tested in the morning.

In a later study, May et al. (2005) used another between-subjects design to test recall in 36 college students and 48 older adults. Participants were once again selected based on their responses on the MEQ, such that the college students were all classified as evening types and all the older adults were classified as morning types. Half of the participants were tested in either the morning or afternoon, with half of each age group at each testing time. Stimuli consisted of 96 words that consisted of five letters, with stems (the first three letters) that could be used to form at least four words. Half of the words were used in the study phase, and the other half were used as controls during the testing portion. In the study phase, participants were presented with 48 word pairs, one at a time. One of the words of each pair was the target and marked with an asterisk. Participants were asked to focus on the target and provide a pleasantness rating for that word. Participants were then tested for implicit memory, followed by explicit recall. In the explicit testing phase, participants were presented with 48 stems; 24 from target words in the study phase and 24 control words not used in the study phase. Participants were instructed to complete the stems using the earlier present words, and that not all stems could be completed using these words. A recall score was calculated by subtracting the number of control words correct by the number of target words correct. Like recognition in the previous study, recall was superior in the morning for the older adults and superior in the afternoon for the college students.

To date, there has been only one published study to examine metacognition in the context of TOD effects. Hourihan and Benjamin (2014) tested 58 college students in the morning (8 a.m. or 9 a.m.) and in the afternoon (3 p.m. or 4 p.m.) using a repeated measures design. They had participants study a list of 20 words and provide a JOL for each word after it was presented. Prior to and after each study session, participants also provided an aggregate JOL of how many words out of 20 they thought they would remember. Participants then performed a 30-second distractor task, followed by a recall test. After another distractor task, participants completed the procedures again with a different list of 20 words. Data for both lists were averaged to produce one outcome measure for each session (i.e., morning and afternoon).

Results reflected the TOD effect for recall that is typically seen in college students: participants recalled significantly more words in the afternoon. Importantly, the post-study aggregate JOLs were significantly higher in the afternoon than in the morning, but there was no TOD effect for the pre-study aggregate JOLs or item-by-item JOLs. Furthermore, there was not a significant TOD effect for resolution (i.e., relative accuracy), although there was a trend, nearing significance, of higher JOL resolution in the afternoon. Still, the study provides some evidence for a TOD effect in JOLs. An important limitation of this study was that 53 percent of participants were classified as intermediate type on the MEQ. Perhaps stronger TOD effects would have been obtained by testing only morning types or evening types.

Exercise and Cognition

TOD is not the only factor that can affect cognitive functions. Exercise can improve mood as well as increase the efficacy of memory processes (Brisswalter, Durand, Delignieres, & Legros, 1995; Colcombe & Kramer, 2003; Labban & Etnier, 2011; Lambourne & Tomporowski, 2010; Roig , Nordbrant, Geertsen, & Nielsen, 2013; Stroth, Hille, Spitzer, & Reinhardt, 2009). Roig et al. (2013) conducted a meta-analysis of 20 studies and reported that acute bouts of exercise produced a moderate-to-large effect on long-term memory (both episodic and semantic). For instance, Coles and Tomporowski (2008) found that exercise improved free recall in a sample of 18 young adults. Using a repeated measure design, participants either pedaled on a stationary bike for 40 minutes, sat on a stationary bike without pedaling, or sat in another room and watched a 40-minute documentary. Participants were tested before and after each condition, and they studied a list of 40 words for each test. Participants were prompted to recall as many words as possible immediately after study and again 12 minutes later, following a distractor task. Recall performance was computed as the difference in amount of words remembered between the immediate and delayed tests. Performance declined for the two non-exercise conditions, but did not decline for the exercise condition. These results suggest that exercise prior to study helps reduce forgetting over time.

To assess whether this effect was due to changes in encoding or retrieval, Labban and Etnier (2011) investigated whether cycling before encoding would improve recall of verbal prose compared to cycling after encoding. They assigned participants to 1 of 3 groups: exercise before study, exercise after study, or a control group (no exercise). In the first group, participants cycled for 30 minutes prior to stimulus presentation, performed a distractor task, and then rested 30 minutes prior to testing. In the second group, participants rested prior to stimulus presentation and then cycled for 30 minutes prior to testing. The control group rested quietly prior to and after stimulus presentation, followed by testing. The stimuli consisted of two short paragraphs. Recall was superior for the group who cycled prior to studying, compared to the exercise after study and control groups, suggesting that the benefit of exercise was derived from pre-study activity.

In addition to cycling, walking prior to study also has been reported to influence memory performance and metacognitive accuracy. Salas, Minakata, and Kelemen (2011) had participants either take a brisk 10-minute walk or sit and watch a 10-minute slideshow prior to studying a list of 30 nouns and providing JOLs. To test for state-dependent effects of memory (i.e., recall better when in the same state during retrieval as during encoding), participants then walked briskly or watched a slideshow for 10 minutes prior to a recall test. Participants who walked prior to study recalled approximately 25 percent more words and had more accurate metacognitive calibration (i.e., bias scores closer to zero) than participants who sat prior to study. However, walking immediately prior to testing did not have a significant effect, nor did a significant state-dependent effect emerge. Finally, neither walking prior to study nor walking prior to testing influenced mean JOL magnitude or resolution compared with sitting. Because JOLs did not differ between conditions, the improved calibration in the walking condition was attributed to the improved recall performance. This finding suggests that participants were not aware of the benefit of walking. In a more controlled laboratory setting, however, Santana and Kelemen (2013) found that acute exercise on a treadmill, at moderate and intense levels, increased both recall and mean item-by-item JOLs compared with a sedentary condition. However, exercise again did not have an impact on JOL resolution. The higher JOLs in both conditions of the Santana and Kelemen study suggest participants were aware of the benefit of exercise. To extend this research on exercise and metacognition, the present study examined how exercise and TOD might interact to influence recall and JOLs.

Exercise and the Time of Day Effect

Two studies have examined the relationship between exercise and the TOD effect on cognition. Bugg, DeLosh, and Clegg (2006) hypothesized that exercise might also influence the TOD effect on working memory in elderly adults. A total of 35 adults aged 61 to 88 were tested both in the morning and afternoon. Participants were assigned to groups based on a median split of scores on a self-report physical activity questionnaire, with 18 participants classified as active and 17 as sedentary. As a measure of working memory, participants were required to keep a running count of the number of small and large boxes presented on a computer screen, one at a time, over 16 trials. During each trial, a total of 10 to 14 boxes were presented. Working memory performance was scored as the proportion of correct counts. Participants with higher levels of daily physical activity had similar working memory performance in the morning and the evening tests, whereas the participants who were considered non-active had a marked decline in their performance in the evening tests compared to the morning tests. Although the researchers did not assess chronotype, the results suggest that the participants were morning-types, consistent with literature. The results also suggest that self-reported physical activity levels did have an influence on the TOD effect in working memory, such that older adults with higher activity did not exhibit the TOD effect. However, an important limitation was the use of self-reported physical activity instead of an exercise manipulation. The use of self-reported activity does not allow for causal inference.

