The Planning in Life and Adapting to Novel Situations (PLANS): An Ecologically Valid Measure of Every-day Planning

by

Evan Charles Knutson

B.S. Psychology, University of Pittsburgh- School of Arts and Sciences, 2013

M.S. Rehabilitation Counseling, University of Pittsburgh- School of Health and Rehabilitation Sciences, 2015

Submitted to the Graduate Faculty of

School of Health and Rehabilitation Science in partial fulfilment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2019

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This dissertation was presented

by

Evan Charles Knutson

It was defended on

March 6, 2019

and approved by

Elizabeth Skidmore, Ph.D, OTR/L, Professor, Occupational Therapy

Lauren Terhorst, Ph.D, Associate Professor, Occupational Therapy

Patricia Arenth, Ph.D, Psychologist, Veterans Administration Pittsburgh Healthcare System

Dissertation Director: Michael McCue, Ph.D, Professor, Rehabilitation Science and Technology

The Planning in Life and Adapting to Novel Situations (PLANS): An Ecologically Valid Measure of Every-day Planning

Evan Knutson, MS, CRC

University of Pittsburgh, 2019

Individuals with brain injuries often experience impairment in executive functioning, including the ability to plan. Planning dysfunction limits the capacity to complete functional planning tasks required for effective independent living and community functioning and is a frequent target of community rehabilitation intervention. Current planning measurement tools have limitations in their ecological validity, and are not sensitive to every-day planning. Research has shown that naturalistic simulated instruments have higher degrees of ecological validity than traditional testing options, and are more predictive of executive functioning limitations observed in real-world contexts.

This dissertation project describes three research studies involved with the development and feasibility of a new naturalistic instrument: The Planning in Life and Adapting to Novel Situations (PLANS). The first study conducted a scoping literature review (n=22) and semi-structured interviews (n=8) with community rehabilitation clinicians to inform the context, tasks, and scoring of the PLANS. Second, an analysis of clinical data from the Community Multiple Errands Test (CoMET), by individuals with cognitive disabilities (n=55) was conducted to assess planning performance errors in community functioning. CoMET performance was significantly related to measures of logical reasoning, planning and verbal fluency, and demonstrated ecological validity but was limited for adaption to new environments. These findings informed the task demands, environments and scoring of the PLANS. A preliminary draft of the PLANS was constructed and surveys (n=4) were completed by content experts to establish PLANS content validity. The validated PLANS was piloted with (n=20) individuals with a traumatic brain injury. The PLANS was feasible for use with individuals with traumatic brain injuries, and the scoring system demonstrated high inter-rater reliability (kappa=.86). The PLANS demonstrated preliminary construct reliability with measures of executive functioning ($p<.05$), logical reasoning ($p<.05$) and planning ($p<.01$). The PLANS also demonstrated ecological validity through correlations with measures of every-day executive dysfunction ($p<.05$). Future work supports investigating the utility and validity of the PLANS in community rehabilitation clinical settings.

Table of Contents

List of Tables

List of Figures

Preface

First, I would like to thank the members of my dissertation committee for being close supporters of my professional development as a researcher and clinician. Dr. Elizabeth Skidmore, Dr. Lauren Terhorst and Dr. Patricia Arenth: I am grateful for your guidance. To Dr. Michael McCue, my committee chair, advisor and mentor, I appreciate all the time we have spent together over the past 5 years. You recognized potential in me that I could not have reached without your guidance, and I have learned so much in the process. You showed unbridled compassion, attention and support in every aspect of my journey over the past 5 years, and I will strive to represent those values in my future endeavors.

I would like to thank the support of the Council on Brain Injury for their support of my research, and awarding my first research grant. Thank you to ReMed of Pittsburgh for your collaboration on research recruitment and clinical guidance on this dissertation project, with special recognition to Regina Lasko and Tina Bunyaratapan. Thank you to the Hiram G. Andrews Center for their clinical and research support. I would like to thank the Clinical Rehabilitation and Mental Health Counseling department for their support, including all the longstanding and new faculty along with Deb Endres, Caitlin Burkett, and Patricia Arenth. I would like to thank those that contributed countless hours helping me with data collection and processing: Andy Dyer, Shreya Singh and Jessica Kersey I wish you the best in the remainder of your continued education.

