PRIME TIME? A LOOK AT THE EFFECTS OF CIRCADIAN MISMATCH ON STEREOTYPE RELIANCE

A Thesis
by
ROBERT ALEXANDER INDIANA MCCLELLAND

Submitted to the Graduate School
at Appalachian State University
in partial fulfillment of the requirements for the degree of
MASTER OF ARTS

May 2016
Department of Psychology

PRIME TIME? A LOOK AT THE EFFECTS OF CIRCADIAN MISMATCH ON STEREOTYPE RELIANCE

A Thesis
by
ROBERT ALEXANDER INDIANA MCCLELLAND
May 2016

APPROVED BY:

David L. Dickinson
Chairperson, Thesis Committee

Andrew R. Smith
Member, Thesis Committee

Hall P. Beck
Member, Thesis Committee

James C. Denniston
Chairperson, Department of Psychology

Max C. Poole, Ph.D.
Dean, Cratis D. Williams School of Graduate Studies

Abstract

PRIME TIME? A LOOK AT THE EFFECTS OF CIRCADIAN MISMATCH ON STEREOTYPE RELIANCE

Robert Alexander Indiana McClelland
B.S., Appalachian State University

Chairperson: David Dickinson

Stereotype-based decisions are formed as the result of employing various heuristics and biases, and they serve as a way to assess ambiguous situations and compensate for limited information processing. Research has demonstrated that during circadian mismatched (non-optimal) periods of the day cognitive resource availability is diminished. This study examined the influence of circadian arousal levels (particularly in mismatched conditions) on the tendency to use stereotypes in decision-making tasks. It was predicted that mismatch between chronotype (individual circadian preference) and time of day would correlate negatively with cognitive resource availability, thus increasing vulnerability to stereotype reliance. Participants were 59 Appalachian State undergraduates. The participants were administered an online survey consisting of the validated reduced Morningness-Eveningness Questionnaire, the Karolinska Sleepiness Scale, the Epworth Sleep Scale, and a stereotyping task. Each subject participated in sessions at two different times of the day, with the sessions occurring approximately one week apart. Though the stereotype priming manipulation failed, results suggest that participants in adverse sleep or circadian states may have still relied on biases or heuristics when assessing guilt.

Acknowledgements

I would like to provide my sincere appreciation to all of the people who supported me throughout the lengthy, and at times trying, process of completing my thesis. First and foremost, I would like to thank my committee chair and mentor, Dr. David Dickinson, for providing me with ideas, support, guidance, and encouragement throughout the entire process. Dr. Dickinson worked tirelessly to assist me in this process and I can say without a doubt that I would not have completed the thesis without his help.

Likewise, I would like to thank the remainder of my committee, Dr. Andrew Smith, and Dr. Hall Beck, for their support. Overall, I could not have asked for a more helpful and dedicated thesis committee. I would also like to thank Dr. Rose Mary Webb for her unwavering support and guidance throughout my thesis and graduate career.

Finally, I would like to express my sincere appreciation to my family, graduate cohort, and all of the countless others without whom I could not have maintained my sanity while completing my thesis.

Table of Contents

Prime Time? A Look at the Effects of Circadian Mismatch on Stereotype Reliance

Robert Alexander Indiana McClelland

Appalachian State University

Abstract

Stereotype-based decisions are formed as the result of employing various heuristics and biases, and they serve as a way to assess ambiguous situations and compensate for limited information processing. Research has demonstrated that during circadian mismatched (non-optimal) periods of the day cognitive resource availability is diminished. This study examined the influence of circadian arousal levels (particularly in mismatched conditions) on the tendency to use stereotypes in decision-making tasks. It was predicted that mismatch between chronotype (individual circadian preference) and time of day would correlate negatively with cognitive resource availability, thus increasing vulnerability to stereotype reliance. Participants were 59 Appalachian State undergraduates. The participants were administered an online survey consisting of the validated reduced Morningness-Eveningness Questionnaire, the Karolinska Sleepiness Scale, the Epworth Sleep Scale, and a stereotyping task. Each subject participated in sessions at two different times of the day, with the sessions occurring approximately one week apart. Though the stereotype priming manipulation failed, results suggest that participants in adverse sleep or circadian states may have still relied on biases or heuristics when assessing guilt.

Keywords: stereotyping, circadian mismatch, heuristics, conjunction fallacy

Circadian Influences on Stereotype Reliance

"Stereotyping" is a word that carries a largely negative connotation. In our society, it has traditionally been associated with bigotry and narrow mindedness. Currently, social rights are a very pertinent political topic. While segregation has been deinstitutionalized and people of all ethnicities are granted the same inherent rights, discrimination is still widespread (Gill, 2015; Sankar, 2014). Notably, recent events have highlighted racially discriminatory practices within certain domains of the justice system (U.S. Department of Justice, Civil Rights Division, 2015), as well as within civil institutions, such as marriage (*U.S. v. Windsor*, 2013). It is highly unpopular to be perceived as prejudiced. However, the truth is that people naturally rely on stereotypes as a method of responding to incomplete or ambiguous information (Rothbart, Fulero, Jensen, Howard, & Birrell, 1978).

In modern society, one area where stereotyping is particularly relevant is within the context of emergency services provision. Many times, during the fulfilment of these services, human lives are placed into the hands of responders in the field. When law enforcement officers are asked to respond appropriately to potentially violent criminals, or when Emergency Medical Service (EMS) workers are asked to prioritize medical treatment, it is important to understand the factors that influence those decisions. Unfortunately, responding to crises is stressful and often uncertain work. This means that responders may not have the time or cognitive resources to critically assess every situation.

Many emergency service workers are also called upon to make life-and-death decisions during all hours of the day (McMenamin, 2007). Therefore, it is essential that society gain a better understanding of the impact of individual differences in time-of-day effects on decision-making processes, including the use of stereotypes. Chronotypes, or

individual differences in time-of-day alertness/sleepiness patterns, often enter into laypersons' explanations of attentiveness (Kruglanski & Pierro, 2008). That is to say, the idea that differences in diurnal preference establish some people as "morning people," and others as "evening people" is well recognized in society (Horne & Östberg, 1975; Lack, Bailey, Lovato, & Wright, 2009). While it is clear that diurnal preferences differ across individuals, it is still somewhat uncertain what some of the larger implications are, especially for individuals working during their non-peak hours of the day. This study aims to expand upon the current body of literature by examining the impact of time-of-day, for individuals with different diurnal preferences, on use of a common decision process. By extension, this study hopes to add to our understanding of how sleepiness may influence the use of stereotypes.

Literature Review

Heuristics and Biases. Stereotypes encompass the broad category of false, or oversimplified, beliefs concerning specific groups of people or things. In order for a belief to be considered a stereotype, it must be connected to all members of a specific subpopulation. Stereotypes are generally subdivided into two distinct categories: illusory correlations and category accentuations (Bodenhausen, 2005). An illusory correlation is the belief that a trait is correlated with a specific group, even though there is no actual statistical difference between that group and the population. On the other hand, a category accentuation is an exaggeration of real between-group differences (Bodenhausen, 2005; Sherman et al., 2009).

Despite the negative connotation that the term "stereotype" caries, these beliefs can be both positive and negative. Commonly the word "prejudice" is closely associated with stereotyping. Prejudice, however, refers specifically to negative stereotypes and is more

commonly what is implied when discussing issues of social rights. In order to better understand stereotypes it is important to first explore some of the less apparent underlying factors behind them.

Over the course of a typical day, the mind is constantly exposed to an abundance of information. Each sensory input: sight, smell, touch, taste, and hearing provides a source of information to the brain about the environment. These stimuli activate associated neural pathways, providing meaningful information and allowing the mind to make informed decisions. In order to more efficiently interpret task-relevant information the brain relies heavily on "mental shortcuts," or heuristics, especially when there is a large amount of situational ambiguity (Tversky & Kahneman, 1974). Many heuristics often promote stereotyped assessments. One such shortcut, the availability heuristic (Tversky & Kahneman, 1974), is employed when determining the probability in uncertain situations. This heuristic is utilized by comparing events and circumstances in the present situation to a mental model formed on the basis of the most salient previous experiences. As such, it is thought to be one of the primary driving forces behind stereotype formation (Rothbart, et al., 1978).

The question arises: what determines the influence of any given past experience on the formation of a stereotype? A study by Rothbart et al. (1978) found that the degree of cognitive load (amount of total mental effort being expended toward tasks held in working memory) during encoding plays an essential role in stereotype formation. Under conditions of low cognitive load, the researchers found that participants observed and encoded characteristics of individual members of a group exclusively. That is, the participants bound features to individuals, rather than to the group as a whole. For example, if a group consisted

of three individuals, (Bill, Sam, and Beth), participants would observe traits of each individual and attribute them specifically to that individual. However, when under a high degree of cognitive load, participants bound those same individual traits to the groups as a whole. Returning to the previous example, if Bill was perceived as lazy, a participant under heavy cognitive load would attribute laziness to the entire group, with Bill as the salient example. This is particularly true when individuals exhibit extreme traits and behaviors.

Another issue presented during encoding is the misrepresentation of relative frequency. When the observer is under a high degree of cognitive load they attribute all instances of an observed trait to the overall group, regardless of selection bias (Rothbart et al., 1978). If, in the previous example, Bill was observed four times and Sam and Beth are only observed once each, attributes associated with Bill would be perceived as being twice as likely as attributes of Sam and Beth together, despite the fact that, according to Bayes' Theorem, this is only the case if the selected member of the group is twice as likely to be Bill (Anderson-Cook, 2014).

The strength of a stereotype is often amplified by what is known as the confirmation bias. The confirmation bias (alternatively referred to as a "hypothesis-confirming bias") refers to the tendency to selectively seek, interpret, and recall only the evidence that confirms one's pre-existing beliefs (Wason, 1960). In a study by Darley and Gross (1983) participants were asked to watch a video of a child answering questions on a standardized exam. In one condition, the child in the video was dressed in ragged, inexpensive clothing, with the goal being to suggest that the child came from a low socioeconomic status (SES) family. In the other condition, the child was wearing nice, expensive-looking clothing, to suggest a high SES family background. Participants in each condition were asked to assess the child's

performance and comprehension level, relative to other children of the same age. In the low SES family condition, participants displayed a strong tendency to rate the child's performance as significantly lower than other children of the same grade-level did, whereas in the high SES condition the opposite was true.

When asked why they assessed the child the way that they did both groups gave examples of behaviors that supported their conclusions. However, neither group mentioned apparent wealth as a factor. It is also worth mentioning that many of the noted behaviors were cited by both groups but framed differently by each (e.g. the low SES group might interpret an expression as "confused," while the high SES group reported the same expression as "thoughtful"). This would indicate that participants formed an initial stereotype-based hypothesis (i.e. SES and academic performance are positively correlated), and then attended to ambiguous behaviors by framing them to support that hypothesis (Darley & Gross, 1983).

Another proclivity that commonly acts to strengthen stereotypes is the conjunction fallacy. The conjunction fallacy is the common belief that the probability of the co-occurrence of two attributes is greater than the probability of the occurrence of either attribute, without the other [this is expressed mathematically as $Pr(A \cap B) > Pr(A) \vee Pr(B)$]. This logically cannot be true, as the actual set of $A \cap$B is a subset of both the actual set of *A* and the actual set of *B* (Tversky & Kahneman, 1983). For example, if an individual watches professional basketball on television and is thereby exposed to many tall African-Americans he/she might form a mental model based on this perception. Acting under the conjunction fallacy, that individual would assess that people are more likely to be both tall and African-American than only African-American or only tall.

In a study by Tversky and Kahneman (1983) participants were given a short biographical summary of a fictional individual, followed by a series of true/false statements about the person. Statements were categorized as either stereotype consistent, stereotype inconsistent, or a combination of the two types of items. For example, the participant would read a paragraph about Linda, an intelligent, single, 31-year-old woman who is interested in social justice. The participant would then respond to a series of propositions, noting the likelihood that each was true. Participants generally reported that stereotype inconsistent propositions (such as *Linda is a bank teller*) were unlikely to be true. They rated stereotype consistent items (such as *Linda is an activist in the feminist movement*) to be far more likely. Finally, participants ranked combined items (such as *Linda is a bank teller and an active feminist*) to be more likely than the inconsistent items alone. In this example, the set of all bank tellers is inclusive of all bank tellers who are active in the feminist movement, as well as all bank tellers who are not. Therefore, it is impossible that Linda is more likely to be both a bank teller and an active feminist than simply a bank teller. This demonstrates that the participants were committing the conjunction fallacy. The stereotype consistent proposition adhered to the mental model that the participants were using to assess likelihood. This increased the perceived probability of any proposition that included that portrayal.

One driving force behind the conjunction fallacy has been called the "representativeness heuristic" (Tversky & Kahneman, 1974). The representativeness heuristic, simply stated, is the tendency to believe that the more related that two events or concepts are then the more likely that they will occur conjointly (Gavanski & Roskos-Ewoldsen, 1991). One method that has been used to measure this effect has been to provide participants with a hypothetical individual (accompanied by several basic personality

characteristics) then ask them to select the individual's career from a list of occupations. For instance, if participants were told about John, (a hardworking, outgoing, and determined man) then were asked to determine his career they would inevitably chose a career that they associated with those traits (e.g. a car salesman) (Tversky & Kahneman, 1974). In many ways, this is operating very similarly to stereotyped assessments. While the representativeness heuristic is thought to be an underlying mechanism of stereotypes, the primary distinction in this case is the source of information. For example, in these studies participants are typically provided with various traits, and then are asked to match people with a group, based on those traits. When people use stereotypes, the direction of the process is typically reversed. Specifically, the "classification" of an individual is observed, and then traits are ascribed to them on the basis of group membership.

