Summer 8-2018

The College Registration Executive Function Task: Ecological Validity and Relationship to Alcohol Use in University Students

Matthew A. Taylor

THE COLLEGE REGISTRATION EXECUTIVE FUNCTION TASK: ECOLOGICAL VALIDITY AND RELATIONSHIP TO ALCOHOL USE IN UNIVERSITY STUDENTS

A Dissertation

Submitted to the School of Graduate Studies and Research

in Partial Fulfillment of the

Requirements for the Degree

Doctor of Psychology

Matthew A. Taylor

Indiana University of Pennsylvania

August 2018

Indiana University of Pennsylvania
School of Graduate Studies and Research
Department of Psychology

We hereby approve the dissertation of

Matthew A. Taylor

Candidate for the degree of Doctor of Psychology

_____ _____
William Meil, Ph.D.
Professor of Psychology, Advisor

_____ _____
David LaPorte, PhD.
Professor of Psychology

_____ _____
Margaret C. Reardon, Ph.D.
Associate Professor of Psychology

ACCEPTED

_____ _____
Randy L. Martin, Ph.D.
Dean
School of Graduate Studies and Research

Title: The College Registration Executive Function Task: Ecological Validity and Relationship to Alcohol Use in University Students

Author: Matthew A. Taylor

Dissertation Chair: Dr. William Meil

Dissertation Committee Members: Dr. David LaPorte
Dr. Margaret C. Reardon

Alcohol use is a major problem in college populations. Executive functioning (EF) problems are predictive of and predicted by alcohol-related problems; however, traditional performance-based measures of executive function (EF) may not accurately reflect individuals' every day functioning. The purpose of this study was to develop an ecologically valid EF task specifically for college students, The College Registration Executive Function Task (CREFT), and to assess its relationship to existing EF measures, as well as its ability to predict alcohol-use and alcohol-related outcomes. CREFT performance was associated with performance on the Tower of London, as well as subscales on self-report measures, including UPPS-P Positive Urgency and BRIEF-A Organization of Materials. Hierarchical regressions indicated that CREFT performance was significantly predictive of binge drinking frequency and added significant predictive value to a model including demographic information, traditional performance-based tasks, and self-report measures such as the BRIEF-A and UPPS-P. Sensation seeking behavior emerged as a significant predictor of each alcohol outcome, including typical drinking rates, binge drinking frequency, alcohol-related problems, and symptoms of alcohol use disorder.

ACKNOWLEDGEMENTS

First, although I'm not sure words can adequately express the gratitude I feel for her, I would like to thank my wife, Amanda Taylor. Amanda has moved with me three times (with a fourth upcoming), supported us financially while I made negative amounts of money, did the bulk of work at home while I was overwhelmed with school work, and patiently endured my periods of graduate school-induced stress and distractedness. Most importantly, however, she has stood by my side through this entire process without complaint. Her selfless, unwavering support, love, and encouragement helped make it possible to earn my doctorate.

Next, I want to acknowledge my mother, Nancy Taylor, for setting the stage for my academic success, encouraging me to pursue graduate school, and showing me that if a single mother working full-time can earn her doctorate, then so can I. Our trips to the library during our "M&M Days" kindled my love of learning, which is certainly necessary if one is to spend over twenty years in school. Because Mom is intimately familiar with this process, I knew that I could count on her to understand and empathize with what I was going through. Her support was always just a phone call away.

The truth is, a degree may be the second most important acquisition from my time at IUP. Perhaps even more significant is the gaining of an incredible group of friends who are better described as family. I want to thank my cohort family for making my time in this program a truly meaningful experience. During our four years in Indiana, we spent many nights at Brunzies, watched the Penguins win two Stanley Cups, held several costume and murder mystery parties, vacationed together in Puerto Rico and Florida, created countless memories, and forged an incredibly tight, lifelong bond. I would also specifically like to thank the first of these friends,

Brian Ferraccio, for sitting at the Uhler hallway table on the first day of class, introducing himself, and quickly becoming one of the closest friends I have ever had.

Finally, thank you to Dr. Meil for capably guiding me through the long, challenging process of writing a dissertation. As an advisor, he was unfailingly knowledgeable, responsive, encouraging, and supportive. I felt confident that I was in good hands each step of the way. In addition, I greatly appreciate all of his help as my teaching mentor and go-to faculty member.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

CHAPTER I

LITERATURE REVIEW

Overview of Executive Functioning

Executive functioning, which has been defined as "those capacities that enable a person to engage successfully in independent, purposive, self-directed, and self-serving behavior" (Lezak, Howieson, Bigler, & Tranel, 2012), is crucial to peoples' abilities to function in everyday life, as greater decrements in executive functions are related to poorer ability to live independently (Hanks, Rapport, Millis, & Deshpande, 1999). Furthermore, individuals with impaired executive functioning have trouble with regard to occupational, academic, and social/interpersonal functioning (Barkley & Murphy, 2010; Diamontopoulou, Rydell, Thorell, & Bohlin, 2007). There are many neuropsychiatric disorders associated with executive functioning difficulties, including schizophrenia (Hutton et al., 1998), substance use (Lundqvist, 2005; Tartar et al., 2003) depression (Moritz et al., 2002), attention-deficit/hyperactivity disorder (Happé, Booth, Charlton, & Hughes, 2006), fetal alcohol syndrome (Rasmussen, 2005), autism spectrum disorder (Happé et al., 2006), and obsessive-compulsive disorder (Moritz et al., 2002), among others.

The construct of executive function is similar to intelligence in that there is no universally agreed-upon definition. There is disagreement as to which cognitive abilities should be included in the concept as well as whether executive function is a unitary trait that can be conceptualized similarly to the idea of "g" in intelligence theory. In addition to the definition above, executive function has also been described as "a set of general-purpose control mechanisms, often linked to the prefrontal cortex of the brain, that regulate the dynamics of human cognition and action" (Miyake & Friedman, 2012). These and other definitions share the notion that executive

functions regulate or control behavior. For this reason, the functions included in the construct are conceptualized as "higher" processes that direct "lower" cognitive processes (Alvarez & Emory, 2006). Specific executive functions often discussed include initiation, planning, inhibition, organization, shifting, and self-monitoring, working memory, and sustained attention (Alvarez & Emory, 2006). Additionally, insight may be considered an aspect of executive function in that knowledge of one's own cognitive abilities is critical to the ability to engage in goal directed behavior (Robinson, Calama, Glascher, Bruss, & Tranel, 2014). In addition to these "metacognitive" executive functions, some have argued that there is also an emotional/motivational aspect of executive function involving the coordination of cognitive processes and emotions (Ardila, 2008).

Rather than just being conceptualized as possible deficits that can occur as a result of brain impairment, executive functions have also been linked to personality traits that vary in normal populations. For example, a factor analysis identified impulsivity as being associated with four NEO-PI-R personality facets (Costa & McCrae, 1992; Whiteside & Lynam, 2001). Other associations have been found between Five Factor Model personality traits and executive functioning. In a sample of older adults, executive functioning was positively associated with Neuroticism and negatively associated with Openness to Experience as well as Agreeableness (Williams, Suchy, & Kraybill, 2010). Additionally, Luu, Collins, and Tucker (2000) found higher Neuroticism correlated with worse performance on measures of response selection/inhibition. Openness to Experience, as measured by the NEO-PI-R, was associated with overall performance on a number of neuropsychological tasks purported to assess functioning of the prefrontal cortex (DeYoung, Peterson, & Higgis, 2005). Further evidence supporting the perspective that executive functions represent traits that vary normally include behavioral

genetics research indicating that executive functioning is very heritable in both children (Engelhardt, Briley, Mann, Harden, & Tucker-Drob, 2015) and young adults (Friedman et al., 2008). Engelhardt et al. (2015) found that an exclusively genetic factor mediated 100% of the variance common to all four executive function domains that they investigated (i.e., inhibition, switching, updating, working memory).

Researchers have proposed models of executive function that place greater emphasis on certain aspects of executive function and less on others. For example, Miyake, Friedman, Emerson, Witzki, and Howerter (2000) have focused on three separate "executive functions," including shifting between tasks or mental sets (shifting), Updating and monitoring of working memory representations (updating), and Inhibition of prepotent responses (inhibition). The first of these, shifting refers to one's ability to shift between multiple tasks, operations, or mental sets (Monsell, 1996). This ability likely involves disengaging from non-relevant task sets and overcoming proactive interference or negative priming in order to perform a new operation (Miyake et al., 2000). For example, one neuropsychological test that is often thought to tap into shifting ability is the Wisconsin Card Sort Test (WCST; Berg, 1948; Grant & Berg, 1948), which requires examinees to sort cards with designs that vary in shape, color, and number into the correct category (color, shape, or number). After a certain number of correct sorts are made, the rule changes, at which point the examinee must disengage from the previous, now-irrelevant sorting rule and discover/maintain the new sorting rule. An individual who continues sorting according to the previous rule is said to be "perseverating," - having difficulty switching mental sets. The second executive function, updating, involves keeping track of incoming information for relevance to the current task and then replacing irrelevant information with newer, more relevant information (Morris & Jones, 1990). A task that requires this updating function is the

tone-monitoring task used by Miyake et al. (2000), which involves listening to a mix of tones of different pitches (low, medium, and high) and responding upon hearing the fourth instance of any tone. The third executive function, inhibition, refers to the ability to suppress dominant/automatic behaviors. This ability is measured using tests such as the Stroop test, in which individuals must inhibit a natural tendency to read a color word and instead name the color of the ink in which the word is printed.

Focusing on the three abilities, Miyake et al. (2000) used confirmatory factor analysis to parse out how these three abilities relate to one another. They found that the three abilities were moderately correlated with each other but also represent clearly separable constructs. Specifically, factor analyses yielded a general EF factor as well as shifting-specific and updating-specific factors. They later concluded that individual differences in executive functions show both unity and diversity in that they represent diverse constructs while still having some underlying commonalities (Miyake & Friedman, 2012). The authors believe that the common EF factor is about one's ability to actively maintain task goals and goal-related information and use that information in order to effectively bias lower-level processing. Shifting-specific factor reflects flexibility, ease of transitioning to new task-set representations, while the authors are less sure about what updating-specific factor taps into. This model has found no evidence for an inhibition-specific factor once general EF is accounted for (Miyake & Friedman, 2012). Of note, a factor analysis of the Delis-Kaplan Executive Function System (D-KEFS; Delis, Kaplan, & Kramer, 2001), a battery of tasks designed to assess key components of executive function within verbal and spatial modalities in children and adults, also yielded similar findings. The test appears to measure conceptual flexibility, monitoring, and inhibition (Latzman and Markon, 2010, Meil et al, 2016).

The research described above indicates that executive function is a construct comprised of a number of cognitive processes, and it is not always clear which should be included. Factor analyses have suggested that inhibition, switching, and updating/monitoring may be three main components of executive functioning. While Miyake's conceptualization of EF is relatively popular, it still has problems. In particular, the factor analyses included few EF measures, and an important aspect of EF, planning, was not included due to the difficulty of operationally defining it.

Executive Functions and the Frontal Lobe

The executive functions have traditionally been associated with the frontal lobes of the brain, specifically the prefrontal cortex (PFC). As Fuster (2001) noted, higher brain areas support functions that are progressively more integrative, and the PFC is the final part of the brain to develop both evolutionarily and ontologically. Fuster (2001) provided an overview of the prefrontal cortex, noting that its general job consists of the "memory, planning, or execution of actions." The PFC is connected to many other brain areas, including the thalamus, basal ganglia, brainstem, and limbic system, making it well-suited to process and integrate sensory, motor, and limbic information.

The PFC is a large area of the brain that houses many important functions. A variety of research approaches including lesion studies, structural imaging, functional imaging, and neuropsychological testing have been used to help to inform the localization of these different functions to the PFC and to specific areas of the PFC. It should be noted that abilities often overlap across areas, and different abilities often activate the same brain regions (Cabeza & Nyberg, 2000; Cabeza, Dolcos, Graham, & Nyberg, 2002).

There are three major regions of the PFC outlined by Fuster (2001), including the orbital (often referred to as "orbitofrontal"), medial, and lateral. The orbital and medial regions are generally involved in emotional behavior. For example, patients with orbital lesions characteristically demonstrate symptoms related to dramatic personality changes such as disinhibtion (including social and moral), impulsivity (leading to high risk behavior), and irritability. Another major function of the PFC, given its connections to the thalamus, which is often viewed as a "sensory relay station" (McCormick & Bal, 1994), is to filter out irrelevant information. The medial region, consisting of the most anterior portion of the cingulate cortex, is involved in attention and emotion. Medial frontal lesions may lead individuals with a loss of spontaneity, sense of apathy, inability to concentrate attention/sustain effort, and a disinterest in the environment. However, the lateral portion of the PFC is the most relevant to the discussion of executive functions in that it "provides cognitive support to temporal organization of behavior, speech, and reasoning" (Fuster, 2001). This temporal organization/integration is crucial for planning and carrying out behaviors, and is critical in novel and complex behaviors. As such, common consequences of lesions to the lateral PFC include cognitive deficits such as inability to formulate and carry out plans and sequences of actions and difficulty to initiate and execute behaviors in an orderly manner. This set of symptoms has been referred to as "dysexecutive syndrome" (Baddeley & Wilson, 1988).

An alternate model of the role of the PFC in executive function, proposed by Stuss and Alexander (2007), posits that the three main types of executive function are initiating/sustaining a response, task setting, and monitoring (as opposed to Miyake et al.'s updating and shifting factors). Basing their theory on neuropsychological test performance of a group of patients with frontal lobe lesions, they view initiation as relying on medial frontal regions of the brain, task

setting as relying on left lateral regions, and monitoring and adjusting task performance involving the right lateral regions. Meanwhile, others have argued that executive function mechanisms are ordered hierarchically, with anterior regions using internally generated information to guide behavior and posterior areas using external information (Christoff & Gabrieli, 2000).

Banich (2009) describes executive functions as comprising a "temporal cascade" of selection processes that occur at distinct areas of the PFC. According to this theory, posterior areas of the dorsolateral PFC are involved in attentional set toward task-relevant processes, the mid-dorsolateral PFC selects among specific representations that are identified as task-relevant. Moving forward, the anterior cingulate is involved in late-stage aspects of selection and evaluation of responses. Activity in the ACC is affected by how well executive control was applied in previous areas.

Metacognitive and Emotional/Motivational Executive Functions

According to Ardila (2008), the prefrontal cortex participates in two broad types of executive function abilities that are separate but closely related. Metacognitive executive functions refer to those abilities that are assessed using traditional neuropsychological measures, including problem solving, abstracting, planning, working memory, etc. Metacognitive executive functions are generally associated with the dorsolateral prefrontal cortex. The second broad type of executive function involves coordinating cognitive processes and emotions – these are the emotional/motivational executive functions. Emotional/motivational executive functions are largely associated with the ventromedial prefrontal cortex and help us to fulfill primitive impulses (e.g., hunger, sex, aggression) using socially acceptable strategies (one may be reminded of Freud's concept of the "ego."). Problems with this ability are seen in brain-injured

individuals who retain formal knowledge of social norms but act in ways that are in contradiction to that knowledge. They are unable to apply the knowledge in real situations.

The idea of the ventromedial portion of the prefrontal cortex being primarily involved in motivational/emotional salience is also reflected in the somatic marker hypothesis (Damasio, 1996). The somatic marker hypothesis was derived from the observations mentioned above that individuals with ventromedial damage often demonstrated intact intelligence and performance on neuropsychological testing but exhibited poor decision-making as well as emotional abnormalities. According to this theory, the ventromedial prefrontal cortex holds memories between the facts that make up a situation (e.g., an outcome for a specific type of response to the situation) and emotions that were paired with the facts in an individual's previous experience. These linkages have the potential to reactivate an emotion, which can then activate an action tendency for the given situation. The factual knowledge related to a given situation is evoked in sensory images in the sensory cortices. Based on previous situations, signals related to the situation act on the ventromedial cortex and re-activate the sensory pattern that describes the appropriate emotion. Damasio notes that this process can result in a conscious emotion or can be unconscious, as the emotions can be represented via a sensory pattern in the brain rather than only in the body. The situation's sensory images are "marked" as good or bad by the somatosensory pattern.

Somatic markers help individuals to make decisions more efficiently by not solely relying on rational, logic-based analysis. For example, if one is standing in the grocery store trying to choose between 50 varieties of cereal, much time could be wasted in taking a cost-benefit analysis approach of each variety (comparing taste, nutrition, etc.); however, a somatic marker will may provide an individual with a positive feeling about a certain cereal, enabling them to

make a quicker decision. In situations in which the future is uncertain and decisions should be influenced by previous experiences, the absence of a somatic marker will result in the decision being based purely on logic. This strategy is slower and more likely to lead to errors due to not considering previous experience. Individuals with ventromedial damage tend to show this pattern of behavior, in addition to a related pattern of decision making involving random and impulsive responding (Clark et al., 2008; Bechara, Tranel, & Damasio, 2000).

Robinson et al. (2014) investigated the neural basis of both metacognitive and emotional/motivational aspects of executive function by administering a series of measures purporting to assess verbal fluency, working memory, cognitive control, and general executive function to participants with documented damage to varying brain areas. Their results replicated a previous study by Glascher et al. (2012) and largely supported the theories of Ardila (2008) and Stuss (2011), with lesions to the dorsolateral prefrontal cortex and anterior cingulate being generally related to metacognitive components of executive function and emotional/motivational aspects (apathy, disinhibition) being generally related to ventromedial dysfunction.

Executive functions are not necessarily specific to the frontal lobes. In fact, Stuss (2011) argued that there are problems with the practice of using the terms "executive dysfunction" and "frontal lobe dysfunction" interchangeably. One of these problems is the vast connectivity of the frontal lobe with non-frontal regions, such as limbic structures (Cummings, 1993), resulting in multiple brain areas being involved in executive function. Previous research has shown the temporal lobes to be linked to inhibition (Simons & Spiers, 2003) and empathy (Vollm et al., 2006) as well as parietal regions being associated with metacognitive functions (Collette et al., 2005). In addition, non-frontal regions, particularly the temporal lobe and sometimes the parietal lobe, were implicated in addition to frontal regions in nearly all of the aspects of executive

functioning Robinson et al. (2014) assessed. Another concern is that much of the research linking specific functions with specific brain areas is based on studies of individuals with diffuse lesions, not limited to the frontal lobes (as in Robinson et al., 2014). Furthermore, Stuss (2011) notes that many injuries result in impairments in executive dysfunction despite any demonstrable injuries to the frontal lobes.

Focusing on patients with focal frontal lesions, Stuss (2011) reviewed studies reporting attentional impairments. Seven basic task types were identified including sustaining, concentrating, sharing, suppressing, switching, preparing, and setting. The processes that might be involved in each task were then analyzed, based on processes defined in Norman and Shallice's (1986) Supervisory Attention System (SAS) model. This is because neuropsychological tasks almost always invoke multiple processes, as will be discussed more later. The SAS model posits that the frontal lobe consists of an attentional system in the anterior region of the frontal lobe involved in executive control of attention, as well as a posterior system concerned with spatial allocation of attention.

Unlike some other approaches, Stuss (2011) identified "executive functions" as consisting of two processes - monitoring and task setting - while classifying metacognition, behavioral/emotional self-regulation, and energization (initiating and sustaining responses) as separate processes. According to their review, patients with right lateral frontal damage demonstrate poor monitoring across tasks, while patients with comparable left lateral damage have problems with task setting. Consistent with other research (Robinson et al., 2014), behavioral/emotional self-regulation impairment was associated with ventromedial cortex damage. Individuals with focal dorsomedial damage demonstrated a pattern of impairment related to energization across different tasks. Impairment in metacognition ability, which

orchestrates the other tasks described, was associated with damage to polar regions; however, the author noted that this process was difficult to disassemble from the effects of ventromedial lesions on behavioral/emotional self-regulation. Overall, the authors conclude that there is no unitary "executive function" and that there are discrete functional categories within the frontal lobes, including energization, executive, emotion/behavioral regulation, and metacognition - each controlled by a specific frontal region (Stuss, 2011); however, it may be argued that these four processes are somewhat arbitrary in that they were defined a priori by the researcher based on theory and then identified in the review.

As Alvarez and Emory (2006) note, certain neuropsychological tests, such as those involving abstract reasoning, are referred to as frontal tasks as a result of studies showing that individuals with frontal lesions perform poorly on them. As alluded to above, however, given the vast connectivity between frontal lobes and other brain regions, as well as the fact that neuropsychological tasks almost always tap into more than one process, this labeling may be misleading. Alvarez and Emory (2006) reviewed lesion and functional neuroimaging studies using three popular executive function tasks: Wisconsin Card Sorting Test (WCST), Phonemic Verbal Fluency, and Stroop task, in order to determine the sensitivity and specificity of these tasks to the frontal lobes.

The majority of studies reviewed found that adults with frontal lobe lesions performed worse than controls on the WCST, while just over half of studies comparing frontal to non-frontal lesions found that participants with frontal lesions performed worse. The remaining studies found no differences when individuals with frontal lobe lesions were compared with diffuse and basal ganglia comparison groups. Five additional studies reviewed showed no significant differences between individuals with frontal and non-frontal lesions. Only two studies

showed no difference in performance when frontal patients were compared to norms; however, these were single-case design studies. Their review of functional neuroimaging studies related to the WCST yielded consistent findings with lesion studies, showing that while a bilaterally intact dorsolateral prefrontal cortex seems to be necessary for normal performance, a number of non-frontal brain regions are activated by the WCST as well. Although these neuroimaging studies tended to have small sample sizes, the authors state the results point to WCST performance activating a neural network involving both frontal and non-frontal areas of the brain.

Alvarez and Emory (2006) also reviewed lesion/neuroimaging studies on Phonemic Verbal Fluency, a task involving naming as many words beginning with a certain letter (usually F, A, or S) within a period of 60 seconds for each trial. Phonemic fluency is a commonly used measure of executive functioning, as individuals with frontal lobe damage tend to exhibit impaired performance on the task relative to semantic verbal fluency, which involves naming as many words as possible within a certain category (Troyer et al., 1998). Of 10 lesion studies comparing individuals with frontal lobe damage to healthy controls, each one found that the frontal group provided significantly fewer words. Further, 8 of 9 studies showed worse performance in individuals with frontal lobe lesions relative to those with non-frontal damage. However, reviewed studies indicated that patients with both frontal and non-frontal lesions performed worse than controls, and diffuse and frontal lobe patients produced similar amounts of words.