An enormous thank you to those that participated in the studies that comprise this dissertation project. Without you this would not be possible, and with your help I intend to benefit the field of rehabilitation research through clinical and research efforts. Thank you to the stake-holders that provided input into this work, and I hope it pays dividends for the community.

Thank you to Mom, Dad and Karen for your love and support, no matter the distance. Thank you to all my friends that kept up with me on this lightning fast path towards research competency, you have the uncanny ability to make me laugh on even the hardest days. Pittsburgh, I would like to thank you for being my home and letting me find myself in a 'big' city, I will miss you but promise to visit often. Thank you to the academic institution where I devoted the last decade of my life, Hail to Pitt!

Finally, I would like to thank my wife Tugce Ozturk for her love and support during this journey. You inspire me to work as hard as I can and to put my whole being into what I care about most. I have learned so much from you, and now I see the world from a better perspective. We have been through two dissertations during our marriage, and (hopefully) this is the last to endure. You are the best support I could have asked for, and I am so excited to continue on to the next part of our journey together in Seattle, WA. Thank you to our puppy-daughter Mochi for hanging out with me while I write and for making every day bright and sunny.

1.0 Statement of the Problem and Specific Aims

According to the Center for Disease Control (CDC), every year approximately 1.7 million individuals experience a Traumatic Brain Injury (TBI) either as isolated injuries or comorbid with other injuries (NCIPC, 2013; Faul, Xu, Wald, & Coronado, 2010). Individuals with TBI frequently exhibit neuropsychological difficulties associated with damage to the pre-frontal lobes, which may result in specific impairments in executive functioning and contributes to complications in goal-directed behavior (Stuss & Alexander, 2000). A critical aspect of goal-directed behavior for persons who have experienced a TBI is functional planning, which involves the ability to cognitively look ahead and identify goals, organize and sequence steps, and execute steps, required for completing tasks in daily life (Morris, Kotitsa & Bramham, 2005; Stuss, 2011b, Ward, 2005).

Traditional performance-based neuropsychological planning instruments show limited relationships to the planning limitations that individuals with TBI encounter in their day to day lives, which limits the ecological validity of these measures (Goldstein, 1996; Chaytor & Schmitter-Edgecombe, 2003). In order to improve the ecological validity of cognitive measurement, research over the past decade has investigated the assessment of cognition using naturalistic instruments, which measure cognition through performance on tasks that resemble those encountered in daily life (Robertson & Schmitter-Edgecombe, 2017; Nalder, Clark, Anderson & Dawson, 2017). The focus of this dissertation involves three integrated studies that develop, validate and pilot a new naturalistic simulated instrument to measure functional planning with individuals with brain injuries: The Planning in Life and Adapting to Novel Situations (PLANS). These studies contribute to the advancement of measuring functional planning in individuals with TBI through the creation of a new naturalistic-simulated instrument that is feasible for use with individuals with TBI.

The first chapter of this dissertation provides a condensed outline of the dissertation's chapters to orient the reader, provide basic description of the chapters' content and clearly present the chapter aims.

The second chapter provides background on executive functioning and planning, TBI and naturalistic instruments. This provides a context for the dissertation by providing the relevant theoretical information and previous research that guides this science. The intersection between executive dysfunction, every-day planning limitations and TBI populations is highlighted, discussing the current limitations in measurement of every-day cognitive functioning and functional planning and how this dissertation addresses these gaps in the literature.

The third chapter describes two components of the PLANS development project: (1) a scoping literature review of published research on naturalistic instruments measuring planning (n=18) and (2) semi-structured interviews with clinicians working in community rehabilitation programs working with individuals with brain injuries (n=8). Through the use of a scoping literature review, previously validated instruments are assessed for instrument properties including: environmental factors, task demands, scoring criteria and psychometric properties. Semi-structured interviews collected stake-holder input on the tasks, environments and functional planning factors that are impacted by TBI in the context of community rehabilitation.

The fourth chapter describes the third component of the instrument development study: (3) an analysis of clinical data from a community-based naturalistic instrument titled the Community Multiple Errands Test (CoMET). Over a two-year period, the CoMET was developed and administered as a clinical tool with individuals with cognitive disabilities to assess their effectiveness and efficiency in problem-solving and planning in a community setting (Knutson et al., 2016). Goals of the study include: CoMET feasibility, preliminary psychometric properties, ecological validity and preliminary findings related to implicit planning demands for the CoMET. The CoMET demonstrated preliminary construct validity and performance on problem-solving and planning tasks conducted in the community were significantly correlated with measures of executive functioning (*p* <.01), planning (*p* <.01) and logical reasoning (*p* <.01). This study contributes to the PLANS development study by providing valuable insight into issues in feasibility and practicality conducting of community-based naturalistic instruments, and support for development of simulated naturalistic instruments.