Gavanski and Roskos-Ewoldsen (1991) propose an alternative driving mechanism behind the conjunction fallacy. Although they acknowledge that the representativeness heuristic is a major contributing factor to the conjunction fallacy, they also posit the idea that people utilize incorrect combination rules for probability assessments. In their study, participants were provided with two statements and given the likelihood of each occurring independently. They were then asked to estimate the likelihood of the two statements occurring conjointly. Finally, each participant completed a short questionnaire aimed at divulging the strategies that were used in assessing the conjoint probability. When two events were perceived to be relatively unrelated participants judged the probability of co-occurrence to be low, even when the probability of each event was relatively high. The opposite was true when the likelihood of each event individually was low, but the two were considered to be connected. Most participants reported that they roughly averaged the

probabilities for each separate statement, and then adjusted that calculation using personal experience to determine whether the two statements could combine into a logical event.

Incorrect combination rules cannot explain the majority of cases where the conjunction fallacy is committed because, in the course of day-to-day life, people are rarely provided with precise estimates of uncertain events. However, it is important to recognize that, even in cases where statistical probabilities are provided (e.g., cultural-sensitivity training seminars), there may be little actual benefit of including those probabilities. Adjusting conjoint probability through incorrect combination rules likely represents another instance where the availability heuristic alters perceptions of base rate probabilities through emphasizing previous experience over explicit counter-attitudinal information.

Cognitive & Motivational Factors. The tendency to rely on stereotypes in decision-making has been shown to be influenced by several factors. These factors can be broken down into two general categories: motivational factors and cognitively limiting factors (Bodenhausen, 1990). Motivational factors are any situational contingencies that increase the investment of the decision maker in the decision scenario. Conversely, cognitively limiting factors include anything that would diminish the decision maker's ability to process task-relevant information.

One major motivational factor is personal involvement. Personal involvement can be conceptualized as the amount of investment, or stake, which is contingent on the individual making an informed decision. An investigation by Erber and Fiske (1984) examined this effect. In their study, participants were brought in to the lab to allegedly complete a "creativity task," which would be assessed for relative value and determine their payout. In each condition, the participants were paired with a confederate, acting as another participant,

with whom they would ostensibly complete different components of the task. In one condition they were told that the payout would be assessed on the basis of the work completed by both group members combined (creating high personal involvement). In the other condition, they were told that their payout would only be based on their own performance (creating low personal involvement). Each participant was given "student evaluations" of the confederate, which either contained information that was consistent or inconsistent with a self-report provided by the confederate. The researchers measured the amount of time that participants spent reading each evaluation. They found that participants who believed that their outcome was to be based, in part, on the performance of the confederate spent far more time reading the inconsistent evaluations than participants who were informed that their outcome would be based solely on their own performance. The amount of time spent reading consistent evaluations was virtually the same across conditions. In the low personal involvement condition participants did not spend a significantly different amount of time between looking at consistent and inconsistent evaluations. This indicates that participants who were more highly involved in the work assessments of the confederates were more likely to spend time and energy considering conflicting information (Erber & Fiske, 1984). This would indicate that when individuals do not have a personal stake in the outcome of a decision they are more likely to ignore information that may conflict with stereotypes that they hold.

Another major motivational factor is whether there are incentives for response accuracy (Bodenhausen, 1990). This factor is primarily concerned with the presence of a quantifiable reward for providing an accurate response set. Neuberg and Fiske (1987) found that when participants were asked to assess the state of an allegedly schizophrenic patient

(from watching a video) they were more likely to describe the patient in terms that were consistent with stereotypes of schizophrenic individuals when there was no reward offered. That is to say, when participants were not offered a direct incentive for providing an accurate assessment, they were less willing to invest cognitive resources in the task and thereby more likely to rely on stereotypes.

One of the earliest factors shown to increase reliance on heuristics is distraction. Festinger and Maccoby (1964) first posited that distraction serves to inhibit internal counter-arguing, making people more receptive to illogical arguments. The context of their study was particularly focused on propaganda, though the principle has since been generalized to a much broader scope. In a study conducted by Petty, Wells, and Brock (1976) participants were asked to read a series of arguments regarding a substantial proposed increase in tuition at their university. One set of students was given a list of logically compelling reasons for the proposed increase, while the other set was given a list of reasons for which it was easy to develop counter-arguments. Half of the participants in each group were asked to watch a monitor during the task, recording any instances where an "X" appeared on the screen. The number of times that the "X" appeared on the screen varied between participants. This was served as a distractor task. After reading through the list of arguments, the participants were asked to rank their agreeability to the idea of increasing tuition. Overall, participants in the easily counter-arguable condition were significantly less likely to indicate agreement with the proposed increase. However, in that condition the agreeability was positively correlated with the instances of distraction. Interestingly, in the logically compelling arguments condition participant agreeability was negatively correlated with instances of distraction. This suggests that the logical arguments lost a degree of their efficacy when participants did not put as

much effort into dissecting them (Petty et al., 1976). This indicates that, when a distractor is present, participants relied more on the ideological belief that "tuition increases are bad," rather than the logically compelling (counter-attitudinal) arguments. These findings imply that weak counter-attitudinal arguments are not very effective at inciting attitude change, regardless of the presence of distractors. However, compelling arguments are effective at inciting attitude change if presented free of distractions. It follows that in uncertain situations, where there are distractors, people will be more likely to discount information that goes against their established beliefs.

In another study that looked at the effect of cognitive impairment, Blair and Banaji (1996) found that stereotype priming occurred when participants' intention to avoid stereotyped decisions was at a baseline level (i.e., they were not instructed to avoid them) and cognitive impairment (via Stimulus Onset Asynchrony) was high. They measured this by priming either a gender-stereotypic, gender-counter-stereotypic, or gender-neutral attribute. Participants were then presented with an unmistakably male or female "target" name (these were pretested), and were asked to indicate if that name corresponded more strongly with males or females. In order to manipulate cognitive impairment, the researchers manipulated the Stimulus Onset Asynchrony (SOA). SOA is a manipulation that varies the amount of time between the presentations of two related stimuli. In this manipulation, more rapid transition between a prime and a target should create higher cognitive constraints (increasing dependency on the prime). The researchers found that when the SOA was high (>350 ms) participants who were instructed to avoid stereotypes were able to respond more accurately than those who were not instructed to avoid stereotypes when the prime was counter-gender-stereotyped to the target name. However, when the SOA was relatively low (<250 ms) all

participants responded in an equally stereotyped response pattern, regardless of intention. The authors conclude that, based on these findings, participants who were given longer to respond were better able to engage controlled processing than participants who were forced to respond more quickly. This resulted in a less stereotyped response pattern for participants who were instructed to avoid the use of stereotypes and were given a longer response duration.

One additional factor that has been shown to limit the use of controlled processing in decision-making is task complexity. Bodenhausen and Lichtenstein (1987) tested this effect by assigning participants to read a booklet detailing a hypothetical criminal trial that involved a (stereotyped) Hispanic male as the defendant, and then asking them to assess either the aggression (a single personality trait) or overall guilt of the defendant. One sample group received approximately 100 items (high task complexity), while the other received around 20 (low task complexity). Each booklet included an equal number of items that were either stereotype consistent or inconsistent for Hispanic males. After a brief distractor task, each participant was asked to defend his/her assessments by providing as many pieces of specific evidence as possible. The authors found that participants in the "high task complexity" conditions tended to score perceived guilt and aggression higher than those in the "low task complexity" conditions. Furthermore, those in the "high task complexity" condition also provided more of the stereotype consistent items as rational.

Circadian Influences. Circadian mismatch effects occur when individuals who have a morning diurnal preference (i.e., "morning types") perform tasks in the evening, or conversely, when those with an evening diurnal preference perform tasks in the morning.

These effects likely influence several aspects of cognition and decision-making, including alertness and willingness to expend mental energy.

Circadian mismatch and cognition. Circadian mismatch has been demonstrated to impair the availability of cognitive resources necessary for completing a wide range of tasks. Maire, Reichert, and Schmidt (2013) propose that circadian rhythms, in conjunction with homeostatic regulation, play a very important role in vigilance (maintenance of attention). According to their research, the endocrine system regulates alertness on an approximate 24-hour basis, during 16 hours of which alertness levels tend to be higher. The regulation is such that vigilance ideally remains relatively stable across the daylight hours, then steeply declines at the onset of night, and then suffers its sharpest decline at the end of the cycle when melatonin levels peak (i.e. in the early morning). Distinctive variants of this cycle are referred to as chronotypes and explain why vigilance levels may be asynchronous between individuals across the day, which implies important variations in alertness/sleepiness at different times of day for different chronotypes.

Chronotypes are formed through the individuation of endogenous biological rhythms. This process, known as entrainment, occurs when the body uses specific environmental cues, called zeitgebers (literally translated as “time-givers”) to determine when biological processes need to perform at peak levels, and when they should be downregulated (Aschoff, 1954). While entrainment creates very distinct chronotypes, these have been shown to remain relatively stable over periods of several years, with some degree of change occurring across the lifespan. Common zeitgebers include: light, temperature, exercise, social interaction, and cognitive engagement (Aschoff, & Pohl, 1978). For example, individuals who, over a prolonged period, are more cognitively active during the evening hours of the

day will naturally shift toward an evening typology. Homeostatic biological processes attempt to adjust these individuals' vigilance cycles, so that they are more vigilant in the evenings, which results in periods of significant performance impairment in the morning and early daytime hours (Maire, Reichert, & Schmidt, 2013). This would indicate that during periods of circadian mismatch the cognitive resources available for high-level thinking are significantly hindered (Dickinson & McElroy, 2012).

In a study by Natale, Alzani, and Cicogna (2003), participants were tested on a series of cognitive measures, including a visual search task, a syllogistic reasoning task, a spatial reasoning task, and a crypto-arithmetic task. Participants were offered financial incentives for accuracy as a motivating factor. Significant differences between individuals in matched and mismatched conditions were only found for the visual search task, which was an operationalization of sustained inhibitory control (an executive cognitive function). Specifically, participants who were in the matched condition (and therefore were more alert) were better able to block out irrelevant perceptual cues, and thus isolate the target items in the search task better than participants in the mismatched condition. These results suggest that some executive control functions may be circadian modulated.

Another study, by Pica, Pierro, and Kruglanski (2014), examined the effects of circadian mismatch on retrieval-induced forgetting (RIF). RIF occurs when explicit retrieval of a memory implicitly interferes with or inhibits recall of other, related memories. Often, these RIF effects are useful for preventing the unintentional recall of less useful information. In their study, the researchers found that, in an associated words retrieval task, participants in a mismatched condition suffered from more moderated RIF effects than those in the matched group. That is, in the mismatched condition, rehearsing one set of words did not inhibit

subsequent recall of another set of words as much as it did for participants who were circadian matched. The authors postulate that this is the result of diminished inhibitory capability, produced by lower cognitive capacity. It should also be noted that RIF has been demonstrated to be highly correlated with working memory capacity (WMC), strengthening the assertion that it is cognitively modulated (Pica, et al., 2014).

Within the context of social heuristics, Kruglanski and Pierro (2008) found that individuals in circadian mismatched conditions were more likely to experience transference effects in social perception. It was observed that, when strangers physically resembled one's significant other, participants assigned traits that were prescribed to their significant other to the complete stranger. Such transference effects have been linked to use of the representativeness heuristic, which, as previously discussed, is a major component of stereotyping.

Decision-making. Circadian influenced cognitive modulation has further been observed in studies examining judgments and decision-making. Decision making research examines the influence of heuristics and biases in uncertain or ambiguous situations.

A study by McElroy and Dickinson (2010) examined correlations between framing effects and circadian rhythms. Participants were assigned to complete a survey at a designated hour within the full 24-hour day cycle. The survey contained the Asian disease problem (Tversky & Kahneman, 1981), followed by a risky and a risk-free policy option. The Asian disease problem is a task in which participants are told of a hypothetical epidemic disease (that, in this example, will kill an estimated 600 people). They are informed that the Centers for Disease Control (CDC) have proposed two possible plans to address the problem. In the first condition, participants are presented with the choice between a plan that is framed

as a certain gain (e.g. “200 people will be saved”) and another plan that is a risky gain (e.g. 33% chance that all 600 people will be saved and 66% chance that none of the 600 people will be saved). In the second condition the first plan is framed as a certain loss (e.g. “400 people will die”), while the second plan is a risky loss (e.g. there is a 33% chance that no one will die and a 66% chance that all 600 people will die). Participants typically display risk-aversion in the gains frame, but risk seeking preferences in the loss-framing of the task (Tversky & Kahneman, 1981). In the study conducted by McElroy and Dickinson (2010), participants at non-optimal times-of-day displayed a higher tendency toward making their decision based on the manner in which the solutions were framed. That is, mismatched participants were more likely to avoid risk in the gains frame, but seek risk in the loss frame. This is consistent with the hypothesis that mismatched participants utilized more automatic (less effortful) processing during these times.