In addition to the narrative review, a meta-analysis of these studies was also conducted (Alvarez and Emory, 2006). The meta-analysis provided more evidence that the WCST is more sensitive to frontal lobe damage than the other tasks and that none of the tasks reviewed appear to be specific to frontal damage. Taken together, the results of the narrative review and meta-

analysis support that the tasks examined are sensitive but not specific to the frontal lobes, suggesting that referring to these measures as "frontal tests" may be misleading. The researchers hypothesize that the three tests are measuring different cognitive processes and that executive functioning may not be a unitary construct.

To summarize, metacognitive aspects of executive function are housed primarily in the dorsolateral PFC (Robinson et al., 2014), while emotional/motivational aspects of executive function are associated with the ventromedial PFC (Ardila, 2008). Given the vast connections between the frontal lobes and other brain structures, measures designed to assess these functions appear to be sensitive but not specific to PFC damage (Stuss, 2011).

Ecological Validity of Executive Functioning Tests

Due to the importance of executive functioning in terms of its impact on individuals' behavior and functional outcomes, measuring executive functioning using performance-based tasks and questionnaires is a routine aspect of neuropsychological evaluations (Rabin, Barr, & Burton, 2005). However, executive functioning is challenging to measure for a number of reasons. As discussed above, there is disagreement on how to define executive functioning as well as which specific cognitive functions should comprise the construct. In addition, there are problems related to construct validity, in which tests purported to tap into particular abilities (e.g., inhibition, planning, problem solving) often require the use of multiple cognitive abilities, such as attention, working memory, etc. (Culbertson & Zilmer, 1998). Similarly, many tests of executive function designed to measure a specific aspect of executive functioning appear to tap into several executive functions (Culberson & Zilmer, 1998; Latzman & Markon, 2010; Miyake et al., 2000). Another challenge of increasing relevance is determining the extent to which measures of executive functioning are ecologically valid.

In the context of neuropsychological assessment, ecological validity has been conceptualized as the "functional and predictive relation between the patient's behavior on a set of neuropsychological tests and the patient's behavior in a variety of real-world settings" (Sbordone & Long, 1996, pg. 16). In other words, ecological validity is the extent to which one's test results correlate with daily functioning. Therefore, for a measure of executive functioning designed to assess planning ability to be considered ecologically valid, poor performance on the task should correlate with real-life planning impairments. The ecological validity of cognitive tests can be examined and conceptualized using two approaches: verisimilitude and veridicality (Franzen & Wilhelm, 1996).

Verisimilitude refers to the degree to which the cognitive demands of a test theoretically resemble cognitive demands of daily tasks in one's environment. This approach generally involves developing novel tests, as opposed to using traditional tests, designed to mimic real-world tasks. An example of this is the Behavioral Assessment of the Dysexecutive Syndrome (BADS; Wilson et al., 1998), which requires examinees to search maps, associate names with faces, remember appointments, organize several activities, etc. The verisimilitude of a task depends not only on the extent to which similar skills required by the task are required in one's daily environment, but also the extent to which the testing environment and methods mimic one's daily environment. For example, although assessment situations often involve an examiner giving verbal directions to an examinee in a one-on-one situation, this may not happen in similar real-world scenarios, perhaps limiting the test's verisimilitude (Gioia & Isquith, 2004).

Veridicality refers to the degree to which existing tests are associated empirically to measures of real-world functioning. Using this approach, for example, researchers may attempt to determine the extent to which the WCST correlates with disability status or informant ratings

of executive functioning. A review of the ecological validity of neuropsychological tests found that measures of executive functioning using the verisimilitude approach tended to correlate more highly with outcome measures than did traditional tests, although there was some mixed evidence (Chaytor & Schmitter-Edgecombe, 2003). A combination of both approaches can also be used, in which a new test is designed to have high verisimilitude but also be tested for veridicality by validating it against outcome measures.

Neuropsychological evaluations now place a major focus on ecological validity. While many neuropsychological tests were developed to assess in the detection and localizations of brain lesions, advances in brain imaging techniques have reduced neuropsychologists' roles in this process - localization of lesions is now one of the least frequent referral questions (Rabin et al., 2005). Instead, referral questions are now often centered on functional questions related to individuals' everyday cognitive abilities, as two individuals with brain lesions in the same areas can experience vastly different functional outcomes. In addition to determination of specific diagnoses, frequent referral questions include rehabilitation/treatment planning, educational planning, forensic determination, assessing capacity to work, and assessing capacity for independent living (Rabin et al., 2005)

In many of these types of cases, the neuropsychologist's report may play a large role in determining important decisions, such as whether an individual loses their independence in some way, how much compensation they are awarded in a forensic case, and whether they can return to work or be placed on disability. In higher education, evaluation reports are sometimes used to determine whether students are eligible to receive academic accommodations such as increased time on exams. Therefore, neuropsychologists have a serious responsibility to be able to answer these questions accurately, using valid assessment tools. In order to do this, the measures

administered during assessments need to be demonstrated to be ecologically valid. However, as noted above, the majority of commonly used neuropsychological tests, including many designed to assess executive functioning, were developed as laboratory measures and were not designed for clinical use, let alone for answering functional questions. Questions remain as to whether the capacity to sort cards into correct categories on the WCST has any meaningful relationship to executive functioning abilities in everyday life. Ecological validity is important not only clinically, but also for research. As Spooner and Pachana (2006) note, neuropsychologists are also engaged in research work, including the assessment of neurologically intact individuals. For example, outcome studies are often conducted that examine changes in cognitive functioning following clinical medical or psychological interventions (e.g., mindfulness training, psychotherapy, medication). Traditional neuropsychological tasks may not provide information about how these interventions have effected individuals' daily functioning.

Ecological validity may be particularly important to study for executive functioning tasks, given the complexity of the construct as well as the findings that many individuals with functional problems will perform within normal limits on performance-based tasks (Stuss & Buckle, 1992). Executive functioning tasks are often designed to tap into one or two aspects of the construct (although this is not always the reality, as discussed earlier), such as the Tower of London assessing planning and the WCST of set-shifting/cognitive flexibility. In other words, they are designed to measure a narrow aspect of a neuropsychological construct, which is then generalized to ability to function in the natural environment, as opposed to assessing real-world adaptive decision-making (Goldberg & Podell, 2000). This strategy could result in missing out on collecting the relevant information required to gain a thorough understanding of one's strengths and weaknesses in the domain of executive functioning (Gioia & Isquith, 2004).

Ardila (2008) argues that because most traditional neuropsychological tests focus on metacognitive executive functions and ignore motivational/emotional ones, the ecological validity of these tests is greatly limited. For example, he argues that tests such as the Wisconsin Card Sorting Test and Tower of London require metacognitive abilities but are emotionally neutral tasks. Therefore, they do not tap into the ability to coordinate cognition with emotion. Although, some may argue that there is an emotional component to the WCST, with examinees being informed after each response whether they are correct or incorrect, which can produce positive or negative emotions – particularly frustration. Nevertheless, such measures primarily assess metacognitive aspects of executive functions and thus may not be accurate estimates of individuals' abilities to function in everyday life, given that real life situations are seldom free of emotional components.

The structured nature of the testing environment also presents a challenge to ecological assessment, given that executive functions are often thought of as being required most in complex, novel, unstructured situations (Gioia & Isquith, 2004). Thus, the structure of the testing situation may be detrimental to critical aspects of executive functioning. In addition, creating a test to be "novel" can be challenging, since a type of task that is novel to one individual may be familiar to another. This also creates difficulties with regard to serial assessment, such as when attempting to measure rehabilitation or progression of symptoms, as the tasks will no longer be novel upon repeat testing.

Challenges in Measuring Ecological Validity

Evaluating a measure's ecological validity presents challenges, many of which were reviewed by Chaytor and Schmitter-Edgecombe (2003). One of these challenges is the environment in which testing takes place. Traditionally, the goal of clinical neuropsychological

assessment is to determine one's level of performance under optimal conditions. This generally involves testing in a distraction-free environment, with an examiner providing instructions, prompts, and neutral feedback regardless of performance. However, a person's daily environment is unlikely to be this distraction-free, structured, and supportive. Thus, conducting validity studies in such environments may not tell researchers much about how an individual can perform in their normal settings. Another challenge in evaluating ecological validity is determining what outcome measures to validate it against. This is particularly difficult given the construct-validity problems associated with executive functioning measures that were discussed earlier. If we do not know what a test is actually measuring, or if that test is measuring multiple constructs, it is difficult to determine what outcome measures we should select to measure ecological validity. This also makes interpreting ecological validity studies difficult. An additional problem outlined by Chaytor and Schmitter-Edgecombe (2003) is that testing gathers only samples of behavior in a single environment, when in actuality, the individuals' behavior is likely to vary across contexts, introducing error into ecological validity research.

Compensatory strategies may also cause problems for determining ecological validity. Individuals with cognitive impairment often develop compensatory strategies for dealing with those impairments. For example, individuals with memory impairment may create a system where they write down important information and set alarms on their smartphones to remind them of appointments. Although they may be able to function normally in their daily environments using these compensatory strategies; these are generally not permitted to be used during the testing process. This limits the ecological validity of tests by causing it to potentially underestimate what people can do in their normal environments. Similarly, neuropsychologists/researchers may not adequately assess examinees' strengths, including such

compensatory mechanisms, supports and environmental resources that might help them to function more successfully in everyday life despite cognitive impairments.

Non-cognitive factors, such as motivation, may also detrimentally influence the ecological validity of tasks. For example, individuals may be more or less motivated to completed neuropsychological measures than to solve problems and complete tasks in their daily environments. Testing is a voluntary process, whereas certain challenges requiring the executive functioning abilities in real life cannot be avoided. Other non-cognitive factors cited include emotional symptoms, motivation for secondary gain in forensic cases, physical problems that can influence test results, premorbid functioning, and environmental demands (Chaytor & Schmitter-Edgecombe, 2003).

Another obstacle to determining ecological validity is that the environmental demands on one person may be very different from the environmental demands on another. For example, some individuals have very cognitively-demanding occupations, filled with situations which require intact executive functioning abilities, such as the ability to plan, organize, and problem-solve in novel, complex situations. Others have jobs that require routine, repetitive tasks, and that may be less demanding of the executive functions. Not only with regard to career, but also in other aspects of daily functioning, all individuals likely have differences with regard to the levels of each aspect of executive functioning that are required of them to function successfully in their environments. Therefore, the extent to which a measure is ecologically valid may vary depending on the examinee.

Despite the challenges outlined above, researchers have examined the ecological validity of traditional performance-based measures as well as self and informant-report questionnaires. Of these studies, the focus here will be on those investigating measures of executive functioning

specifically. Overall, many of the tests discussed here were validated on populations with brain lesions in general, as opposed to those with focal frontal lobe damage. Mitchell and Miller (2008) examined the ability of four subtests from the D-KEFS (Trail Making Test-4, Tower Test, Verbal Letter Fluency, and Design Fluency) to predict functional status in older adults, as measured by a clinician-rated scale of activities of daily living (ADLs). A composite score of these four measures was moderately correlated with functional status ($r = .66$). However, hierarchical regression analyses indicated that the Trail-Making Test 4 score was the only variable that accounted for a significant amount of the variance beyond that accounted for by depression, years of education, and the other executive functioning tasks used.

Kibby, Schmitter-Edgecombe, and Long (1998) compared the ecological validity of the WCST and a learning/memory task, the California Verbal Learning Test (CVLT), in a small sample of participants diagnosed with closed head injuries. The CVLT involves learning a list of words read by the examiner over multiple trials. The test also contains short-delay, long-delay, and recognition trials. The authors used a traditional scale of return to work as well as a job performance scale created for the purposes of the study. Neither the WCST nor the CVLT were significantly related to return to work. Two measures of the CVLT significantly correlated with the job performance scale, while the WCST perseverative error score did not. The WCST perseverative responses score did predict occupational status. Overall, the researchers concluded that the results supported their hypothesis that measures of memory functioning would predict job performance better than the WCST.

The ecological validity of the Hayling and Brixton tests, two tasks that are now part of the BADS, were studied in a sample of participants with brain injuries (Odhuba, van der Broek, & Johns, 2005). The Hayling test is a sentence completion task designed to measure response

initiation/suppression; the Brixton test is a rule-detection and rule-following task, requiring examinees to view a series of pages with multiple unfilled circles and one filled circle, and then predict the position of the filled circle on each subsequent page. Scores on these tests were correlated with measures of disability associated with dysexcutive impairments, level of handicap, and psychosocial problems associated with frontal lobe damage. Given the modest correlations between the executive functioning tasks and the outcome measures used, the researchers concluded "it is not possible to make inferences about everyday behaviour on the basis of these tests alone." Another study using the Hayling and Brixton tests, in addition to the Zoo Map and Key search sub-tests from the BADS battery, did not find significant correlations with informant ratings on the Dysexecutive Questionnaire in a sample of individuals whom had sustained severe brain injuries and showed CT scan evidence of frontal lesions (Wood & Liossi, 2006). This finding provides more evidence that individuals with documented brain impairment and difficulties in everyday life may still perform well on performance-based measures.

Ready, Stierman, and Paulsen (2001) found that traditional tests of neuropsychological functioning did not correlate significantly with self-reported everyday social behavior in undergraduate students; however, this could reflect problems with self-report questionnaires, as opposed to the ecological validity of the tests themselves.

In a study of substance dependent individuals, poly-substance users (compared to matched controls) showed widespread deficits on the Behavioral Assessment of Dysexecutive Syndrome (BADS; Wilson et al., 1998), a performance-based test designed to be high in verisimilitude. The BADS is comprised of six subtests, designed to examine multiple aspects of executive functioning including cognitive flexibility, problem solving, planning, temporal judgment, and behavioral regulation. It also contains a supplemental questionnaire, the

Dysexutive Questionnaire (DEX), designed to provide an informant-report of everyday behaviors related to dysexecutive syndrome. This measure also predicted everyday problems related to apathy, disinhibition, and executive dysfunction. However, there were no group differences with regard to performance on the WCST, providing evidence that the WCST lacks ecological validity in substance dependent individuals.

In another study assessing the validity of the BADS, Norris and Tate (2000) administered the test along with a measure of role functioning and several traditional neuropsychological measures, including the COWAT and WCST. They found that only one subtest from the BADS (Zoo Map) correlated significantly with the role functioning scale. The Zoo Map subtest and the COWAT were the only tests to correlate significantly with the BADS dysexecutive questionnaire, with the Zoo subtest correlating in an unexpected direction. In their narrative review of the ecological validity of neuropsychological tests, Chaytor and Schmitter-Edgecombe (2003) concluded that in the domain of executive functioning, verisimilitude tests such as the BADS tended to show more consistent relations to outcome measures than traditional tests such as the WCST. Additionally, they noted that self-report measures tended have little relationship to neuropsychological test performance.

In addition to the BADS, other performance-based tests have been developed specifically for the purposes of being ecologically valid measures of executive functioning, using the verisimilitude approach. One of the first of these tests was the Multiple Errands Test (MET; Shallice & Burgess 1991), which was designed to mimic daily life situations more accurately than traditional executive functioning tasks. The purpose of this task is for participants to carry out a number of tasks within a short period of time in novel locations, where unexpected events can occur. In the original study using the MET, hospital patients were handed a card with tasks

written on it, including six simple tasks to be completed at a local shopping center – six simple tasks (e.g., buy a loaf of bread), a seventh task telling them to be at a certain place in 15 minutes, and the final tasks involving obtaining and writing down multiple pieces of information. The participants were told to complete the tasks as quickly as possible. Multiple versions of this test have since been used by researchers, including a virtual-reality based version (Raspelli et al., 2012). Different versions are necessary, because the test is site-specific; the tasks have to take place in a local shopping center, hospital, etc. Each version of the test consists of 12 tasks similar to those outlined above, as well as a set of rules to follow. One of the rules is that participants are not allowed to speak to the examiner. Thus, participants must engage in purposeful and goal-directed behavior while inhibiting any impulses to talk to the examiner (Dawson et al., 2009).

The original article using the MET found that a small sample of just 3 individuals with diagnosed executive dysfunction performed poorly on the task compared to controls (Shallice & Burgess, 1991). On a simplified version of the MET, a sample of individuals with brain injuries made more rule-breaks, task failures, and total errors than healthy controls, even after controlling for general intelligence (Alderman, Burgess, Knight, & Henman, 2003). Performance was then compared to everyday dysexecutive problems as assessed by informant ratings on the Dysexecutive Questionnaire (DEX) to provide further evidence of ecological validity. Participants with many rule-breaking errors showed more severe executive memory symptoms (e.g., confabulating, temporal sequencing deficits) in everyday life, while those who committed task failure errors demonstrated more negative affect-related symptoms on the DEX. In another study, participants with acquired brain injury (stroke or TBI; not specific to the prefrontal cortex) performed worse on a version of the MET than matched controls (Dawson et al., 2009). In this study, researchers also examined the associations between MET performance and 4 informant-

reported outcome measures, including the DEX. Among individuals with strokes, the MET did not correlate significantly with the DEX total score. However, rule adherence on the task correlated with health-related quality of life in stroke survivors. In TBI survivors, total errors on the MET correlated significantly with the DEX and adjustment in daily life.

Several other verisimilitude tasks designed to measure executive function have been developed, most of which have limited research supporting their validity at this point. One such test is the Executive Function Performance Task (EFPT; Baum, Morrison, Hahn, & Edwards, 2003), which assesses the level of support needed for individuals to be successful in four daily life tasks including preparing meals, managing medications, using the telephone, and paying bills (Baum et al., 2008). Significant differences were found between participants with mild to moderate stroke and healthy controls on the EFPT (Baum et al., 2008). Doherty, Barker, Denniss, Jalil, and Beer (2015) recently developed a computerized cooking task requiring participants to virtually cook multiple items requiring different cooking times, with the goal of having all items finish cooking at the same time without burning or going cold. A study examining the ecological validity of another verisimilitude test, the Virtual Library Task (but not four other traditional executive function tests), was able to differentiate well between participants with TBI and healthy controls, and it was also predictive of everyday executive functioning, as measured by the DEX (Renison, Ponsford, Testa, Richardson, & Brownfield, 2012). Another task, the Virtual Action Planning-Supermarket, is a virtual reality program that assesses deficits in instrumental activities of daily living. In a sample of 5 individuals diagnosed with schizophrenia, difficulties on the virtual supermarket-shopping task translated to difficulties in real life grocery shopping (Aubin, Beliveau, & Klinger, 2015).

Although the tests described above are high in verisimilitude, there are some other problems associated with them. Most previous research conducted on these measures included small sample sizes and were not specific to prefrontal cortex lesions. Furthermore, the MET may not be practical for many clinicians and researchers, given that a new version must be developed at each site. The time and cost in developing the task and establishing its validity and reliability before using it will not be feasible for many. In addition, the MET (excluding the virtual version) requires examiners to take participants outside of the office/laboratory and into a real-world shopping center or similar place. In clinical settings, this may be unrealistic for logistical reasons and also poses confidentiality problems. Virtual-reality tasks may also be cost-prohibitive. Another potential problem with many of these tasks is that they may not be sensitive enough to pick up on subtle executive functioning deficits in populations without brain lesions. As will be discussed in more detail below, there is much evidence that individuals using drugs or at risk for drug use have some underlying executive functioning deficits. Tasks assessing activities of daily living are unlikely to identify such subtle impairments, given that executive functions are most taxed when undergoing novel and challenging tasks, as opposed to overlearned skills.

Finally, questionnaires assessing executive functioning are often considered to be inherently ecologically valid, since they inquire directly about individuals' functioning in daily life. Consequently, there is little research directly investigating their ecological validity. Rather, such questionnaires, especially the DEX (Wilson et al., 1998), are often used as outcome measures when determining the ecological validity of performance-based tasks (e.g., Alderman et al., 2003; Chaytor, Schmitter-Edgecome, & Burr, 2006; Renison et al., 2014; Wilson et al., 1998). Other similar measures include the Brief Rating Inventory of Executive Function – Adult

Version (BRIEF-A; Roth, Isquith, & Gioia, 2005) and the Frontal Systems Behavioral Scale (FrSBe; Grace & Malloy, 2000).

The DEX, as mentioned earlier, is a supplemental part of the BADS battery. The questionnaire consists of 20 items designed to sample problems commonly associated with dysexecutive syndrome (Wilson et al., 1998), including problems related to emotional/personality changes, motivational changes, behavioral changes, and cognitive changes. As it is assumed to be inherently ecologically valid, it was used to validate the performance-based tasks included in the BADS (Wilson et al., 1998)

The BRIEF is a questionnaire designed to provide information on everyday executive functioning skills within eight domains including *inhibit*, *shift*, *emotional control*, *initiate*, *working memory*, *plan*, *organize*, and *self-monitor* (Gioia & Isquith, 2004). It features self- and informant-report versions for both adults and children. The measure was designed to be ecologically valid from a verisimilitude approach, and the authors also provide some evidence of veridicality. Specifically, correlations have been found between BRIEF scales and academic achievement (Mahone, Koth, Cutting, Singer, & Denckla, 2001) and both injury severity and family functioning in children with TBI (Mangeot, Armstrong, Colvin, Yeates, & Taylor, 2002).

The Frontal Systems Behavior Scale (FrSBe) is a brief rating scale designed to provide a measure of behavior disturbances associated with damage to frontal systems. Individuals rate themselves or are rated by an informant with regard to behaviors both before and after an injury. The test consists of three subscales (Apathy, Disinhibtion, and Executive Dysfunction), and a confirmatory factor analysis provided support for these subscales (Stout, Ready, Grace, Malloy, & Paulsen, 2003). Support for the ecological validity of this measure includes FrSBe scores

being predictive of community integration in TBI survivors (Reid-Arndt, Nehl, & Hinkebein, 2007).

Although the research may suggest that performance-based measures have less to offer than self-report questionnaires when trying to predict everyday functioning, relying solely on self-report measures, however, is also problematic in that they will only be ecologically valid insofar as the respondents are providing accurate reports. Chaytor and Schmitter-Edgecombe (2003) concluded in their review that self-report measures are inferior to informant and clinician-reports of everyday cognitive performance, at least in patients with brain impairment. Self-reports may be inaccurate for a number of reasons, ranging from positive or negative impression management to lack of insight or brain impairment itself (e.g., memory difficulties).

Executive Functioning in Alcohol and Drug Use

One way in which the ecological validity of a measure can be assessed is the extent to which it is predictive of problems. If multiple tests are similarly predictive, then those tests may be valid for predicting a problem/disorder. Substance use is one such problem known to be associated with executive dysfunction (Lundqvist, 2005; Tartar et al., 2003). According to the National Survey on Drug Use and Health, 8.4 percent of people aged 12 and older in 2014 reported being current marijuana users (Center for Behavioral Health Statistics and Quality, 2015). An estimated 25.2 percent of the United States population use tobacco. In 2014, 52.7 percent of individuals aged 12 and older reported using alcohol within the past month, while 6.2 percent were heavy alcohol users in the past month (Center for Behavioral Health and Quality, 2015). The significant amount of Americans who currently use drugs is meaningful, because drug use is associated with a number of harmful correlates such as psychiatric disorders and impairments in social and role functioning (Compton, Thomas, Stinson, & Grant, 2007).