The fifth chapter presents a synthesis of the results from the clinician input and scoping literature review and the CoMET used inform the development of the PLANS. The construction of the first draft of the PLANS is described, which is used in a content validation study. Content experts in TBI and every-day cognition (n=4) completed content validity surveys investigating the clarity, relevance and sufficiency of PLANS materials. The PLANS demonstrated adequate content validity, and minor changes were made to the PLANS before piloting.

The sixth chapter describes a pilot study that assesses the feasibility of the PLANS for use with individuals with TBI (n=20) and investigates PLANS preliminary reliability and validity with this population. This study establishes preliminary psychometric properties for the PLANS, including: inter-rater reliability, internal consistency, construct validity & ecological validity.

The seventh chapter provides a summary and discussion of results of the three studies, and provides implications for the use of the PLANS in clinical and research applications. This discussion will highlight the limitations of the dissertation studies and examine the future directions for this work.

A depiction of the study flow of this dissertation can be referenced in Figure 1. The graphic depicts a funnel containing the three sources of information used to develop the PLANS. These development components are synthesized and produce a first draft of the PLANS which underwent a validation study. Finally, a validated PLANS was used in a pilot study.

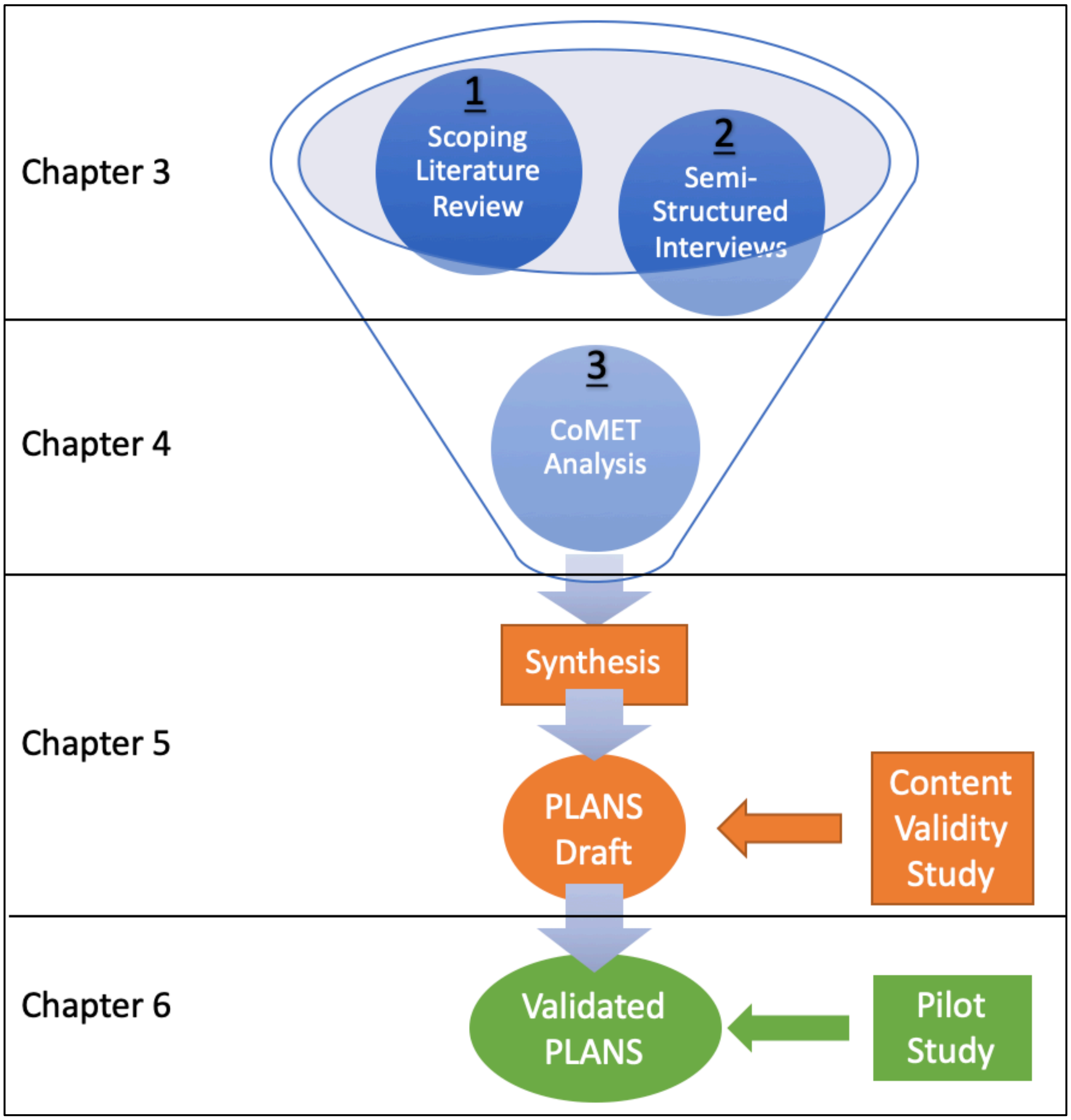


Figure 1 Dissertation Study Flow

The specific aims of chapters three through six are provided below:

3. The Development of the PLANS
    a. Conduct a scoping literature review of naturalistic planning instruments gathering research evidence for development study
        i. Scoring components used by naturalistic planning instruments
        ii. Tasks used by naturalistic planning instruments

        iii. Environments used by naturalistic planning instruments
    b. Conduct semi-interviews with clinicians working with brain injury populations gathering clinical evidence for development study
        i. Functional planning components valued in clinical contexts
        ii. Tasks impacted by functional planning limitations
        iii. Environments impacted by functional planning limitations
4. Analysis of the CoMET by Individuals with Cognitive Disabilities
    a. Investigate CoMET performance by individuals with cognitive disabilities
    b. Investigate CoMET psychometrics including construct validity, ecological validity and internal consistency