A 1997 study by Gordon, examined the influence of circadian rhythms on predisposition toward reliance on illusory correlations. Participants who were tested at non-optimal (mismatched) times displayed a significantly higher tendency to form illusory correlations. It was further demonstrated that the participants with a high need for personal structure were more likely to form illusory correlations, likely due to a drive to categorize information (Gordon, 1997).

Finally, the research most closely associated with the present study is Bodenhausen’s (1990) investigation, which specifically examined circadian influences on stereotyping. Participants were randomly assigned to complete a survey at either 9 a.m., 3 p.m., or 8 p.m. At the beginning of the survey, participants completed the Morningness-Eveningness Questionnaire (MEQ) and then read a short description of an alleged misconduct case at their

university. The participants were assessed for chronotype based on a median split of their MEQ scores (i.e., half of the participants were designated "morning types" and the other half were designated "evening types"). Morning types were considered to be "circadian matched" in the 9 a.m. session and "circadian mismatched" in the 8 p.m. session. The opposite was true for evening types. The 3 p.m. session served as a control session for both chronotypes. In one condition the description of the hypothetical perpetrator identified him/her as a member of a sub-group that previous measures had established as being stereotyped as more likely to commit the particular offense. In the other condition, this information was excluded. Participants were asked to rate how likely they felt that the suspect was guilty on a Likert scale. The study showed that morning types were more likely to assess guilt (i.e., respond in a more stereotype-consistent pattern) during the afternoon and evening than in the morning and that, similarly, evening types were more likely to assess guilt during the morning and afternoon than in the evening.

Present Study

Study Overview. The primary goal of this study was to provide a more detailed analysis of the impact of circadian typology, time-of-day (and perhaps, by extension, sleepiness) on the use of stereotypes. There is a significant gap in the literature concerning the more polarized hours of the day (i.e. between 11:00 p.m. and 6:00 a.m.), where several studies have shown that the most pronounced sleepiness or decision effects occur (Adan et al., 2012; Dickinson & McElroy, 2010; Maire, et al., 2013; McElroy & Dickinson, 2010).

While Bodenhausen (1990) provided a substantial foundation for the current literature on circadian influences on judgments (Dickinson & McElroy, 2012; Kruglanski & Pierro, 2008; Maire, et al., 2013; McElroy & Dickinson, 2010; Natale, Alzani, & Cicogna, 2003;

Pica, et al., 2014) there are some prominent methodological issues with the study. Perhaps the most noticeable issue is in regards to the use of a median split to define circadian typology. Smith et al. (2002) demonstrated that, in a cross-cultural sample of 1,749 young-adult participants, the majority were classified as intermediate types (exhibiting no strong preference for morningness or eveningness). Based on other studies examining circadian preferences in college samples (Fabbri, Mencarelli, Adan, & Natale, 2013; Kruglanski & Pierro, 2008) it is evident that there is a strong bias toward evening typology relative to morning typology, with only around 10% or less of young adults classified as morning-types. Therefore, using a median split for such samples would incorrectly categorize many intermediate types as morning type subjects. Incorrectly labeling these intermediate types as morning types implies that the study actually used what amounts to a control group (not truly mismatched in the evening, nor matched in the morning). This reality suggests that results from any study using the median split methodology are more likely "time-of-day" studies than chronotype or circadian mismatch studies.

To address this issue in the current study, participants were categorized using the standardized scoring system for the reduced Morningness-Eveningness Questionnaire (rMEQ), established in the literature. Specifically, chronotype is divided into three levels: Morning, Intermediate, and Evening Types, which are assigned based on where participants' scores fall within the range of possible rMEQ values. Theoretically, there is no predicted main effect for the rMEQ score because there is no ex ante hypothesis as to how diurnal preference alone would influence decision-making or the use of stereotypes. However, since a key hypothesis is that off-peak (i.e., circadian misaligned) TOD decisions are more likely

the product of heuristics, it was predicted that there would be a significant interaction effect between TOD and rMEQ category.

Hypotheses.

Hypothesis 1. When participants are responding at less circadian-optimal times, they will rely more on stereotypes.

Based on the reviewed literature, circadian mismatch appears to act as a cognitive inhibitor in many instances (Dickinson & McElroy, 2012; Kruglanski & Pierro, 2008; Natale, et al., 2003; Pica, et al., 2014). It is also well established that when cognitive resources are inaccessible, and/or there is insufficient motivation to be accurate, there is a greater tendency to rely on stereotypes (Blair & Banaji, 1996; Bodenhausen & Lichtenstein, 1987; Bodenhausen, 1990; Erber & Fiske, 1984; Festinger & Maccoby, 1964; Neuberg & Fiske, 1987; Petty, et al., 1976). More directly, circadian mismatch has been demonstrated to increase reliance on several heuristics and biases that are thought to underlie stereotyping (Bodenhausen, 1990; Gordon, 1997; McElroy & Dickinson, 2010). Hypothesis 1 builds off of this body of literature and testing Hypothesis 1 helped establish whether circadian mismatch increases the tendency to rely on stereotypes.

Hypothesis 2. Collapsing across chronotype, stereotype reliance will be higher in the evening sessions than in the morning sessions, due to accumulated sleep debt.

While hypothesis 1 follows directly from the previous literature reviewed above, hypothesis 2 stems from the fact that the longer that people are awake the more homeostatic pressure they feel to sleep. Conversely, during periods of sleep this homeostatic pressure declines. This process is referred to as the sleep homeostat, which interacts with our bodies' circadian rhythms (Borbély, 1982). Research has shown that as homeostatic sleep pressure

mounts cognitive performance tends to decline (Van Dongen & Dinges, 2005). Morning, intermediate and evening types should all have been considerably more tired during the evening sessions (after extended periods of wakefulness), as opposed to in the late morning or early afternoon, when all of them should have more recently woken up. This pressure should result in diminished cognitive performance and thereby higher stereotype reliance (Lust, 2013). Thus, a secondary hypothesis that follows from hypothesis 2 is that circadian matched evening-types (i.e., evening-types making a decision in the evening) will exhibit more reliance on stereotypes than circadian matched morning-types, because of the daytime buildup of homeostatic sleep pressure.

Hypothesis 3. Participants who are sleepier will be more likely to use stereotypes when assessing guilt.

The third hypothesis is, in many ways, an extension of hypothesis 2. However, it is meant to account for chronic sleep-debt, as well as other factors, outside of normative biological sleep pressure, which may be more prevalent in college populations (Lund, Reider, Whiting, & Prichard, 2010). In order to measure sleepiness, the Epworth Sleep Scale (ESS) and Karolinska Sleepiness Scale (KSS) were used, in conjunction with self-reported sleep questions (for additional details, please refer to the Measures section).

Method

This study received the approval of the Appalachian State Institutional Review Board (IRB) on November 06, 2015 (Appendix K).

Participants

Participants in this study were 59 Appalachian State University undergraduates, recruited through an online participant database (SONA). Of those participants, 75.2% were

female; 86.3% identified as white/Caucasian, 8.5% identified as black/African American, and 5.1% identified as Hispanic/Latino. Participants ranged from 18-27 years of age (M = 19.15, SD = 1.50). All participants were compensated with Experiential Learning Credits (ELCs), which count as partial course credit in many undergraduate psychology courses. Participants who completed both administrations were awarded three ELCs, while participants who only completed the first week were given one ELC. Additionally, participants who completed the entire protocol were entered into a drawing for a financial reward (in the form of cash).

Design

The current study employed a quasi-experimental alternate forms (repeated measures) design. The central outcome variable in the study was participants' stereotype dependence, which is constructed with subject ratings of perceived guilt in each stimulus received (for an explanation of this measure, please refer to the Materials section). The first component of our primary predictor of interest (i.e., circadian mismatch) was circadian preference, a between-subjects measure. This was assessed using participants' scores on the reduced Morningness Eveningness Questionnaire (for a full description, please refer to the Materials section). Participants completed this instrument during both administrations of the survey. The second major predictor was the time frame in which participants completed the survey. This was experimentally manipulated by randomly assigning participants to complete the survey during one of 6 two-hour windows each week. The time frame for the second week administration was intentionally set to be either 6 hours ahead of or behind the first week administration (depending on whether the first week time frame was toward the beginning or end of the 12-hour study window). Research by Smith, et al. (2002) suggests that, regardless of chronotype, this 6-hour difference between administration times should provide each

participant with one time frame that is definitively more "optimal" for him/her than the other (for a graphical depiction of this, see figure derived from Smith, et al.'s 2002 investigation in Appendix A). In this way, the repeated measures element of the time-frame separation imposed on each subject across manipulations allowed me to code one of the administrations for each subject as occurring during a more "good time" than the other, given the subject's specific chronotype. Participants' stereotype reliance was compared within subjects, between administrations.

Materials

reduced Moningness-Eveningness Questionnaire (rMEQ). The rMEQ (Adan & Almirall, 1991) is a validated instrument, which is used for circadian typology. A recent meta-analysis by Adan et al. (2012) has revealed that the full-length MEQ (Horne, & Östberg, 1975) demonstrates high construct validity and has a test-retest reliability of .88-.89, over a 3-month period (Larsen, 1985; Neubauer, 1992). However, the majority of the variance comes from a significantly smaller subset of items. Further, the MEQ is a multidimensional measure, diminishing its value by creating a composite score. Therefore, it has become common practice to use the rMEQ. The rMEQ is a five item "pure" measure of morningness (Di Milia, Adan, Natale, & Randler, 2013), with scores ranging from 4-24, where higher values are indicative of a stronger preference for morningness. The rMEQ correlates with the original MEQ from .69 to .90 (Adan, et al., 2012). The rMEQ was used to assign participants to one of three diurnal preference categories using standard rMEQ cutoff scores: Morning Types (assessed as individuals scoring 18 or higher on the rMEQ), Intermediate Types (assessed as individuals scoring between 12 and 17 on the rMEQ) and Evening Types (assessed as individuals scoring 11 or less on the rMEQ). These cutoffs are

established such that participants who have a large amount of variance in their responses, or employ a fence-sitting response strategy should be classified as intermediate types (Caci, Deschaux, Adan, & Natale, 2009). It is worth noting that originally the MEQ was designed to identify individuals with a strong morningness/eveningness preference, while those in the middle of the scale were simply referred to as "indeterminate" (Horne, & Östberg, 1975). As noted in the Results section, the correlation between rMEQ scores for each participant, across the two administrations, was positive and statistically significant, which indicates good test-retest reliability.

Karolinska Sleepiness Scale (KSS). The KSS is a commonly-used measure of participants' in-the-moment, or current-state, sleepiness. This instrument asks participants to rank their subjective alertness on a nine-point Likert scale (Åkerstedt, & Gillberg, 1990). The KSS has been validated using both performance (psychomotor vigilance) and electroencephalogram (EEG) data (Kaida, et al., 2006).

Epworth Sleep Scale (ESS). The ESS is a diagnostic sleep-pattern instrument that is designed to measure chronic daytime sleepiness. The scale consists of eight items that ask participants how likely they typically are to fall asleep while engaging in various everyday tasks, such as reading before bed, or as a passenger in a car (Johns, 1991). The scale was originally designed as a lower-cost alternative to the multiple sleep latency test (MLST), and has since been found to possess higher construct validity (Johns, 2000) for assessing Excessive Daytime Sleepiness (EDS). Patients who have been diagnosed with a wide-variety of sleep disorders, including obstructive sleep apnea, narcolepsy, idiopathic hypersomnia, and insomnia all tend to have significantly higher scores on the ESS (Johns, 1991).

Additional Sleep Measures. Some additional self-reported sleep measures were also elicited, such the number of hours slept the previous night, the average number of hours slept each night over the previous week, subjective belief of the number of hours of sleep needed for optimal performance, and total number of hours since the participant was last asleep.

Stereotype-Reliance Measure (SRM). The SRM was developed specifically for use in the present study. However, it is heavily based on the instrument established by Bodenhausen (1990). This task served as the basis for determining the impact of stereotype-consistent primes on participants' judgements. The SRM presents participants with a series of two alleged criminal misconduct cases. In each case participants are provided with a basic report of the supposed offense and then asked to assess how likely they perceive the suspected perpetrator to be guilty. The reports give a series of basic details, including the nature of the offense and evidence that suggests that the individual might or might not be guilty. In each case the "evidence" provided should not be sufficient to establish indisputable guilt or innocence. In addition, participants are presented with a photograph of the alleged perpetrator's face, displayed above the evidence on the computer screen. In the stereotype-consistent condition the individual pictured is a member of a group stereotyped as being more likely to commit the purported offense (for an example, see Appendix B). In the other condition, the facial profile is of an individual who is not a member of that stereotype group. In the multivariate analysis, Stereotype Reliance was conceptualized by looking at the impact of priming on the perceived guilt measure.