Deficits in executive functioning are significantly associated with drug use. Specifically, there is evidence for drug use contributing to executive dysfunction and for pre-existing executive dysfunction serving as a predictor for future drug use (Meil, LaPorte, & Stewart, 2012; Nigg et al., 2006; Peeters, Vollebergh, Wiers, & Field, 2014; Tartar et al., 2003; Wilens et al., 2011). With regard to alcohol, chronic abuse is associated with impairment in frontal lobe functioning, as evidenced by brain imaging studies and neuropsychological testing (see review in Lyvers, 2000). Several studies have linked neuroanatomical findings with neuropsychological test results in alcohol abuse (e.g., Akine et al., 2007; Noel et al., 2001). For example, decreased gray matter was correlated with decreased performance on the Trail Making Test and WCST (Chanraud et al., 2007).

Although the emphasis of this study will be the effects of alcohol use, other drugs have also been shown to detrimentally influence executive functioning. Nicotine dependence is also associated with executive dysfunction (Billeux et al., 2010). Chronic heavy smokers performed significantly more poorly than non/light and moderate smokers on the WCST test, but not on other neuropsychological measures (Razani, Boone, Lesser, & Weiss, 2004). Individuals with ADHD and substance use disorder were more likely to have deficits in executive functioning, as defined by impaired performance on 2/6 executive functioning measures, than individuals with ADHD whom did not have substance use disorder (Wilens et al., 2012). In that study, similar results were found for non-ADHD participants. Nicotine dependence has also been linked to self-reported executive dysfunction on all three FrSBe domains (Spinella, 2003).

Research on marijuana use has yielded similar findings. Gruber, Sager, Dahlgren, Racine, and Lukas (2012) compared the performances of early-onset (began using marijuana before the age of 16) and late-onset chronic marijuana users as well as non-users on a variety of

neuropsychological measures. Marijuana users overall committed significantly more perseverative errors as well as failure to maintain set errors on the WCST compared to non-users, with early-onset users performing significantly worse than late-onset users. Bolla, Brown, Eldreth, Tate, & Cadet (2002) studied the extent to which neurocognitive deficits persist in marijuana users after a period of abstinence (28 days) in addition to whether these deficits follow a dose-response curve. There was a significant dose-related effect, with more joints smoked per week associated with lower performance on measures of executive functioning as well as visuoperception, psychomotor speed, and manual dexterity. In particular, large effect sizes in WCST performance were found when increasing marijuana use.

Another existing question is the extent to which these executive functioning deficits predate the existence of substance use. Multiple literature reviews have concluded that in addition to drug use contributing to executive function impairment, pre-existing impairment is in fact a variable that influences individuals' risk of developing substance dependence (Meil, LaPorte, & Stewart, 2012; Peeters et al., 2015). As noted earlier, Nigg et al. (2006) found childhood impairments on a response inhibition task was predictive of adolescent problem drinking and drug use in adolescents at-risk for substance abuse. Similar results have been found with regard to behavioral disinhibition and later substance use disorder (Tartar et al., 2003). Response inhibition and working memory impairments in adolescence appear to predict the age at which one first drinks (Peeters et al., 2015). One study did not find any differences in maturation of executive functioning when comparing six groups of adolescents varying in amount/frequency of alcohol consumption (Boelema et al., 2015). In college students, facets of impulsivity – namely sensation seeking and positive urgency – are predictive of increases in drinking frequency and

negative outcomes associated with alcohol use, respectively (Cyders, Flory, Rainer, & Smith, 2009).

The use of alcohol is of particular concern in college student populations. Although college students tend to use less illicit drugs than their noncollege peers, college students report higher annual and 30-day use of alcohol as well as more occasions involving becoming intoxicated and binge drinking (Johnston, O'Malley, Bachman, Schulenberg, & Miech, 2015). Binge drinking in college has been linked with careless/reckless behaviors in college as well as increased risk for alcohol abuse and dependence after college (Jennison, 2004; Vik, Carrello, Tate, & Field, 2000).

As described above, substance use can result in neuropsychological impairments, particularly in the domain of executive functioning. The focus of this paper, however, is on prediction of substance use. There is some evidence that ecologically-valid measures are more predictive than less-ecologically valid tasks. The BADS, but not the WCST, was able to discriminate substance-dependent individuals from controls (Verdejo-Garcia & Perez-Garcia, 2007). In another study, Meil et al. (2016) administered the D-KEFS, FrSBe, Sensation Seeking Scale V (SSS-V) along with measures of alcohol, tobacco, and marijuana use to a sample of undergraduate students. The D-KEFS was not predictive of substance use. However, FrsBE disinhibition scores as well as sensation seeking predicted alcohol use frequency, and all three FrSBE subscales (i.e., Apathy, Executive Dysfunction, and Disinhibition) were predictive of tobacco use frequency. Sensation seeking was the only unique predictor of marijuana use frequency and also predicted alcohol related problems as well as cigarette dependence. Overall, when compared to a performance-based measure, ecologically-valid measures emerged as stronger predictors both of frequency of use and dependence.

Summary

In sum, there is concern that traditional measures of EF may not adequately reflect or predict individuals' abilities in everyday life. Tests designed specifically with ecological validity in mind may be better able to detect these everyday problems as well as other difficulties associated with executive dysfunction. For example, EF deficits are predictive of substance abuse, and ecological measures have been more sensitive in identifying these early deficits. While ecologically valid EF tests for general adult populations exist, there are currently no measures specifically designed to be ecologically valid in college students. Existing ecological measures have some drawbacks when used for the purpose of identifying impairments in executive dysfunction to predict other problems such as substance use. Some previously developed measures are impractical in that they require separate versions to be developed for each location they are administered and also involve taking an individual out into the community, which is not always realistic for clinicians and researchers. In addition, batteries such as the BADS, while sensitive with regard to detecting significant impairments, may be insensitive in detecting minor executive dysfunction that may precede the development of substance use/abuse. While questionnaires designed to be ecologically-valid have proven useful for these purposes, individuals with executive dysfunction may provide inaccurate self-reports due to a lack of insight into their difficulties (Chaytor & Schmitter-Edgecombe, 2003). Consequently, it is useful to have both self/informant-report questionnaires as well as performance-based tasks that are valid for detecting impairment.

Current Study

The purpose of the current study was to develop an ecologically valid performance-based task of executive functioning that is sensitive to normal variation as well as executive

dysfunction in college students, titled the College Registration Executive Function Task (CREFT). The test was designed with verisimilitude in mind and was also tested for veridicality based on its correlations with other measures of EF, as well as its ability to predict alcohol use in a sample of undergraduate students. In order to determine the extent to which it is ecologically valid, performance on the task was compared to scores on other brief performance-based tasks (i.e., WCST, Tower of London) as well as self-report measures of EF, including the BRIEF-A and UPPS Impulsive Behavior Scale (UPPS-P; Lynam, Smith, Whiteside, & Cyders, 2006).

As described above, alcohol use is common in college students, and chronic/heavy use is associated with neuropsychological impairment as well as school-related problems. Therefore, an additional goal of this study was to investigate the test's ability to predict excessive drinking and alcohol-related problems in college students. To this end, the Daily Drinking Questionnaire (DDQ; (Collins, Parks, & Marlatt, 1985), Alcohol Use Disorders Identification Test (AUDIT; Saunders, Aasland, Babor, De la Fuente, & Grant, 1993) as well as the Rutgers Alcohol Problem Index (RAPI; White & Labouvie, 2000) were administered.

The following hypotheses were made for the current study:

1) Based on theoretical alignment between the CREFT and aspects of EF designed to be assessed by the other measures, CREFT scores (i.e., total errors, perseverative errors, execution time, and total time) were expected to yield the highest correlations with the following measures: BRIEF-A Plan/Organize, Inhibit, Task Monitor and Global Executive Composite (GEC), scales, UPPS-P Positive Urgency and Sensation Seeking subscales, and Tower of London execution time and total correct scores.

2) CREFT execution time and total error scores, BRIEF-A GEC and Inhibition, and UPPS-P Positive Urgency and Sensation Seeking subscales were expected to be predictive of typical drinking rates and time spent drinking per week, as measured by the DDQ.

3) CREFT execution time and total error scores, BRIEF-A GEC and Inhibition, and UPPS-P Positive Urgency and Sensation Seeking subscales were expected to predict binge drinking frequency.

4) CREFT execution time and total error scores, BRIEF-A GEC and Inhibition, and UPPS-P Positive Urgency and Sensation Seeking subscales were expected to be predictive of excessive drinking, as measured by the AUDIT total score.

5) CREFT execution time and total error scores, BRIEF-A GEC and Inhibition, and UPPS-P Positive Urgency and Sensation Seeking subscales was expected to be predictive of alcohol-associated problems, as measured by the RAPI total score.

CHAPTER II

METHODS

Participants

The initial sample for this study consisted of 199 undergraduates enrolled at Indiana University of Pennsylvania. These participants were recruited from the subject pool, which consists of Introductory Psychology students. Sona software was used to allow potential participants to register to participate in the study during specific time slots in which the experimenter was available. To be included in the study, participants had to be native English speakers between the ages of 18-24 years old. Participants were excluded if they had experienced any diagnosed concussions or loss of consciousness within the previous three months. Twelve participants were eliminated from the study due to invalid BRIEF results, reducing the number of participants to 187. Of these, 123 (65.8%) identified as female, 63 (33.7%) identified as male, and 1 did not specify gender (See Table 1). Seventy-eight percent of the sample identified as White Non-Hispanic, 15% as African-American, 3.7% as Hispanic or Latino, 1.6% as Asian or Pacific Islander, and 1.1% as "other." The majority of participants were in their Freshman year of college (77.5%), while 15.5% were Sophomores, 3.7% were Juniors, and 2.1% were Seniors.

Table 1

Individual Characteristics as Frequency and Percentage of Sample

Characteristic	Frequency	Percentage
Gender		
Male	63	33.7
Female	123	65.8
Not Reported	1	0.5
Ethnicity		
White Non-Hispanic	146	78.1
African American	28	15
Hispanic or Latino	7	3.7
Asian or Pacific Islander	3	1.6
Other	2	1.1
Not Reported	1	0.5
Current Year in College		
Freshman	145	77.5
Sophomore	29	15.5
Junior	7	3.7
Senior	4	2.1
Super Senior (5+ years)	1	0.5
Not Reported	1	0.5

Procedures

Data collection took place in computer labs/classrooms at Indiana University of Pennsylvania. After signing a consent form, participants were asked to take a seat in front of a computer and turn their cell phones off. They then completed a brief demographics questionnaire before being administered a series of computerized performance-based EF measures, including the CREFT, the WCST, and the ToL. Participants then completed two self-report measures of EF, including the BRIEF-A and UPPS-P. Lastly, participants completed the DDQ, AUDIT, and RAPI, measures of typical alcohol consumption amount/frequency, excessive drinking, and alcohol-associated problems, respectively, in addition to a question assessing frequency of binge drinking.

Measures

Executive Functioning

College registration executive function task. The College Registration Executive Function Task (CREFT) was designed by William Meil and Matthew Taylor as an ecologically-valid performance-based measures of executive functioning to be used with college students. The task was constructed to simulate the process of registering for college courses based on a hypothetical major - the idea for which was based on anecdotal evidence that students often have difficulties with course registration, which is a process that likely involves EF. Over a number of years spent advising undergraduate students, Meil observed that students frequently present with questions and confusion regarding registering for courses, as they struggle to determine which courses they need to register to complete requirements for their majors and general studies curricula. The CREFT is considered to be high in verisimilitude, since it is modeled on a real-world task of relevance to college students.

As noted earlier, EF ability is generally considered to be most used when facing novel situations and tasks, as opposed to routine ones. This may pose a challenge with regard to the sensitivity of tests high in verisimilitude. To mitigate this problem, participants will be provided with a relatively obscure hypothetical major that most students would be unfamiliar with – Fermentation Sciences. Using a major that already exists at IUP would be problematic in that participants who actually have that major – and consequently experience with the requirements for that major - would have an advantages over others. In other words, the task would not be equally novel for everyone. Fermentation Science is an applied science that provides instruction in using micro-organisms as agents in the production of beer, wine, and other fermented foods. Curricula generally include courses in biology, chemistry, business, as well as specialized fermentation sciences courses. Fermentation Sciences degrees are currently offered at only a few universities in the United States. The course requirements for this major are adapted from the Fermentation Sciences program at Appalachian State University (http://programsofstudy.appstate.edu/programs/departments/Fermentation-Sciences-Program).

The CREFT is primarily computerized administered online via Qualtrics software (Provo, UT, 2016), although examinees are also provided with hard copies of certain materials discussed below. The participants are provided with the following instructions via the computer:

"On the following pages, you will be asked to register for 5 courses for one college semester (Fall semester of your Junior year). For this task, you are majoring in Fermentation Sciences. You are seeking to take the classes necessary to help you complete your major and General Studies requirements, while also balancing other responsibilities such as your work schedule. Please register for the appropriate courses based on all of these requirements."

In addition to these instructions, participants are informed as part of the task instructions that they have been provided with materials which contain information necessary to completing these tasks, including hard copies of a hypothetical transcript of courses that they have already taken (Appendix A), an uncompleted checklist of requirements to complete a major in Fermentation Sciences (Appendix B), an uncompleted checklist of general studies requirements (Appendix C), and a hypothetical work schedule for the semester in which participants are registering for courses (Appendix D). Participants are given no further information or prompting from the experimenter on how to best complete the task. If participants ask questions about how to complete the task or inquire as to whether their answers are correct (other than technical questions/issues), the experimenter's response is, "I cannot tell you how to complete the task – just do your best."

In selecting courses to register for, participants need to consider times they have available (based on their work schedule), courses they have already taken (based on their transcripts), and courses they need to take, as well as prerequisites for courses (based on major and General Studies requirement checklists).

After clicking through the screen with the task's instructions, the next screen instructs participants to "Please select your first course for Fall – Junior year." They then select a course from a dropdown list of 60 possible choices, each of which contains the course number, section, and time (e.g., PSYC 101 – MWF 9:00-9:50 AM). Five of the choices are considered correct responses, in that they meet the criteria described above. The rest are distractors that may be incorrect for one or more reasons (i.e., prerequisites not fulfilled, course offered at a time that conflicts with work schedule, course already taken). After selecting a correct choice, participants are then be prompted to select their second course, using the same list of possible responses,

minus the correct choice just selected. If an incorrect response is selected, they are prompted with an additional statement: "This course is not appropriate. Please select another option for your first course." To increase difficulty and assess the participants' ability to problem solve after receiving feedback, the specific reason why the course is an incorrect option is not given. If a participant makes 8 mistakes, they are then be taken to the next question and instructed to select their next course. The test will proceed in this manner until the participant has correctly selected their fifth course or committed the maximum number of errors for the fifth course.

Several scores are recorded from this measure. The first, Total Time, is the total time taken to complete the task. The Total Time score is a complex DV, as it was unclear exactly what it might be measuring. It might be expected that an individual with excellent EF ability could quickly determine what required courses fit into the schedule, resulting in a quick total time score. On the other hand, one might take a significant amount of time to plan/organize all of the information before selecting courses, resulting in a longer total time but also suggesting good EF ability. Alternatively, the Total Time score may be misleading without also considering the number of errors committed, as two participants with significant discrepancies in errors committed could have very similar Total Time scores. Therefore, the Total Time score may be less informative than other dependent measures. However, it was still analyzed in order to determine if it had significant predictive validity or relationships with the other variables.

The second CREFT score recorded is Total Errors, representing the sum of errors made on the task. A high Total Error score may be particularly indicative of lack of ability to monitor performance on a task and then appropriately modify one's strategy or approach. The Total Error score also features working memory components, as participants must keep in mind the multiple factors that would determine whether a specific course is an appropriate selection.

The third CREFT score, Execution Time, is defined by the amount of time between participants' first and final responses on the task. Theoretically, this dependent variable should serve as a measure of planning ability, as the most efficient way to approach this task is likely to use the information provided to plan out the entire semester, as opposed to one course at a time. If one were to approach the task in this manner, then once the first course is selected, the participant should be able to select the other courses fairly quickly, resulting in a short execution time. On the other hand, participants who take a different approach to the task by choosing one course, then determining the next course that they need to sign up for, and so on, will be expected to have a longer Execution Time. Like the Total Time score, however, this dependent variable is complex and could be misleading if one were to quickly respond to all of the items, despite making many errors, resulting in a short Execution Time.

In addition to the quantitative outcome measures above, qualitative information regarding participants' approach to completing the task was recorded. After selecting their fifth course, participants were given a series of eight items that ask "What is the (*first, second, etc.)* approach you took in determining what courses to register?" They were then provided with several possible answers, such as "Completed the Major requirement checklist based on your transcript to determine what major classes you needed to take," "Completed the General Studies requirements based on your transcript," Looked at your work schedule to determine which times you had available to sign up for courses," "Selected courses randomly without considering information provided," etc. There was also an "other" option in which they could type in any approaches not shown.

Wisconsin card sorting test. The WCST (Berg, 1948; Grant & Berg, 1948) is the most commonly used measure of EF according to a survey of neuropsychologists (Rabin et al., 2005).

It is sensitive to frontal lobe impairment (e.g., Alvarez & Emory, 2006). The purpose of the task "is to assess the ability to form abstract concepts, to shift and maintain set, and to utilize feedback." (Strauss, Sherman, & Spreen, 2006, p. 526). A computerized version of the short form (WCST-64) of the original WCST was administered via Millisecond's Inquisit Web software (www.millisecond.com), which features a library of psychological tests that can be administered online. The WCST, as described earlier, requires examinees to sort a deck of cards with designs that vary in shape, color, and number under the correct card within a row of four stimulus/key cards, which also vary in shape, color, and number. The first key card consists of a red triangle, the second consists of two green stars, the third consists of three yellow crosses, and the fourth consists of four blue circles. After the examinee sorts each card, he/she is then given feedback as to whether his response was correct or incorrect. Thus, the examinee must determine what rule to use to sort the cards (i.e., shape, color, or number). After a certain number of correct sorts are made, the rule changes (without the examinee knowing), at which point the examinee must disengage from the previous, now-irrelevant sorting rule and discover/maintain the new sorting rule. In the original version, the test ends after the examinee has completed either all 6 categories or 128 trials. Normative data is provided for individuals aged 5 years, 5 months to 89 years (Heaton et al., 1993). The test yields several scores, including number of categories completed, trials to complete first category, perseverative responses, perseverative errors, percent perseverative errors, failure to maintain set, percent conceptual level responses, and learning to learn. Of these scores, categories completed and perseverative errors may be considered the most diagnostically useful (Lezak et al., 2012). Because of this, as well as the redundancy of some of the scores (Strauss et al., 2006), categories completed and perseverative errors were the only scores analyzed in this study. In order to reduce the total time of the test battery, the short form

of the measure was used. This version consists of 64 trials, as opposed to 128. The short form of the WCST tends to yield results that are similar to the long form (Lezak et al., 2012), with the short form exhibiting correlations above 0.7 to corresponding scores on the long form (Strauss et al., 2006).

Tower of London. The ToL (Shallice, 1982) is one of the most widely-used measures of EF (Rabin et al., 2005). It is being included in the current study not only because of its popularity, but also because it is generally considered primarily a planning task, as participants must plan ahead to determine what moves they need to make in order to solve the problems using the fewest possible moves. The CREFT is expected to require significant planning ability as well; therefore, the ToL is an appropriate task to use to assess its construct validity. Successful performance also requires working memory, response inhibition, and visuospatial memory (Lezak et al., 2012).

A computerized version of Shallice's original ToL was administered via Millisecond software. On this test, participants are presented with three vertical pegs of descending heights, with three colored beads arranged in a fixed, initial position on two of the three pegs. On each trial, participants must rearrange these beads from their initial position to a new, predetermined position on one or more pegs. Participants are provided with a list of rules at the beginning of the task: only one bead may be moved at a time, each bead may be moved only from peg to peg, and only a specified number of beads may be left on each peg at a time (i.e., the first peg can hold three beads, the second peg can hold two beads, and the third peg can hold just one). Trials are only scored as correct if completed in the minimum number of moves necessary, and three attempts are allowed for each problem. The task consists of 12 trials of increasing difficulty.

The first score to be analyzed from this measure is "Total Score". The maximum possible total score is 36, as participants receive three points for an item if they solve it within the first attempt, two if they solve it on the second attempt, and one point if they solve it on the third attempt. The second score analyzed was "Execution Time," Execution Time is defined as the total amount of time taken for each item to be solved, minus the amount of time lapsed before the first move was made on each item.

UPPS-P impulsive behavior scale. Impulsivity is an important aspect of EF that plays a role in alcohol use and alcohol-related problems in college students (Cyders et al., 2009). A computerized version of the UPPS-P. (Lynam, et a., 2006; Appendix E) a self-report measure of impulsivity based on the original UPPS Impulsive Behavior Scale published by Whiteside and Lynam (2001), was administered via Millisecond software. The UPPS was developed in response to the heterogeneous conceptualizations of the construct of impulsivity in the psychological literature. Whiteside and Lynam (2001) conducted a factor analysis of several measures of impulsivity as well as facets of the NEO-PI-R (a personality measure based on the Five Factor Model of Personality) that relate to impulsivity in order to clarify the multi-faceted nature of impulsivity. Factor analyses identified four distinct personality facets associated with impulsivity, including (lack of) Premeditation (tendency to delay action in favor of careful thinking and planning), Urgency (tendency to commit rash or regrettable actions as a result of intense negative affect), Sensation Seeking (tendency to seek excitement and adventure), and (lack of) Perseverance (ability to remain with a task until completion and avoid boredom). After identifying these four factors, Whiteside and Lynam then created the UPPS by selecting items to measure each of the factors, resulting in a 45-item scale of impulsivity. The revised version, the UPPS-P, is a 59-item scale that assesses the four original factors as well as a fifth factor –

Positive Urgency, which measures rash actions related to intense positive moods (Cyders et al., 2007).

Studies have provided evidence for the construct validity of the UPPS, as the measure demonstrates relationships with several forms of psychopathology thought to be associated with poor impulsive control (Whiteside, Lynam, Miller, & Reynolds, 2005). Magid and Colder (2007) found that college students elevated on the traits of premeditation and sensation seeking engaged in high levels of alcohol consumption, while urgency and perseverance were related to alcohol-related problems, regardless of amount of alcohol use. Because the UPPS-P conceptualizes impulsivity as personality facets rather than psychopathology, it may be better equipped to pick up on normal variations in impulsivity that may be found in high-functioning college populations, as opposed to other measures of executive functioning that are designed specifically for clinical use.