Potter and Keeling (2005) did utilize an exercise manipulation to investigate the effect of moderate activity on memory throughout the day in middle-aged shift workers. Over two consecutive days, they tested word recall in 31 participants, twice at four different times of the day: 9:30 a.m., 12:30 p.m., 3:30 p.m., and 6:30 p.m. In the experimental condition, a fast-paced, 10-minute walk followed by a 15- to 30-minute recovery period preceded half of the sessions on the first day and the other half of the sessions on the second day. Thus, each participant was tested once without exercise and once with exercise at each time-slot. A different word list was used for each session, and each session consisted of five trials in which participants were presented with the same list and tested for recall. There was a main effect of exercise, such that participants recalled more words during the sessions in which they exercised prior to testing. Planned comparisons indicated a TOD effect between the 9:30 a.m. and 3:30 p.m. sessions. Within each session, a benefit of repeated study and test was observed such that participants progressively recalled more words on each successive presentation of the list. However, participants recalled increasingly more words per trial during exercise sessions than during no-exercise sessions. It was worth noting, however that Potter and Keeling did not assess individual differences in chronotype. The average age of the participants was 32 years old, so it is likely there was a mixture of chronotypes included. It is not clear whether the results would change if chronotype were controlled.

In sum, research suggests that (a) people may have different levels of cognitive performance (e.g., memory) throughout the day (May et al., 2005; May et al., 1993; Yoon et al, 2000; West et al. 2002), (b) moderate exercise can improve performance on memory tasks (Coles & Tomporowski, 2008; Labban & Etnier, 2011, Salas et al., 2011, Santana & Kelemen, 2013), and (c) exercise may offset the TOD effect on memory in middle-age adults (Potter & Keeling, 2005) and older adults (Bugg et al., 2006). In addition, there is no extent literature on the combined effects of exercise and TOD on metacognition. The current study extended previous research in two major ways: it controlled for chronotype in examining the relationship between exercise and TOD (and extended past research to young adults), and it examined the possible interaction of exercise and TOD in metacognition.

CHAPTER II

Purpose of Study, Research Questions, Hypothesis and Method

The purpose of this study was to explore the effect of TOD and exercise on JOLs and recall. Although there is emerging research on the effects of exercise on JOLs (e.g., Salas, Minakata, & Kelemen, 2011; Santana & Kelemen, 2013), only one study has explored the effect of TOD on JOLs (Hourihan & Benjamin, 2014) andthere is no published research that has examined the combined effect of TOD and exercise on JOLs ;;. Additionally, although there is extensive research on the separate effects of exercise and TOD on recall, very little research has examined the effects of them together. Furthermore, the one study that has examined the effects of TOD and exercise on recall did not control for individual chronotype (Potter & Keeling, 2005).

Two research questions were posited to address these gaps in the literature. First, what effects do TOD and exercise have on JOLs? Second, would exercise moderate the TOD effect in recall for evening-type young adults? Based on the work of Hourihan and Benjamin (2014) and Santana and Kelemen (2013), the first general hypothesis was that both TOD and activity would influence all three types of JOLs (i.e., pre-study aggregate JOLs, item-by-item JOLs, and post-study aggregate JOLs). Although Hourihan and Benjamin observed TOD effects only on post-study aggregate JOLs, the present study included a more homogenous sample of evening chronotypes and so the TOD manipulation was expected to produce more robust effects across all types of JOLs. Thus, Hypothesis 1 was that main effects would emerge for both TOD and activity, such that JOLs would be higher in the afternoon sessions than in the morning sessions (cf.

Hourihan & Benjamin, 2014), and JOLs would be higher in the exercise conditions than in the sedentary conditions (cf. Santana & Kelemen, 2013).

The second major question in this research was whether or not moderate exercise would offset the TOD effect in recall for evening-type young adults. Based on the findings of Potter and Keeling (2005), Hypothesis 2 was that there would be an interaction between TOD and activity on recall scores. Specifically, participants should have better recall in the afternoon than in the morning during the sedentary conditions, (a standard TOD effect), whereas there should be no TOD effect in the exercise conditions. Furthermore, it was expected that there would be a main effect of activity, such that participants should recall more words during the exercise conditions than they would during the sedentary conditions (cf. Santana & Kelemen, 2013).

Method

Participants

The recruitment of participants and all other procedures of the study were approved by the Texas State University Institutional Review Board. Participants were recruited through the psychology department human subjects pool at Texas State University, via email and through advertisements on electric signboards. Students who signed up through the human subjects pool received course credit for the session and monetary compensation for completing the subsequent test sessions, whereas the remaining students were paid for all session. Students were asked not to participate if they had a medical condition that made them unable to exercise at a moderate intensity level. A total of 129 volunteers were screened for participation in the study: only students identified as evening-type by the MEQ were invited to continue to the four experimental sessions of the study. In all, 46 students were identified as evening-types on the MEQ and 35 volunteered for the experimental sessions. Three individuals failed to complete all sessions, so 32 participants completed the study. Descriptive statistics for gender and age are shown in Table 1. The dependent measures were examined for gender differences, and the effects were consistent by gender. Therefore all variables were collapsed across gender in the reported analyses.

Design

The study consisted of 5 sessions on separate days. The first session was a screening and practice session, and the final four sessions were testing sessions. The four testing sessions were administered using a 2 X 2 repeated-measures design. The two factors were testing TOD (morning versus afternoon) and type of activity (exercise versus sedentary). All sessions lasted approximately 45 minutes. Morning sessions began at either 8 a.m. or 9 a.m. and afternoon sessions began at either 3 p.m. or 4 p.m. The exercise conditions included 10 minutes of moderate exercise prior to testing, whereas the sedentary conditions included 10 minutes of a sedentary activity prior to testing.

The order of the testing sessions (morning-sedentary, morning-exercise, afternoon-sedentary, and afternoon-exercise), was meant to be counterbalanced across participants such that each condition would be included first, second, third, and fourth an equal number of times, with the order of study lists constant to ensure each list was used an equal number of times for each condition. However scheduling errors resulted in an unequal number of conditions across days. Separate chi-square goodness-of-tests were performed to determine whether the occurrence of each condition differed significantly for each day of testing. Chi-square values ranged from 0.25 to 3.25, and none of the tests were significant at a .05 alpha level. Thus, conditions did not vary significantly across days.

Apparatus

The exercise manipulation used a Lifespan TR4000i treadmill and a Polar FT60 heart-rate monitor. The treadmill was used to induce a moderate level of exercise. For this study, a moderate level of exercise was defined as 55 percent of the participant's predicted maximum heart rate. Maximum heart rate for each participant was calculated as 208 minus the participant's age multiplied by 0.7 (Tanaka, Monahan, & Seals, 2001). In addition, the treadmill recorded heart rate, speed, total number of steps, distance walked, and number of calories burned. The heart-rate monitor was worn on a chest strap, which provided feedback to control the treadmill speed in order to reach and maintain a moderate level of exercise for 10 minutes. A BodyMedia SenseWear armband was used to measure participant's physical activity for 2 hours prior to and during each session. These data were not used to address the two hypotheses and so they are not discussed further.