Each stereotype used has been validated by previous research. It is also worth noting that visual images have been used to successfully elicit stereotype primes in other studies (Brown, Coyne, Barlow, & Qualter, 2010; Dasgupta, & Greenwald, 2001; Eberhardt, Goff,

Purdie, & Davies, 2004), but not specifically in the context of guilt assessment. The groups used for stereotypes are: African Americans, college athletes, Hispanic Americans, and heterosexual males. The African American individual has been charged with assault (Dixon & Maddox, 2005), the college athlete with underage drinking (Ashmore, Del Boca, & Beebe, 2002), the Hispanic American with breaking and entering (Welch, Payne, Chiricos, & Gertz, 2011), and the heterosexual male with domestic abuse (Hamby & Jackson, 2010). In each case, the stereotype-inconsistent face belongs to a Caucasian male, except in the case of the domestic-abuse condition, in which it belongs to a Caucasian female. In the "college athlete" condition participants are provided with a Caucasian male facial portrait. In this case, the information that the individual pictured is a student athlete is provided in the "evidence" statements. By using pictures of the alleged criminals' faces, rather than racially suggestive names (as used Bodenhausen's 1990 investigation) the goal was to create a more powerful prime, as well as eliminate any racial ambiguity that use of a name may create. Facial portrait stimulus images have been obtained from the "Face Place," a database composed of 200 individuals of varying ethnicity, constructed by researchers at the Michael J. Tarr Center for the Neural Basis of Cognition and Department of Psychology, at Carnegie Melon University. This database is available via a Creative Commons Attribution-Noncommercial-Share Alike 3.0 Unported License, which makes it publicly available for non-commercial applications, without the express consent of the original licensor.

In order to minimize potential confounds in the choice of facial photographs and also minimize incidences of corner solutions in the assessment of likely crime guilt (i.e., everyone thinks alleged perpetrator is guilty or not guilty due to the set of contrived evidence being too strong or too weak), an initial pilot study was run, using participants from the Amazon.com

Mechanical Turk database. The pilot study presented participants with multiple items of "evidence" for each "crime," in order to assist in balancing the perceived weight and severity of each item, respectively. Each evidentiary item was assessed using a 5-point Likert scale, with 1 representing "strongly indicative of innocence" and 5 representing "strongly indicative of guilt." The evidentiary items that were selected for use in the main investigation average a score of approximately 2.95 for each "case," which should indicate that overall neither innocence nor guilt is strongly established.

The images used for facial stimuli were likewise prescreened by pilot testing, both for attractiveness and for interpreted ethnic identity. Participants were presented with a number of racially diverse facial portraits from the Face Place database. For each portrait, they were asked to assess attractiveness, using a five point Likert scale, ranging from "very unattractive" to "very attractive." They were also asked to select the race of the individual pictured from a drop-down box, and rate their confidence in this selection, from 0% to 100% certainty. Faces being used for each "crime" were matched on attractiveness (such that there was not a statistically significant difference between scores, $p > .05$), and all of the images used were identified with at least 75% confidence as the correct ethnicity by at least 80% of the participants surveyed. This should have helped to eliminate stimuli that were potentially racially ambiguous.

For the SRM, the assessment of perceived likelihood of guilt was measured using a sliding scale, with values ranging from 0-100. Participants were instructed to respond to this item "promptly and accurately." In the final multivariate analysis, Stereotype Reliance is designed to be measured by evaluating the impact of priming on the perceived guilt measure.

Procedure

Participants were directed to complete an online survey, developed through Qualtrics, within a pre-selected time frame. Each participant was assigned to complete the survey twice, with each administration taking place approximately one week apart. Each time frame spanned a 2-hour period, beginning at 22:00 and running through 10:00 the following morning. Time stamps on each survey completion were checked to assess subject compliance. This provided 6 two-hour sessions per day, which should have allowed for relatively discernable variation in circadian arousal for each chronotype (Smith et al., 2002). Each participant was randomly assigned to one of the six time frames for the first session. The second session time frame was then assigned to the one 3 sessions (or 6 hours) apart from the first (for clarification, see Appendix C). This should have generated a substantive difference in circadian mismatch and likely cognitive deficit level differences between sessions, and allows for the construction of an indicator variable that equals one for the session at a more optimal time of day given that subject's chronotype. For all participants the Underage Drinking and Breaking & Entering stimuli were shown during the first administration, and the Assault and Domestic Abuse stimuli were shown during the second (in that order). However, only half of the participants saw the Underage Drinking stimulus primed (while the other half saw the Breaking and Entering primed) during the first week. During the second week, this was counterbalanced, such that the participants who saw the Underage Drinking stimulus primed in the first week saw the Domestic Abuse stimulus primed in the second week (and visa-versa for the other half of participants).

The survey started with an Informed Consent page, where subjects were required to indicate their consent in order to complete the rest of the survey. Initial survey questions

gather subject data on common demographic variables such as age, sex, ethnicity, etc. Next, participants completed the rMEQ, KSS, ESS and other self-reported sleep questions. After finishing the sleep-oriented items, participants were presented with the SRM task. The crime reports used in the SRM varied for each session, such that each participant saw a total of four unique reports (two per administration). Participants were also asked to report how severe they interpreted the supposed crimes to be on a scale of 1-100, with 1 representing "Not at all" and 100 representing "Most severe crime possible." This item was originally included to be used as an independent variable in the regression analysis that would control for differences in legal attitudes toward each offense. For example, participants might have been more sympathetic toward someone accused of a minor alcohol or drug offense, as opposed to domestic violence. Therefore, they might have been less likely to perceive guilt in the drug or alcohol offense condition.

Next, participants responded to a number of other legally oriented questions. These questions were designed to conceal the target variables of the investigation, and are of no conceptual value. Finally, the participants responded to a suspicion check, at the end of the second session. The suspicion check asked participants what they believe the study was attempting to examine. This was assessed in order to exclude any participants whom may have correctly guessed that stereotyping was a key component in the study.

Results

In order to test the priming manipulation, a Factorial ANOVA, that included stimulus (i.e. crime) and priming as factors, was used to compare mean perceptions of guilt. A significant ANOVA model was found, F (7, 227) = 5.25, $p < .001$, $\eta^2 = .15$. The results showed that the perceived probability of guilt was significantly different between crimes, F

(3, 231) = 11.22, $p < .001$, $\eta^2 = .14$. However, probability of guilt was not significantly different for crimes that were stereotype primed (versus not primed), F (1, 233) = .48, p = .58, $\eta^2 < .01$, and there was not a significant interaction between stimulus and priming F (3, 231) = .87, p = .46, $\eta^2 = .01$.

Post hoc tests revealed that the average perceived guilt for Underage Drinking (M = 65.25, SD = 26.35, 95% CI [58.50, 72.01]) was significantly greater than the average perceived guilt for Assault (M = 38.21, SD = 26.35, 95% CI [31.40, 45.03]), M_{diff} = 27.04, t (58) = 5.42, $p < .001$, d = 1.05, and for Domestic Abuse (M = 49.36, SD = 26.58, 95% CI [42.48, 56.24]), M_{diff} = 15.70, t (58) = 3.46, $p < .001$, d = .60. That said, it did not significantly differ from the average guilt rating for Breaking and Entering (M = 57.28, SD = 26.12, 95% CI [50.53, 64.04]), M_{diff} = 7.85, t (58) = 1.56, p = .12, d = .26. The mean perceived guilt for Breaking and Entering was significantly greater than for Assault, M_{diff} = 19.12, t (58) = 4.10, p = .001, d = .79, but was not significantly greater than for Domestic Abuse, M_{diff} = 7.85, t (58) = 1.89, p = .38, d = .32. Finally, the post hoc analysis revealed that participants were not significantly more likely to assess the suspect of being guilty in the Domestic Abuse case than in the Assault case, M_{diff} = 11.27, t (58) = 3.04, p = .10, d = .45. For a detailed representation of these findings, please refer to Appendix F.

In order to ensure that participants were not aware of the priming manipulation (and thusly compensating), each participant's suspicion check was carefully looked at. The item was not mandatory, and many participants elected not to respond. However, of those who did (approximately 60%) only two participants indicated a belief that the study may have been looking at stereotyped/prejudiced responding. These participants were excluded from further analysis. Though we did not find a main effect of priming on guilt ratings, it would

be premature to conclude that the stereotype priming was a failed manipulation. It is possible that an interaction effect may be present, whereby the priming manipulation was effective when administered at a suboptimal time of day (for example) but actually worked in reverse at the more optimal time of day. This possibility will be explored more thoroughly in the multivariate regression analysis

In addition to further examining the impact of the priming manipulation, the multivariate analysis will also allow for some exploratory analyses that may help identify other factors that influence participants' guilt assessments. Additionally, it will be discussed how the assessment of guilt stimulus may still be used to test the impact of adverse sleep states (i.e., circadian mismatch, high KSS, high ESS) on use of a distinct heuristic that was not initially considered but may still relate to stereotype use.

Before analyzing circadian-related variables, a bivariate correlation was run to compare participants' rMEQ scores between weeks. Consistent with the literature, the scores were highly correlated, $r^2 = .84$, $p < .001$. This indicates that participants responded consistently on this item between administrations. In order to test the effect of circadian mismatch, a categorical proxy variable for "optimal time" was constructed. This proxy variable is favored over using the rMEQ scores (or categories) in the analysis for two reasons: First, the distribution of rMEQ scores is highly skewed toward low rMEQ scores ($M = 12.28$, $SD = 3.52$, 95% CI [11.83, 12.73]) (see Appendix D for histogram of rMEQ scores), which significantly hinders the ability of the data to represent the full continuum of preferences. Second, the selection of the time frame for session two was explicitly intended to create a more optimal and less optimal time-of-day session for each participant. This proxy variable dichotomously classified each administration as being at the more or less

optimal time for each participant. As such, each participant had one administration that was classified as "*More Optimal Time*=1" and one that was classified as "less optimal (i.e., *More Optimal Time*=0)."

As alluded to earlier, the *More Optimal Time* indicator variable is based on each participant's assigned rMEQ category, in conjunction with the time frame of the administration. The scoring values were assigned using the graph by Smith et al. (2002) (Appendix A) as a reference and, as a result, the *More Optimal Time* variable is scored individually for each subject. Using the Smith et al (2002) data for typical alertness at different times for different chronotypes is beneficial for our purposes. Specifically, the data in their study are based on a large sample (of over 1700 participants), assessing alertness over less-constrained weekend day/nights only for times when typically awake. As such, their data are more likely to reflect typical expected alertness ratings without the confound of stimulant or stimulant intake (to combat extreme sleepiness due to taking the survey when not typically awake). Because we explicitly manipulate time-of-day for decision-making, our data are more likely to include such confounds. The construction of this *More Optimal Time* variable for the present study in conjunction with KSS scores allowed us to separately estimate the impact of current state sleepiness from typical sleepiness in comparable individuals at the same time frame, which is more reflective of the underlying time-of-day optimality construct we attempted to examine. In order to provide a visualization of how the *More Optimal Time* variable was constructed, an example of the scoring template has been added to the figure in Appendix A. A histogram of Probability of Guilt scores broken down between more and less optimal time-of-day is provided in Appendix E.

Additionally, a proxy "sleep deprivation" variable was calculated. This variable was designed to provide a rudimentary look at chronic sleep restriction, which has been shown to overly represented in college populations (Lund, et al., 2010; Regestein, et al., 2010; Tsai & Li, 2004). This variable, *Sleep Deprivation*, was calculated by taking the average amount of nightly sleep that participants reported getting over the past week and subtracting it from their reported optimal amount of sleep. For this variable, higher values indicate higher levels of self-reported sleep deprivation. While *Sleep Deprivation* is only a proxy for objectively measured sleep deprivation, it is difficult to measure sleep debt outside of self-reported measures (Lund, et al., 2010). This variable will at least allow for an examination of the impact of this highly common adverse state on decision outcomes.

Interestingly, the crimes that, on the surface seem more severe also appeared to produce lower perceptions that the suspect was guilty. In order to look at differences in crime severity between cases, a One-Way ANOVA was used. The ANOVA showed that the crimes did, in fact, have differing severity ratings, F (3, 231) = 43.26, $p < .001$, $\eta^2 = .37$.

Once again, a post hoc analysis was used to examine differences between each specific crime. This revealed that Underage Drinking (M = 37.45, SD = 29.01, 95% CI [30.11, 44.79]) was, on average, rated as being less severe than Breaking and Entering (M = 75.25, SD = 16.87, 95% CI [70.98, 79.52]), M_{diff} = 37.80, t (58) = 11.57, $p < .001$, d = 1.59, Assault (M = 70.26, SD = 23.20, 95% CI [64.39, 76.13]), M_{diff} = 32.82, t (58) = 7.92, $p < .001$, d = 1.25, and Domestic Abuse (M = 81.79, SD = 18.76, 95% CI [77.04, 86.54]), M_{diff} = 44.35, t (58) = 9.66, $p < .001$, d = 1.82. The only other significant difference in crime severity was between Assault and Domestic Abuse, M_{diff} = 11.53, t (58) = 3.30, $p < .001$, d = .55. Meanwhile, Domestic Abuse was marginally significantly more severe than Breaking

and Entering, M_{diff} = 6.54, t (58) = 1.88, p = .07, d = .37. Breaking and Entering and Assault were not significantly different from each other, M_{diff} = 4.99, t (58) = 1.42, p = .16, d = .25.