5. Finalization and Validation of the PLANS
    a. Synthesize empirical evidence from scoping literature review, semi-structured interviews and CoMET analysis to inform development
        i. Identify the scoring components, tasks, environment, functional planning definition, and administration context for the PLANS
    b. Construct the PLANS materials
    c. Establish content validity for the PLANS
6. Feasibility & Preliminary PLANS Psychometrics
    a. Investigate feasibility of the PLANS for use with individuals with TBI
    b. Investigate preliminary PLANS psychometrics including: inter-rater reliability, construct validity and ecological validity

2.0 Background

2.1 Executive Functioning

2.1.1 Executive Functioning Overview

In the fields of neuropsychology and rehabilitation, executive functioning is a topic receiving increasing attention and research effort (Alvarez & Emory, 2006). Executive functioning is a cognitive construct that refers to the psychological processes involved in conscious control of thought, and that determine goal-directed behavior and purposeful action (Zelazo & Müller, 2002; Lezak, 1982). These include the ability to: formulate goals, initiate behavior, anticipate the consequences of actions, plan and organize behavior, and to monitor behavior and adapt to novel tasks (Cicerone et al., 2000, Lezak, 2004). The study of executive functioning is a complex undertaking, as executive functions are an umbrella term used to capture the various cognitive skills that organize, guide, and execute thought, emotions and behavior.

2.1.1.1 Executive Functioning Definition

There is not a central agreed-upon definition for executive functioning, and researchers have generated several definitions to capture this construct (Jurado & Rosselli, 2007; Banich, 2009). The scope of executive functions is wide, and may refer to cognitive processes or behavioral and emotional regulation (Zelazo & Müller, 2002). Logic and cognitive components are considered “cold” executive functions and are based in cognition and logic, contrasted with “hot” components associated with regulation of emotions and social decision-making (Chan Shum, Toulopoulou & Chen, 2008). Impairment in either cold or hot executive functioning components can have substantial effects on people’s every-day lives, however this dissertation directs its attention toward investigating cold components of executive functioning. Studies of the neuropsychology of executive functions investigate the cognitive processes that determine goal-related behavior including: the ability to formulate goals, initiate behavior, plan and organize, and monitor behavior (Cicerone et al., 2000; Jurado & Rosselli, 2007; Lezak, 2004).