Instruments

The Morningness-Eveningness Questionnaire (MEQ; Horne & Ostberg, 1976) is a 19-item survey. Responses were summed to indicate the respondent's chronotype. Possible scores range from 16-86: scores of 41 or less indicated an evening type person, which was the inclusion criterion.

The Activation-Deactivation Adjective Checklist (ADACL) is a 20-item questionnaire designed to measure two dimensions of arousal: energy and tension (Thayer, 1978). Each item consists of one adjective and asks the participants to identify the extent to which that particular adjective represents their current mood. The ADACL can be partitioned into 4 subscales, with 5 adjectives assigned to each subscale. The subscales are Energy, Tiredness, Tension, and Calmness. These subscales have shown to be sensitive to exercise, such that exercise produces increases in Energy and decreases in Tiredness and Calmness (Ekkekakis, Hall, & Petruzzello, 2005). Therefore the ADACL was included as an exercise manipulation check.

A brief questionnaire was developed that asked participants to report the length and quality of the their previous night's sleep, any exercise they performed in the last 12 hours, as well whether they ingested any nicotine any caffeine that day (see Appendix A).

Stimuli used for the JOL and recall tasks were five lists of 30 words (see Appendix B). Each list consisted of 30 concrete English nouns. The nouns for each list were selected from the MRC Psycholinguistic database (Coltheart, 1981) and chosen based on number of letters (5-7), concreteness (500-700) and imageability (500-700). E-Prime 2.0 software was used to present stimuli and record JOL responses (Psychology Software Tools, 2012). List 1 was used during practice trials and Lists 2 thru 5 were used for the experimental trials. Each list was only used on the corresponding day of testing (e.g., List 3 on Day 3), so that participants were presented the lists in the same order.

Procedures

Testing during Session 1 (i.e., the screening session) occurred between 11 a.m. and 1 p.m.. During this session, the researcher obtained informed consent, administered the MEQ, and asked participants to complete a memory and JOL test using List 1. This testing was done to familiarize the participants with the procedures of the experiment. While the potential participants were being tested, the researcher scored the MEQ and identified those who scored less than 42 (i.e., evening-types). Those who did not meet criteria or did not want to further participate were dismissed and the researcher then scheduled the remaining participants for the four testing sessions.

Testing for Sessions 2-5 was conducted on separate days, with at least 24 hours between each session. Upon arrival each day, participants completed the prior-activity questionnaire, followed by the ADACL. Testing procedures were identical for all sessions. However, the pre-testing activity was manipulated across sessions. During the sedentary sessions, participants completed a pre-testing activity that consisted of watching a slideshow of various landscapes for 13 minutes. During the exercise sessions, pre-testing activity consisted of an exercise manipulation in which participants walked on a treadmill for 10 minutes at a moderate intensity after a 3-minute warm up period. This activity manipulation was the same manipulation as used by Santana and Kelemen (2013).

Prior to the participant getting on the treadmill, the experimenter explained how to put on the heart rate monitor and left the room while the participant put on the strap. During this time, the experimenter calculated the age-predicted maximum heart rate for the participant using the formula from Tanaka et al. (2001). The treadmill used feedback from the heart rate monitor to adjust its speed to achieve and maintain the participant’s heart rate at 55% of the calculated maximum. Once participants were finished on the treadmill, the experimenter left the room to allow the participant to remove the chest strap. After completing the slideshow or treadmill activity, participants again completed the ADACL and then began the testing portion of the session.

Testing consisted of a study and JOL phase, followed by the recall phase. The study and JOL phase used E-Prime to present the stimuli and record participant JOL responses. First, participants were informed that they would study 30 words and asked how many words, from 0 to 30, they thought they would remember about 10 minutes later. The participants' response to this prompt was used as the pre-study aggregate JOL.

Next, each of the 30 words was presented; one at a time, for 6 seconds each. After each word was presented, the participants were asked to rate how likely they would be to recall the word approximately 10 minutes later on a scale from 0 (definitely will not recall) to 100 (definitely will recall). The participants' responses to these prompts were used as the mean item-by-item JOL.

After all 30 words were presented and item-by-item JOLs recorded, participants were asked, how many words, from 0 to 30, they thought they would remember 5 minutes later. The participants' response to this prompt was used as the post-study aggregate JOLs.

After completing the study and JOL phase, participants completed a distractor task for 5 minutes. For the distractor task, participants were be asked to write a name on a piece of paper and then write a different name that starts with the last letter of the previous name. They then repeated the steps with the new name. They were asked to continue to this process and not repeat any names. The participants were instructed to draw a line and start over if they felt they were unable to come up with a name based on the previous name. After the completion of the 5-minute distractor task, participants were given a blank piece of paper and were asked to write as many of the words they could recall from that session's study and JOL phase for 3 minutes. The number of words correctly recalled was used as the measure of recall.

Participants were then paid $10, reminded of their next session, and dismissed. On the last day of testing, participants were informed of the purpose of the study and the expected results, and paid an additional $10 bonus if they completed all sessions as scheduled.

Statistical Analyses

For all analyses, alpha was set at a .05 significance level. Effect sizes are reported as partial Eta squared. To test the hypotheses, separate 2 X 2 repeated-measures Analyses of Variance (ANOVAs) were computed on each dependent measure. Both within-subject factors (TOD: morning and afternoon; type of activity: exercise or sedentary) were the same for each analysis. The dependent variables were pre-study aggregate, post-study aggregate, and mean item-by-item JOLs for Hypotheses 1. It was predicted that there would be a main effect of TOD and activity for all three JOLs, with higher JOLs in the afternoon and after exercise. Recall was the dependent variable for Hypothesis 2. It was predicted that there would an interaction between TOD and activity, such that exercise would offset the TOD effect in recall.

To test for a TOD effect in isolation from the effect of exercise, a planned comparison paired samples *t*-test was conducted on recall scores between the morning sedentary condition and afternoon sedentary condition. It was expected the recall would be significantly higher in the afternoon sedentary session. As a manipulation check of exercise, a planned comparison paired samples *t*-test was conducted on morning recall scores between the sedentary and exercise conditions. Recall was expected to be significantly higher in the exercise conditions than in the sedentary conditions.

Two additional 2 X 2 repeated-measures ANOVAs were computed to explore the combined effects of TOD and activity on metacognitive accuracy. Again, both within-subject factors (TOD: morning and afternoon; type of activity: exercise or sedentary) were the same for each analysis. The dependent variables were calibration and resolution, respectively. Calibration was assessed using bias scores for item-by-item JOLs. To calculate bias scores, the participants' recall scores were converted to percentages and then subtracted from the mean item-by-item JOLs. The resulting bias scores were the measure of calibration. In order to calculate resolution, a dichotomous recall variable was created (coded as 0 = not recalled and 1 = recalled) for each of the 30 words. A Goodman-Kruskal Gamma correlation was then calculated between each item-by-item JOL and this new recall variable. The resulting Gamma coefficient was used as a measure of metacognitive resolution.