In order to investigate a pure time-of-day effect on perceived guilt, a One-Way ANOVA was used. Times were sorted into their time frame categories (with 1 corresponding to the 22:00-24:00 time frame and 6 corresponding to the 8:00-10:00 time frame). A significant ANOVA model was produced, F (235, 5) = 2.75, p = .02, η^2 = .06. A post hoc analysis revealed that participants responding between 22:00 and 24:00 (M = 41.63, SD = 26.70, 95% CI [32.00, 51.25]) rated the suspects as being significantly less (one tail) likely to be guilty than participants responding between 4:00 and 6:00 (M = 63.62, SD = 27.32, 95% CI [54.09, 73.15]), M_{diff} = 21.99, p = .01, d = .81. Participants responding between 0:00 and 2:00 (M = 45.45, SD = 29.03, 95% CI [35.90, 54.99]) were likewise significantly less (one tail) likely to find the suspects guilty than participants who responded between 4:00 and 6:00, M_{diff} = 18.17, p = .03, d = .65. There were not any other statistically significant differences in participants' perceived guilt ratings between time frames.

Before conducting a multivariate analysis that incorporated sleep variables, a correlation matrix between these variables was produced. The purpose of this matrix was to help identify factors with a high degree of concept overlap. This analysis found several significant correlations between sleep variables. Specifically, *KSS* scores were significantly correlated with *Sleep Deprivation*, r (232) = .24, p = .001, as well as with *More Optimal Time*, r (232) = -.33, p < .001. Intuitively, it makes sense that participants who are more sleep deprived and participants who are responding at worse times would be more tired. Unsurprisingly, *KSS* was also significantly correlated with both the number of hours of sleep the night prior, r (232) = -.19, p = .002, and the number of hours since the participant was

last asleep, r (232) = -.12, p = .04. Participants' average amount of nightly sleep over the last week was highly correlated with their sleep level the night prior, r (232) = .53, p < .001. This indicates that the nights when participants completed the survey were likely representative of a "typical night" for them over the last week. The average weekly sleep score was also correlated with participants' *Epworth* scores, r (232) = -.13, p = .03, as well with *Sleep Deprivation*, r (232) = -.32, p < .001. This seems to suggest that participants who sleep less each night, on average, tend to be more chronically daytime sleepy, and are also more sleep deprived. The number of hours since participants were last asleep was highly correlated with the *More Optimal Time* measure, r (232) = .46, p < .001, which suggests that many participants in the less optimal times were asleep very shortly before taking the survey. Finally, the amount of sleep that participants received the night prior was correlated with whether they were assigned to a *More Optimal Time*, r (232) = .26, p < .001. This finding makes sense, given that participants who were assigned to the early morning hours (which were nearly universally considered to be less optimal) likely went to bed before taking the survey and reported the number of hours that they slept before waking to take the survey. For the full correlation matrix, please refer to Appendix G. Based on the strength of these correlations, participants' average sleep over the last week, sleep last night, and number of hours since last asleep were excluded from the multivariate analysis.

In order to look at the unique impact of each adverse sleep state (*More Optimal Time=0, KSS, Epworth, Sleep Deprivation*) on *Probability of Guilt*, while holding all other variables constant, a series of three multivariate regression models were estimated (Appendix H). Each elicited *Probability of Guilt* score was treated as a separate observation in each regression, with the error term clustered by participant, in order to account for multiple

observations per participant (i.e., 4 total *Probability of Guilt* scores per participant). Though not reported, we also estimated each model with an interaction variable *More Optimal Time*Prime,* in order to test for the possibility of a stereotype priming effect that may have worked in opposite directions for the key *More Optimal Time* measure. In no instance is this interaction variable even marginally significant ($p > .10$ for all models), and so we consider that the stereotype prime manipulation was not effective.

In Model 1, r (223) = .30, r^2 = .09, p = .004, all independent variables were assumed to have a linear impact on *Probability of Guilt*. Based on this assumption, both *Epworth* scores, b = 1.69, p = .01, 95% CI [.53, 2.85], and *More Optimal Time*, b = -8.87, p = .04, 95% CI [-16.98, -.76] were found to be significant predictors of participants' guilt assessments. Specifically, the estimates in model 1 indicate that both of these adverse sleep states—higher *Epworth*, and *More Optimal Time*=0—lead to increased assessments of *Probability of Guilt.*

The second model, r (220) = .32, r^2 = .10, $p < .001$, incorporated quadratic terms for the continuous-measure sleep-related variables, as the impact of these variables may be assumed to differ depending on the degree of the adverse sleep state—this was more of an exploratory analysis. In this model, *Epworth* scores were again a significant (linear only) predictor, b = 4.43, p = .05, 95% CI [.17, 8.69]. Meanwhile, *More Optimal Time* remained a significant predictor as well, b = -8.56, p = .05, 95% CI [-16.75, -.37].

The final model, r (215) = .37, r^2 = .14, $p < .001$, incorporated interaction effects between the *More Optimal Time* and the other sleep-related variables, in addition to the quadratic terms of model 2. The idea behind estimating the interaction effects is to test the hypothesis that adverse sleep states (e.g., high *Epworth* sleepiness) may be compounded by

suboptimal circadian response times. In this model, *Epworth* was only a marginally significant predictor of perceptions of guilt, $b = 4.29$, $p = .07$, 95% CI [-.63, 9.21]. Model 3 estimates also indicate that *KSS* scores are a new significant predictor of guilt ratings, $b = 17.20$, $p = .01$, 95% CI [4.48, 29.92], and the significant negative coefficient on the *KSS* quadratic term, $b = -1.04$, $p = .04$, 95% CI -1.99, -.09] indicates that increased state-level sleepiness appears to increase *Probability of Guilt* ratings at a decreasing rate. For a graph depicting this non-linear relationship, see Appendix I. While this non-linear effect is only marginally significant, it is still potentially interesting from an exploratory perspective, and seems to indicate that there may be a ceiling effect for the relationship between *KSS* and *ProbGuilt.* Largely due to the inclusion of the *More Optimal Time* interaction effects, the main effect for *More Optimal Time* was not estimated to be significant in model 3, $b = 25.25$, $p = .13$, 95% CI [-12.34, 62.84]. However, there was a significant interaction between *More Optimal Time* and *KSS*, $b = -5.94$, $p = .03$, 95% CI [-11.24, -.64]. The sign of this significant interaction term indicates that decisions during suboptimal times-of-day may only increase *Probability of Guilt* ratings for those subjects who are at high levels of self-reported *KSS*-sleepy at the time of the decision. While speculative, this may be due to the fact that some subjects may engage in compensatory strategies to overcome sleepiness during the task administration at a suboptimal time (e.g., drink caffeine before completing survey). Thus, only those subjects who did not engage in such strategies, and who therefore are most sleepy at the time of the decision task, are adversely affected by the bad time-of-day. For a graph depicting this interaction, please refer to Appendix I.

In order to look at the relationships between *Epworth*, *KSS*, *More Optimal Time*, and *Sleep Deprivation* and *ProbGuilt,* without controlling for other predictors, a simple

regression model was run using each sleep variable predictor. In the regression with *Epworth,* it was a significant predictor, r (232) = .20, r^2 = .04, p = .02, b = 1.41, p = .01, 95% CI [.44, 2.38]. In the regression with *KSS,* it was a significant predictor, r (232) = .10, r^2 = .01, p = .04, b = 1.70, p = .04, 95% CI [.10, 3.30]. In the regression with *More Optimal Time,* it was a significant predictor, r (232) = .20, r^2 = .04, p = .02, b = -8.73, p = .02, 95% CI [-15.49, -1.97]. Finally, in the regression with *Sleep Deprivation,* it was not a significant predictor, r (232) = .05, r^2 = .00, p = .96, b = .03, p = .96, 95% CI [-1.07, 1.13].

Overall, three major factors emerged as significant predictors of perceptions of guilt, and there is a consistent theme regarding adverse sleep state effects found in multivariate estimations (and simple regressions). It would appear that when participants were experiencing higher levels of chronic daytime (Epworth) sleepiness, they were more likely to assume that the suspects were guilty, holding constant the case evidence presented. The final measure, which was the proxy measure for the more optimal decision time-of-day, sought to look at the effects of circadian mismatch. Based on this study, the results indicated that suboptimal times-of-day produced increased assessments of guilt. Based on the interaction between more optimal times and in-the-moment sleepiness it appears that this effect is most prominently found when participants are both sleepy, and responding at a less optimal time. The consistent theme across all three models is that adverse sleep states, of various sorts, seem to increase perceptions of guilt in this study. Based on the failed priming manipulation this effect cannot be attributed to the stereotype-bias as originally conceived. However, it may be representative of a different heuristic approach to decision making (more on this in the Discussion).

Discussion

A diverse body of literature has shown that cognitive resource availability may be modulated by time of day and variations in diurnal (circadian) preference (Bodenhausen, 1990; Dickinson & McElroy, 2012; Gordon, 1997; Kruglanski and Pierro, 2008; Maire, et al., 2013; Pica, et al., 2014). It is also well established that when faced with ambiguous decisions people tend to rely on mental "shortcuts" or heuristics (Tversky & Kahneman, 1974), particularly when cognitive limitations are imposed (Bodenhausen, & Lichtenstein, 1987; Petty, et al., 1976; Rothbart, et al., 1978). Stereotyping is a biased decision making strategy that results from reliance on multiple heuristics (Bodenhausen, 2005). Research has shown that stereotyping may be more prevalent for individuals making decisions during times that are less optimal, based on their individual circadian preferences (Bodenhausen, 1990; Gordon, 1997). While these effects appear prominent, there are pertinent concerns with the methodology employed in prior studies (for details, refer to Present Study). Additionally, there is a significant gap in the literature surrounding circadian effects during the late evening and early morning hours. For these reasons, this study sought to provide a more robust methodology and to extend the findings of prior research to these more valenced hours.

Based on prior research, it was hypothesized that participants would rely more on stereotypes in the following instances: when they were more circadian mismatched, during the later hours of the day, and when they were sleepier. Unfortunately, without a successful induction of the stereotype primes, the results cannot be applied directly to these hypotheses. That said, there was evidence to suggest that Epworth, KSS, and circadian mismatch (via optimal time proxy) significantly predict perceptions of suspect guilt, when the presented

evidence is ambiguous or inconclusive. Based on literature linking circadian mismatch with biased decision-making, these findings may be indicative of participants employing other biases when assessing guilt.

Attempting to determine why the priming manipulation was unsuccessful, a number of theories emerge. While many studies have employed a similar mechanism to induce priming (Abraham, & Appiah, 2006; Chiao, Heck, Nakayama, & Ambady, 2006; Eberhardt, Goff, Purdie, & Davies, 2004) none of the studies examined presented the visual stimuli simultaneously with information (evidence) that was relevant to the decision being posed. A meta-analysis of masked priming induction by Van den Bussche, Van den Noortgate, and Reynvoet (2009) provides some key insights into factors that may have potentially detracted from the strength of the priming used. According to their findings, the literature supports that visual priming is strongest when the response format is likewise visual. That is to say, priming effects appear to be influenced by encoding specificity. This means that semantic primes are more likely to influence concept-based responding, while visual primes will be more effective for visual-recognition tasks (Schacter, Dobbins, & Schyner, 2004; Ziegler, Ferrand, Jacobs, Rey, & Grainger, 2000). This literature suggests that semantic concepts are often not evoked by visual primes unless participants are specifically asked to link the visual prime to the concept that it represents (for example, “what ethnic group did the individual pictured belong to?”). While the present study was not intended to utilize a masked priming manipulation, it may have effectively simulated one, by asking participants to rely on the (task-relevant) evidence, drawing attention away from the facial stimuli. Lending support to this argument, research has shown that presenting unmasked facial primes in conjunction

with other task-related information tends to produce inconsistent priming effects (Banse, 2001).

Another potential issue raised by the meta-analysis is with target repetition. Specifically, the literature suggests that when a concept is primed with multiple stimuli the prime will have a larger effect (Kaschak, Loney, & Borreggine, 2006; Van den Bussche, et al., 2009). In the present study, each priming stimulus was only presented one time during the survey, and was accompanied by the evidentiary items (which were likely perceived as being more task relevant). As such, it is possible that the prime was not induced very strongly, if at all.

Another explanation for the failure of the priming manipulation comes from a study by Sommers, and Ellsworth (2000). In their investigation, the authors found that White mock-jurors were more cognitively aware of defendant race than Black mock-jurors. When rendering verdicts, White mock-jurors were more aware of (and thereby were more likely to suppress) their racial prejudices than Black mock-jurors. In the investigation, race was a highly salient feature of the defendant, which the authors posit as an explanation of this phenomenon (Sommers & Ellsworth, 2000). It is possible that the recent increased social awareness of racial discrimination in the legal system (Bell, Funk, Joshi, & Valdivia, 2016) may have similarly increased the perceptual salience of race as a biasing factor during the decision task. However, overall the results of the suspicion check in the survey fail to lend support to this theory.