Brief rating inventory of executive function – Adult. The BRIEF-A (Roth et al., 2005) is a standardized self-report measure designed to measure adults' executive functioning abilities in their everyday environment. It consists of 75 items that comprise nine theoretically and empirically-derived scales measuring different aspects of executive functioning, including Inhibit, Shift, Emotional Control, Self-Monitor, Initiate, Working Memory, Plan/Organize, Task Monitor, and Organization of Materials. Two indexes (the Behavioral Regulation Index; BRI and Metacognition Index) as well as an overall summary score – the Global Executive Composite (GEC) are also derived from the clinical scales. The BRIEF-A features validity scales designed to measure inconsistent responding, unusually infrequent responding, and unusually negative responding. The test has high internal consistency as well as test-retest reliability (Roth et al., 2005). The professional manual (Roth et al., 2005) provides thorough information regarding the

validity of the test, including its convergent validity with related measures. For example, significant correlations were found between BRIEF-A and FrSBe scores.

Alcohol/Drug Use and Alcohol-Related Problems

Daily drinking questionnaire. A modified, computerized version of the DDQ (Collins et al., 1985) was administered to assess frequency and amount of alcohol consumption during a typical week (Appendix F). Participants used a sliding scale to indicate how many standard drinks they consumed for each day of a typical week over the last 30 days as well as how many hours they spent drinking each day of a typical week. Participants were provided with a diagram indicating how much beer, wine, and liquor constitutes a standard drink.

Binge drinking. Participants were presented with a single question via Qualtrics assessing frequency of binge drinking. The question asked, "In the past two weeks, how many times have you consumed five or more standard drinks (males) or four or more standard drinks (females) over a two-hour period?" Response options included "0," "1," "2," "3," "4," and "5 or more."

Nicotine/marijuana use. Participants answered two questions via Qualtrics assessing frequency of nicotine and marijuana use, respectively. For each question, response options included "daily," "weekly," "monthly," "less than monthly," and "never."

Alcohol use disorders identification test. The AUDIT (Saunders et al., 1993; Appendix G) was developed as part of a World Health Organization project as a screening instrument for hazardous and harmful alcohol consumption, drinking behavior, and alcohol-related problems. The AUDIT consists of 10 questions assessing frequency/amount of use, dependence, and alcohol-related problems. Possible scores range from 0 to 40, with a score of 8 or more indicating a strong likelihood of hazardous or harmful alcohol consumption. This cut-off score

was chosen due to findings that among a large sample of individuals attending primary health care facilities, 92% of those diagnosed as having hazardous or harmful alcohol use had an AUDIT score of 8 or more, while 94% of those with non-hazardous use achieved scores of less than 8 (Saunders et al., 1993). A review of studies using the AUDIT found that the measure demonstrates high sensitivity and specificity (often higher than other self-report screening measures) to harmful drinking, in addition to adequate internal consistency and test-retest reliability (Reinert & Allen, 2002). Research has established its validity for use with college student populations (Kokotailo et al., 2004).

Rutgers alcohol problem index. For the current study, the 18-item version of the RAPI, a self-administered screening tool for assessing adolescent problem drinking, was administered (White & Labouvie, 2000; Appendix H). The original version of the measure (White & Labouvie, 1989) consisted of 23-items designed to assess problems drinking by asking about problems related to alcohol use, as opposed to frequency/amount of use. On the original scale, some of the items could be affected by gender bias and were therefore removed to create the shorter version. The scale possesses high reliability (.92). The 18-item version correlates highly with the 23-item version (r = .99; White & Labouvie, 2000).

Statistical Analyses

In order to investigate the construct and ecological validity of the CREFT and compare it to other performance-based tasks (see hypotheses 1-3 above), Pearson product-moment correlation coefficients were computed comparing CREFT outcome measures with other EF measures (i.e., ToL, WCST, BRIEF, and UPPS-P). Regression analyses were used to determine the extent to which the various measures of EF used in this study predicted scores on measures of alcohol consumption, binge drinking, and alcohol-related problems. In order to investigate the

EF measures' abilities to predict excessive drinking and alcohol-associated problems, four hierarchical multiple regressions analyses were conducted, using the DDQ, binge drinking, AUDIT, and RAPI as dependent variables, respectively. For each regression, demographic information (i.e., gender and age) were entered as the first step. Ethnicity was not analyzed as a predictor variable, due to the homogenous nature of the study sample. Scores on performance-based EF tasks (WCST-64 and ToL) were entered as the second step. The third step consisted of the self-report EF measures, including the BRIEF-A and UPPS-P scales specified in hypotheses 2-4. Finally, the CREFT was entered as step 4 in order to determine if it adds any predictive validity beyond that of the other measures.

CHAPTER III

RESULTS

Descriptive Statistics

Executive Functioning

One hundred and eighty-seven participants completed the CREFT. Three dependent variables were analyzed from the CREFT, including total time to complete the task (Total Time), Execution time (i.e., time to first item response subtracted from total time), and Total Errors. Means and standard deviations of each variable are presented in Table 2. The Total Time score, as measured in seconds, had a mean of 1069, with a standard deviation of 436. The mean Execution Time score was 754 seconds, with a standard deviation of 320 seconds. The Total Errors score had a mean of 14.8, with a standard deviation of 8.3. Thus, there was large variability in all of these measures. There were no significant gender differences with regard to performance on this task. Fifty-one percent of participants reported that the first step they took on the task was completing the Major Requirements checklist. As the following steps in determining which courses to select, participants were then most likely to look at their hypothetical work schedule to determine available times, look at the dropdown list of courses and examine each to determine if appropriate, complete the General Studies checklist, and check to see if course times interfered with courses already registered for, respectively.

Means and standard deviations for performance-based EF measure scores, including the WCST and ToL, are presented in Table 2. For the WCST, data were available for 181 participants (six participants did not complete the task due to computer/technical errors). The mean number of Categories Completed was 2.8 (range 0-5), with a standard deviation of 1.15. The mean number of Perseverative Errors was 5.3, with a standard deviation of 2.6. Based on an

independent samples t-test, females made significantly more perseverative errors than males; t (179) = -2.35; p = .02. Overall, participants made fewer perseverative errors relative to a study aiming to create demographically-corrected norms for the WCST-64, which found that Caucasian adults made a Mean of 7.6 perseverative errors, while African-American adults made a mean of 11.0 perseverative errors (Norman et al., 2012).

ToL data were analyzed for 181 participants, with a mean Total Score of 28.7 (SD = 6.61). Overall, these scores are slightly lower than those obtained by Krikorian, Barok, and Gay (1994), who obtained a Total Score of 33.2 (SD = 2.1), in a normative sample of undergraduate students using a non-computerized version of the task. There was a significant gender difference for ToL Total Score, with males scoring higher than females; t (179) = 2.86, p = .005. The mean Execution Time score was 94.5 seconds (SD = 31.7).

As noted above, 12 participants were excluded from further analysis due to exceeding the cut-off scores on embedded validity indices on the BRIEF-A. Means and standard deviations for BRIEF-A subscales and composite scores are included in Table 2 below. There was a significant gender difference on the BRIEF-A Emotional Control subscale, with females reporting higher scores; t (184) = -5.44, p = .002. There were no other significant gender differences in BRIEF-A scores. Internal consistency data, as determined by Cronbach's alpha statistics, are provided for each BRIEF-A scale in Table 2.

With regard to the UPPS-P, scores were available for 181 participants. For this sample, the mean Positive Urgency subscale score was 26.69 (SD = 8.7), and the Sensation Seeking subscale mean was 33.38 (SD = 7.06). Men endorsed higher levels of Sensation Seeking behavior than women; t (179) = 2.86, p = .005. See Table 2 for UPPS-P internal consistency data as determined by Cronbach's alpha statistics.

Table 2

Means and Standard Deviations for Executive Functioning Measures

Measure	Mean	Standard Deviation	Cronbach's Alpha
CREFT			
Total Time (in seconds)	1069	436	
Execution Time (in seconds)	754	320	
Total Errors	14.8	8.3	
WCST			
Categories Completed	2.81	1.15	
Perseverative Errors	5.33	2.6	
ToL			
Total Score	28.7	6.61	
Execution Time	94.5	31.7	
BRIEF-A (T-Scores)			
Inhibit	56.8	10.3	.67
Shift	56.3	10	.60
Emotional Control	54.2	11.4	.85
Self-Monitor	52.6	10.3	.69
Initiate	55.8	10.2	.71
Plan/Organize	54.2	9.2	.78
Working Memory	58.2	11.2	.78
Task Monitor	54.7	10.3	.64
Organization of Materials	50	9.4	.80
Metacognition Index	55.1	10.2	.91
Behavior Regulation Index	56.1	10.2	.90

Global Executive Composite	56.1	9.8	.95
UPPS-P			
Positive Urgency	26.69	8.7	.90
Sensation Seeking	33.38	7.06	.82

Alcohol/Drug Use and Alcohol-Related Problems

Means and standard deviations for alcohol and drug-related measures are provided in Table 3. As part of the demographic questionnaire, a single question was used to assess nicotine use, marijuana use, and binge drinking frequency. Of the 187 participants included in analysis, 138 (73.8%) did not report or endorse use of nicotine. Twenty-two participants (11.8%) reported using nicotine less than monthly, 10 (5.3%) reported monthly use, 7 (3.7%) reported weekly use, and 8 (4.3%) reported daily use. Ninety-nine participants (52.9%) endorsed never using marijuana, 42 (22.5%) endorsed less than monthly use, 11 participants (5.9%) reported monthly use, 19 (10.2%) reported weekly use, and 13 (7%) use daily. With regard to binge drinking, 81 participants (43.3%) did not endorse binge drinking in the past two weeks, as defined by having consumed five or more drinks for males or four or more drinks for females over a two-hour period. Thirty-five (18%) participants engaged in binge drinking once in the two week period prior to the study, 21 (11.2%) reported two occasions of binge drinking, 20 (10.7%) reported three occasions, 13 (7%) reported four occasions, and 15 (8%) reported five or more occasions of binge drinking. Men endorsed significantly more binge drinking episodes than females; $t(182) = 2.02$, $p = .04$.

On the DDQ, participants reported a mean of 7.37 drinks per week ($n = 184$; $SD = 9.22$; range = 0-59) and a mean of 6.69 hours spent drinking on a typical week ($SD = 9.28$). Men reported a significantly higher amount of total drinks over the last month compared to females; $t(183) = 4.61$, $p < .001$. For the AUDIT ($n = 185$), the mean score was 6.07 ($SD = 5.65$), with males ($M = 7.83$) reporting significantly more symptoms than females ($M = 5.16$); $t(183) = 3.11$, $p = .002$. The scale had a high level of internal consistency, as determined by a Cronbach's alpha of .83. Thirty-six percent of respondents met or exceeded the AUDIT's suggested cut-off

score of eight for classification of harmful drinking. A Chi-Square Test of Independence revealed that males were significantly more likely than females to meet this cut-off score ($p = .04$). In fact, more than half of the males in the study (52%) reported harmful drinking levels, compared to 28 percent of females. Finally, the RAPI was administered as a measure of problematic drinking. The scale had a high level of internal consistency (Cronbach's alpha = .87). Respondents on the RAPI ($n = 185$) had a mean score of 6.81 ($SD = 8.16$) (higher numbers represent higher levels of problematic drinking). An independent samples T-test comparing male and female RAPI total scores was not significant; $t(183) = 1.53$, $p = .129$).

Table 3

Means and Standard Deviations for Alcohol and Drug-Related Measures

Measure	Mean	Standard Deviation
DDQ		
Drinks Per Week	7.37	9.22
Hours Drinking Per Week	6.69	9.28
AUDIT Total Score	6.07	5.65
RAPI Total Score	6.81	8.16

Hypothesis 1: Correlations Between Executive Functioning Measures

It was hypothesized that CREFT scores (i.e., total errors, execution time, and total time) were expected to yield the highest correlations with the following measures: BRIEF-A Plan/Organize, Inhibit, Task Monitor and Global Executive Composite (GEC), scales, UPPS-P Positive Urgency and Sensation Seeking subscales, and Tower of London Execution Time and Total Score. To test this, Pearson's product-moment correlations were run to assess the relationship between CREFT outcome measures (i.e., total errors, total time, and execution time) and performance on EF measures including the Tower of London, WCST, BRIEF-A, and UPPS-P. For presentation purposes, although all scales and subscales of measures were included in the correlation analysis, Table 4 below includes only correlation data for variables specified above. Preliminary analyses showed the relationships between these variables to be linear. Not all variables were normally distributed, as assessed by Shapiro-Wilk's test ($p > .05$). There was a significant, moderate negative correlation between CREFT Total Errors and ToL Total Score, $r = -.31$ ($p < .001$). Lower ToL Total Scores indicate poorer performance on the task. There was a small negative correlation between CREFT Execution Time (lower time suggests better EF performance) and UPPS Positive Urgency subscale ($r = -.15$; $p = .047$). There was also a small positive correlation between CREFT Execution Time and BRIEF Organization of Materials ($r = .15$; $p = .043$). There were no other significant correlations between CREFT outcomes and EF measures. Thus, the hypothesis was partially supported, as the CREFT was significantly correlated with UPPS-P Positive Urgency and ToL Total Score, but not with other variables specified in the hypothesis.

Table 4

Pearson's Correlations for CREFT Scores and EF Variables Specified in Hypothesis 1

Measure	1	2	3	4	5	6	7	8	9	10	11
1. CREFT Total Errors	–	-.23**	-.29**	-.002	-.02	-.08	-.05	.09	.07	-.31**	.09
2. CREFT Total Time		–	.42**	-.06	-.07	-.05	-.06	-.02	-.03	.01	-.09
3. CREFT Execution Time			–	-.001	-.05	-.05	.00	-.15*	-.07	.07	-.07
4. BRIEF-A Plan/Organize				–	.55**	.63**	.80**	.43**	-.07	-.10	.06
5. BRIEF-A Inhibit					–	.55**	.73**	.41**	.08	.11	-.05
6. BRIEF-A Task Monitor						–	.79**	.40**	-.01	.03	.01
7. BRIEF-A GEC							–	.46**	-.09	.08	.06
8. UPPS-P Positive Urgency								–	.18*	.05	-.02
9. UPPS-P Sensation Seeking									–	.08	-.003
10. ToL Total Score										–	.22**
11. ToL Execution Time											–

Note. * $p \leq .05$, ** $p \leq .01$

Hypothesis 2: Prediction of DDQ Drinks Per Week by EF Measures

CREFT execution time and total error scores, BRIEF-A GEC and Inhibition, and UPPS-P Positive Urgency and Sensation Seeking subscales were hypothesized to be predictive of typical drinking rates, as measured by the DDQ. This hypothesis was partially supported. A hierarchical multiple regression was run to determine if the addition of traditional performance-based executive function measures (i.e., WCST and ToL), then self-report measures of impulsivity and executive function (i.e., UPPS-P Positive Urgency and Sensation Seeking; BRIEF-A Inhibition and GEC), and then finally CREFT, improved the prediction of DDQ typical drinks per week over and above demographic information (i.e., gender and age) alone. After two outliers were removed from the analysis, this regression included data from 170 participants, with a mean number of drinks per week of 7.02 (SD = 8.17). See Table 5 for full details on each regression model. There were linear relationships between DDQ Drinks Per Week and the independent variables collectively, as well as at least approximate linearity between DDQ Drinks Per Week and each individual predictor, as assessed by scatterplots.

The initial model, consisting of age and gender was statistically significant, R squared = .09, $F(2, 167) = 7.87$, $p < .001$., with gender ($p < .001$) being the lone significant predictor of DDQ Drinks Per Week. The full model (model 4) including all of the above variables was statistically significant, R squared = .19, $F(13, 156) = 2.81$, $p = .001$; adjusted R squared = .11. The addition of performance-based EF task scores to the prediction of DDQ Drinks Per Week (Model 2) did not lead to a significant increase in R squared ($p = .614$) beyond age and gender alone. The addition of self-report EF scores (Model 3) lead to a statistically significant increase in R squared of .07, $F(4, 159) = 3.27$, $p = .014$. Significant predictors in Model 3 included gender ($p = .002$) and UPPS-P Sensation Seeking ($p = .046$). The addition of CREFT scores (Model 4)

to the prediction of Drinks Per Week did not result in a statistically significant increase in R squared, $p = .265$. Thus, the hypothesis was partially supported, as the overall model was predictive of DDQ drinks per week, and self-report measures increased the predictive value of the model over demographics and performance-based tasks alone. However, as noted above, CREFT scores did not increase the model's predictive value.

Table 5

Hierarchical Multiple Regression Analyses Predicting DDQ Drinks Per Week From Demographic Variables, Performance-Based EF Measures, Self-Report EF Measures, and CREFT

DDQ Drinks Per Week								
	Model 1		Model 2		Model 3		Model 4	
Variable	B	β	B	β	B	β	B	β
Age	.14	0.02	.13	.02	.05	.01	.29	.04
Gender	**-5.01**	**-.29**	**-5.18**	**-.30**	**-4.32**	**-.25**	**-4.37**	**-.25**
ToL TS			.04	.04	.03	.03	.01	.01
ToL ET			.00	.002	.002	.01	.01	.02
WCST Cat			-.20	-.03	-.50	-.07	-.35	-.05
WCST PE			.39	.13	.43	.14	.37	.12
Inhibit					.06	.08	.06	.07
GEC					-.003	-.004	-.01	-.01
UPPS-P PU					.14	.15	.15	.15
UPPS-P SS					**.18**	**.16**	.17	.15
CREFT Errors							-.13	-.13
CREFT TT							-.004	-.24
CREFT ET							.004	.17
R^2	.09		.10		.17		.19	
F	7.87**		3.05**		3.23**		2.81**	
ΔR^2	.09		.02		.07		.02	
ΔF	7.87**		.67		3.24*		1.34	

Note. N = 169. * $p < .05$, ** $p < .01$. ToL TS = ToL Total Score; ToL ET = ToL Execution Time; WCST = WCST Categories Completed; WCST PE = WCST Perseverative Errors; Inhibit = BRIEF-A Inhibit; GEC = BRIEF-A GEC; UPPS-P PU = UPPS-P Positive Urgency; UPPS-P SS = UPPS-P Sensation Seeking; CREFT Errors = CREFT Total Errors; CREFT TT = CREFT Total Time; CREFT ET = CREFT Execution Time.

Hypothesis 3: Prediction of Binge Drinking by EF Measures

CREFT execution time and total error scores, BRIEF-A GEC and Inhibition, and UPPS-P Positive Urgency and Sensation Seeking subscales were expected to be predictive of binge drinking frequency. A second hierarchical regression was run, which included the same predictor variables but used self-reported binge drinking as an outcome variable to determine if traditional performance-based executive function measures (i.e., WCST and ToL), then self-report measures of impulsivity and executive function (i.e., UPPS-P and BRIEF-A), and then finally CREFT, improved the prediction of binge drinking over and above demographic information (i.e., gender and age) alone. There was linearity as assessed by partial regression plots and a plot of studentized residuals against the predicted values. There was homoscedasticity, as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values. There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1. There was one participant with a leverage value greater than 0.2, and this case was excluded from analysis. The distribution was somewhat non-normal, as assessed by Q-Q Plot; however, regression analyses are fairly robust to non-normality. For the 171 participants included in this regression, the mean number of binge drinking occurrences over the past two weeks was 1.46 (SD = 1.7). See Table 6 for full details on each regression model.

Models 1 and 2 were not statistically significant. Model 3 was statistically significant (R squared = .14, $F(10, 160) = 2.19$, $p = .021$. The only significant predictor in Model 3 was UPPS-P Sensation Seeking ($p = .03$). The full model (model 4) including all of the above variables was statistically significant, R squared = .12, $F(13, 157) = 2.45$, $p = .005$; adjusted R squared = .10. For Model 4, CREFT Total Time was a significant predictor of binge drinking ($p = .013$), as was CREFT Total Errors ($p = .039$).

The addition of performance-based EF task scores to the prediction of binge drinking (Model 2) did not lead to a significant increase in R squared ($p = .88$) above demographics alone. The addition of self-report EF scores (Model 3) led to a statistically significant increase in R squared of .09, $F(4, 164) = 4.19$, $p = .003$. The addition of CREFT scores (Model 4) to the prediction of binge drinking led to a statistically significant increase in R squared of .05, $F(3, 157) = 3.04$, $p = .031$. Overall, Hypothesis Three was partially supported, particularly with regard to UPPS-P and CREFT scores being significant predictors of binge drinking.

Table 6

Hierarchical Multiple Regression Analyses Predicting Binge Drinking From Demographic Variables, Performance-Based EF Measures, Self-Report EF Measures, and CREFT

Binge Drinking								
	Model 1		Model 2		Model 3		Model 4	
Variable	B	β	B	β	B	β	B	β
Age	.02	.01	.01	.01	.02	.01	.10	.06
Gender	-.51	-.14	**-.58**	**-.16**	-.39	-.11	-.40	-.11
ToL TS			-.01	-.04	-.02	-.07	-.02	-.09
ToL ET			-.003	-.05	-.002	-.04	-.002	-.04
WCST Cat			-.01	-.003	-.08	-.05	-.01	-.004
WCST PE			.03	.05	.04	.07	.02	.03
Inhibit					.03	.17	.02	.15
GEC					.01	.05	.01	.06
UPPS-P PU					.01	.07	.01	.05
UPPS-P SS					**.04**	**.18**	.04	.15
CREFT Errors							**-.04**	**-.18**
CREFT TT							**-.001**	**-.36**
CREFT ET							.001	.17
R^2	.02		.03		.12		.17	
F	1.79		.80		2.19*		2.45**	
ΔR^2	.02		.01		.09		.05	
ΔF	1.79		.31		4.19**		3.04*	

Note. * $p < .05$, ** $p < .01$. ToL TS = ToL Total Score; ToL ET = ToL Execution Time; WCST = WCST Categories Completed; WCST PE = WCST Perseverative Errors; Inhibit = BRIEF-A Inhibit; GEC = BRIEF-A GEC; UPPS-P PU = UPPS-P Positive Urgency; UPPS-P SS = UPPS-P Sensation Seeking; CREFT Errors = CREFT Total Errors; CREFT TT = CREFT Total Time; CREFT ET = CREFT Execution Time.