The specific cognitive skills that comprise executive functioning are not clearly understood. The cognitive domains that have been proposed to make up the executive functions are varied, and there are competing theoretical models used to describe these domains and how they work together (Jurado & Rosselli, 2007). Some of these cognitive domains have received substantial research attention, and have independent theories about their functions i.e. planning, multi-tasking, and goal formulation (Morris & Ward, 2004; Salvucci & Taatgen, 2008; Kruglanski, Shah, Fishbach & Friedman, 2018). Another reason the molecular components of executive functioning are misunderstood is that executive functioning measures are correlated with other measures of cognitive functioning (Obonsawin et al., 2002). Attention, working memory, reasoning and processing speed are integral to executive functioning (Duncan, 1995; Salthouse, 2005; Wood & Liossi, 2007; Roca et al., 2010). However, counter-evidence suggests that these are aspects of general intelligence, which does not have a strong relationship to executive functions (Ardila, Pineda & Rosselli, 2000; Friedman, Miyake, Corley, Young, DeFries, & Hewitt, 2006; Stuss & Alexander, 2000). Measures of fluid intelligence appear to have stronger relationships to executive functioning than crystalized intelligence (Zook, Davalos, DeLosh & Davis, 2004; Friedman et al., 2016).

Inhibition, self-regulation and self-monitoring are involved with executive functioning as well, and control the regulation of behavior and emotion (Barkley, 1997; Barkley, 2001; Hofman, Schmeichel & Baddeley, 2012). The role of executive functions on inhibition and regulation is also tied to the processing of emotions and regulation of emotional response (Schmeichel & Tang, 2015). Despite the differences between various cognitive domains that fall under executive functions, each can be linked to the frontal lobe and its importance in these functions (Stuss, 2011a). Meta-cognitive skills such as self-awareness, theory of mind and insight are also theorized to be aspects of executive functioning (Prigatano & Schacter, 1991; Toglia & Kirk, 2000). The frontal lobes are believed to control meta-cognitive functions, and individuals with frontal lobe damage often display executive functioning impairment and anosognosia (Shallice, 1982; Stuss & Benson, 1986; Robertson & Schmitter-Edgecombe, 2015).

2.1.1.2 Models of Executive Functions

Due to the complexity of executive functioning science, its theoretical underpinnings are the subject of continued debate (Jurado & Rosselli, 2007). One argument states the executive functions are a single underlying ability that controls all its related sub processes (a unified theory) and the other states that the sub processes are distinct but related processes (a fractionated theory). Early research efforts treated executive functioning as a unidimensional concept, as studied in individuals with injuries to their pre-frontal cortex (Harlow, 1848; Shallice, 1982; Shallice 1988). These clinical presentations were attributed to an "executive dysfunction syndrome" which resulted in a complex symptom presentation including: difficulties in planning, cognitive flexibility, concept formation, self-regulation, use of feedback and self-awareness.

One of the most prominent and well-researched models of executive functioning is the Supervisory Attentional System (Norman & Shallice, 1980; Shallice 1988). This model outlines that perceptual information that we receive is processed cognitively and activates multiple schemas. Regulation and action on these schemas are controlled by two processes: Contention Scheduling and the Supervisory Attentional System (SAS). In the presence of routine demands, contention scheduling suppresses unrelated schemas and organizes and sequences relevant schemas to reach a goal. When a schema is activated multiple times for similar set of circumstances, this reinforces the schema selection and creates a habitual response (Hommel, Ridderinkhof & Theeuwes, 2002). The SAS is considered a unidimensional executive control that enables planning, decision making, cognitive estimation, problem solving and error correction and attentional shifting (Norman & Shallice, 1986; Shallice 1982). In novel situations, the SAS over-rides contention scheduling processes and engages the cognitive skills required to safely and appropriately confront the novel task demands. Shallice and Burgess used the SAS as an explaining factor for errors and inefficiencies in planning observed by a sample of individuals with traumatic brain injuries solving everyday problems in their Multiple Errands Test (MET) (Shallice & Burgess, 1991b).

Other researchers expanded this theory, suggesting that the executive functions represented a set of interrelated sub-processes, which could be separately classified and measured. (Burgess, Alderman, Evans, Emslie & Wilson, 1998; Baddeley & Wilson, 1988; Miyake et al., 2000). Dissociation in error types between different neuropsychological tests of executive function suggests the fractionation into different executive sub-processes that are able to be quantitatively studied (Chan et al., 2008). Neuro-imaging of frontal lobes during executive-functioning tests find that multiple distinct areas are activated during completion of these tasks, and supports that multiple functions are handled by the frontal lobes and connected systems (Stuss, 2011a).