Despite the failure to induce a stereotype priming effect, it is evident that other factors may have influenced participants' perceptions of guilt. One theory is that participants may have been relying more on the availability heuristic when they were circadian

mismatched and/or sleepy. Research suggests that when participants are sleep deprived they tend to rely more on information that they have been previously exposed to than new (and potentially more relevant) information (Dickinson, Drummond, & Dyche, 2015). This theory is supported by the data suggesting that participants were more likely to assess guilt for crimes that they likely perceive to occur at higher base rates. For example, studies show that underage drinking prevalence rates are particularly high in college populations (Wechsler, et al., 2002; Wechsler, et al., 2003). Further evidence shows that college students tend to overestimate the amount of alcohol consumed by their peers (Perkins, Haines, & Rice, 2005), which may be linked to use of the availability heuristic. Specifically, the theory is that college students more easily recall examples of excessive drinking behavior in their peers than examples of moderation and sobriety (Carey, Borsari, Carey, & Maisto, 2006). Based on this information, participants who relied more on this perceived base rate for alcohol consumption in their peers would have assessed a higher likelihood of guilt for the underage drinking suspect than for a suspect involved in a crime which the participants had not had as much exposure to (such as domestic abuse or assault). It is worth noting that, while this framework accounts for the mean differences between alleged crimes, it was not directly tested by asking participants how prevalent they felt that each crime is.

Another potential explanation for higher perceived guilt when participants are mismatched and/or sleepy is that they are falling victim to confirmation bias. Specifically, the prompt for each stimulus states that the suspect has been "charged" with the alleged crime. This prompt may be creating an initial impression that the suspect is in fact guilty. It is also possible that participants hold a pre-existing belief (analogous to a stereotype) that people who are charged with crimes are more likely guilty than not. Research supports that

jurors are likely subject to pre-trial influences that lead to assumptions of guilt prior to the presentation of case evidence (Rassin, Eerland, & Kuijpers, 2010). On this basis, any additional pieces of information that might have suggested guilt (particularly when perceived prior to ambiguous evidence) might have pre-biased participants toward assumptions of guilt. Research suggests that participants who deliberately process uncertain information are less likely to perceive that information in accordance with their pre-existing beliefs (Hernandez & Preston, 2013). In other words, people who spend time and energy to process all of the information available are more likely to utilize a Bayesian framework and not simply default to the outcome that was initially implied. By this rationale, participants who responded during more optimal time frames, and/or were less sleepy may have been able to more carefully examine and process the presented evidence and discount it on the basis of ambiguity.

Limitations and Future Research

One limitation of the study is that there was a strong sampling bias in regards to the sleep behaviors of the participants. The literature suggests that the sleeping habits of college students, while typical of young adults, are poorly representative of the general population as a whole (Lund, et al., 2010). Therefore, the results may not generalize well to actual juror pools, who likely have a lower instance rate of chronic sleep issues. It is also worth noting that jurors are not tasked with rendering verdicts during the extreme late evening/early morning hours. Rather, this research may be more applicable to Emergency Service responders, particularly when those responders work long shifts, or alternate between day and night shifts frequently.

Perhaps the most pertinent limitation is that the posited explanations for biased responding are purely speculative. Without intentionally manipulating availability and pre-decisional attitudes, conclusions cannot be drawn. It is also worth mentioning that higher perceptions of probability of guilt do not necessarily equate to more biased responses. It is possible that individuals who rated the probability of suspect guilt as lower were relying in bias more than those who rated it as higher, or were simply utilizing a different set of biases. With that in mind, these insights are aimed at informing and providing direction for further research on sleep-related biases in decision-making.

While the stereotype priming failed, preventing the original hypotheses from being tested, the results of this study do provide some key insights into sleep-related factors that may be particularly influential toward biasing judgements and decisions during the more extreme hours of the day. The study has also provided evidence to support that when decisions are made at more optimal times (accounting for chronotypes) people are less likely to presume guilt. This finding may indirectly lend some support to the idea that people are more biased when circadian mismatched. As a whole, these findings open the door to several new outlets of research, and emphasizes the impact of adverse sleep states beyond simple time of day effects.

References

Abraham, L., & Appiah, O. (2006). Framing news stories: The role of visual imagery in priming racial stereotypes. *The Howard Journal of Communications, 17*, 183-203.

Adan, A., & Almirall, H. (1991). Horne & Östberg morningness-eveningness questionnaire: A reduced scale. *Personality and Individual Differences, 12*, 241-253.

Adan, A., Archer, S. N., Hidalgo, M., Di Milia, L., Natale, V., & Randler, C. (2012). Circadian typology: A comprehensive review. *Chronobiology International: The Journal of Biological & Medical Rhythm Research, 29*, 1153-1175.

Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience, 52*, 29-37.

Anderson-Cook, C. M. (2014). Unraveling Bayes' theorem. *Quality Progress, 47*, 52-54.

Aschoff, J. (1954). Zeitgeber der tierischen Tagesperiodik. *Naturwissenschaften, 41*, 49-56.

Aschoff, J., & Pohl, H., (1978). Phase relations between a circadian rhythm and its zeitgebers within the range of entrainment. *Naturwissenschaften, 65*, 80-84.

Ashmore, R. D., Del Boca, F. K., & Beebe, M. (2002). 'Alkie,' 'Frat Brother,' and 'Jock': Perceived types of college students and stereotypes about drinking. *Journal of Applied Social Psychology, 32*, 885-907.

Banse, R. (2001). Affective priming with liked and disliked persons: Prime visibility determines congruency and incongruency effects. *Cognition & Emotion, 15*, 501-520.

Bell, L. A., Funk, M. S., Joshi, K. Y., & Valdivia, M. (2016). Racism and White Privilege. *Teaching for Diversity and Social Justice*, 133-178.

Blair, I. V., & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology, 70*, 1142-1163.

Bodenhausen, G. V. (1990). Stereotypes as judgmental heuristics: Evidence of circadian variations in discrimination. *Psychological Science, 1,* 319-322.

Bodenhausen, G. V. (2005). The role of stereotypes in decision-making processes. *Medical Decision Making, 25,* 112-118.

Bodenhausen, G. V., & Lichtenstein, M. (1987). Social stereotypes and information-processing strategies: The impact of task complexity. *Journal of Personality and Social Psychology, 52,* 871-880.

Borbély, A. A. (1982). A two process model of sleep regulation. *Human Neurobiology, 1,* 195-204.

Brown, S. L., Coyne, S. M., Barlow, A., & Qualter, P. (2010). Alcohol-related image priming and aggression in adolescents aged 11–14. *Addictive Behaviors, 35,* 791-794.

Caci, H., Deschaux, O., Adan, A., & Natale, V. (2009). Comparing three morningness scales: age and gender effects, structure and cut-off criteria. *Sleep Medicine, 10,* 240-245.

Carey, K. B., Borsari, B., Carey, M. P., & Maisto, S. A. (2006). Patterns and importance of self-other differences in college drinking norms. *Psychology of Addictive Behaviors, 20,* 385-393.

Chiao, J. Y., Heck, H. E., Nakayama, K., & Ambady, N. (2006). Priming race in biracial observers affects visual search for Black and White faces. *Psychological Science, 17,* 387-392.

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology, 44,* 20-33.

Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology, 81,* 800-814.

Dickinson, D. L., & McElroy, T. (2010) Rationality around the clock: Sleep and time-of-day effects on guessing game responses. *Economics Letters, 108,* 245-248.

Dickinson, D. L., & McElroy, T. (2012) Circadian effects on strategic reasoning. *Experimental Economics, 15,* 444-459.

Dickinson, D. L., Drummond, S. P., & Dyche, J. (2015). Voluntary sleep choice and its effects on Bayesian decisions. *Behavioral Sleep Medicine*, 1-13.

Di Milia, L., Adan, A., Natale, V., & Randler, C. (2013). Reviewing the psychometric properties of contemporary circadian typology measures. *Chronobiology International, 30,* 1261-1271.

Dixon, T. L., & Maddox, K. B. (2005). Skin tone, crime news, and social reality judgments: Priming the stereotype of the dark and dangerous black criminal. *Journal of Applied Social Psychology, 35,* 1555-1570.

Eberhardt, J. L., Goff, P. A., Purdie, V. J., & Davies, P. G. (2004). Seeing black: race, crime, and visual processing. *Journal of Personality and Social Psychology, 87,* 876-893.

Erber, R., & Fiske, S. T. (1984). Outcome dependency and attention to inconsistent information. *Journal of Personality and Social Psychology, 47,* 709-726.

Fabbri, M., Mencarelli, C., Adan, A., & Natale, V. (2013). Time-of-day and circadian typology on memory retrieval. *Biological Rhythm Research, 44,* 125-142.

Festinger, L., & Maccoby, N. (1964). On resistance to persuasive communications. *The Journal of Abnormal and Social Psychology, 68,* 359-366.

Gavanski, I., & Roskos-Ewoldsen, D. R. (1991). Representativeness and conjoint probability. *Journal of Personality and Social Psychology, 61*, 181-194.

Gill, J. S. (2015). Permissibility of colour and racial profiling. *Western Journal of Legal Studies, 5*, 1-16.

Gordon, R. A. (1997). The moderation of distinctiveness-based illusory correlation: The impact of circadian variations and personal need for structure. *The Journal of Social Psychology, 137*, 514-526.

Hamby, S., & Jackson, A. (2010). Size does matter: The effects of gender on perceptions of dating violence. *Sex Roles, 63*, 324-331.

Hernandez, I., & Preston, J. L. (2013). Disfluency disrupts the confirmation bias. *Journal of Experimental Social Psychology, 49*, 178-182.

Horne, J. A., & Östberg, O. (1975). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology, 4*, 97-110.

Johns, M. W. (1991). A new method for measuring daytime sleepiness: The Epworth sleepiness scale. S*leep, 14*, 540-545.

Johns, M. W. (2000). Sensitivity and specificity of the multiple sleep latency test (MSLT), the maintenance of wakefulness test and the Epworth sleepiness scale: failure of the MSLT as a gold standard. *Journal of Sleep Research, 9*, 5-11.

Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., & Fukasawa, K. (2006). Validation of the Karolinska sleepiness scale against performance and EEG variables. *Clinical Neurophysiology, 117*, 1574-1581.

Kaschak, M. P., Loney, R. A., & Borreggine, K. L. (2006). Recent experience affects the strength of structural priming. *Cognition, 99*, B73-B82.

Kruglanski, A. W., & Pierro, A. (2008). Night and day, you are the one: On circadian mismatches and the transference effect in social perception. *Psychological Science, 19*, 296-301.

Larsen, R. J. (1985). Individual differences in circadian activity rhythm and personality. *Personality and Individual Differences, 6*, 305-311

Lack, L., Bailey, M., Lovato, N., & Wright, H. (2009). Chronotype differences in circadian rhythms of temperature, melatonin, and sleepiness as measured in a modified constant routine protocol. *Nature and Science of Sleep, 1*, 1-8.

Lund, H. G., Reider, B. D., Whiting, A. B., & Prichard, J. R. (2010). Sleep patterns and predictors of disturbed sleep in a large population of college students. *Journal of Adolescent Health, 46*, 124-132.

Lust, S. A. (2013). To sleep, perchance to decide: The effect of sleep deprivation on error reactivity. *Dissertation Abstracts International, 73*, 1-46.

Maire, M., Reichert, C. F., & Schmidt, C. (2013). Sleep-wake rhythms and cognition. *Journal of Cognitive and Behavioral Psychotherapies, 13*, 133-170.

McElroy, T., & Dickinson, D. L. (2010). Thoughtful days and valenced nights: How much will you think about the problem? *Judgment and Decision Making, 5*, 516-523.

McMenamin, T. M. (2007). Time to work: recent trends in shift work and flexible schedules. *Monthly Labor Review, 130*, 3-15.

Natale, V., Alzani, A., & Cicogna, P. (2003). Cognitive efficiency and circadian typologies: A diurnal study. *Personality and Individual Differences, 35*, 1089-1105.

Neubauer, A. C. (1992). Psychometric comparison of two circadian rhythm questionnaires and their relationship with personality. *Personality and Individual Differences, 13,* 125-131.

Neuberg, S. L., & Fiske, S. T. (1987). Motivational influences on impression formation: Outcome dependency, accuracy-driven attention, and individuating processes. *Journal of Personality and Social Psychology*, *53,* 431.

Perkins, H. W., Haines, M. P., & Rice, R. (2005). Misperceiving the college drinking norm and related problems: a nationwide study of exposure to prevention information, perceived norms and student alcohol misuse. *Journal of Studies on Alcohol*, *66,* 470-478.

Petty, R. E., Wells, G. L., & Brock, T. C. (1976). Distraction can enhance or reduce yielding to propaganda: Thought disruption versus effort justification. *Journal of Personality and Social Psychology*, *34*, 874.

Pica, G., Pierro, A., & Kruglanski, A. W. (2014). Effect of circadian rhythms on retrieval-induced forgetting. *Cognitive Processing, 15*, 29-38.

Rassin, E., Eerland, A., & Kuijpers, I. (2010). Let's find the evidence: An analogue study of confirmation bias in criminal investigations. *Journal of Investigative Psychology and Offender Profiling, 7,* 231-246.

Regestein, Q., Natarajan, V., Pavlova, M., Kawasaki, S., Gleason, R., & Koff, E. (2010). Sleep debt and depression in female college students. *Psychiatry Research*, *176,* 34-39.

Rothbart, M., Fulero, S., Jensen, C., Howard, J., & Birrell, P. (1978). From individual to group impressions: Availability heuristic in stereotype formation. *Journal of Experimental Social Psychology, 14*, 237-255.