Hypothesis 4: Prediction of AUDIT Total Score by EF Measures

CREFT execution time and total error scores, BRIEF-A GEC and Inhibition, and UPPS-P Positive Urgency and Sensation Seeking subscales were expected to be predictive of excessive drinking, as measured by the AUDIT total score. To establish whether the addition of performance-based EF measures, then self-report EF measures, and finally the CREFT improved prediction of problematic drinking over demographics alone, a third hierarchical multiple regression was conducted using the AUDIT Total Score as the dependent variable. One hundred and seventy-two participants were included in this regression. There was linearity as assessed by partial regression plots and a plot of studentized residuals against the predicted values. There was heteroscedasticity, as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values. There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1. There were no studentized deleted residuals greater than three standard deviations, no leverage values greater than 0.2, nor values for Cook's distance above 1. The assumption of normality was met, as assessed by Q-Q Plot. For participants included in this analysis (n = 172), the mean AUDIT Total Score was 6.04 (SD = 5.62). Thirty-five percent of respondents met the recommended cut-off score for harmful drinking. See Table 7 for full details on each regression model.

The initial model was statistically significant, R squared = .04 $F(2, 169) = 3.73$, $p = .026$. In the initial model, gender ($p = .008$) was a significant predictor. The full model (Model 4) including all of the above variables was statistically significant, R squared = .25 $F(13, 158) = 4.04$, $p < .001$; adjusted R squared = .19.

The addition of performance-based EF task scores to the prediction of AUDIT Total Score (Model 2) did not lead to a significant increase in R squared ($p = .568$) beyond

demographics alone. The addition of self-report EF scores (Model 3) led to a statistically significant increase in R squared of .18, $F(4, 161) = 9.7$, $p < .001$, with BRIEF Inhibit ($p = .016$), UPPS-P Sensation Seeking ($p = .005$), and UPPS-P Positive Urgency ($p = .01$), in addition to gender, predicting AUDIT Total Score. The addition of CREFT Total Time, Execution Time, and Total Errors scores (Model 4) to the prediction of AUDIT Total Score was not significant, $p = .645$. Thus, the hypothesis was partially supported in that self-report EF measures were predictive of AUDIT Total Scores beyond performance-based tasks, while CREFT scores were not.

Table 7

Hierarchical Multiple Regression Analyses Predicting AUDIT Total Score From Demographic Variables, Performance-Based EF Measures, Self-Report EF Measures, and CREFT

AUDIT Total Score								
	Model 1		Model 2		Model 3		Model 4	
Variable	B	β	B	β	B	β	B	β
Age	.00	.00	.003	.001	-.02	-.004	.09	.02
Gender	**-2.44**	**-.21**	**-2.55**	**-.22**	-1.52	-.13	-1.52	-.13
ToL TS			.03	.03	.01	.01	.002	.003
ToL ET			.01	.04	.01	.06	.01	.06
WCST Cat			-.06	-.01	-.40	-.08	-.30	-.06
WCST PE			.26	.12	.30	.14	.27	.13
Inhibit					**.14**	**.25**	**.13**	**.25**
GEC					-.04	-.07	-.04	-.06
UPPS-P PU					**.14**	**.21**	**.14**	**.21**
UPPS-P SS					**.17**	**.21**	**.16**	**.20**
CREFT Errors							-.05	-.08
CREFT TT							-.002	-.14
CREFT ET							.001	.06
R^2	.04		.06		.24		.25	
F	3.73*		1.73		5.12**		4.04**	
ΔR^2	.04		.02		.18		.09	
ΔF	3.73*		.74		9.67**		.56	

Note. N = 172. * $p < .05$, ** $p < .01$. ToL TS = ToL Total Score; ToL ET = ToL Execution Time; WCST = WCST Categories Completed; WCST PE = WCST Perseverative Errors; Inhibit = BRIEF-A Inhibit; GEC = BRIEF-A GEC; UPPS-P PU = UPPS-P Positive Urgency; UPPS-P SS = UPPS-P Sensation Seeking; CREFT Errors = CREFT Total Errors; CREFT TT = CREFT Total Time; CREFT ET = CREFT Execution Time.

Hypothesis 5: Prediction of RAPI Total Score by EF Measures

CREFT execution time and total error scores, BRIEF-A GEC and Inhibition, and UPPS-P Positive Urgency and Sensation Seeking subscales were expected to be predictive of alcohol-associated problems, as measured by the RAPI total score. A hierarchical multiple regression was run to determine if the addition of traditional performance-based executive function measures (i.e., WCST and ToL), then self-report measures of impulsivity and executive function (i.e., UPPS-P and BRIEF-A), and then finally CREFT, improved the prediction of problematic drinking, as assessed by RAPI Total Score, over and above demographic information alone. There was linearity as assessed by partial regression plots and a plot of studentized residuals against the predicted values. There was homoscedasticity, as assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values. There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1. There were no studentized deleted residuals greater than three standard deviations, no leverage values greater than 0.2, and values for Cook's distance above one. The assumption of normality was met, as assessed by Q-Q Plot. The mean for RAPI Total Score after three outliers were removed (n = 169) = 6.84 (SD = 8.22). See Table 8 for full details on each regression model.

Neither Model 1 nor Model 2 significantly predicted RAPI Total Score. Model 3, including demographics, performance-based EF measures, and self-report EF measures, was statistically significant, R squared = .21 $F(10, 158) = 4.17$, $p < .001$; adjusted R squared = .16. In Model 3, UPPS-P Positive Urgency ($p = .004$) and UPPS-P Sensation Seeking ($p = .021$) were significant predictors. The full model (Model 4) including all of the above variables was statistically significant, R squared = .21, $F(13, 155) = 3.18$, $p < .001$; adjusted R squared = .144. The addition of performance-based EF task scores to the prediction of RAPI Total Score (Model

2) did not lead to a significant increase in R squared. The addition of self-report EF scores, including UPPS-P Positive Urgency and UPPS-P Sensation Seeking (Model 3) led to a statistically significant increase in R squared of .19, $F(4, 158) = 9.21$, $p < .001$. The addition of CREFT scores (Model 4) to the prediction of RAPI was not significant. Thus, this hypothesis was also partially supported, as some of the predicted measures, particularly UPPS-P Positive Urgency and UPPS-P Sensation Seeking, significantly predicted RAPI Total Score.

Table 8

Hierarchical Multiple Regression Analyses Predicting RAPI Total Score From Demographic Variables, Performance-Based EF Measures, Self-Report EF Measures, and CREFT

RAPI Total Score								
	Model 1		Model 2		Model 3		Model 4	
Variable	B	β	B	β	B	β	B	β
Age	-.24	-.03	-.22	-.03	-.28	-.03	-.23	-.03
Gender	-2.24	-.13	-2.28	-.13	-1.05	-.06	-1.02	-.06
ToL TS			.06	.05	.04	.03	.04	.03
ToL ET			-.02	-.06	-.01	-.05	-.02	-.06
WCST Cat			.11	.02	-.43	-.06	-.37	-.05
WCST PE			.20	.06	.26	.08	.24	.08
Inhibit					.09	.11	.08	.10
GEC					.08	.10	.09	.10
UPPS-P PU					**.24**	**.25**	**.24**	**.24**
UPPS-P SS					**.21**	**.18**	**.21**	**.18**
CREFT Errors							-.03	-.03
CREFT TT							-.001	-.03
CREFT ET							-.001	-.03
R^2	.02		.02		.21		.21	
F	1.34		.67		4.17**		3.18**	
ΔR^2	.02		.01		.19		.002	
ΔF	1.34		.34		9.21**		.13	

Note. N = 169. * $p < .05$, ** $p < .01$. ToL TS = ToL Total Score; ToL ET = ToL Execution Time; WCST = WCST Categories Completed; WCST PE = WCST Perseverative Errors; Inhibit = BRIEF-A Inhibit; GEC = BRIEF-A GEC; UPPS-P PU = UPPS-P Positive Urgency; UPPS-P SS = UPPS-P Sensation Seeking; CREFT Errors = CREFT Total Errors; CREFT TT = CREFT Total Time; CREFT ET = CREFT Execution Time.

CHAPTER IV

DISCUSSION

Ecological Validity of CREFT

One purpose of the current study was to develop an ecologically valid performance-based task of executive functioning, the CREFT, and determine its ecological validity by comparing it to existing EF measures, including a self-report measure previously established to possess high ecological validity. To this end, Pearson's Correlations were computed after administering the measures to a sample of undergraduate students. It was expected that the CREFT would be more ecologically valid than traditional performance-based EF measures, given that it closely mimics a real world task that college students engage in. Therefore, it was hypothesized that it would show significant correlations with the BRIEF-A. More specifically, because CREFT is thought to require aspects of EF including planning, organization, monitoring, and inhibition, hypothesis one predicted that CREFT scores would correlate most highly with BRIEF-A Inhibit, Plan/Organize, Task Monitor, and Global Executive Composite scales. In addition, based on expected theoretical alignment between the measures, CREFT was hypothesized to correlate with UPPS-P Positive Urgency, UPPS-P Sensation Seeking, and ToL Execution Time and Total Scores**.** This hypothesis was partially confirmed, as there was only one significant correlation between CREFT and the BRIEF-A, and that correlation was a positive relationship between CREFT Execution Time and BRIEF Organization of Materials, a relationship that was not predicted. There were, however, significant correlations between CREFT Total Errors and ToL Total Score, and CREFT Execution Time and UPPS-P Positive Urgency.

The finding of few relationships between CREFT and BRIEF-A scores is consistent with previous research showing little relationship between self-report and performance-based

measures of EF. For example, Ready et al. (2013) administered a series of neuropsychological tasks and self-report EF measures to college students, finding no significant correlations. A review of six studies investigating ecological validity of EF tests found no significant relationships between performance-based measures - both traditional tests such as the WCST and those designed with higher verisimilitude, such as the BADS - and self-report. However, there were significant relationships between EF measures and everyday abilities as assessed by informant and clinician ratings (Chaytor & Schmitter-Edgecome, 2003). Although they examined less commonly used tasks, Wood and Liossi (2006) similarly obtained minimal correlations between performance-based and self-report EF measures in a sample of severely brain injured individuals, providing further evidence of the lack of ecological validity of performance-based measures. Hagen et al. (2016) examined the ability of "hot" and "cold" EF tasks, as well as the BRIEF-A, to predict membership in control or substance use disorder groups. They also discovered a lack of association between the self-report and performance measures, concluding that the two types of tasks may be tapping different aspects of EF. Specifically, performance-based measures may assess specific aspects of EF, while self-reports assess more global, applied EF ability (Isquith, Gioia, & Roth, 2013). On the other hand, some studies have higher ecological validity in measures focused on verisimilitude. Verdejo-Garcia and Perez-Garcia (2007) found that the BADS (but not the WCST) significantly predicted self-reported apathy, disinhibition, and executive dysfunction, as measured by the FrSBe.

As expected, CREFT Total Errors was correlated with the ToL Total Score, indicating that as individuals made fewer errors on CREFT, they tended to make more correct answers on the ToL. The ToL is largely a task of planning ability, as items are only scored as correct if the solution is achieved within the minimal possible number of moves, requiring participants to plan

ahead, and CREFT likely makes significant demands on planning ability as well. In addition to planning demands, both tasks feature rules that participants must keep in mind in order to make correct responses. The ToL's rules stipulate that larger beads may not be placed on smaller beads, that only one bead can be moved at a time, and that each peg can only hold a certain number of beads. The CREFT's rules require participants to register for classes that they have the required prerequisites for, have not already taken, and do not interfere with their work schedule. This finding, in combination with evidence of discriminant validity (a lack of correlations with other measures and domains of EF) suggests that CREFT may be an effective measure of planning ability.

The modest but significant negative correlation between CREFT Execution Time (i.e., time between first and final responses on the task). and UPPS-P Positive Urgency confirms part of the study's first hypothesis and suggests that slower execution time is associated with increased levels of positive urgency (i.e., tendency to engage in rash actions related to positive moods). This relationship was predicted due to the theoretical assumption that CREFT scores would be related to planning and impulsivity/inhibition and previous research suggesting that self-reported positive urgency may be associated with lower performance on EF tasks involving inhibition (Gay, Rochat, Billieux, d'Acremont, & Van der Linden, 2008).

The unexpected finding of a modest correlation between CREFT Execution Time and BRIEF-A Organization of Materials, although not hypothesized, makes theoretical sense. This correlation shows that as self-reported ability to organize materials required for tasks in everyday life improves, so does CREFT Execution Time. As part of the administration of the CREFT, participants were provided several pages containing information including transcripts of previous semesters, required coursework, and a work schedule. In order to select the correct courses,

participants needed to incorporate all of this information. It follows logically that individuals who are very organized may take longer to organize all of the information provided at the beginning of the task and perhaps finding multiple courses to fit into the schedule, as opposed to jumping into the task in a less organized way. One question to consider is why there was no association between Total Errors, the clearest indicator of performance level on the CREFT, and BRIEF-A Organization of Materials. One might think that that being well-organized would be associated with making fewer errors, but this was not the case according to the data. This suggests that while organizing the handouts might lead to more efficient performance, as measured by Execution Time, there may be multiple ways of organizing or approaching the task while still completing it accurately.

Overall, these results suggest that CREFT may not possess the expected ecological validity, at least as defined by its relationship to an ecologically valid self-report EF measure. It was thought that, as opposed to traditional performance-based tasks such as the WCST and ToL, the CREFT may be more ecologically valid due to having being designed to mimic a real world task most college students must complete. In fact, the CREFT appears to be the first performance-based EF task designed specifically for university students. However, there was still a lack of associations between the CREFT and BRIEF-A. One possible explanation for this result is that performance-based tasks assess quite a narrow range of individuals' functioning, whereas self-report measures such as the BRIEF-A assess broader domains of everyday EF ability. As Chaytor and Schmitter-Edgecombe (2003) note, one of the challenges of measuring ecological validity is that neuropsychological testing involves gathering a small sample of behavior in a single environment, and this sample may not be representative of individuals' behavior, which can greatly vary across time and settings.

Another similar potential contributor to the findings is that, despite the CREFT mimicking a real-world task, the test was still administered in an artificial test environment. As discussed earlier in the study, one problem with EF tests is that they are administered in structured and artificial testing environments, where individuals are monitored by the examiner and given at least basic instructions – thus limiting their representativeness of real world functioning. Although the current study made an effort to leave the task as open-ended as possible, it was still somewhat artificial. Individuals completed only one task at a time in a relatively quiet environment without distractions. Thus, it still may not have approximated real world situations in which people must often perform under distracting and unsupportive conditions, as discussed in Chaytor and Schmitter-Edgecome (2003). Furthermore, the CREFT administration was unlike the real process of scheduling courses in that students are able to consult with advisors when registering.

Relatedly, students' performance may have been affected by the fact that they were given hypothetical majors, schedules, and classes from which to choose. Registering for their college courses in real life situations may be more stressful, given that their choices have actual consequences. Furthermore, participants may be more motivated when registering for an actual college semester. All of these factors could reduce the ecological validity of the measure.

The pattern of correlations between CREFT and other EF measures also provides preliminary support for the hypothesis that CREFT is effectively a measure of planning ability, as it correlated strongly with an established planning task but not with other measures.

Predictive Validity of CREFT for Alcohol Measures

It was hypothesized that CREFT Execution time and total error scores, BRIEF-A GEC and Inhibition, and UPPS-P Positive Urgency and Sensation Seeking subscales would be

predictive of typical drinking rates, as measured by the DDQ Drinks Per Week. However, UPPS-P Sensation Seeking was the only one of these variables to predict weekly drinking rates. Gender, though, was the best predictor of weekly drinking rates. This is consistent with previous research showing that college-aged males tend to use more alcohol than females (Johnston et al., 2014). Traditional performance-based tasks did not increase the predictive value of our model. Therefore, at least in this study, typical drinking rates are better predicted by demographic information alone (i.e., gender), rather than performance-based EF measures. Including self-report EF measures, in particular the UPPS-P Sensation Seeking subscale, did increase the ability to predict drinking rates. Past research has established relationships between sensation seeking and alcohol use frequency as well as alcohol related problems (Cyders et al., 2009; Meil et al., 2016). None of the predicted scores from the BRIEF-A emerged as significant predictors in the model.

While gender was a significant predictor of weekly drinking rates, it did not predict self-reported rates of binge drinking. A possible explanation for this finding is that, while males overall drink more alcohol than females, binge drinking rates are defined differently for males and females in order to control for overall drinking rates. For males, binge drinking was defined as "five or more drinks within a two-hour period;" while for females, binge drinking was defined as "four or more drinks within a two-hour period.". The study found that UPPS-P Sensation Seeking did emerge as a significant predictor of binge drinking.

CREFT Total Time and Total Error scores also significantly predicted binge drinking. One possible explanation for this finding is that their may be similar personality and EF factors influencing both CREFT performance and binge drinking behavior. Specifically, CREFT is a difficult task, with good performance requiring a patient, deliberate approach, likely in addition

to adequate frustration tolerance. On the other hand, binge drinking behavior may be undertaken by individuals who are more impulsive and less able to regulate frustration in stressful/frustrating situations. This is the first example in this study that a performance-based EF measure higher in verisimilitude did a better job of predicting an outcome than traditional EF tasks and self-report measures, such as the BRIEF-A. Generally, performance-based tasks have been less predictive of alcohol-related outcomes (e.g., Meil et al., 2016; Verdejo-Garcia & Perez-Garcia, 2007). Thus, the CREFT is uniquely positioned in being a performance-based EF task specifically designed for college students with the ability to predict outcomes such as binge drinking. This is important, given that binge drinking is a major issue in college populations (Jennison, 2004)

In this study, gender emerged as a predictor for symptoms of alcohol use disorder, and performance-based measures, including the ToL and WCST, did not. This is consistent with Hagen et al. (2016), who found that gender and education predicted membership into a group of individuals diagnosed with substance use disorders better than performance-based EF measures. A regression model including self-report measure led to an increase in ability to predict symptoms of alcohol use disorder. In particular, the hypothesized UPPS-P scales and BRIEF-A Inhibit were significant predictors. However, CREFT measures were not predictive. The study also found that UPPS-P Positive Urgency and UPPS-P Sensation Seeking scores were the only significant predictors of problematic drinking, as measured by the RAPI.

Overall, the findings indicate that CREFT did not demonstrate increased ecological validity as compared to traditional performance-based measures of EF. In addition, with the exception of the prediction of binge drinking, the CREFT performed no better than these measures when it comes to predicting alcohol use and alcohol-related problems. These results also speak to the possible differences in what is being assessed by performance-based and self-

report tasks. Ready et al (2010) studied both types of tasks with regard to their ability to predict different types of behaviors. Their results indicated that performance tasks were predictive of achievement and work-related behavior, while, similar to the current study, self-reports were predictive of substance use, disinhibited, and risk-taking behavior. Thus, they concluded that both types of measures are ecologically valid but associated with different types of outcomes.

Instead, it appears that the UPPS-P, specifically UPPS-P Sensation Seeking subscale, was the most consistently predictive of alcohol-related outcome measures. These results are similar to those of Meil et al. (2016), who examined the ability of ecologically-based measures to predict substance use and found that a self-report measure of sensation seeking was predictive of both alcohol and marijuana use frequency.

Previous research has also supported the notion of a relationship between traits of impulsivity based in the UPPS model (including sensation seeking) and alcohol use outcomes. A meta-analysis conducted by Coskunpinar, Dir, and Cyders (2013) found that all UPPS model traits, with the exception of Positive Urgency, were significantly predictive of drinking quantity (medium effect sizes). Also consistent with the current study, the meta-analysis found that sensation seeking showed the largest effect size for predicting binge drinking. With regard to problematic alcohol consumption, Negative Urgency and Positive Urgency had the largest effects, with Sensation Seeking still being a significant predictor. Thus, the current data provides more evidence supporting the overall relationship between impulsivity, and sensation seeking specifically, as a predictor of alcohol use and alcohol-related problems.

Limitations

As noted earlier, some specific problems with the current research include limitations of the CREFT, including its lack of validity index as well as its administration using Qualtrics, a

program designed for surveys rather than neuropsychological tests. A computer program specifically designed for the task may have been more intuitive and user friendly for participants. In addition, although the CREFT was designed to closely approximate a real-world task, it was still administered in an artificial testing environment using hypothetical majors and work schedules. This limitation, as well as the fact that the test gathered a small sample of behavior that may not be representative of individuals' overall behavior, mimics problems with performance-based neuropsychological testing in general.

Despite a relatively large sample size, the sample obtained in this study may not accurately reflect the entire population of undergraduate students, as the sample was quite homogenous with regard to ethnicity. In addition, the nature of the subject pool from which participants were gathered meant that the majority of the participants were Freshman. Thus, while the study was aimed at investigating college students in general, the data collected is most generalizable to first-year college students.

An additional concern related to the composition of the study sample was the possibility of variation in frontal lobe development between the younger and older participants, given research showing that maturation of the frontal lobe continues into adulthood (Romine & Reynolds, 2010). In other words, the 18-year-old participants' brains may be less developed overall than participants in their early twenties, which could have created differences in EF performance and distorted conclusions drawn from the study. To address this, the correlational and hierarchical regression analyses were run again, excluding the seven participants aged between 22-24, in order to determine if removing older participants would change the results of the study. There were no significant changes in the study's results after this analysis.

Similarly, because the sample was composed of college students, the study mostly examined high-functioning individuals. College students, for the most part, are likely to have a certain level of EF ability; therefore, the sample may have a more restricted range of EF ability relative to a sample of the general population. Thus, stronger findings may be obtained by comparing groups, such as college students with same-aged peers not attending college, or students in good academic standing with those on academic probation or mandatorily enrolled in university drug and alcohol management programs.

The lack of means to measure the motivation and effort level of participants is another limitation. Of the measures administered, only the BRIEF-A had imbedded validity scales. Data from participants that scored at the indicated cut-off on the BRIEF-A were completely eliminated from analysis, as it was assumed that these participants may have also responded noncredibly on other measures. The CREFT, in its current state, lacks any objective validity index. Anecdotally, it was observed during data collection that several participants simply responded rapidly and randomly, attempting to complete the task as quickly as possible. Given the range in possible options, this approach will almost always result in a large number of errors. However, without an objective validity index, the frequencies of noncredible responding as well as the impact of noncredible responding on study results remain unclear.

Instead of being administered in a counterbalanced order, EF measures were administered in the same order for each participant. Given the amount of tasks participants were asked to complete, there was potential for fatigue effects to influence results, particularly with regard to performance on tasks administered later in the study.

While participants were asked about frequency of marijuana use, this information was not controlled for in the analyses. Given that marijuana use has dose-related detrimental effects on

EF ability (Bolla et al., 2002), marijuana use could have influenced findings. Moreover, participants were not asked about use of other substances other than alcohol, tobacco, and marijuana. It is possible that substance use, other than alcohol, impacted the results. For example, some participants may have denied alcohol use but showed impaired EF ability associated with abuse of other substances.