There were two exercise manipulation checks conducted. First, treadmill data (hear rate, speed, steps, distance, and calories burned) were analyzed with paired-sample *t*-tests to ensure there were no differences between the morning and afternoon exercise sessions. The second manipulation check analyzed changes in ADACL arousal scores. Difference scores for all four subscales of the ADACL (i.e., Energy, Tension, Tiredness, and Calmness) were calculated to assess changes in arousal as a function of activity. To create the difference scores, pre-manipulation ADACL scores were subtracted from post-manipulation ADACL scores. Thus, a difference score of zero indicates no change in arousal, while positive and negative scores indicate an increase and decrease, respectively. The difference scores for each subscale were compared to zero using one-samples *t*-tests. It was expected that the difference score would we significantly different than zero.

CHAPTER III

Results

<u>Metacognition</u>

Pre-study Aggregate Judgments of Learning

Pre-study aggregate JOLs were only analyzed for 28 participants because two participants did not provide a response during at least one session and responses for another two participants were not recorded due to incomplete E-Prime files. Consistent with Hypothesis 1, pre-study aggregate JOLs tended to be higher in the afternoon (M = 15.88, SD = 5.51) compared with the morning (M = 14.03, SD = 5.67) (see Table 2). A two-way repeated-measures ANOVA revealed that TOD had a significant effect on pre-study aggregate JOLs, $F(1, 27) = 5.47$, $p = .027$, partial $\eta^2 = .168$. Prior to studying, participants predicted they would remember more words in the afternoon than they did in the morning, suggesting they were expecting TOD differences in their performance. Neither activity nor the time x activity interaction had a significant effect on pre-study aggregate JOLs, $F(1, 27) = 2.89$, $p = .10$, partial $\eta^2 = .097$ and $F(1, 27) = 1.07$, $p > .05$, partial $\eta^2 = .038$, respectively.

Item-by-item Judgments of Learning

Item-by-item JOLs also tended to be higher in the afternoon (M = 55.20, SD = 18.16) compared with the morning (M = 52.50, SD = 20.46) (see Table 2). A two-way repeated-measures ANOVA confirmed that TOD again had a statistically significant effect, $F(1, 31) = 4.40$, $p = .044$, partial $\eta^2 = .124$. Thus, item-by-item JOL magnitude was significantly higher in the afternoon than in the morning. Neither activity nor the time x activity interaction had a significant effect on mean item-by-item JOLs, $F(1, 31)$ =

3.42, $p > .074$, partial $\eta^2 = .099$ and $F(1, 31) = 1.97$, $p = .17$, partial $\eta^2 = .06$, respectively. As before, these results supported Hypothesis 1 for TOD, but not for activity.

Post-study Aggregate Judgments of Learning

Post-study aggregate JOLs tended to be higher in both the afternoon and after exercise (see Table 2). As before, a two-way repeated-measures ANOVA was conducted on post-study aggregate JOLs. In this case, both TOD and activity produced significant main effects on post-study aggregate JOLs, $F(1, 28) = 12.99$, $p = .001$, partial $\eta^2 = .317$ and $F(1, 28) = 14.71$, $p = .001$, partial $\eta^2 = .344$, respectively. After studying, participants predicted they would remember more words in the afternoon ($M = 13.69$, $SD = 4.21$) than they did in the morning ($M = 12.17$, $SD = 3.59$), and they predicted better memory after walking on the treadmill ($M = 13.61$, $SD = 4.55$) than after being sedentary ($M = 12.25$, $SD = 3.19$). In other words, they expected that both TOD and activity would influence their recall. The interaction between TOD and exercise did not have a significant effect on post-study aggregate JOLs, $F(1, 28) = 1.07$, $p = .13$, partial $\eta^2 = .080$.

Metacognitive Accuracy

Calibration (i.e., bias scores) did not differ across all sessions (see Table 2). Mean bias scores were analyzed using a two-way repeated-measures ANOVA. Neither TOD, activity, nor the TOD x activity interaction had a significant effect on bias scores, $F(1, 31) = 2.50$, $p < .124$, partial $\eta^2 = .075$, $F(1, 31) = .00$, $p > .05$, partial $\eta^2 = .00$, and $F(1, 31) = 2.80$, $p = .105$, partial $\eta^2 = .083$, respectively. In addition, bias scores for each condition were compared to zero using one-sample t-tests to check for over and under confidence. Bias scores did not differ significantly from zero, all t-values < 1.67, and all *p*-values > .05, suggesting that participants' item-by-item JOLs were well calibrated: overall, participants were not significantly over or under confident.

Resolution (i.e., Gamma coefficients) also did not differ across all sessions (see Table 2). Gamma coefficients were analyzed using a two-way repeated-measures ANOVA. Like calibration, neither TOD, activity, nor the time x activity interaction had a significant effect on resolution, $F(1, 30) = 1.61$, $p > .05$, partial $\eta^2 = .051$, $F(1, 30) = .95$, $p > .05$, partial $\eta^2 = .031$, and $F(1, 30) = .08$, $p > .05$, partial $\eta^2 = .003$, respectively.

<u>Recall</u>

Table 2 shows that recall scores displayed a similar pattern as pre-study aggregate JOLs, with more words recalled in the afternoon ($M = 52.92$, $SD = 18.38$) compared with the morning ($M = 46.77$, $SD = 18.88$). To test Hypothesis 2, a two-way repeated-measures ANOVA was conducted on recall. TOD had a significant effect on recall, $F(1, 31) = 9.90$, $p = .004$, partial $\eta^2 = .242$. Recall scores were significantly higher in the afternoon than in the morning, thus replicating the TOD effect for evening types. However, neither activity nor the time x activity interaction had a significant effect on recall, $F(1, 31) = 1.65$, $p > .05$, partial $\eta^2 = .05$ and $F(1, 31) = .97$, $p > .05$, partial $\eta^2 = .03$, respectively.

Planned paired-sample *t*-tests were also conducted. There was no significant difference in recall scores between the morning sedentary condition ($M = 46.77$, $SD = 17.53$) and morning exercise conditions ($M = 46.77$, SD 20.43), $t\ (31) = 0.00$, $p > .05$, suggesting that the exercise manipulation did not improve recall performance at participants' nonpeak times. To check the effect of exercise at peak times, an exploratory paired-sample t-test was conduct between the two afternoon conditions. Afternoon recall after exercise (M = 55.31, $SD = 19.82$) was higher than afternoon recall after sedentary

(M = 50.53, SD = 16.79), and approached significance, $t(31) = -1.80$, $p = .082$. In the other planned comparison, the morning sedentary condition ($M = 46.77$, $SD = 17.53$) and afternoon sedentary condition ($M = 50.52$, $SD = 16.79$) were not significantly different, $t(31) = -1.19$, $p > .05$, failing to confirm a TOD effect in the sedentary condition, although the means were in the predicted direction.