Recent literature supports the notion of a multi-dimensional model of executive functioning, and several theoretical models have been proposed (Jurado & Rosselli, 2007). Almost all these models share a set of important cognitive processes hypothesized to be included under executive functioning such as goal formation, planning, executing goal-directed plans and self-monitoring. Lezak (1982) proposed a popular conceptual model of executive functioning from a neuropsychological theoretical scope. The executive functions are used to describe how an individual completes a task and whether they complete it at all (Lezak, 1982). Lezak's conceptualization of executive functioning is closely related to Luria's model of Executive Functioning, and expands upon the original domains of anticipation, planning, execution and self-monitoring (Luria, 1966). The popular reference "Neuropsychological Assessment" describes four major categories of Executive Functions: volition, planning, purposive action and effective performance (Lezak, 2004). These categories can be grouped by first looking ahead in a problem-space (volition and planning) and second by being able to act on those intentions in an effective way (purposive action and effective performance). This model states that in order to understand the nature of executive dysfunction, that all four of the categories of executive functioning must be examined in order to identify the specific nature of the executive dysfunction (Lezak, 2004).

2.1.2 The Role of Planning in Executive Functioning

Within the various models of executive functioning, planning is a cognitive construct muddled in theoretical duality. Simultaneously it is considered an integral aspect within executive functioning models and also as a stand-alone cognitive domain worthy of its own theoretical models, and an abundance of research exists from either position (Ward, 2005).

Planning often requires individuals to set goals, organize and sequence steps and self-monitor; each of these "planning components" could be considered an independent facet of higher-order cognition (Ward, 2005). Adding further complication, both executive functioning skills and planning are necessary in order to complete novel or complex tasks encountered in daily life. This creates inherent connections between planning and other aspects of executive functioning that influence performance on every-day tasks.

When investigating planning, a primary way to differentiate between planning and other cognitive domains is through careful definition of planning, and even more careful consideration of the demands of a task. Whether or not planning is involved in completion of a task is influenced by the demands that are placed on the task or situation that requires planning. The complexity of the demands, the level of structure present, and the provision of a desired goal-state can all impact the planning process (Ormerod, 2005).

2.1.3 Defining Every-day Planning: Functional Planning

A general definition describes planning as "the identification and organization of the steps and elements needed to carry out an intention or achieve a goal…" (Lezak, 2004). Planning is required in scenarios that require organizing and sequencing several steps that are necessary to reach a goal. Two different conceptualizations of planning are used depending on the methodology selected to measure (Morris & Ward, 2004). The first is well-defined planning, where "puzzles" or "board game-like" instruments are used to measure planning. These tests are well defined in their rules and procedures, and have high structure in the test design. It has been argued that well-defined planning measures do not elicit “real-world” planning because they do not require participants to look ahead in the same way required in daily life (Davies, 2004). The second conceptualization is ill-defined planning, where individuals complete several errand tasks while following set of rules and instructions. These tasks have higher inherent ecological validity, and introduce elements of complexity analogous to the real world. The use of low-structure tasks encourages the “looking ahead” cognitive processes that are required in daily life, and are more representative of planning required in the real world at the cost of lower experimental control (Omerond, 2004).

When investigating ill-defined planning, the process can be separated into two stages: a plan generation stage where a mental representation is stored and evaluated and possible actions are selected, and a plan execution stage in which a generated plan is retrieved and carried out (Hayes-Roth & Hayes-Roth, 1979; Chevignard et al., 2000; Gilhooly, 2005). Both of these stages can be viewed in the context of every-day tasks, where planning can occur both before a task has been started and planning can occur in the middle of a task as steps are completed. Planning that occurs during plan execution is required when a plan needs adjusting in response to barriers impeding plan execution, which of referred to as “opportunistic planning” (Newell & Simon, 1972).