Another possible limitation of the study is the modest reliability (Cronbach's alphas below .70) of some BRIEF-A subscales (i.e., Inhibit, Shift, and Self-Monitor), which could limit the validity of findings involving these scales. However, the majority of measures included in the study had high reliability, with Cronbach's alphas of greater than .80.

Future Research

Future studies involving the CREFT could focus on refining the measure and learning more about the nature of the scores and domains it measures. It turned out to be quite difficult for participants, given that the five-item task yielded a mean of over 14 errors. Many participants commented that the CREFT was difficult and frustrating. For example, one participant stated "I dread doing this in real life. It was very annoying to do again." Another stated, "I really was kind of lost. I'm a freshman so I struggle doing this without the help of my advisor…I guessed most of the time." It may be beneficial to more closely examine what specifically was so difficult about the task, both through subjective report and objective analysis, such as by determining what error types participants are most likely to make, as well as the domain(s) of executive functioning involved in those error types. There also may have been aspects of the CREFT administration that added to the difficulty. For example, one participant suggested that providing a blank paper schedule to fill in would have made the task easier (participants only entered their courses on the computer). Another potential contributor to the perceived difficulty and

participants' frustration may be the uniform item difficulty. The WCST, which is similar to the CREFT in that the structure of the test does not allow for progressive item difficulty, has been criticized for being too frustrating, at least when used with severely-impaired individuals (Smith-Seemiller, Franzen, & Bowers, 1997). Unlike the CREFT, many neuropsychological tests of EF, including the ToL, begin with easier items and gradually increase in complexity.

More analysis and refinement of what CREFT outcome measures are actually measuring and which are most meaningful could also be useful. In the Method section, it was noted that it was unclear what CREFT Total Time and Execution Time scores would be measuring. Although longer time to complete tasks is generally indicative of poorer performance on EF tests, the opposite appears to be the case here. There was a significant negative correlation between CREFT Total Errors and CREFT Total Time. It appears that individuals who rushed through the task made more errors, while more accurate performance was related to taking one's time working through the task. However, in some cases, longer Total Time could be indicative of difficulty completing the task. Similarly, Execution Time was expected to be a measure of planning ability, with the idea that the most efficient approach to the task would be to read the instructions and spend a significant amount of time determining five courses to register, and then beginning to submit one's answers. Thus, for individuals approaching the task this way, the Execution Time (i.e., the time between first response on the task and completion of the task) would be much shorter than for those taking a less organized approach to the task. However, it is not clear whether this was the case in our current study. It is possible that someone could have short or long Execution/Total times, yet still make many errors on the task. It appears that Total Errors was the clearest indicator of task performance. One potential solution would be to create a

CREFT score measuring the ratio of Total Time to Total Errors. Analyzing such a score may be a more effective way of examining participants' efficiency and level of performance on the task.

Another potential goal of further research on the CREFT would be to establish normative data, perhaps even stratifying norms according to age and college class, allowing for greater ability to compare results with other measures and across studies. Finally, an important refinement of the CREFT would be to create an independent computer program to run and score the task. The current version of the task uses Qualtrics, which is designed primarily for surveys, rather than tasks with specific correct answers. As a result, scoring had to be done by hand. Creating a stand-alone software program for the CREFT would improve the usability for both researchers and participants.

Additionally, a maximum of eight errors could be made prior to the participant being moved on to the next item, regardless of obtaining a correct response or not. This number was selected, somewhat arbitrarily, in order to allow for a reasonable possible range of errors without becoming overly frustrating for participants. In addition, limitations in Qualtrics survey software played a role in this decision. Further refinement of the CREFT should address whether the amount of errors allowed on each item should be adjusted.

As noted earlier, the CREFT did not include a measure of validity, and several participants appeared to be randomly selecting courses during the task. Validity indicators designed to flag variations of non-credible responding, whether due to lack of motivation or malingering, is an important future refinement of the task.

The current study relied on self-report to gather information on alcohol use and alcohol-related problems. Overall, research has supported self-report "quantity/frequency" measures of alcohol use to be generally reliable and valid, despite a tendency toward under-reporting at

higher levels of alcohol use (Del Boca & Darkes, 2003; Northcote & Livingston, 2011). Nonetheless, self-report can be flawed for numerous reasons, including inaccuracy of memories or hesitancy to report drinking habits due to embarrassment or fear of consequences. It is possible that this distorted some results of the study. Therefore, future studies may wish to gather objective information on alcohol-related problems, perhaps by comparing CREFT scores of students involved in alcohol and drug management programs to a control group of undergraduates.

Future research studying ecological validity of EF in samples of college students may want to explore the use of other operational definitions of everyday EF ability. In addition to self-report measures, informant-reports, such as from family members, teachers, or other sources could be used. While some studies, such as the current one, have used comparisons with self-report measures such as the BRIEF-A, others have examined the ability of measures to predict objective aspects of functioning, such as employment status (Kibby et al., 1998). In a high functioning sample of college students, such ability to work or complete activities of daily living would not constitute sufficient measures of everyday executive functioning ability. Perhaps other outcome measures may be more appropriate for college students, such as data related to academic success, including attendance rates and course grades.

Clinical Implications

Alcohol use is a major problem in college student populations, with students engaging in higher rates of alcohol use, intoxication, and binge drinking than non-college peers (Johnston, O'Malley, Bachman, Schulenberg, & Miech, 2015). As noted earlier, binge drinking behavior is associated with not only reckless behavior and poorer grades while in college, but also increased risk of alcohol related problems after college (Jennison, 2004; Vik, Carrello, Tate, & Field,

2000). Furthermore, there is substantial evidence that poor EF is both predictive of and worsened by substance abuse (Meil, LaPorte, & Stewart, 2012; Nigg et al., 2006; Peeters, Vollebergh, Wiers, & Field, 2014; Tartar et al., 2003; Wilens et al., 2011). Thus, it is important for psychologists to have appropriate, ecologically valid means of assessing EF in this population in order to inform diagnostic decisions and treatment recommendations. While self-report measures may be the most ecologically valid measures according to this study, their validity is dependent on accurate self-report. As noted earlier, these measures may be inaccurate for a number of reasons (e.g., positive or negative impression management, memory difficulties, lack of insight). Therefore, they should not be solely relied upon. In general, a thorough assessment should, then, use a combination self-report and performance-based measures, which are less susceptible to those problems.

To date, the CREFT appears to be the only performance-based EF task with a focus on verisimilitude designed specifically for college students. The current study suggests that it is uniquely predictive of binge drinking – more so than the BRIEF-A, suggesting that it has utility as part of a more comprehensive neuropsychological testing battery. In addition, CREFT has an advantage over self-report measures in that it allows clinicians to obtain qualitative data based on observations of how test-takers approach the task. As discussed earlier, some participants rushed through the task, while others were more methodical. Some participants quickly became frustrated, while others did not appear to be. There were also many different ways of organizing and integrating the information provided to participants, such as by first completing the major and general studies requirements checklist to determine which courses they still needed, first ruling out courses that interfered with their work schedule, etc. These different ways of approaching and responding to the task can be a source of valuable clinical information.

This study also provided more evidence that aspects of impulsivity and disinhibition, particularly sensation-seeking behavior, are important to assess for. Perhaps the most consistent finding in this study was sensation seeking behavior being predictive of alcohol outcomes, including typical drinking rates, binge drinking frequency, and symptoms of alcohol dependence. Sensation seeking behaviors may need to be a focus of intervention at both individual and programmatic levels. Psychologists working with individuals with substance abuse problems may want to consider including interventions aimed at reducing or managing sensation seeking behavior. Similarly, university alcohol/substance management programs could consider ways to target outreach to address the role of sensation seeking and impulsivity in substance use.

Conclusion

The purpose of the study was to develop an ecologically valid EF task specifically for college students, and to assess its correlations with existing EF measures, as well as its ability to predict outcomes related to alcohol use. As expected, CREFT performance was associated with performance on the ToL, with both tasks thought to involve the planning aspect of EF. CREFT scores were also related to BRIEF-A Organization of Materials, which was not hypothesized but made sense theoretically given that CREFT involved organization of several pages of materials. Although CREFT was hypothesized to be correlated more highly with BRIEF-A due to being designed with ecological validity in mind, the lack of other associations between the measures was consistent with previous research suggesting that self-report and performance-based EF tasks are measuring different things (Hagen et al., 2016, Isquith et al., 2013). Sensation seeking emerged as a significant predictor of each alcohol outcome, including typical drinking rates, binge drinking frequency, alcohol-related problems, and symptoms of alcohol use disorder. Notably, CREFT performance was predictive of binge drinking, while BRIEF-A scores were not.

Several important conclusions can be drawn from this study. First, the current research is the first to attempt to develop an ecologically valid, performance EF task to be used specifically with college students. Second CREFT shows promise as a task of planning ability, given its correlations with an established planning measure and initial evidence of discriminant validity. Third, although CREFT showed few correlations with other measures, it is a strong predictor of binge drinking, suggesting that it has clinical utility in assessing college students with possible EF and/or substance problems. Finally, results provided further evidence of the importance in assessing and targeting interventions/outreach toward sensation seeking and impulsivity, given that it was consistently predictive of alcohol outcomes. Future research could focus on further refinement and validation of CREFT, including examination of outcome measures and specific aspects of EF being tapped into by the task. Future studies may also want to replicate this research using a sample of students referred to university alcohol management/outreach programs.

References

Akine, Y., Kato, M., Muramatsu, T., Umeda, S., Mimura, M., Asai, Y., ... & Suhara, T. (2007). Altered brain activation by a false recognition task in young abstinent patients with alcohol dependence. *Alcoholism: Clinical and Experimental Research*, *31*(9), 1589-1597. doi:10.1111/j.1530-0277.2007.00453.x

Alderman, N., Burgess, P. W., Knight, C., & Henman, C. (2003). Ecological validity of a simplified version of the multiple errands shopping test. *Journal of the International Neuropsychological Society*, *9*(1), 31-44. doi:10.1017/S13556177039100046

Alvarez, J. A, & Emory, E. (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review*, *16*(1), 17–42. doi:10.1007/s11065-006-9002-x

Ardila, A. (2008). On the evolutionary origins of executive functions. *Brain and Cognition*, *68*(1), 92–99. doi:10.1016/j.bandc.2008.03.003

Aubin, G., Béliveau, M. F., & Klinger, E. (2015). An exploration of the ecological validity of the Virtual Action Planning–Supermarket (VAP-S) with people with schizophrenia. *Neuropsychological Rehabilitation*, 1-20. doi:10.1080/09602011.2015.1074083

Baddeley, A., & Wilson, B. (1988). Frontal amnesia and the dysexecutive syndrome. *Brain and Cognition*, *7*(2), 212-230. doi: https:10.1016/0278-2626(88)90031-0

Barkley, R. A., & Murphy, K. R. (2010). Impairment in occupational functioning and adult ADHD: The predictive utility of executive function (EF) ratings versus EF tests. *Archives of Clinical Neuropsychology, 25*(3), 157-173. doi:10.1093/arclin/acq014

Banich, M. T. (2009). Executive function: The search for an integrated account. *Current Directions in Psychological Science*, *18*(2), 89–94. doi: 10.1111/j.1467-8721.2009.01615.x

Baum, C. M., Connor, L. T., Morrison, T., Hahn, M., Dromerick, A. W., & Edwards, D. F. (2008). Reliability, validity, and clinical utility of the Executive Function Performance Test: A measure of executive function in a sample of people with stroke. *American Journal of Occupational Therapy*, *62*(4), 446-455. Retrieved from http://ajot.aota.org/pdfaccess.ashx?url=/data/journals/ajot/930098

Baum, C. M., Morrison, T., Hahn, M., & Edwards, D. F. (2003). *Test manual: Executive Function Performance Test*. St. Louis, MO: Washington University.

Bechara, A., Tranel, D., & Damasio, H. (2000). Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain*, *123*(11), 2189–2202. doi: 10.1093/brain/123.11.2189

Berg, E. A. (1948). A simple objective technique for measuring flexibility in thinking. *The Journal of General Psychology*, *39*(1), 15-22. doi:10.1080/00221309.1948.9918159

Billieux, J., Gay, P., Rochat, L., Khazaal, Y., Zullino, D., & Van der Linden, M. (2010). Lack of inhibitory control predicts cigarette smoking dependence: Evidence from a non-deprived sample of light to moderate smokers. *Drug and Alcohol Dependent, 112*, 164-167. doi:10.1016/j.drugalcdep.2010.06.006

Boelema, S. R., Harakeh, Z., Van Zandvoort, M. J. E., Reijneveld, S. A., Verhulst, F. C., Ormel, J., & Vollebergh, W. A. M. (2015). Adolescent heavy drinking does not affect maturation of basic executive functioning: Longitudinal findings from the TRAILS study. *PLoS ONE*, *10*(10), 1–15. doi:10.1371/journal.pone.0139186

Bolla, K. I., Brown, K., Eldreth, D., Tate, K., & Cadet, J. L. (2002). Dose-related neurocognitive effects of marijuana use. *Neurology*, *59*(9), 1337-1343. doi:10.1212/01.WNL.0000031422.66442.49

Burgess, P. W., Alderman, N., Evans, J., Emslie, H., & Wilson, B. A. (1998). The ecological validity of tests of executive function. *Journal of the International Neuropsychological Society : JINS*, *4*(October), 547–558. doi:10.1017/S1355617798466037

Cabeza, R., Dolcos, F., Graham, R., & Nyberg, L. (2002). Similarities and differences in the neural correlates of episodic memory retrieval and working memory. *Neuroimage*, *16*(2), 317-330. doi:10.1006/nimg.2002.1063

Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience*, *12*(1), 1-47. doi:10.1162/08989290051137585

Center for Behavioral Health Statistics and Quality. (2015). Behavioral health trends in the United States: Results from the 2014 National Survey on Drug Use and Health, 64. Retrieved from http://www.samhsa.gov/data/sites/default/files/NSDUH-FRR1-2014/NSDUH-FRR1-2014.pdf\nhttp://www.samhsa.gov/data/

Chanraud, S., Martelli, C., Delain, F., Kostogianni, N., Douaud, G., Aubin, H. J., ... & Martinot, J. L. (2007). Brain morphometry and cognitive performance in detoxified alcohol-dependents with preserved psychosocial functioning. *Neuropsychopharmacology*, *32*(2), 429-438. doi:10.1038/sj.npp.1301219

Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology Review*, *13*(4), 181-197. doi:10.1023/B:NERV.0000009483.91468.fb

Chaytor, N., Schmitter-Edgecombe, M., & Burr, R. (2006). Improving the ecological validity of executive functioning assessment. *Archives of Clinical Neuropsychology*, *21*(3), 217–227. doi:10.1016/j.acn.2005.12.002

Christoff, K., & Gabrieli, J. D. (2000). The frontopolar cortex and human cognition: Evidence for a rostrocaudal hierarchical organization within the human prefrontal cortex. *Psychobiology*, *28*(2), 168-186. doi: 10.3758/BF03331976

Clark, L., Bechara, A., Damasio, H., Aitken, M. R. F., Sahakian, B. J., & Robbins, T. W. (2008). Differential effects of insular and ventromedial prefrontal cortex lesions on risky decision-making. *Brain*, *131*(5), 1311–1322. doi:10.1093/brain/awn066

Collette, F., Olivier, L., Van der Linden, M., Laureys, S., Delfiore, G., Luxen, A., & Salmon, E. (2005). Involvement of both prefrontal and inferior parietal cortex in dual-task performance. *Cognitive Brain Research*, *24*(2), 237-251. doi:10.1016/j.cogbrainres.2005.01.023

Collins R. L., Parks, G. A., & Marlatt, G. A. (1985). Social determinants of alcohol consumption: The effects of social interaction and model status on the self-administration of alcohol. Journal of Consulting and Clinical Psychology, *53*, 189-200. doi: 0022-006X/85/00.75

Compton, W. M., Thomas, Y. F., Stinson, F. S., & Grant, B. F. (2007). Prevalence, correlates, disability, and comorbidity of DSM-IV drug abuse and dependence in the United States: results from the national epidemiologic survey on alcohol and related conditions. *Archives of General Psychiatry*, *64*(5), 566-576. doi:10.1001/archpsyc.64.5.566

Coskunpinar, M. A., Dir, A. L., & Cyders, M. A. (2013). Multidimensionality in impulsivity and alcohol use: A meta-analysis using the UPPS model of impulsivity. *Alcohol Clin Exp Res, 37*(9), 1441-1450. doi:10.1111/acer.12131

Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment*, *4*(1), 5. Retrieved from

https://www.researchgate.net/profile/Paul_Costa3/publication/232463584_Normal_Personality_Assessment_in_Clinical_Practice_The_NEO_Personality_Inventory/links/550ed48d0cf2ac2905ad119a

Culbertson, W. C., & Zillmer, E. A. (1998). The construct validity of the Tower of London DX as a measure of the executive functioning of ADHD children. *Assessment*, *5*(3), 215-226. doi:10.1177/107319119800500302

Cummings, J. L. (1993). Frontal sub-cortical circuits and human behaviour. *Archives of Neurology, 50*, 873-880. doi:10.1001/archneur.1993.00540080076020

Cyders, M. A., Smith, G. T., Spillane, N. S., Fischer, S., Annus, A. M., & Peterson, C. (2007). Integration of impulsivity and positive mood to predict risky behavior: development and validation of a measure of positive urgency. *Psychological Assessment*, *19*(1), 107. doi:10.1037/1040-3590.19.1.107

Cyders, M. A., Flory, K., Rainer, S., & Smith, G. T. (2009). The role of personality dispositions to risky behavior in predicting first-year college drinking. *Addiction, 104*(2), 193-202. doi:10.1111/j.1360-0443.2008.02434.x

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *351*(1346), 1413–1420. doi:10.1098/rstb.1996.0125

Dawson, D. R., Anderson, N. D., Burgess, P., Cooper, E., Krpan, K. M., & Stuss, D. T. (2009). Further development of the Multiple Errands Test: Standardized scoring, reliability, and ecological validity for the Baycrest version. *Archives of Physical Medicine and Rehabilitation*, *90*(11), S41-S51. doi:10.1016/j.apmr.2009.07.012

Del Boca, F. K., & Darkes, J. (2003). The Validity of self-reports of alcohol consumption: State of the science and challenges for research. *Addiction, 98*(2), 1-12. doi:10.1046/j.1359-6357.2003.00586.x

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan executive function system (D-KEFS)*. Psychological Corporation.

DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2005). Sources of openness/intellect: Cognitive and neuropsychological correlates of the fifth factor of personality. *Journal of personality*, *73*(4), 825-858. doi:10.1111/j.1467-6494.2005.00330.x

Diamantopoulou, S., Rydell, A., Thorell, L. B., & Bohlin, G. (2007). Impact of executive functioning and symptoms of Attention Deficit Hyperactivity Disorder on children's peer relations and school performance. *Developmental Neuropsychology, 32*(1), 521-542. doi:10.1080/87565640701360981

Doherty, T. A., Barker, L. A., Denniss, R., Jalil, A., & Beer, M. D. (2015). The cooking task: Making a meal of executive functions. *Frontiers in Behavioral Neuroscience*, *9*, 1-10. doi:10.3389/fnbeh.2015.00022

Engelhardt, L. E., Briley, D. A., Mann, F. D., Harden, K. P., & Tucker-Drob, E. M. (2015). Genes unite executive functions in childhood. *Psychological Science*, *26*(8), 1151-1163. doi:10.1177/0956797615577209

Franzen, M. D. & Wilhelm, K. L. (1996). Conceptual foundations of ecological validity in neuropsychological assessment. In R. J. Sbordone & C. J. Long (Eds.). *Ecological validity in neuropsychological assessment* (91-112). CRC Press.

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General*, *137*(2), 201. doi:10.1037/0096-3445.137.2.201

Fuster, J. M. (2001). The prefrontal cortex—an update: Time is of the essence. *Neuron*, *30*(2), 319-333. doi:10.1016/S0896-6273(01)00285-9

Gioia, G. A., & Isquith, P. K. (2004). Ecological assessment of executive function in traumatic brain injury. *Developmental Neuropsychology*, *25*, 135-158. doi:10.1080/87565641.2004.9651925

Gläscher, J., Adolphs, R., Damasio, H., Bechara, A., Rudrauf, D., Calamia, M., ... & Tranel, D. (2012). Lesion mapping of cognitive control and value-based decision making in the prefrontal cortex. *Proceedings of the National Academy of Sciences*, *109*(36), 14681-14686. doi:10.1073/pnas.1206608109

Goldberg, E., & Podell, K. (2000). Adaptive decision making, ecological validity, and the frontal lobes. *Journal of Clinical and Experimental Neuropsychology*, *22*(1), 56-68. doi:10.1076/1380-3395(200002)22

Grace, J., & Malloy, P. F. (2000). *Frontal systems behavior scale: professional manual*. Psychological Assessment Resources, Incorporated.

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, *38*(4), 404. doi:10.1037/h0059831

Gruber, S. A., Sagar, K. A., Dahlgren, M. K., Racine, M., & Lukas, S. E. (2012). Age of onset of marijuana use and executive function. *Psychology of Addictive Behaviors*, *26*(3), 496. doi:10.1037/a0026269.Age

Hagen, E., Erga, A. H., Hagen, K. P., Nesvag, S. M., McKay, J. R., Lundervold, A. J., & Walderhaug, E. (2016). Assessment of executive function in patients with substance use disorder: A comparison of inventory and performance-based assessment. *Journal of Substance Abuse Treatment, 66*, 1-8. doi:10.1016/j.jsat.2016.02.010

Hanks, R. A., Rapport, L. J., Millis, S. R., & Deshpande, S. A. (1999). Measures of executive functioning as predictors of functional ability and social integration in a rehabilitation sample. *Archives of Physical Medicine and Rehabilitation, 80*(9), 1030-1037. doi:10.1016/s0003-9993(99)90056-4

Happé, F., Booth, R., Charlton, R., & Hughes, C. (2006). Executive function deficits in autism spectrum disorders and attention-deficit/hyperactivity disorder: Examining profiles across domains and ages. *Brain and Cognition*, *61*(1), 25-39. doi:10.1016/j.bandc.2006.03.004

Heaton, R. K. (1993). Wisconsin card sorting test: computer version 2. *Odessa: Psychological Assessment Resources*.

Hutton, S. B., Puri, B. K., Duncan, L. J., Robbins, T. W., Barnes, T. R. E., & Joyce, E. M. (1998). Executive function in first-episode schizophrenia. *Psychological Medicine*, *28*(02), 463-473.

Inquisit 5 [Computer software]. (2016). Retrieved from http://www.millisecond.com.