Exercise Manipulation Check

Treadmill

The participants' exercise data by session are shown in Table 3. The data recorded by the treadmill included heart rate (beats per minute; $M = 109.88$, $SD = 3.40$), speed (miles per hours; $M = 2.33$, $SD = 1.09$), total number of steps ($M = 941.32$, $SD = 289.66$), distance walked (miles; $M = 0.43$, $SD = .18$), and number of calories burned ($M = 32.46$, $SD = 16.47$). To ensure that exercise intensity was consistent across TOD, paired-sample *t*-tests were conducted between the morning and afternoon sessions. There were no significant differences by TOD (all *t*s < 1.0 and *p*s > .05), confirming that exercise intensity did not vary significantly between TOD.

Self-Reported Arousal

Descriptive statistics for all administrations of the ADACL are shown in Table 4. Participants reported higher levels of energy in the afternoon ($M = 11.03$, $SD = 2.56$) compared to the morning ($M = 8.52$, $SD = 2.76$) and higher levels of tiredness in the morning ($M = 13.52$, $SD = 3.77$) compared to the afternoon ($M = 10.73$, $SD = 2.68$). As previously mentioned, difference scores for all four subscales of the ADACL (i.e., Energy, Tension, Tiredness, and Calmness) were calculated to assess changes in arousal as a function of activity. As expected, energy scores tended to increase, while tiredness and calmness scores tended to decrease in the exercise conditions (see Table 5). An opposite pattern occurred in the sedentary condition, such that energy scores tended to decrease, while tiredness and calmness scores tended to increase. Tension scores were stable across conditions and are not discussed further.

As a manipulation check on the activity effects, the difference scores in Table 5 were compared to zero using one-sample t-tests. In the morning sedentary conditions, energy significantly decreased, whereas tiredness and calmness significantly increased, $t(31) = -4.00$, $p < .001$, $t(31) = 5.62$, $p < .001$, and $t(31) = 3.00$, $p < .001$, respectively. The same pattern emerged in the afternoon sedentary conditions: energy significantly decreased, whereas tiredness and calmness significantly increased, $t(31) = -4.69$, $p < .001$, $t(31) = 3.39$, $p < .001$, and $t(31) = 2.56$, $p < .01$. In the morning exercise conditions, energy significantly increased, whereas tiredness and calmness significantly decreased, and tension did not significantly change $t(31) = 3.14$, $p < .01$, $t(31) = -2.77$, and $p < .01$, $t(31) = -3.12$, $p < .01$, respectively. Finally, in the afternoon exercise conditions, calmness significantly decreased, while energy and tiredness did not significantly change, $t(31) = -2.83$ $p < .01$, $t(31) = 2.01$, $p = .053$, and $t(31) = -0.13$, $p > .05$, respectively. Overall, these results suggest that the exercise manipulation did produce nearly all of the expected changes in self-reported arousal.

CHAPTER IV

Discussion

This study investigated the combined influence of TOD and exercise on metacognition and recall and posed two research questions: What would be the effects of TOD and activity on metacognition, and would exercise offset the TOD effect on recall? This section discusses the answers to these two questions, the limitations of the study, and future directions.

Metacognition

This study investigated the influence of TOD and exercise on metacognition. Based on prior research, a hypothesis was formed for metacognitive monitoring and addressed three aspects: pre-study aggregate JOLs, item-by-item JOL and post-study aggregate JOLs. Two measures of metacognitive accuracy (calibration and resolution) also were assessed. Based on previous research, it was hypothesized that there would be a main effect of TOD and activity; such that JOLs would be higher in the afternoon and also after exercise. This hypothesis was fully supported in post-study aggregate JOLs and partially supported for TOD in pre-study aggregate and mean item-by-item JOLs.

Metacognitive Monitoring

A significant TOD effect emerged for all three types of JOLs in the present study. In terms of pre-study aggregate JOLs, Hourihan and Benjamin (2014) did not observe a TOD effect, so the current positive findings contrast their null results. This inconsistency may be due to differences in samples between the two studies. Hourihan and Benjamin did not screen participants based on chronotype, whereas only evening-type participants were included in the present study. Thus, it appears that even pre-study aggregate JOLs can be sensitive to TOD when individuals are tested during their non-preferred vs. preferred times (e.g., evening-type individuals tested in the morning vs. afternoon). In contrast to the TOD findings, the activity manipulation produced no significant differences in pre-study aggregate JOLs.

Item-by-item JOLs showed the same pattern of results. A significant effect of TOD emerged, whereas Hourihan and Benjamin (2014) failed to observe such an effect. This difference again may be due to the sole inclusion of evening-type young adults in the present study. Using the Nelson and Narens (1990) model, this finding can be interpreted that at the meta-level, participants' metacognitive model for memory contains the idea that their memory is better in the afternoon, thus they predict higher recall in the afternoon. In contrast to Santana and Kelemen's (2013) findings, activity failed to produce a significant main effect. Although the interaction between TOD and activity was also non-significant, exploratory analyses suggested that exercise can influence the TOD effect in at least one case: paired sample *t*-tests indicated a TOD effect in the sedentary conditions, which were significantly higher in the afternoon compared to the morning $t(31) = -2.24$, $p < .05$, but no difference emerged between morning vs. evening sessions during the exercise conditions. Although there was no main effect for activity, or a significant interaction with TOD, this suggests that exercise did offset the TOD effect in item-by-item JOLs.

The study also examined the effect of TOD and activity on post-study aggregate JOLs. Hypothesis 1 predicted both a TOD and exercise effect in post-study aggregate JOLs. In this case, the results fully supported Hypothesis 1 and were consistent with the findings of both Hourihan and Benjamin (2014) and Santana and Kelemen (2013). The

TOD effect sizes in this study are much larger than those reported by Hourihan and Benjamin, whose effects may have been attenuated by non-evening type participants.

Overall, TOD had a significant effect on all types of JOLs, such that evening-type, young adults predicted higher recall in the afternoon than the morning. Why did they predict higher recall in the afternoon? One possibility is that participants actually do recall more words in the afternoon (e.g., Hourihan& Benjamin, 2014). Alternatively, it could be because they felt more aroused in the afternoon and they inferred that they would perform better because they felt better. In the current study, both recall and ADACL Energy scores were higher in the afternoon. However, ADACL Energy scores were also higher in the exercise conditions than the sedentary conditions, but only post-study JOLs showed ADACL changes after exercise. If increased arousal were to increase JOL magnitude, then there should have been an increase in all JOLs after exercise (to correspond with the observed increases in arousal). Therefore, it is more likely that the participants were aware that they would perform better in the afternoon rather than simply misattributing increased arousal for better memory.

Metacognitive Accuracy

Both TOD and activity had no effect on calibration and resolution. The one past study in this area reported a TOD effect in resolution that approached significance (p = .063) and null results for calibration (Hourihan & Benjamin, 2014). Thus, in the present study, null results were obtained again despite the possibly enhanced TOD effect in metacognitive monitoring due to the exclusively evening-type sample. In this study, resolution was numerically higher in the afternoon, matching Hourihan and Benjamin's findings, but the findings did not approach significance. Overall, there is little evidence in

Hourihan and Benjamin or the current study that TOD can significantly influence metacognitive accuracy.