Sankar, G. (2014). Discriminatory policy among the undergraduate students towards racism and white privilege in America. *Educational Research and Reviews*, *9*, 743-749.

Schacter, D. L., Dobbins, I. G., & Schnyer, D. M. (2004). Specificity of priming: A cognitive neuroscience perspective. *Nature Reviews Neuroscience, 5*, 853-862.

Sherman, J. W., Kruschke, J. K., Sherman, S. J., Percy, E. J., Petrocelli, J. V., & Conrey, F. R. (2009). Attentional processes in stereotype formation: A common model for category accentuation and illusory correlation. *Journal of Personality and Social Psychology, 96*, 305-323.

Smith, C. S., Folkard, S., Schmieder, R. A., Parra, L. F., Spelten, E., Almiral, H., & ... Tisak, J. (2002). Investigation of morning-evening orientation in six countries using the preferences scale. *Personality and Individual Differences, 32*, 949-968.

Sommers, S. R., & Ellsworth, P. C. (2000). Race in the courtroom: Perceptions of guilt and dispositional attributions. *Personality and Social Psychology Bulletin, 26*, 1367-1379.

Tsai, L. L., & Li, S. P. (2004). Sleep patterns in college students: Gender and grade differences. *Journal of Psychosomatic Research, 56*, 231-237.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124-1131.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*, 453-458.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90,* 293-315.

U.S. Department of Justice, Civil Rights Division. (2015). Investigation of the Ferguson Police Department. Retrieved from http://www.justice.gov/sites/default/files/opa/press-releases/attachments/2015/03/04/ferguson_police_department_report.pdf

US v. Windsor, 133 S. Ct. 2675, 570 U.S. 12, 186 L. Ed. 2d 808 (2013).

Van den Bussche, E., Van den Noortgate, W., & Reynvoet, B. (2009). Mechanisms of masked priming: a meta-analysis. *Psychological Bulletin,135,* 452-477.

Van Dongen, H. P., & Dinges, D. F. (2005). Sleep, circadian rhythms, and psychomotor vigilance. *Clinics in Sports Medicine, 24,* 237-249.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology, 12,* 129-140.

Wechsler, H., Lee, J. E., Nelson, T. F., & Kuo, M. (2002). Underage college students' drinking behavior, access to alcohol, and the influence of deterrence policies: Findings from the Harvard School of Public Health College Alcohol Study. *Journal of American College Health, 50,* 223-236.

Wechsler, H., Nelson, T. E., Lee, J. E., Seibring, M., Lewis, C., & Keeling, R. P. (2003). Perception and reality: a national evaluation of social norms marketing interventions to reduce college students' heavy alcohol use. *Journal of Studies on Alcohol, 64,* 484-494.

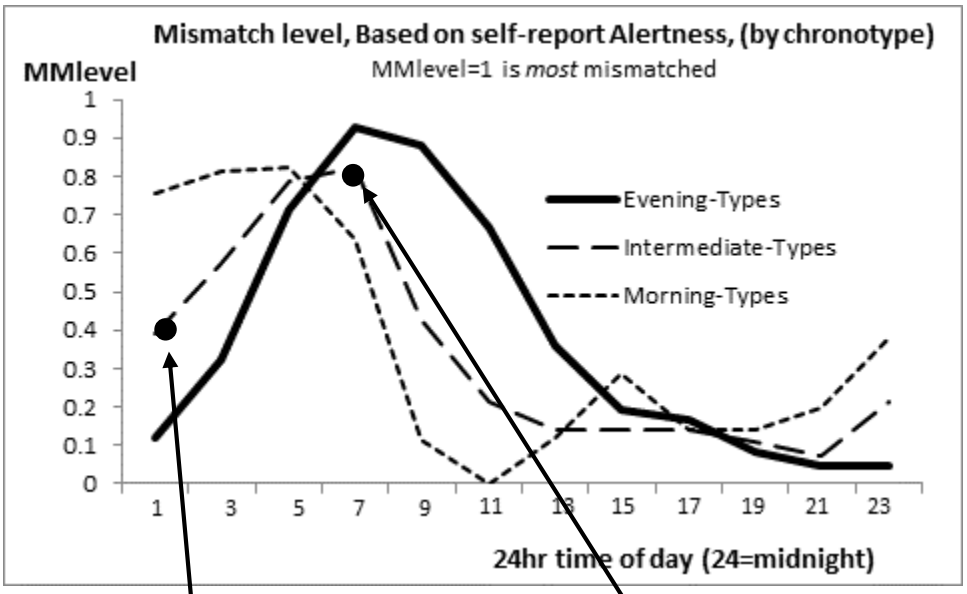
Welch, K., Payne, A. A., Chiricos, T., & Gertz, M. (2011). The typification of Hispanics as criminals and support for punitive crime control policies. *Social Science Research, 40,* 822-840.

Ziegler, J. C., Ferrand, L., Jacobs, A. M., Rey, A., & Grainger, J. (2000). Visual and phonological codes in letter and word recognition: Evidence from incremental priming. *The Quarterly Journal of Experimental Psychology A, 53*, 671– 692.

Appendix A

Self-Reported Alertness Data by Chronotype from Smith et al. (2002)

(scores normalized to a [0,1] scale, with Mismatch level (MMlevel) equal to zero for highest alertness ratings and MMlevel=1 for lowest alertness ratings in Smith et al data)



1:00 AM: *More Optimal Time*=1 for an Intermediate Type

7:00 AM: *More Optimal Time*=0 for an Intermediate Type

Appendix B

Example of a Stereotype-Consistent Stimulus

The individual pictured has been charged with consuming alcohol, while under the legal drinking age. Please carefully look over the information below, then continue.

- The individual is a student athlete.

- The individual has not previously been charged with underage drinking.

- The individual was attending a party, where alcoholic beverages were served.

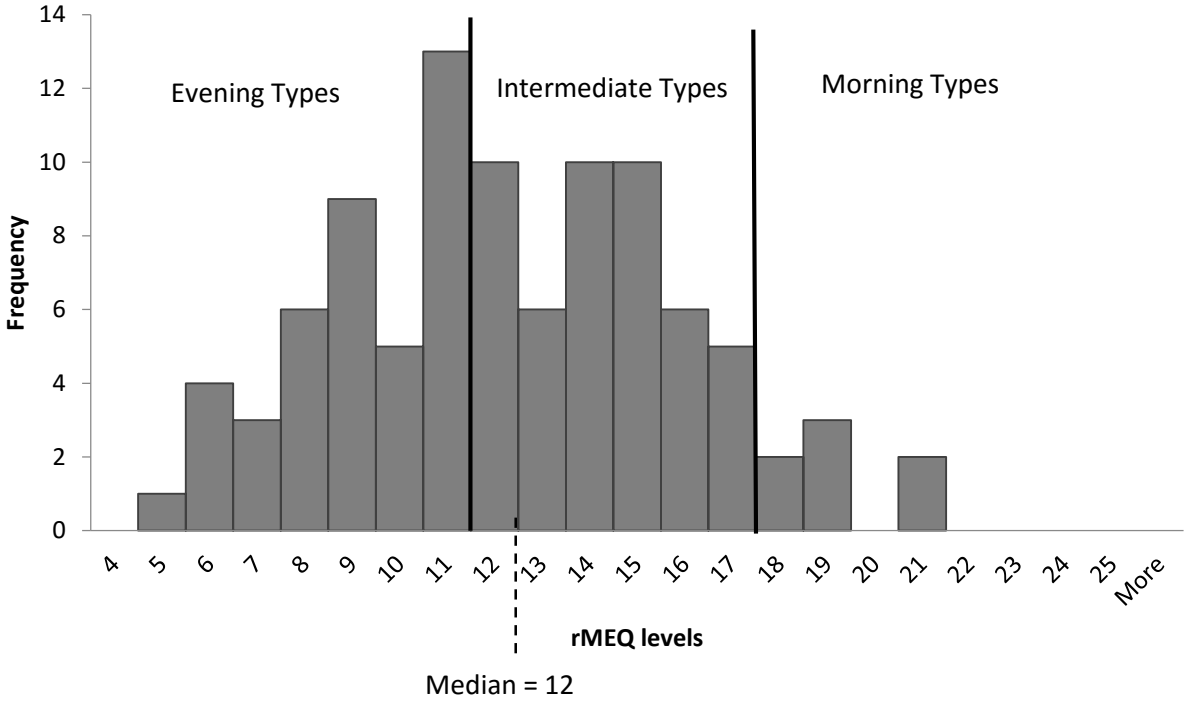
- The individual appeared to be impaired, but blames the impairment on not getting enough sleep.

Appendix C

Corresponding Participant Time Frames

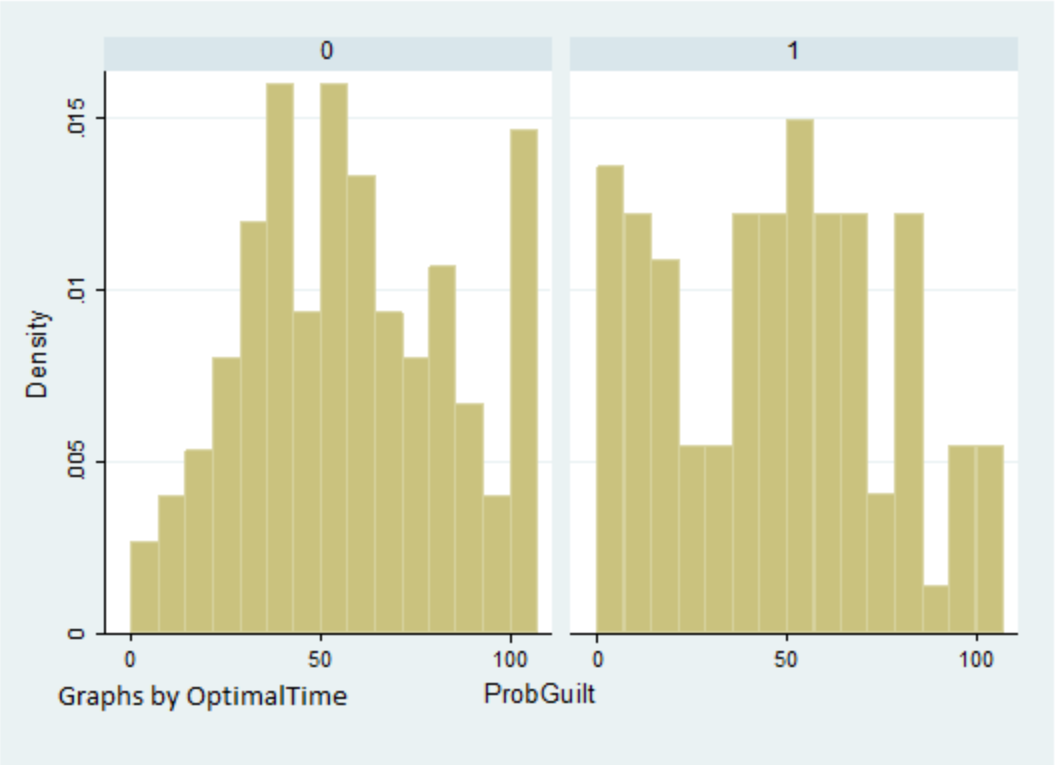
Week 1 Timeframe	Week 2 Timeframe
22:00-24:00	4:00-6:00
24:00-2:00	6:00-8:00
2:00-4:00	8:00-10:00
4:00-6:00	22:00-24:00
6:00-8:00	24:00-2:00
8:00-10:00	2:00-4:00