Isquith, P. K., Roth, R. M., & Gioia, G. (2013). Contribution of rating scales to the assessment of executive functions. *Applied Neuropsychology: Child, 2*, 125-132 doi:1080/21622965.2013.748389

Jennison, K. M. (2004). The short□term effects and unintended long□term consequences of binge drinking in college: A 10□year follow□up study. *The American Journal of Drug and Alcohol Abuse*, *30*(3), 659-684. doi:10.1081/ADA-200032331

Johnston, L. D., O'Malley, P. M., Bachman, J. G., Schulenberg, J. E., & Miech, R. A. (2014). Monitoring the FUTURE: College students & adults ages 19 – 55. *Monitoring the Future*, *2*. Retrieved from https://files.eric.ed.gov/fulltext/ED514370.pdf

Kibby, M., Schmitter-Edgecombe, M., & Long, C. J. (1998). Ecological validity of neuropsychological tests: Focus on the California Verbal Learning Test and the Wisconsin Card Sorting Test. *Archives of Clinical Neuropsychology*, *13*(6), 523–534. doi:10.1016/S0887-6177(97)00038-3

Krikorian, R., Bartok, J., & Gay, N. (1994). Tower of London Procedure: A Standard Method and Developmental Data. *Journal of Clinical and Experimental Neuropsychology, 16*(2), 840-850. doi:10.1080/01688639408402697

Kokotailo, P. K., Egan, J., Gangnon, R., Brown, D., Mundt, M., & Fleming, M. (2004). Validity of the Alcohol Use Disorders Identification Test in college students. *Alcoholism: Clinical and Experimental Research, 28*(6), 914-920. doi:10.1097/01.alc.0000128239.87611.f5

Latzman, R. D., & Markon, K. E. (2010). The factor structure and age-related factorial invariance of the Delis-Kaplan Executive Function System (D-KEFS). *Assessment*, *17*(2), 172-184. doi:10.1177/1073191109356254

Lezak, M. L., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment.* New York, NY: Oxford.

Lundqvist, T. (2005). Cognitive consequences of cannabis use: Comparison with abuse of stimulants and heroin with regard to attention, memory and executive functions. *Pharmacology Biochemistry and Behavior, 81*(2), 319-330. doi:10.1016/j.pbb.2005.02.017

Luu, P., Collins, P., & Tucker, D. M. (2000). Mood, personality, and self-monitoring: Negative affect and emotionality in relation to frontal lobe mechanisms of error monitoring. *Journal of Experimental Psychology: General*, *129*(1), 43. doi:10.1037/0096-3445.129.1.43

Lynam, D. R., Smith, G. T., Whiteside, S. P., & Cyders, M. A. (2006). The UPPS-P: Assessing five personality pathways to impulsive behavior. *West Lafayette, IN: Purdue University*.

Lyvers, M. (2000). " Loss of control" in alcoholism and drug addiction: A neuroscientific interpretation. *Experimental and Clinical Psychopharmacology*, *8*(2), 225. doi:10.1037/1064-1297.8.2.225

Magid, V., & Colder, C. R. (2007). The UPPS Impulsive Behavior Scale: Factor structure and associations with college drinking. *Personality and Individual Differences*, *43*(7), 1927-1937. doi:10.1016/j.paid.2007.06.013

Mahone, E. M., Koth, C. W., Cutting, L., Singer, H. S., & Denckla, M. B. (2001). Executive function in fluency and recall measures among children with Tourette syndrome or ADHD. *Journal of the International Neuropsychological Society*, *7*, 102-111. doi:10.1017/S1355617701711101

Mangeot, S., Armstrong, K., Colvin, A. N., Yeates, K. O., & Taylor, H. G. (2002). Long-term executive function deficits in children with traumatic brain injuries: Assessment using the Behavior Rating Inventory of Executive Function (BRIEF). *Child Neuropsychology*, *8*(4), 271-284. doi:10.1076/chin.8.4.271.13503

McCormick, D. A., & Bal, T. (1994). Sensory gating mechanisms of the thalamus. *Current opinion in neurobiology*, *4*(4), 550-556. doi:10.1016/0959-4388(94)90056-6

Meil, W. M., LaPorte, D., & Stewart, P. (2012). Substance dependence as a neurological disorder. In K. S. Chen (Ed.), *Advanced Topics in Neurological Disorders*. InTech. Available from: http://www.intechopen.com/books/advanced-topics-in-neurological-disorders/substance- dependence-as-a-neurological-disorder

Meil, W. M., LaPorte, D. J., Mills, J. A., Sesti, A., Collins, S. M., & Stiver, A. G. (2016). Sensation seeking and executive deficits in relation to alcohol, tobacco, and marijuana use frequency among university students: Value of ecologically based measures. *Addictive Behaviors*. doi:10.1016/j.addbeh.2016.06.014

Miech, R. A., Bachman, J. G., & Schulenberg, J. E. (2015). Monitoring the FUTURE 2015 overview key findings on adolescent drug use. *Monitoring the Future.* Retrieved from http://www.monitoringthefuture.org

Mitchell, M., & Miller, L. S. (2008). Prediction of functional status in older adults: The ecological validity of four Delis-Kaplan Executive Function System tests. *Journal of Clinical and Experimental Neuropsychology*, *30*(6), 683–690. doi:10.1080/13803390701679893

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, a H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. doi:10.1006/cogp.1999.0734

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. doi:10.1177/0963721411429458

Monsell, S. (1996). Control of mental processes. *Unsolved mysteries of the mind: Tutorial essays in cognition*, 93-148.

Moritz, S., Birkner, C., Kloss, M., Jahn, H., Hand, I., Haasen, C., & Krausz, M. (2002). Executive functioning in obsessive–compulsive disorder, unipolar depression, and schizophrenia. *Archives of Clinical Neuropsychology*, *17*(5), 477-483. doi:10.1093/arclin/17.5.477

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British journal of psychology*, *81*(2), 111-121. doi:10.1111/j.2044-8295.1990.tb02349.x

Nigg, J. T., Wong, M. M., Martel, M. M., Jester, J. M., Puttler, L. I., Glass, J. M., … Zucker, R. a. (2006). Poor response inhibition as a predictor of problem drinking and illicit drug use in adolescents at risk for alcoholism and other substance use disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *45*(4), 468–475. doi:10.1097/01.chi.0000199028.76452.a9

Noël, X., Paternot, J., Van der Linden, M., Sferrazza, R., Verhas, M., Hanak, C., ... & Verbanck, P. (2001). Correlation between inhibition, working memory and delimited frontal area blood flow measured by 99MTC–bicisate spect in alcohol–dependent patients. *Alcohol and Alcoholism*, *36*(6), 556-563. doi:10.1093/alcalc/36.6.556

Norman, M. A., Moore, D. J., Taylor, M., Franklin, D., Cysique, L., Ake, C., … & Heaton, R. K. (2011). Demographically corrected norms for African Americans and Caucasians on the Hopkins Verbal Learning Test-Revised, Brief Visuospatial Memory Test-Revised, Stroop Color and Word Test, and Wisconsin Card Sorting Test 64-Card Version. *Journal of*

Clinical and Experimental Neuropsychology, 33(7), 793-804. doi:10.1080/13803395.2011.559157

Norman, D. A., & Shallice, T. (1986). Attention to action. In *Consciousness and self-regulation* (pp. 1-18). Springer US. doi: 10.1007/978-1-4757-0629-1_1

Norris, G., & Tate, R. L. (2000). The Behavioural Assessment of the Dysexecutive Syndrome (BADS): Ecological, Concurrent and Construct Validity. *Neuropsychological Rehabilitation*, *10*(1), 33–45. doi:10.1080/096020100389282

Northcote, J., & Livingston, M. (2011). Accuracy of self-reported drinking: Observational verification of "last occasion" drink estimates of young adults. *Alcohol and Alcoholism, 46*(6), 709-713. doi:10.1093/alcalc/agr138

Odhuba, R. A., van den Broek, M. D., & Johns, L. C. (2005). Ecological validity of measures of executive functioning. *The British Journal of Clinical Psychology*, *44*, 269–278. http://doi.org/10.1348/014466505X29431

Peeters, M., Janssen, T., Monshouwer, K., Boendermaker, W., Pronk, T., Wiers, R., & Vollebergh, W. (2015). Weaknesses in executive functioning predict the initiating of adolescents' alcohol use. *Developmental Cognitive Neuroscience*, 1–8. doi:10.1016/j.dcn.2015.04.003

Peeters, M., Vollebergh, W. A. M., Wiers, R. W., & Field, M. (2014). Psychological changes and cognitive impairments in adolescent heavy drinkers. *Alcohol and Alcoholism*, *49*(2), 182–186. doi:10.1093/alcalc/agt162

Qualtrics (computer software). (2016). Provo, UT. Retrieved from http://www.qualtrics.com.

Rabin, L. A., Barr, W. B., & Burton, L. A. (2005). Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, and APA

Division 40 members. *Archives of Clinical Neuropsychology*, *20*(1), 33–65. doi:10.1016/j.acn.2004.02.005

Rasmussen, C. (2005). Executive functioning and working memory in fetal alcohol spectrum disorder. *Alcoholism: Clinical and Experimental Research*, *29*(8), 1359-1367. doi:10.1097/01.alc.0000175040.91007.d0

Raspelli, S., Pallavicini, F., Carelli, L., Morganti, F., Pedroli, E., Cipresso, P., ... & Riva, G. (2012). Validating the Neuro VR-based virtual version of the Multiple Errands Test: preliminary results. *Presence: Teleoperators and Virtual Environments*, *21*(1), 31-42. doi:10.1162/PRES_a_00077

Razani, J., Boone, K., Lesser, I., & Weiss, D. (2004). Effects of cigarette smoking history on cognitive functioning in healthy older adults. *The American Journal of Geriatric Psychiatry*, *12*(4), 404-411. doi:10.1097/00019442-200407000-00008

Ready, R. E., Stierman, L., & Paulsen, J. S. (2001). Ecological validity of neuropsychological and personality measures of executive functions. *The Clinical Neuropsychologist*, *15*(3), 314-323. doi:10.1076/clin.15.3.314.10269

Reid-Arndt, S. A., Nehl, C., & Hinkebein, J. (2007). The Frontal Systems Behaviour Scale (FrSBe) as a predictor of community integration following a traumatic brain injury. *Brain Injury*, *21*(13-14), 1361-1369. doi:10.1080/02699050701785062

Reinert, D. F., & Allen, J. P. (2002). The alcohol use disorders identification test (AUDIT): a review of recent research. *Alcoholism: Clinical and Experimental Research*, *26*(2), 272-279. doi:10.1111/j.1530-0277.2002.tb02534.x

Renison, B., Ponsford, J., Testa, R., Richardson, B., & Brownfield, K. (2012). The ecological and construct validity of a newly developed measure of executive function: the Virtual

Library Task. *Journal of the International Neuropsychological Society*, *18*(03), 440-450. doi:10.1017/S1355617711001883

Robinson, H., Calamia, M., Gläscher, J., Bruss, J., & Tranel, D. (2014). Neuroanatomical correlates of executive functions: A neuropsychological approach using the EXAMINER battery. *Journal of the International Neuropsychological Society : JINS*, *20*(1), 52–63. doi:10.1017/S135561771300060X

Roth, R. M., Isquith, P., & Gioia, G. A. (2005). *Behavior rating inventory of executive function--adult version*. Lutz, FL: Psychological Assessment Resources.

Saunders, J. B., Aasland, O. G., Babor, T. F., De la Fuente, J. R., & Grant, M. (1993). Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption□II. *Addiction*, *88*(6), 791-804. doi:10.1111/j.1360-0443.1993.tb02093.x

Sbordone, R. J., & Long, C. (1996). *Ecological validity of neuropsychological testing*. CRC Press.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *298*(1089), 199-209. doi:10.1098/rstb.1982.0082

Shallice, T. I. M., & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, *114*(2), 727-741. doi:10.1093/brain/114.2.727

Simons, J. S., & Spiers, H. J. (2003). Prefrontal and medial temporal lobe interactions in long-term memory. *Nature Reviews Neuroscience*, *4*(8), 637-648. doi:10.1038/nrn1178

Smith-Seemiller, L., Franzen, M. D., & Bowers, D. (1997). Use of Wisconsin Card Sorting Test short forms in clinical samples. *The Clinical Neuropsychologist, 11*, 421-427. doi:10.1093/arclin/16.5.489

Spinella, M. (2003). Relationship between drug use and prefrontal□associated traits. *Addiction Biology*, *8*(1), 67-74. doi:10.1080/1355621031000069909

Spooner, D. M., & Pachana, N. A. (2006). Ecological validity in neuropsychological assessment: a case for greater consideration in research with neurologically intact populations. *Archives of Clinical Neuropsychology*, *21*(4), 327-337. doi:10.1016/j.acn.2006.04.004

Stout, J. C., Ready, R. E., Grace, J., Malloy, P. F., & Paulsen, J. S. (2003). Factor analysis of the frontal systems behavior scale (FrSBe). *Assessment*, *10*(1), 79-85. doi:10.1177/1073191102250339

Strauss, E., Sherman, E., & Spreen, O. (2006). *A compendium of neuropsychological tests*. New York, NY: Oxford.

Stuss, D. T. (2011). Functions of the frontal lobes: Relation to executive functions. *Journal of the International Neuropsychological Society*, *17*(5), 759–765. doi:10.1017/S1355617711000695

Stuss, D. T., & Alexander, M. P. (2000). Executive functions and the frontal lobes: A conceptual view. *Psychological Research, 63*, 289–298. doi:1007/s004269900007

Stuss, D. T., & Buckle, L. (1992). Traumatic brain injury: Neuropsychological deficits and evaluation at different stages of recovery and in different pathologic subtypes. *The Journal of Head Trauma Rehabilitation*, *7*(2), 40-49. Retrieved from https://journals.lww.com/headtraumarehab

Tarter, R. E., Kirisci, L., Mezzich, A., Cornelius, J. R., Pajer, K., Vanyukov, M., . . . Clark, D. (2003). Neurobehavioral disinhibition in childhood predicts early age at onset of substance use disorder. *American Journal of Psychiatry, 160*(6), 1078-1085. doi:10.1176/appi.ajp.160.6.1078

Troyer, A. K., Moscovitch, M., Winocur, G., Alexander, M. P., & Stuss, D. (1998). Clustering and switching on verbal fluency: The effects of focal frontal-and temporal-lobe lesions. *Neuropsychologia*, *36*(6), 499-504. doi:10.1016/S0028-3932(97)00152-8

Vendrell, P., Junqué, C., Pujol, J., Jurado, M. A., Molet, J., & Grafman, J. (1995). The role of prefrontal regions in the Stroop task. *Neuropsychologia*, *33*(3), 341-352. doi:0.1016/0028-3932(94)00116-7

Verdejo-García, A., & Pérez-García, M. (2007). Ecological assessment of executive functions in substance dependent individuals. *Drug and Alcohol Dependence*, *90*(1), 48–55. doi:10.1016/j.drugalcdep.2007.02.010

Vik, P. W., Carrello, P., Tate, S. R., & Field, C. (2000). Progression of consequences among heavy-drinking college students. *Psychology of Addictive Behaviors*, *14*(2), 91. doi:10.1037/0893-164X.14.2.91

Völlm, B. A., Taylor, A. N., Richardson, P., Corcoran, R., Stirling, J., McKie, S., ... & Elliott, R. (2006). Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. *Neuroimage*, *29*(1), 90-98. doi:10.1016/j.neuroimage.2005.07.022

White, H. R., & Labouvie, E. W. (1989). Towards the assessment of adolescent problem drinking. *Journal of Studies on Alcohol*, *50*(1), 30-37. doi:10.15288/jsa.1989.50.30

White, H. R., & Labouvie, E. W. (2000). Longitudinal trends in problem drinking as measured by the Rutgers Alcohol Problem Index. *Alcoholism: Clinical and Experimental Research*, *24*(5 Suppl 1).

White, H. R., Labouvie, E. W., & Papadaratsakis, V. (2005). Changes in substance use during the transition to adulthood: A comparison of college students and their noncollege age peers. *Journal of Drug Issues, 35*, 281-305. doi:10.1177/002204260503500204

Whiteside, S. P., & Lynam, D. R. (2001). The Five Factor Model and impulsivity: Using a structural model of personality to understand impulsivity. *Personality and Individual Differences, 30*(4), 669-689. doi:0.1016/S0191-8869(00)00064-7

Whiteside, S. P., Lynam, D. R., Miller, J. D., & Reynolds, S. K. (2005). Validation of the UPPS impulsive behaviour scale: a four□factor model of impulsivity. *European Journal of Personality*, *19*(7), 559-574. doi:10.1002/per.556

Wilens, T. E., Martelon, M., Fried, R., Petty, C., Bateman, C., & Biederman, J. (2012). Do executive function deficits predict later substance use disorders among adolescents and young adults? *Journal of the American Academy of Child and Adolescent Psychiatry*, *50*(2), 141–149. doi:10.1016/j.jaac.2010.11.010.Do

Williams, P. G., Suchy, Y., & Kraybill, M. L. (2010). Five-factor model personality traits and executive functioning among older adults. *Journal of Research in Personality*, *44*(4), 485-491. doi:10.1016/j.jrp.2010.06.002

Wilson, B. A., Evans, J. J., Emslie, H., Alderman, N., & Burgess, P. (1998). The development of an ecologically valid test for assessing patients with a dysexecutive syndrome. *Neuropsychological Rehabilitation*, *8*(3), 213–228. doi:10.1080/713755570

Wood, R. L., & Liossi, C. (2006). The ecological validity of executive tests in a severely brain injured sample. *Archives of Clinical Neuropsychology*, *21*(5), 429–437. doi:10.1016/j.acn.2005.06.014

Appendix A

CREFT Transcript

Your Transcript (courses you have already taken)

Term: Freshman Year – Fall Semester:

Course	Grade	Credits Received
UCO 1200	A	3
ART 2022	B	3
R C 1000	A	3
HP 1105	A	2
MAT 1025	A	4

Term: Freshman Year – Spring Semester:

Course	Grade	Credits Received
FER 1000	B	3
MUS 2022	B	3
CHE 1101	B	3
CHE 1110	C	1
BIO 1801	B	4
HIS 1101	A	3

Term: Sophomore Year – Fall Semester:

Course	Grade	Credits Received
ECO 2030	A	3
R C 2001	A	3
CHE 1102	B	3
CHE 1120	B	1
HIS 1102	A	3
BUS 1050	A	3

Term: Sophomore Year – Spring Semester:

Course	Grade	Credits Received
FER 3000	A	4
CHE 2101	B	3
CHE 2102	B	1
ACC 1050	A	3
CHE 2210	B	3
CHE 2211	C	1

Appendix B

CREFT Major Requirements

2015-2016

Bachelor of Science (BS)
Degree Code 118A

Program of Study for
FERMENTATION SCIENCES

I. GENERAL EDUCATION CURRICULUM.....44

CHE 1101, 1110, 1102, 1120, and MAT 1110 will fulfill General Education requirements. Some other general education requirements may be double-counted in the major with departmental approval. Please see your advisor for information.

II. MAJOR REQUIREMENTS *(Not including 15 s.h. already counted in I, above)* **.....73-75**

2.0 major GPA is required for graduation. Major GPA calculation includes all courses in the major department, plus any other courses under II.

A. Fermentation Science (18 semester hours)

FER 1000	_____ (3)	Principles of Fermentation Sciences
FER 2000	_____ (1)	Social Implications of Fermented Beverages
FER 3200	_____ (3)	Facility Design and Operation **[WID]** *(Pre: RC 2001)*
FER 4300	_____ (3)	Sensory Analysis of Wine and Beer **[CAP]** *(Pre: FER 3200, 4100 or FER/CHE 4200, STT 2810)*

Choose two of the following:

FER 3000	_____ (4)	Viticulture: Vine Physiology and Vineyard Establishment *(Pre: BIO 1801)*
FER 4100	_____ (4)	Wine Production and Analysis *(Pre: CHE 2210 & 2211)*
CHE/FER 4200	_____ (4)	Brewing Science and Analysis *(Pre: CHE 2210 & 2211)*

B. Chemistry (19 semester hours)

CHE 1101/1110	_____ (4)	Introductory Chemistry I & Introductory Chemistry I Lab
CHE 1102/1120	_____ (4)	Introductory Chemistry II & Introductory Chemistry II Lab *(Pre: CHE 1101 & 1110)*
CHE 2101	_____ (3)	Fundamentals of Organic Chemistry *(Pre: CHE 1102 & 1120; Co: CHE 2102)*
CHE 2102	_____ (1)	Fundamentals of Organic Chemistry Lab *(Pre: CHE 1102 & 1120; Co:2101)*
CHE 2210	_____ (3)	Quantitative Analysis *(Pre: CHE 1102 & 1120; Co: CHE 2211)*
CHE 2211	_____ (1)	Quantitative Analysis Lab *(Co: CHE 2210)*
CHE 4580	_____ (3)	Biochemistry I *(Pre: CHE 2101 or 2202)*

C. Biology (11 semester hours)

BIO 1801	_____ (4)	Biological Concepts I *(Co: CHE 1101)*
BIO 2600	_____ (3)	Cell Biology *(Pre: BIO 1801, CHE 1102)*
BIO 3308	_____ (4)	Microbiology *(Pre: CHE 1102 & 1120; BIO 1801; 2400/2600/2700)*

D. Mathematics and Economics (10 semester hours)

MAT 1110	_____ (4)	Calculus with Analytic Geometry I *(Pre: MAT 1025 w/min grade of C-)*
STT 2810	_____ (3)	Introduction to Statistics *(Pre: MAT 1010)*
ECO 2030	_____ (3)	Principles of Microeconomics *(may be used in Gen Ed depending on choices)*

E. Focus area courses (12 - 14 semester hours)

NUT 1202 _____ (3) Basic Food Science **AND** NUT 3210_____ (3) Beverage Management

Choose two of the following:

SD 3100	_____ (3)	Principles of Agroecology	S D 3150_____	(4) Soil & Soil Fertility Management
SD 4200	_____ (4)	Ecologically-Based Pest Mgmt *(Pre: SD 3100)*	TEC 3605_____	(3) Sustainable Resrce Mgmt *(Pre: TEC 2029)*
TEC 3606	_____ (3)	Sust Water & Wastewater Tech *(Pre: TEC 2029)*	TEC 4700_____	(3) Biofuels Tech *(Pre: TEC 3638; Sr. stndg)*

F. Business courses (15 semester hours)

BUS 1050	_____ (3)	Introduction to Business	ACC 1050_____	(3) Survey of Accounting
ENG 3700	_____ (3)	Technical Writing **[WID-ENG]** *(Pre: ENG 2001, 3090)*	MKT 3050_____	(3) Principles of Marketing *(Pre: ECO 2030; 57 earned hours*

Choose one of the following:

HOS 2000	_____ (3)	Survey of the Hospitality & Tourism Industry	MGT 3010_____	(3) Survey of Management *(Pre: Jr standing)*
ENT 3060	_____ (3)	Opportunity & Entrepreneurship *(Pre: Jr standing)*		

III. MINOR (optional) Students are encouraged to pursue a minor in one of the following areas: Chemistry, Biology, Marketing, or Entrepreneurship.