Recall

Hypothesis 2 predicted that there would be a main effect of exercise on recall, and a significant interaction between TOD and exercise. The interaction was expected such that a TOD effect in recall would emerge when the participants sat sedentary prior to testing, but that the morning decrement would not occur when they exercised prior to testing. Hypothesis 2 was not supported. Overall, there was a significant TOD effect on recall scores, such that participants recalled a higher proportion of words in the afternoon compared to the morning, which supports the findings of Hourihan and Benjamin (2014). However, recall was not significantly different between morning testing and afternoon testing when participants did not exercise, although recall scores were higher in the afternoon. One possible explanation could be related to insufficient counterbalancing: more participants were tested in a sedentary condition in the afternoon on days 2 and 3 than on days 4 and 5, and the opposite pattern occurred for the sedentary condition in the morning. If there was a practice effect due to insufficient counterbalancing, this could have led to the afternoon/sedentary recall scores being lower than expected and the morning/sedentary scores being higher than expected. This would explain the lack of the expected effect for activity. Furthermore, when participants exercised prior to testing in the morning, their performance was just as low as when they were sedentary prior to testing. An alternative explanation may be formed by combining the lack of TOD in sedentary conditions and lack of exercise effect in the morning: The observed TOD effect simply may be attributed to participants performing better after they exercised in the afternoon.

Overall, it was a surprise that exercise did not have a main effect on participants' recall: there was no significant difference between participants' recall when they exercised prior to study compared to when they were sedentary prior to study. This null finding is inconsistent with the published findings of Salas et al. (2011) and the unpublished findings of Santana and Kelemen (2013), despite using the same equipment and manipulation as in the latter study. In fact, the treadmill data and changes in ADACL scores in the current study are quite similar to those reported by Santana and Kelemen, yet the manipulation did not produce a significant effect in this case. However, Santana and Kelemen tested very few participants early in the morning. In the present study, mean recall scores during the afternoon were numerically higher ($p = .08$) after exercising compared to being sedentary, whereas in the morning, average recall scores did not differ. This lack of exercise effect in the morning is most likely why an overall exercise effect did not emerge in the present study and may account for the difference between studies.

Another possibility is that the intensity of the exercise may not have reached moderate levels. The Centers for Disease Control (CDC) defines a moderate intensity as equaling 3 to 5.9 metabolic equivalents (METs), and recommends walking at a "brisk" pace of 3 to 4.5 miles per hour to attain a moderate intensity level (Buchner, 2010). Participants in this study walked at a pace slower than that recommended by the CDC. On the other hand, walking at 91 to 111 steps per minutes has been equated to 3 METs (Marshall et al., 2009). On average, the current sample fell within this range. In addition, there was an increase in subjective arousal, as evidenced by the ADACL scores, which suggests that the exercise manipulation was not a complete failure.

Limitations

Two major limitations of this study are the incomplete counterbalancing of condition order and the small sample size. The study was designed to achieve partial counterbalancing of condition by day, such that each condition would have been run on each day the same number of times. However, as previously mentioned, scheduling errors resulted in an unequal presentation of conditions across days. Although the frequency of conditions across days did not very significantly, it is possible that the unequal counterbalancing may impact the results. Because the list of words was held constant for each day of testing, practice effects may have occurred and list difficulty may have varied across days. Collecting additional data to balance the ordering of conditions would be helpful in addressing these concerns and to increase the sample size, but unfortunately the additional data required (approximately 25 percent more participants) is beyond the scope of the current project.

The two activity conditions used in this study also may have limited this study. Firstly, the exercise manipulation may not have been adequate to achieve a moderate intensity. The study would have benefitted by also including a more vigorous exercise condition, which would have prevented speculation that the current manipulation was insufficient to produce an effect. Secondly, the sedentary condition may not have been the best control for activity, because arousal decreased significantly during the sedentary conditions rather than remaining constant. Because the sedentary condition resulted in a decrease in arousal, any differences in recall could have been attributed to a performance decrease in the sedentary condition, rather than a performance increase in the exercise condition. The study could have attempted to utilize a different control condition, such as standing on the treadmill. However, Coles and Tomporowski (2008) have shown that passively sitting on an exercise machine had the same effect on performance as sitting at a computer. Coles and Tomporowski also showed that recall differences between exercise and non-exercise conditions can be attributed to reductions in the non-exercise conditions, rather than increases from exercise. Baseline measures of recall at the morning and afternoon times would have been appropriate comparisons and should be considered in any future design.

Finally, some variations in testing environment were unavoidable. In order to collect the data in a single semester, multiple participants were tested at the same time in up to three separate rooms of different sizes, and only one room had a window. The blinds in the windowed room were kept closed in an attempt to minimize lighting differences between rooms. Also, because only one of the rooms contained a treadmill, participants were not always tested in the same room across sessions. Additionally, four researchers were used to conduct the testing and different researchers tested the participants across sessions. However, researchers followed a script to standardize, as much as possible, the interactions with participants.

Conclusion and Future Directions

The results of this study indicate that TOD has a significant effect on JOLs and recall, such that evening-type, young adults predicted better recall and in fact recalled more words in the afternoon compared to in the morning. This has implications for academic performance young evening-type college students. For example, it may be beneficial for evening-type students to avoid enrolling in early morning classes. However, this is not always practical, nor altogether avoidable. The present study suggests that these students would have to possibly invest more time and effort in a morning class to achieve the same result compared to an afternoon class. Fortunately, students in this study appeared to be aware of this TOD effect because it was reflected in their JOLs (i.e., monitoring); future research may wish to examine whether it can be reflected in their study activity (i.e., control). For example, students may adjust their study habits according to TOD, perhaps by studying words longer during a morning session compared with an afternoon session.

To further extend these findings on the TOD effect in metacognition, future research may also expand to broader samples in regard to age and chronotype. Specifically, it is recommended to investigate whether the current findings generalize to morning-type, older adults. Older adults' cognitive functioning tends to decline throughout the day (Yoon et al., 2000). A question that has not been asked is whether they are aware of this decline at the meta-level. Would older adults predict higher recall in the morning compared to the afternoon and would more beneficial effects of exercise occur?

Although the activity manipulation did not have an effect on recall, important insights emerged in regard to arousal on two fronts. First, the present study provides evidence that sitting passively at a computer, for as little as 13 minutes, can reduce energy and increase tiredness. Not only can this impact classroom performance, these findings extend to any environment in which people are subjected to lengthy visual presentations. Instructors should be aware of the impact that the presentation method has on students' arousal and try to avoid prolonged periods of audience passivity. This effect warrants more research. Future studies can investigate just how long it takes for arousal to decrease in this sedentary condition. More importantly, research should center on how to avoid this arousal decrement. Research could control the nature of the stimuli to ascertain whether or not the drop in arousal is a function of the slideshow itself, or a result of sitting passively. Would the decline occur if participants were prompted for input periodically? Second, the study provides evidence that a brief 10-minute walk has a positive effect on arousal. Regardless of the TOD, participants had more energy and felt less tired after walking on the treadmill. A brisk walk around the block might be the remedy of choice, as one finds his or her energy waning.