The complexities between executive functioning, planning and every-day functioning can be addressed by establishing a new term to represent the cognitive skill of planning in every-day life. Functional planning is defined as: the ability to cognitively look ahead and identify goals, organize and sequence steps, and execute steps required for completing tasks in daily life (Morris, et al., 2005). Functional planning is necessary in open-ended goal-directed behavior, and facilitates adapting and overcoming the complex, unstructured, novel problems as seen in every-day life. Ill-structured planning measurement and ecologically valid measurement benefit from a distinct term to represent the cognitive processes required to complete real-world planning tasks.

2.1.3.1 Challenges with Isolating Functional Planning

A challenge in conceptualizing functional planning as a term used in ecologically valid measurement is that functional planning can be understood in terms of a cognitive process or as an activity that one completes in daily life. In example, planning is the cognitive process involved in organizing, sequencing, and executing a multi-step action. Additionally, planning can be seen as an activity present in daily life, such as planning an event, planning a meal, or planning a weekly schedule. The terms Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL) are used by rehabilitation scientists and clinicians as a classification of real-world tasks (Kovar & Lawton, 1994). ADL are defined as basic self-care tasks, and include activities such as bathing, eating or dressing. On the other hand, IADL are tasks considered to be more complex than ADL and are associated with measures of higher-order cognitive functioning (Cahn-Weiner, Boyle & Malloy, 2002). Examples of IADL include domains such as: shopping, home establishment & maintenance, financial management and meal preparation & cleanup (Lawton & Brody, 1969).

Where planning is typically unnecessary while completing individual routine ADL, planning is a necessity in IADL, as these tasks require more cognitive attention and activate the SAS (Cicerone & Tupper, 1991). The SAS is suggested to be required in five types of situations: 1) Ones that involve planning or decision making 2) ones that involve error correction or troubleshooting 3) ones where responses are not well-learned or contain novel sequences of actions 4) ones judged to be dangerous or technically difficult and 5) ones that require overcoming a strong habitual response or resisting temptation (Shallice & Burgess, 1991a). Without planning, IADL would be completed inefficiently, incorrectly, or would remain unfinished in the face of an unforeseen obstacle (Morris, et al., 2005). Since functional planning is required for successful completion of IADL, there is significant variability in the types of real-world tasks that can involve planning. Challenges in completing IADL can limit an individual's abilities in self-care, independent living skills and vocational functioning, therefore identification of functional abilities in IADL are often a target in rehabilitation intervention (Jefferson, Paul, Ozonoff & Cohen, 2006; Rabinowitz & Levin, 2014).

2.1.4 Planning Measurement

While the term functional planning has not been used heretofore (as it is unique to this dissertation), assessment of functional planning is an important part of assessment involved in community rehabilitation with individuals with traumatic brain injuries (Cicerone, Levin, Malec, Stuss & Whyte, 2006; Chua, Ng, Yap & Bok, 2007). As opposed to functional planning, researchers and clinicians measure these concepts using well-defined planning instruments and concepts. Current measurement strategies for measurement of functional planning include: neuropsychological tests, clinical rating scales and in-vivo observation (Lewis, Babbage & Leathem, 2011). Each of these measurement strategies will be reviewed, and their limitations will be discussed.

2.1.4.1 Neuropsychological Testing

Neuropsychological testing is considered a gold-standard used for measuring executive functioning resulting from TBI (Lezak, 2004; Bennett, Ong & Ponsford, 2005). Many neuropsychological tests have theoretical basis in the SAS theory of executive functioning and also require planning abilities, such as the Six-Elements Test, the Hotel Test and the Greenwich Test (Wilson, Alderman, Burgess, Emslie & Evans 1996; Manly, Hawkins, Evans, Woldt & Robertson, 2002; Burgess, Veitch, Costello & Shallice, 2000). One of the most ubiquitous planning tests is the "Tower of London" (ToL), a puzzle-based planning tests which requires individuals to solve a puzzle in the fewest amount of moves while following a rule. Shallice states the ToL requires: formulation of a schema, identification of sub goals required, sequencing the sub-goals, and holding the sub-goals in working memory until goal completion (Norman & Shallice, 1980). The ToL has a large quantity of empirical research examining its relationships between frontal lobe and executive functioning (Shum, Gill, Banks, Maujean, Griffin & Ward, 2009). Another planning test based off the ToL is the Delis-Kaplan Executive Functioning System (D-KEFS) Tower Test. The D-KEFS Tower Test is an example of a well-defined test that measures problem-solving and planning (Delis, Kaplan & Kramer, 2001). The test involves a series of 9 problems of increasing difficulty where an individual must generate a sequence of moves that will achieve a specified end state while making as few moves as possible and while following rules. The Tower Test is inspired by the Tower of London test, originally developed by Shallice (1982) as a tool to study Supervisory Attentional System theory of executive control in individuals with frontal lobe damage.