Appendix D
Distribution of rMEQ Scores with Cutoffs



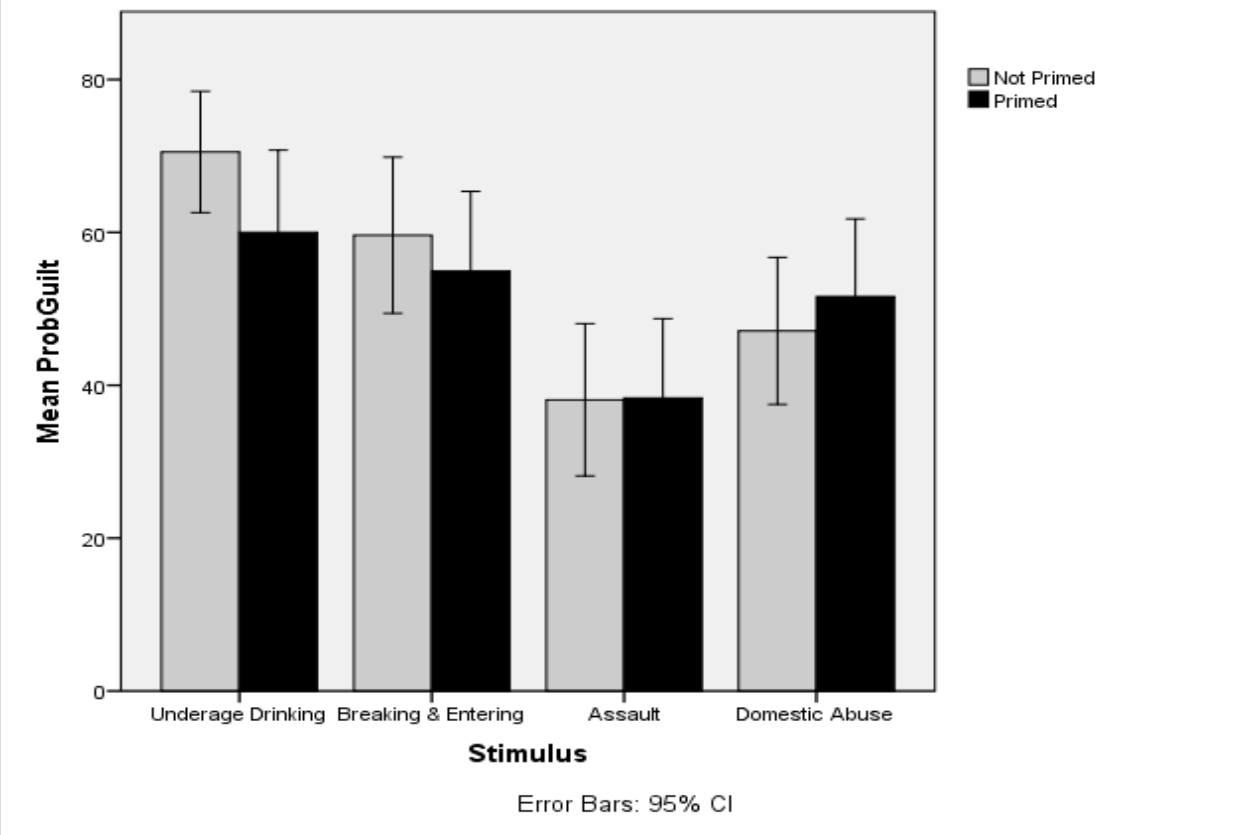
Appendix E

Distribution of ProbGuilt scores for More and Less Optimal



Appendix F

Graph Depicting Mean Differences in Perceived Guilt



Appendix G

Correlation Matrix for Sleep Measures

	KSS	Sleep Deprive	Optimal Time = 1	Epworth	Avg. Last Week	Sleep Last Night	Since Last Slept
KSS	1						
Sleep Deprive	.24*	1					
Optimal Time = 1	-.33*	-.01	1				
Epworth	-.08	.07	.10	1			
Avg. Last Week	.04	-.32*	.10	-.13*	1		
Sleep Last Night	-.19*	.00	.26**	-.04	.53**	1	
Since Last Slept	-.12*	.03	.46**	.02	.00	.08	1

* *Significant at p < .05, two-tailed*

** *Significant at p < .001, two-tailed*

Appendix H

Model Estimates for Multiple Regressions of Probability of Guilt Assessments.

The regression models estimated include clustered error terms by subject, to account for the multiple observations for each subject.

(estimations performed using STATA software panel data command options)

Dependent Variable=*Probability of Guilt*

Variable	**Model 1 Coefficient (SE)**	**Model 2 Coefficient (SE)**	**Model 3 Coefficient (SE)**
Constant	41.36 (13.02) **	18.38 (19.45)	-24.52 (29.33)
Prime (Primed = 1)	-3.69 (3.35)	-3.66 (3.38)	-3.67 (3.43)
Sex (Female = 1)	-.67 (6.34)	-1.11 (6.40)	-3.63 (6.29)
Minority (Minority = 1)	2.52 (5.56)	2.59 (5.86)	.41 (5.63)
KSS	1.00 (1.28)	6.23 (5.88)	17.20 (7.73) **
Crime Severity	.05 (.08)	-.04 (.08)	-.04 (.08)
Epworth (ESS)	1.70 (.70) **	4.43 (2.59) **	4.29 (2.99) *
Sleep Deprivation	.18 (.56)	.29 (1.51)	.37 (1.52)
Evening Type (Yes = 1)	.91 (4.94)	.27 (4.94)	6.54 (8.39)
Morning Type (Yes = 1)	3.78 (11.45)	3.82 (10.90)	-.26 (10.76)
More Optimal Time (Yes = 1)	-8.87 (4.93) **	-8.55 (4.98) **	25.25 (22.85)
KSS^2		-.45 (.54)	-1.04 (.58) **
$Epworth^2$		-.16 (.15)	-.21 (.17)
Sleep $Deprivation^2$		-.02 (.07)	-.07 (.09)
More Optimal Time*KSS			-5.94 (3.23) **
More Optimal Time*Epworth			1.06 (1.28)
More Optimal Time*Sleep Deprivation			.82 (1.71)
More Optimal Time * Morning Type			23.16 (19.33)
More Optimal Time * Evening Type			-7.01 (10.01)
Model Pearson's r	.30	.32	.37
Model r^2	.09	.10	.14
Model χ^2 *significance (p)*	.004	<.001	<.001

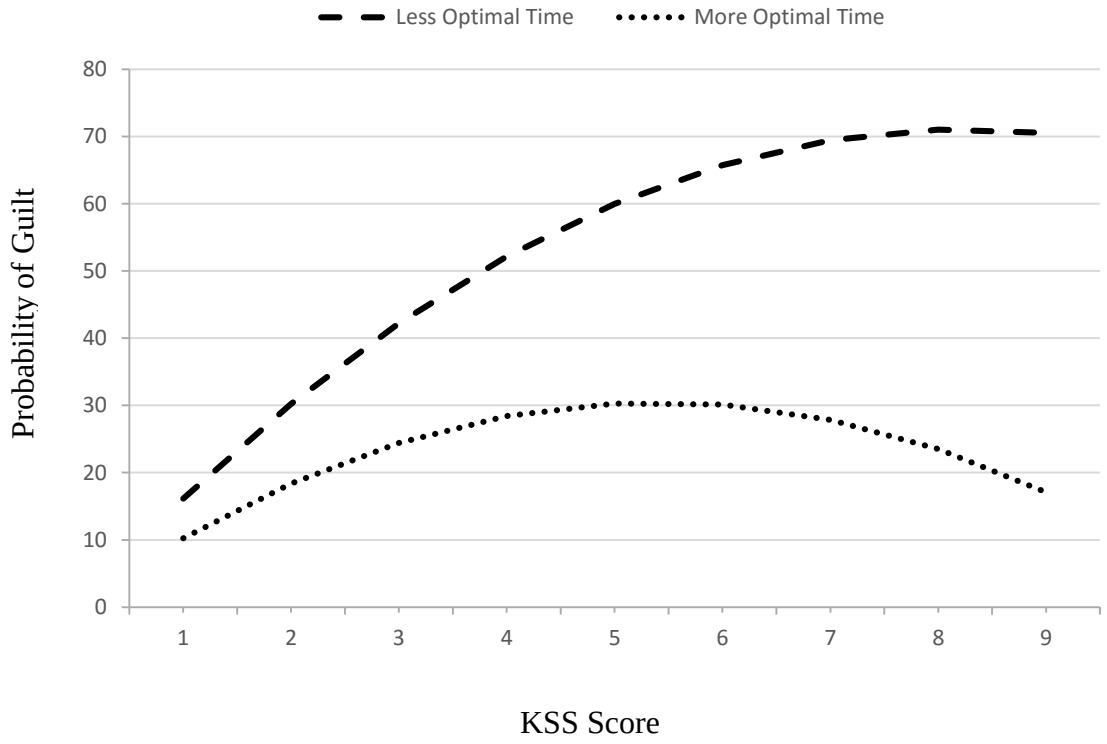
**marginally significant at p < .10, one-tailed*

*** significant at p < .05, one-tailed*

Appendix I

Graph Depicting the Interaction Effect Between Optimal Time and KSS scores on Probability of Guilt Rating

Graph based on significant coefficient estimates from Model 3 of Appendix H. Forecast shows Probability of Guilt ratings for participants at varying reported levels of in-the-moment sleepiness, separated into more and less optimal time frames, while holding all other variables constant at zero.



Appendix J

Participant Informed Consent Agreement

Informed Consent:

This study will be used to collect data, as a part of a Master's thesis research project. This investigation is designed to look at the effects of time of day and sleep on decision making. Many college students have non-traditional sleep schedules and research into this area may lead to a better understanding of decisions that they make during certain hours of the day.

The data collected will be stored completely anonymously. The survey settings will be configured to delete any individuating information collected (such as IP Address) and any other personal information will be removed from the data, and subsequently deleted once your participation is compensated.

Participation should take anywhere from 15-30 minutes and is completely voluntary. **You may elect to discontinue participation at any point.** Refusal to participate or discontinuation will not result in any penalty, nor will it incur the loss of any benefit to which you are otherwise entitled.

Society as a whole will benefit from this study in terms of new knowledge gained regarding how time of day influences decision making.

You may be asked to complete the survey at a time in which you would normally be asleep or resting. This could potentially result in you not sleeping for as long as you traditionally would on the days that you participate in this study.

Failure to complete the survey during the designated time-frame may constitute a failure to comply with the study protocol. If this is problematic you may elect to discontinue.

You will not be paid for your participation in this study. However, you can earn [up to] 2 ELC credits for your participation. If you complete this study both weeks, you will also be entered into a random prize drawing for a chance to win $50. There will be multiple winners and all participants who complete both weeks will have an equal chance of winning.

There are other research options and non-research options for obtaining extra credit or ELC's. One non-research option to receive 1 ELC is to read an article and write a 1-2 page paper summarizing an article and your reaction to the article. More information about this option can be found at: psych.appstate.edu/research. You may also wish to consult your professor to see if other non-research options are available.

For additional information, questions, or complaints related to this questionnaire, contact Dr. David Dickinson, Department of Economics, Appalachian State University, at dickinsondl@appstate.edu or Robert McClelland, Department of Psychology, Appalachian State University, at mcclellandra@appstate.edu.

This research project has been approved on November 6, 2015 by the Institutional Review Board (IRB) at Appalachian State University. This approval will expire on November 5, 2016 unless the IRB renews the approval of this research.

For additional information, you may contact the IRB at (828) 262-2692 or email at irb@appstate.edu

- ○ I understand and agree to participate (and I am at least 18 years old)
- ○ I do not agree to participate and do not wish to continue

Appendix K

Appalachian State Institutional Review Board Study Approval

Appalachian
STATE UNIVERSITY

INSTITUTIONAL REVIEW BOARD
Office of Research Protections
ASU Box 32068
Boone, NC 28608
828.262.2692
Web site: http://researchprotections.appstate.edu/
Email: irb@appstate.edu

To: Robert McClelland
Psychology
EMAIL

From: Dr. Lisa Curtin, Institutional Review Board Chairperson
Date: November 6, 2015
RE: Notice of IRB Approval by Expedited Review (under 45 CFR 46.110)
Study #: 16-0080
Study Title: Circadian Influences on Stereotype Reliance
Submission Type: initial
Expedited Category: 7. Research on Group Characteristics or Behavior, or Surveys, Interviews, etc.

Approval Date: November 6, 2015
Expiration Date of Approval: November 5, 2016

The Institutional Review Board (IRB) approved this study for the period indicated above. The IRB found that the research procedures meet the expedited category cited above. IRB approval is limited to the activities described in the IRB approved materials, and extends to the performance of the described activities in the sites identified in the IRB application. In accordance with this approval, IRB findings and approval conditions for the conduct of this research are listed below.

The IRB determined that this study involves minimal risk to participants.

Approval Conditions:

Appalachian State University Policies: All individuals engaged in research with human participants are responsible for compliance with the University policies and procedures, and IRB determinations.

Principal Investigator Responsibilities: The PI should review the IRB's list of PI responsibilities. The Principal Investigator (PI), or Faculty Advisor if the PI is a student, is ultimately responsible for ensuring the protection of research participants; conducting sound ethical research that complies with federal regulations, University policy and procedures; and maintaining study records.

Modifications and Addendums: IRB approval must be sought and obtained for any proposed modification or addendum (e.g., a change in procedure, personnel, study location, study instruments) to the IRB approved protocol, and informed consent form before changes may be implemented, unless changes are necessary to eliminate apparent immediate hazards to participants. Changes to eliminate apparent immediate hazards must be reported promptly to the IRB.

Approval Expiration and Continuing Review: The PI is responsible for requesting continuing review in a timely manner and receiving continuing approval for the duration of the research with human participants. Lapses in approval should be avoided to protect the welfare of enrolled participants. If approval expires, all research activities with human participants must cease.

Prompt Reporting of Events: Unanticipated Problems involving risks to participants or others; serious or continuing noncompliance with IRB requirements and determinations; and suspension or termination of IRB approval by an external entity, must be promptly reported to the IRB.

Closing a study: When research procedures with human subjects are completed, please complete the Request for Closure of IRB review form and send it to irb@appstate.edu.

Vita

Robert Alexander Indiana McClelland was born in Asheville, North Carolina, to James and Terri McClelland. He graduated from Buncombe County Early College in May of 2012. The following August, he began studies at Appalachian State University and, in May of 2014, was awarded a Bachelor of Science in Psychology. In the fall of 2014, Mr. McClelland accepted a research assistantship in Psychology at Appalachian State University and began work on a Master of Arts in Experimental Psychology. He is scheduled to complete his M.A. in May of 2016.

Mr. McClelland is a member of Phi Theta Kappa and Psi Chi. He enjoys reading, hiking, and travel in his free time.