Table 1.

Participant Characteristics by Gender

	Frequency	Age	BMI
Female	23	20.00 (1.92)	22.76 (0.8)
Male	9	19.33 (1.23)	24.53 (1.8)
Total	32	19.80 (1.75)	23.25 (4.32)

Note. Main entries are means and entries in parenthesis are standard deviations.

Table 2

Descriptive Statistics for JOL Magnitude, JOL accuracy, and Recall

	Condition			
	Morning		Afternoon	
Dependent Measure	Exercise	Sedentary	Exercise	Sedentary
JOLs				
Pre-Study Aggregate	13.36 (5.10)	13.11 (4.82)	14.64 (5.09)	13.57 (4.70)
Item-by-Item	54.78 (21.19)	50.22 (19.77)	55.38 (19.68)	55.01 (18.22)
Post-Study Aggregate	12.62 (4.18)	11.72 (3.18)	14.86 (4.85)	12.69 (3.34)
Bias	8.01 (27.19)	3.44 (22.64)	0.07 (23.14)	4.49 (21.86)
Gamma	0.29 (.37)	0.36 (.29)	0.36 (.32)	0.40 (.25)
Recall	46.77 (20.43)	46.77 (17.53)	55.31 (19.82)	50.52 (16.79)

Note. Main entries are means and entries in parenthesis are standard deviations.

Table 3

Exercise Data as a Function of Time of Day

Measure	Overall	Morning	Afternoon
Heart rate (bpm)	109.88 (3.40)	110.07 (3.82)	109.67 (2.96)
Speed (mph)	2.33 (1.09)	2.26 (1.17)	2.39 (1.01)
Total steps	941.32 (289.66)	922.94 (318.43)	960.29 (260.56)
Distance (miles)	0.43 (.18)	0.42 (.20)	0.44 (.17)
Calories burned	32.46 (16.47)	32.44 (17.59)	32.48 (15.52)

Note. Main entries are means and entries in parenthesis are standard deviations

Table 4

Pre- and Post-Manipulation ADACL Subscale Scores by Condition

	Condition			
Arousal	Morning		Afternoon	
Subscale	Sedentary	Exercise	Sedentary	Exercise
Pre-Energy	8.75 (3.27)	8.28 (3.42)	11.00 (3.88)	11.06 (3.62)
Post-Energy	6.44 (2.12	10.16 (3.27)	8.00 (2.29)	12.53 (3.01)
Pre-Tension	6.13 (1.58)	6.53 (2.75)	6.97 (3.12)	7.16 (3.12)
Post-Tension	5.94 (1.44)	7.06 (3.08)	6.38 (2.49)	6.72 (2.50)
Pre-Tiredness	12.90 (4.20)	14.13 (4.63)	11.13 (4.74)	10.34 (3.90)
Post-Tiredness	16.63 (3.18)	12.22 (4.44)	14.13 (3.84)	10.25 (3.88)
Pre-Calmness	13.13 (2.90)	12.63 (3.18)	12.28 (3.71)	12.31 (2.66)
Post-Calmness	14.72 (2.79)	10.78 (2.83)	14.13 (3.00)	10.63 (2.42)

Note. Main entries are means, entries in parenthesis are standard deviations

Table 5

Pre- and Post-Manipulation Difference in ADACL Scores by Condition

Arousal Subscale	Morning: Sedentary	Morning: Exercise	Afternoon: Sedentary	Afternoon: Exercise
Energy	-2.31 (3.27)***	1.88 (3.38)**	-3.00 (3.62)***	1.47 (4.13)
Tension	-0.19 (1.00)	0.53 (1.95)	-0.59 (2.75)	-0.44 (2.55)
Tiredness	3.72 (3.74)***	-1.91 (3.89)**	3.00 (5.00)**	-0.09 (4.11)
Calmness	1.59 (3.00)**	-1.84 (3.34)**	1.84 (4.08)*	-1.69 (3.37)**

Note. Main entries are means, entries in parenthesis are standard deviations

* indicates significance < .05

** indicates significance < .01

*** indicates significance < .001

APPENDIX SECTION

APENDIX A

Sleep Quality, Exercise, and Stimulant Ingestion Questionnaire

1. How many hours of sleep did you get *last night*? Please write down a number:_________

2. Please circle the most appropriate rating for the ***quality*** of sleep you got *last night* from 1 ("not at all restful") to 7 ("very restful") on the scale below:

Not at all restful Very restful
1----------2----------3----------4----------5----------6-----------7
3. Please circle the most appropriate rating for how ***rested*** you ***feel right now*** from 1 ("not at all rested") to 7 ("very rested") on the scale below:

Not at all restful Very restful
1----------2----------3----------4----------5----------6-----------7
4. Please circle any activities below that you have completed in the last 24 hours (you may circle more than one):
a. Running e. Cross-training
b. Aerobics f. Outdoor hiking
c. Weight training g. Power Walking
d. Swimming h. Other Exercise:___________________

5. Have you consumed any food or drinks with caffeine today? ________

If so, please describe:
If so, how long ago in minutes did you last consume caffeine?________
6. Have you smoked any cigarettes today? ________

If so, how many?_______
If so, how long ago in minutes did you smoke your last cigarette?_______

APENDIX B

List 1	**List2**	**List 3**	**List 4**	**List 5**
banker	ankle	basin	animal	avenue
beetle	apple	blossom	balloon	banana
candle	blanket	bucket	bullet	biscuit
cement	bubble	cabin	closet	builder
clothes	button	cigar	corner	capitol
dentist	camera	collar	daisy	choir
dollar	ceiling	dinner	doctor	circle
engine	curler	driver	elbow	cotton
human	dancer	family	flower	drizzle
lemon	dresser	garden	gorilla	father
linen	essay	hockey	helmet	garbage
magnet	forest	hotel	husband	grocer
miner	harbor	island	jacket	jewel
movie	insect	kettle	library	leather
ocean	jersey	letter	mirror	liquid
officer	lawyer	locker	mother	missile
partner	lettuce	metal	needle	morning
piano	maple	monkey	onion	nickel
picture	money	music	parcel	painter
potato	muscle	office	people	person
referee	novel	pickle	powder	pupil
silver	parade	poster	rabbit	spider
speaker	pepper	puppy	rocket	sunburn
teacher	quarter	salad	shovel	table
toaster	record	stable	station	tiger
tongue	ruler	tennis	ticket	tomato
uncle	saucer	toilet	tourist	tulip
walnut	soldier	tunnel	turtle	umpire
whistle	thunder	velvet	village	valley
zipper	wrapper	water	wedding	wallet

REFERENCES

Bassili, J. J. & Kelemen, W. L. (2014, November). *How do Time of Day and Activity Levels Influence Memory and Metacognitive Accuracy?* Poster presented at the Annual Meeting of the Psychonomic Society, Long Beach, CA.