IV. ELECTIVES (taken to total 122 hours for the degree) 3-5

2 semester hours of free electives must be outside the major discipline. **122**

Electives recommended by department:

MGT 3170 (3) Creativity & Design *(Pre: MGT 3060)*, ENT 3190 (3) Int'l Entrepreneurship *(Pre: ENT 3060)*, ENT 4650 (3) Venture Creation *(Pre: FIN 3680; ENT 3060; MKT 3050; "C" in WID)*, CHE 2526 (1) Chemical Safety *(Pre: 1101/1110)*, HOS 3700 (3) Hospitality Management Operations I *(Pre: HOS 2000)*, HOS 4040 (3) Destination Management *(Pre: WID w/min grade C)*

Appendix C

CREFT General Studies Requirements

APPALACHIAN STATE UNIVERSITY GENERAL EDUCATION 2015-2016 **(44 Semester Hours)** Updated 2-26-15 http://generaleducation.appstate.edu

FIRST YEAR SEMINAR
___UCO 1200 or HON 1515 or WRC 1103 (3 of 6 s.h. will count here)
or ___UCO MET* (0 s.h.)
*open to students who transfer at least 30 s.h. AND are classified as transfer students by Admissions. **The 30 s.h. must be completed prior to enrollment at ASU.** Students using this option will have 41 s.h. of General Education.

WRITING ACROSS THE CURRICULUM (6 s.h. required)
___R C 1000 or LLC 1000 or WRC 1103 (3 of 6 s.h. will count here)
___R C 2001 or WRC 2001

MAJOR REQUIREMENTS (hours count in major requirements)
___Junior Writing in the Discipline ("WID" on major program of study)
___Senior Capstone Experience ("CAP" on major program of study)

QUANTITATIVE LITERACY (4 s.h. required)
1 s.h. courses (for students who transfer 3 s.h. of QL coursework): ___MAT 1005, ___STT 1805
If you have transferred 3 hours of QL coursework, please check with your advisor to see if your intended major requires a course which can count for the final QL hour.
3 s.h. courses (These courses count toward, but do not fully complete, this requirement.):
___ECO 2100, ___STT 1810, ___STT 2810, ___STT 3820
4 s.h. courses (These courses fully complete this requirement.): ___MAT 1010, ___MAT 1020, ___MAT 1025, ___MAT 1030, ___MAT 1110, ___C S 1445, ___STT 2820

WELLNESS LITERACY (2 s.h. required)- Choose from the following:
1 s.h. courses: ___, ___P E 1530-1549, P E 1700-1877
2 s.h. courses: ___H P 1105, ___HED 1000, ___DAN 1400, ___DAN 1410, ___DAN 1420, ___DAN 1430, ___DAN 2400, ___DAN 2410, ___DAN 2420, ___DAN 3280, ___DAN 3480, ___DAN 3580, ___MSL 1101
3 s.h. courses (additional hour may count as elective): ___DAN 4460, ___HPE 4320, ___NUT 2202, ___P E 1718, ___P E 1768, ___P E 1769

SCIENCE INQUIRY requires 8 semester hours from one theme (underlined). Courses in themes marked with an * must be taken sequentially. Check the course descriptions for any pre- and/or co-requisites.

☐Biology in Society: BIO ___1201, ___1202, AND ___1203
OR (for transfers with credit) BIO ___1201/1204 AND ___1202/1205
☐The Blue Planet: GLY___1104 AND ___1105
* ☐Chemistry Connections to Our Changing World: CHE ___ 1101/ 1110 AND ___1102/1120
☐Global Environmental Change: ___BIO 1103, ___GHY 1011, ___GHY 1012, ___GLY 1103, ___GLY 1104
* ☐How Things Work: PHY ___ 1101 AND ___ 1102
☐Life, Earth and Evolution: ___ANT 1430 AND ___GLY 1102
* ☐The Physics of Our Technological World: PHY ___ 1103 AND ___ 1104
☐Physics of Self Expression: PHY ___ 1101, ___1810, ___1812, ___1814
* ☐Physics with Calculus: PHY ___ 1150 AND ___ 1151
☐Restless Planet: Earth, Environment and Evolution: GLY ___1101, ___1102, ___1103
* ☐Voyages Through the Cosmos: AST ___ 1001 AND ___1002

NOTE: A 1 hour lab option **may** be available for students who have transferred in a 3 hour science lecture course. Please contact the Office of General Education at (828) 262-2028 or gened@appstate.edu for more information.

DESIGNATIONS (3 s.h. of each required); may be taken in Integrative Learning Experience or Liberal Studies Experience:
___Fine Arts (FA), ___Historical Studies (HS),
___Literary Studies (LS), ___ Social Sciences (SS)

INTEGRATIVE LEARNING EXPERIENCE: Complete 9 semester hours from a single theme. Students must take courses from at least two discipline prefixes in the chosen theme with the exceptions of "Appalachian Mountains: Community, Culture, and Land" and "Experiencing Inquiry: How to Ask Questions."

☐**American Culture: Past and Present:** ___COM 3130, ___ENG 2120 (LS), ___HIS 2525 (HS), ___IDS 2000, ___REL 3110, ___S W 2020 (SS)
☐**Appalachian Mountains: Community, Culture, and Land:** ___A S/MUS 2016 (FA), ___A S 2200 (LS), ___A S/GLY 2301, ___HIS 3726 (HS), ___SOC 3710 (SS)
☐**Critical Consciousness: Learning for Equity and Justice:** ___C I 2250, ___C I 2350, ___FDN 2150, ___FDN 2250 (LS), ___FDN 2350 (SS) (Both C I and FDN prefixes are required.)
☐**Cultivating Creative Expression:** ___ART 2022 (FA), ___ENG 2360 (LS), ___MUS 2022 (FA), ___THR 2022 (FA)
☐**Experiencing Inquiry: How to Ask Questions:** ___WRC 2201 (LS), ___WRC 2202 (HS), ___WRC 3203 (FA)
☐**Expressions of Culture:** ___COM 2124 (SS), ___DAN 2020 (FA), ___DAN 2030 (FA), ___LLC 2050, ___THR 2020 (FA)
☐**From Empire to Globalization:** ___ANT 1415 (SS), ___ENG 2040 (LS), ___GLS 2000, ___GWS 2525, ___HIS 1400 (HS), ___P S 2120
☐**How We Know What We Know About the Past: Method, Evidence, Knowledge:** ___ANT 1420 (SS), ___GLY 1842, ___HIS 2312 (HS), ___HIS 2320 (HS), ___LLC 2045 (LS)
☐**How We Tell Stories:** ___C I/ITC 2010, ___ENG 2170, ___IDS 3210, ___LLC 2025 (LS), ___PHL 1502, ___REL 2020 (LS), ___THR 2005 (FA), ___THR 3640 (FA)
☐**Imagination, Innovation, and Meaning:** ___DAN 2010 (FA), ___HIS 1110 (HS), ___MUS 2018 (FA), ___PHL 2013, ___THR 2010 (FA)
☐**Intersections: Race, Class, and Gender:** ___ANT 2420 (SS), ___ENG 2130 (LS), __GWS 2421, ___MUS 2023 (FA), ___PHL 3050, ___P S 3410, ___SOC 2050 (SS), ___S W 2615
☐**Las Américas:** ___ANT 2300 (SS), ___GHY 3014 (SS), ___GLS 3020 (HS), ___HIS 2301 (HS), ___HIS 2302 (HS), ___LLC 2040 (LS), ___REL 3100, ___WRC 2400 (FA)
☐**Revolutions: Social and Political:** ___ART 2019 (FA), ___HIS 1501 (HS), ___HIS 2340 (HS), ___IDS 3650, ___MUS 2015 (FA), ___ PHL 3030, ___SOC 1100 (SS), ___THR 2017 (FA)
☐**Social Relations Across Contexts:** ___COM 2121, ___HIS 1120 (HS), ___PSY 1100, ___PSY 2213 (SS), ___SOC 1110 (SS),
☐**Sustainability and Global Resources:** ___FCS 2110, ___FER 1000, ___GHY 1010, ___IDS 3010, ___PHL 2015, ___PHY 1830, ___S D 2400, ___TEC 2029 (SS)
☐**War and Peace:** ___ANT 2222 (SS), ___GLS 2350, ___HIS 3158 (HS), ___HIS 3823 (HS), ___P S 4225 (SS), ___REL 3170, ___SOC 3800 (SS)

LIBERAL STUDIES EXPERIENCE: Complete 12 semester hours from at least three discipline prefixes.

☐**Apparel Design and Merchandising (ADM):** ___ 1000 (SS)
☐**Anthropology (ANT):** ___2100 (SS), ___2235 (SS), ___2400 (SS), ___2430 (SS), ___2700 (SS), ___2800 (SS)
☐**Appalachian Studies (A S):** ___2020, ___2025, ___2411 (SS)
☐**Arabic (ARB):** ___1050
☐**Art (ART):** ___2011 (FA), ___2016 (FA), ___2030 (FA), ___2130 (FA)
☐**Astronomy (AST):** ___1050
☐**Chinese (CHN):** ___1050
☐**Communication (COM):** ___2105, ___3300, ___3315
☐**Criminal Justice (C J):** ___3450 (SS)
☐**Dance (DAN):** ___3430 (FA), ___3435 (FA)
☐**Economics (ECO):** ___2030 (SS), ___2040 (SS), ___2620 (SS)
☐**English (ENG):** ___2030 (LS), ___2050 (LS), ___2060 (LS), ___2350 (LS) , ___ENG/S D 3715 (LS)
☐**Family & Consumer Sciences (FCS):** ___2103 (SS), ___2111
☐**Finance (FIN):** ___2860
☐**French (FRE):** ___1050 OR ___1060 (3 of 6 s.h. count in General Education)
☐**Gender, Women's, and Sexuality Studies (GWS):** ___2600
☐**Geography (GHY):** ___1020 (SS), ___1040 (SS)
☐**German (GER):** ___1050 OR ___1060 (3 of 6 s.h. count in General Education)
☐**History (HIS):** ___1101 (HS), ___1102 (HS), ___1130 (HS), ___1200 (HS), ___1700 (HS), ___3728 (HS)
☐**Industrial Design (IND):** ___2012 (FA)
☐**Interdisciplinary Studies (IDS):** ___IDS/WRC 2302, ___ 3025, ___ 3250
☐**Japanese (JPN):** ___1050
☐**Latin (LAT):** ___1050
☐**Music (MUS):** ___2011 (FA), ___2014 (FA), ___2611 (HS) (2 s.h.), ___2612 (HS) (2 s.h.), ___2613 (FA), ___2616 (FA), ___3611 (HS) (2 s.h.)
☐**Nutrition (NUT):** ___2351 (SS)
☐**Philosophy (PHL):** ___1000, ___1100, ___1501, ___1503, ___2000, ___3000 (HS), ___3013, ___3015, ___3020, ___3200 (HS), ___3550, ___3600
☐**Photography (PHO):** ___2032
☐**Planning (PLN):** ___2410 (SS)
☐**Political Science (P S):** ___1100 (SS), ___1200 (SS), ___2130 (SS)
☐**Portuguese (POR):** ___1050
☐**Psychology (PSY):** ___1200 (SS)
☐**Recreation Management (R M):** ___2100, ___2140 (HS)
☐**Religion (REL):** ___1010 (LS), ___1100 (SS), ___1110, ___1115, ___1120, ___2010 (LS), ___2030 (LS), ___2110 (HS), ___2120 (HS), ___2130, ___3120
☐**Russian (RSN):** ___1050
☐**Sociology (SOC):** ___1000 (SS), ___2020 (SS), ___2850 (SS), ___3100 (SS)
☐**Spanish (SNH):** ___1050 OR ___1060 (3 of 6 s.h. count in General Education)
☐**Sustainable Development (S D):** ___S D/ENG 3715 (LS)
☐**Technology (TEC):** ___2601
☐**Theatre (THR):** ___2300 (FA), ___2610 (FA), ___3730 (FA)
☐**Watauga Residential College (WRC):** ___1104 (6 s.h.) (This course is available only to Watauga Residential College students), ___WRC/IDS 2302, ___2401, ___2403 (LS), ___3401 (LS), ___3402, ___3403 (FA), ___3665 (LS)

Appendix D

CREFT Work Schedule

Work Schedule at College Bookstore Work Schedule (5 minute walk from all classes)

Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**
9:00 AM-1:00 PM	3:00 PM-9:00 PM	10:00 AM-2:00 PM	3:00 PM-9:00 PM	Off

Appendix E

UPPS-P Impulsivity Scale

UPPS-P

Below are a number of statements that describe ways in which people act and think. For each statement, please indicate how much you agree or disagree with the statement. If you **Agree Strongly** circle **1**, if you **Agree Somewhat** circle **2**, if you **Disagree somewhat** circle **3**, and if you **Disagree Strongly** circle **4**. Be sure to indicate your agreement or disagreement for every statement below. Also, there are questions on the following pages.

	Agree Strongly	Agree Some	Disagree Some	Disagree Strongly
1. I have a reserved and cautious attitude toward life.	1	2	3	4
2. I have trouble controlling my impulses.	1	2	3	4
3. I generally seek new and exciting experiences and sensations.	1	2	3	4
4. I generally like to see things through to the end.	1	2	3	4
5. When I am very happy, I can't seem to stop myself from doing things that can have bad consequences.	1	2	3	4
6. My thinking is usually careful and purposeful.	1	2	3	4
7. I have trouble resisting my cravings (for food, cigarettes, etc.).	1	2	3	4
8. I'll try anything once.	1	2	3	4
9. I tend to give up easily.	1	2	3	4
10. When I am in great mood, I tend to get into situations that could cause me problems.	1	2	3	4
11. I am not one of those people who blurt out things without thinking.	1	2	3	4
12. I often get involved in things I later wish I could get out of.	1	2	3	4
13. I like sports and games in which you have to choose your next move very quickly.	1	2	3	4
14. Unfinished tasks really bother me.	1	2	3	4
15. When I am very happy, I tend to do things that may cause problems in my life.	1	2	3	4
16. I like to stop and think things over before I do them.	1	2	3	4
17. When I feel bad, I will often do things I later regret in order to make myself feel better now.	1	2	3	4
18. I would enjoy water skiing.	1	2	3	4
19. Once I get going on something I hate to stop.	1	2	3	4
20. I tend to lose control when I am in a great mood.	1	2	3	4
21. I don't like to start a project until I know exactly how to proceed.	1	2	3	4

Please go to the next page

	Agree Strongly	Agree Some	Disagree Some	Disagree Strongly
22. Sometimes when I feel bad, I can't seem to stop what I am doing even though it is making me feel worse.	1	2	3	4
23. I quite enjoy taking risks.	1	2	3	4
24. I concentrate easily.	1	2	3	4
25. When I am really ecstatic, I tend to get out of control.	1	2	3	4
26. I would enjoy parachute jumping.	1	2	3	4
27. I finish what I start.	1	2	3	4
28. I tend to value and follow a rational, "sensible" approach to things.	1	2	3	4
29. When I am upset I often act without thinking.	1	2	3	4
30. Others would say I make bad choices when I am extremely happy about something.	1	2	3	4
31. I welcome new and exciting experiences and sensations, even if they are a little frightening and unconventional.	1	2	3	4
32. I am able to pace myself so as to get things done on time.	1	2	3	4
33. I usually make up my mind through careful reasoning.	1	2	3	4
34. When I feel rejected, I will often say things that I later regret.	1	2	3	4
35. Others are shocked or worried about the things I do when I am feeling very excited.	1	2	3	4
36. I would like to learn to fly an airplane.	1	2	3	4
37. I am a person who always gets the job done.	1	2	3	4
38. I am a cautious person.	1	2	3	4
39. It is hard for me to resist acting on my feelings.	1	2	3	4
40. When I get really happy about something, I tend to do things that can have bad consequences.	1	2	3	4
41. I sometimes like doing things that are a bit frightening.	1	2	3	4
42. I almost always finish projects that I start.	1	2	3	4
43. Before I get into a new situation I like to find out what to expect from it.	1	2	3	4
44. I often make matters worse because I act without thinking when I am upset.	1	2	3	4
45. When overjoyed, I feel like I can't stop myself from going overboard.	1	2	3	4

Please go to the next page

	Agree Strongly	Agree Some	Disagree Some	Disagree Strongly
46. I would enjoy the sensation of skiing very fast down a high mountain slope.	1	2	3	4
47. Sometimes there are so many little things to be done that I just ignore them all.	1	2	3	4
48. I usually think carefully before doing anything.	1	2	3	4
49. When I am really excited, I tend not to think of the consequences of my actions.	1	2	3	4
50. In the heat of an argument, I will often say things that I later regret.	1	2	3	4
51. I would like to go scuba diving.	1	2	3	4
52. I tend to act without thinking when I am really excited.	1	2	3	4
53. I always keep my feelings under control.	1	2	3	4
54. When I am really happy, I often find myself in situations that I normally wouldn't be comfortable with.	1	2	3	4
55. Before making up my mind, I consider all the advantages and disadvantages.	1	2	3	4
56. I would enjoy fast driving.	1	2	3	4
57. When I am very happy, I feel like it is ok to give in to cravings or overindulge.	1	2	3	4
58. Sometimes I do impulsive things that I later regret.	1	2	3	4
59. I am surprised at the things I do while in a great mood.	1	2	3	4

Appendix F

Daily Drinking Questionnaire

For the next questions, indicate the requested drinking rates for <u>each day of a TYPICAL week in the last 30 days.</u> If no alcohol was consumed, enter zero.

How many standard drinks did you consume?

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
Monday																				
Tuesday																				
Wednesday																				
Thursday																				
Friday																				
Saturday																				
Sunday																				

How many hours did you spend drinking <u>each day of a TYPICAL week in the last 30 days?</u>

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
Monday																				
Tuesday																				
Wednesday																				
Thursday																				
Friday																				
Saturday																				
Sunday																				

Appendix G

Alcohol Use Disorders Identification Test

The Alcohol Use Disorders Identification Test: Self-Report Version

PATIENT: Because alcohol use can affect your health and can interfere with certain medications and treatments, it is important that we ask some questions about your use of alcohol. Your answers will remain confidential so please be honest. Place an X in one box that best describes your answer to each question.

Questions	0	1	2	3	4	
1. How often do you have a drink containing alcohol?	Never	Monthly or less	2-4 times a month	2-3 times a week	4 or more times a week	
2. How many drinks containing alcohol do you have on a typical day when you are drinking?	1 or 2	3 or 4	5 or 6	7 to 9	10 or more	
3. How often do you have six or more drinks on one occasion?	Never	Less than monthly	Monthly	Weekly	Daily or almost daily	
4. How often during the last year have you found that you were not able to stop drinking once you had started?	Never	Less than monthly	Monthly	Weekly	Daily or almost daily	
5. How often during the last year have you failed to do what was normally expected of you because of drinking?	Never	Less than monthly	Monthly	Weekly	Daily or almost daily	
6. How often during the last year have you needed a first drink in the morning to get yourself going after a heavy drinking session?	Never	Less than monthly	Monthly	Weekly	Daily or almost daily	
7. How often during the last year have you had a feeling of guilt or remorse after drinking?	Never	Less than monthly	Monthly	Weekly	Daily or almost daily	
8. How often during the last year have you been unable to remember what happened the night before because of your drinking?	Never	Less than monthly	Monthly	Weekly	Daily or almost daily	
9. Have you or someone else been injured because of your drinking?	No		Yes, but not in the last year		Yes, during the last year	
10. Has a relative, friend, doctor, or other health care worker been concerned about your drinking or suggested you cut down?	No		Yes, but not in the last year		Yes, during the last year	
					Total	

Appendix H

Rutgers Alcohol Problem Index

{Module Name} Module

Rutgers Alcohol Problem Index

Agency Name: _____ **Site Name:** _____

ID #: _____ **Date:** __ __ / __ __ / __ __ __ __

Different things happen to people when they are drinking ALCOHOL, or as a result of their ALCOHOL use. Some of these things are listed below. Please indicate how many times each has happened to you during the last 3 years while you were drinking alcohol or as the result of your alcohol use.

How many times did the following things happen to you while you were drinking alcohol or because of your alcohol use during the last 3 years?	**Never**	**1-2 times**	**3-5 times**	**6-10 times**	**More than 10 times**
1. Got into fights, acted bad, or did mean things.	❑0	❑1	❑2	❑3	❑4
2. Went to work or school high or drunk.	❑0	❑1	❑2	❑3	❑4
3. Caused shame or embarrassment to someone.	❑0	❑1	❑2	❑3	❑4
4. Neglected your responsibilities.	❑0	❑1	❑2	❑3	❑4
5. Relatives avoided you.	❑0	❑1	❑2	❑3	❑4
6. Felt that you needed more alcohol than you used to use in order to get the same effect.	❑0	❑1	❑2	❑3	❑4
7. Tried to control your drinking by trying to drink only at certain times of day or certain places.	❑0	❑1	❑2	❑3	❑4
8. Had withdrawal symptoms, that is, felt sick because you stopped or cut down on drinking.	❑0	❑1	❑2	❑3	❑4
9. Noticed a change in your personality.	❑0	❑1	❑2	❑3	❑4
10. Felt that you had a problem with school.	❑0	❑1	❑2	❑3	❑4
11. Tried to cut down on drinking.	❑0	❑1	❑2	❑3	❑4
12. Suddenly found yourself in a place that you could not remember getting to.	❑0	❑1	❑2	❑3	❑4
13. Passed out or fainted suddenly.	❑0	❑1	❑2	❑3	❑4
14. Had a fight, argument, or bad feelings with a friend.	❑0	❑1	❑2	❑3	❑4
15. Kept drinking when you promised yourself not to.	❑0	❑1	❑2	❑3	❑4
16. Felt you were going crazy.	❑0	❑1	❑2	❑3	❑4
17. Felt physically or physiologically dependent on alcohol.	❑0	❑1	❑2	❑3	❑4
18. Was told by a friend or neighbor to stop or cut down drinking.	❑0	❑1	❑2	❑3	❑4

Reference: White HR; Labouvie EW. Towards the assessment of adolescent problem drinking. J Stud Alcohol 50:30-37, 1989.