Benjamin, A. S., & Diaz, M. (2008). Measurement of relative metamnemonic accuracy. In J. Dunlosky & R. A. Bjork (Eds.), Handbook of memory and metamemory (pp. 73_94). New York, NY: Psychology.

Brisswalter, J., Durand, M., Delignieres, D., & Legros, P. (1995). Optimal and non-optimal demand in a dual task of pedalling and simple reaction time: effects on energy expenditure and cognitive performance. *Journal of Human Movement Studies*, *29*(1), 15-34.

Bugg, J., DeLosh, E., & Clegg, B. (2006). Physical activity moderates time-of-day differences in older adults' working memory performance. *Experimental Aging Research, 32*(4), 431-446. doi:10.1080/03610730600875833

Buchner, David M., (2010). Health benefits of physical activity. In David R. Brown, Gregory W. Heath, and Sarah Levin Martin, (eds), *Promoting physical activity: a guide for community action (*pp. 3-20). Champaign, IL: Human Kinetics.

Colcombe, S. J., & Kramer, A. F. (2003). Fitness effects on the cognitive function of older adults: A meta-analytic study. *Psychological Science, 14*(2), 125–130.

Coles, K., & Tomporowski, P. D. (2008). Effects of acute exercise on executive processing, short-term, and long-term memory. Journal of Sports Sciences, 26, 333-344.

Coltheart, M (1981) The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology,* 33(4), 497-505.

Dunlosky, J. & Metcalfe, J. (2009). *Metacognition*. Thousand Oaks, CA: Sage.

Ebbinghaus, H. (1913). *Memory: A contribution to experimental psychology* (No. 3). Teachers college, Columbia University. Retrieved from https://play.google.com/store/books/details?id=CIN9AAAAMAAJ

Ekkekakis, P., Hall, E. E., & Petruzzello, S. J. (2005). Evaluation of the circumplex structure of the Activation Deactivation Adjective Check List before and after a short walk. Psychology of Sport and Exercise, 6(1), 83-101.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34, 906-911.

Horne, J. & Ostberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology, 4*, 97-110.

Hourihan, K. L., & Benjamin, A. S. (2014). State-based metacognition: How time of day affects the accuracy of metamemory. *Memory*, *22*(5), 553-558.

Kleitman, N. (1939). *Sleep and wakefulness as alternating phases in the cycle of existence, by Nathaniel Kleitman.* Chicago, Ill., The University of Chicago press.

Kraemer, S., Danker-Hopfe, H., Dorn, H., Schmidt, A., Ehlert, I., & Herrmann, W. M. (2000). Time-of-day variations of indicators of attention: performance, physiologic parameters, and self-assessment of sleepiness. *Biological Psychiatry*, *48*(11), 1069-1080.

Labban, J. D., & Etnier, J. L. (2011). Effects of acute exercise on long-term memory. *Research Quarterly for Exercise and Sport*, *82*(4), 712-721.

Lambourne, K., & Tomporowski, P. (2010). The effect of exercise-induced arousal on cognitive task performance: A meta-regression analysis. *Brain Research*, *1341*12-24. doi:10.1016/j.brainres.2010.03.091

Larson, C. B. (Ed.) (2009). *Metacognition: New research developments.* New York: Nova Science Publishers.

Marshall, S. J., Levy, S. S., Tudor-Locke, C. E., Kolkhorst, F. W., Wooten, K. M., Ji, M., ... & Ainsworth, B. E. (2009). Translating physical activity recommendations into a pedometer-based step goal: 3000 steps in 30 minutes. *American journal of preventive medicine*, *36*(5), 410-415.

May, C. P., Hasher, L., & Foong, N. (2005). Implicit memory, age, and time of day: Paradoxical priming effects. *Psychological Science, 16* (2), 96-100.

May, C. P., Hasher, L., & Stoltzfus, E. R. (1993). Optimal time of day and the magnitude of age differences in memory. *Psychological Science, 4,* 326-330.

Metcalfe, J., & Shimamura, A. P. (1994). *Metacognition: knowing about knowing.* J. Metcalfe and A. P. Shimamura (eds.). Cambridge, MA: MIT Press.

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological bulletin*, *95*(1), 109.

Nelson, T. O., & Narens, L. (1990). Metacognition: A theoretical framework and new findings. *The Psychology of Learning and Motivation*, *26*, 125-141.

O'Shea, M. V. (1901). Aspects of mental economy. *Bulletin of the University of Wisconsin: Science series*, *2*. Retrieved from https://play.google.com/store/books/details?id=GEw9AQAAMAAJ

Potter, D., & Keeling, D. (2005). Effects of moderate exercise and circadian rhythms on human memory. *Journal Of Sport & Exercise Psychology*, *27*, 117-125.

Roig, M., Nordbrandt, S., Geertsen, S. S., & Nielsen, J. B. (2013). The effects of cardiovascular exercise on human memory: a review with meta-analysis. *Neuroscience & Biobehavioral Reviews*, *37*(8), 1645-1666.

Salas, C. R., Minakata, K., & Kelemen, W. L. (2011). Walking before study enhances free recall but not judgment-of-learning magnitude. *Journal of Cognitive Psychology*, *23*, 507-513.

Santana, A. & Kelemen, W. L. (2013). *How varying levels of acute exercise influence cognitive functions*. Unpublished manuscript, Department of Psychology, Texas State University, San Marcos, TX.

Schmidt, C., Collette, F., Cajochen, C., & Peigneux, P. (2007). A time to think: circadian rhythms in human cognition. *Cognitive Neuropsychology*, *24*(7), 755-789.

Schwarz, N. (2015). Metacognition. In M. Mikulincer, P. R. Shaver, E. Borgida, J. A. Bargh, M. Mikulincer, P. R. Shaver, ... J. A. Bargh (Eds.) , *APA handbook of personality and social psychology, Volume 1: Attitudes and social cognition* (pp. 203-229). Washington, DC, US: American Psychological Association. doi:10.1037/14341-006

Stroth, S., Hille, K., Spitzer, M., & Reinhardt, R. (2009). Aerobic endurance exercise benefits memory and affect in young adults. *Neuropsychological Rehabilitation*, *19*(2), 223-243.

Tanaka, H., Monahan, K. D., & Seals, D. R. (2001). Age-predicted maximal heart rate revisited. *Journal of the American College of Cardiology*, *37*(1), 153-156.

Thayer, R. E. (1978). Toward a psychological theory of multidimensional arousal (activation). Motivation and Emotion, 2, 1-34.

West, R., Murphy, K. J., Armilio, M. L., Craik, F. I., & Stuss, D. T. (2002). Effects of TOD on age differences in working memory. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *57*, 3-10.

Yoon, C., May, C. P., & Hasher, L. (2000). Aging, circadian arousal patterns, and cognition. In D. C. Park & N. Schwarz (Eds.), *Cognitive aging: A primer* (pp. 151-171). New York, NY: Psychology Press.