A contemporary criticism of neuropsychological testing is that many tests and testing batteries demonstrate mixed ecological validity (Long, 1996, Burgess et al., 2006, Chan et al., 2008). While neuropsychological tests have strengths in assessing isolated cognitive functions and diagnosing cognitive disorders, they do not always have strong predictive capabilities of cognitive functioning in every-day contexts (Burgess et al., 2006). Neuropsychological testing demands often target a narrow set of cognitive constructs and use heavily structured settings and demands, which may not be sensitive to deficits in executive functioning that impact daily functioning (Burgess et al., 2006). Additionally, planning tests show weak relationships to limitations that individuals with TBI experience in their day to day lives, which limits the value of instrument results for independent living and community rehabilitation services (Goldstein, 1996; Chaytor & Schmitter-Edgecombe, 2003, Mitchell & Miller, 2008).

2.1.4.2 Clinical Rating Scales

Measurement of planning can also be conducted by survey instruments, such as the those that measure executive functioning such as the Dysexecutive Questionnaire (DEX; Wilson, 1998), Behavior Rating Inventory of Executive Functions (BRIEF; Gioia, Isquith, Guy & Kenworthy, 2000), the Comprehensive Executive Function Inventory (CEFI; Naglieri & Goldstein, 2013), and Delis-Rating of Executive Functions (D-REF; Delis, 2012) or cognitive failures such as the Cognitive Failures Questionnaire (CFQ; Broadbent, Cooper, FitzGerald & Parkes, 1982).

Anosognosia resulting from TBI can limit the ability to accurately self-identify limitations in day-to-day functioning, and reduces the usefulness of self-report measures for measuring functional planning in individuals with moderately severe brain injuries (Hart, Whyte, Kim & Vaccaro, 2005; Prigatano, 2005). Clinical rating scales have also shown to measure cognition differently than performance based instruments, where completing questionnaires may target goal-related performance and performance based instruments may capture cognitive efficiency (Toplak, West, Stanovich, 2013). There are many scales used to assess IADL completion, however these measures typically assess capacity or level of assistance required for

IADL completion, and do not assess functional cognition during completion of these tasks (Moore, Palmer, Patterson & Jeste, 2007).

2.1.4.3 In-Vivo Observation

In-vivo behavioral observation refers to a specific form of behavioral assessment where an individual is observed in their natural context (Thompson, Felce & Symons, 2000). Behaviors are observed by one or more trained raters, and coded based on an a priori classification system related to the behavior(s) of interest. In-vivo behavioral observation of performance in real-world situations can be used to infer executive functioning through categorizing errors committed on real-world tasks.

There are several limitations to in-vivo observation that reduce its usefulness for measuring executive functioning and planning. First, the plan generation phase of functional planning is typically a covert cognitive event, and is difficult to assess using in-vivo methods. During plan execution, it is difficult to rely solely on clinical judgement, as inter-rater agreement for in-vivo observation can be limited depending on the training and discipline of raters (Wildman, Erikson, & Kent, 1975; Bernardin & Buckley, 1981; Murphy & Balzer, 1989). Clinicians have difficulty identifying pathological task errors in real-world performance, especially when the individuals being rated exhibit mild executive dysfunction (Bottari, Iliopoulos, Shun & Dawson, 2014). Depending on the in-vivo situation, individuals may employ learned strategies or accommodations to facilitate task completion, which can complicate the measurement of executive functioning limitations. Additionally, observing familiar or routine task completion can limit the generalizability of in-vivo observations, as there is less dependence on executive functioning in routine task completion.

2.2 Traumatic Brain Injury

A traumatic brain injury is a disruption in the normal functioning of the brain as the result of a bump, blow or jolt to the head or a penetrating head injury (Centers of Disease Control and Prevention [CDC], 2014). The leading causes of TBI include falls, motor vehicle accidents, external forces such as being struck or the head striking against, and assaults (Langlois, Ruthland-Brown, & Thomas, 2004). The mechanism of injury does not indicate the severity of the functional limitations resulting from TBI nor clinical presentation (Im, Hibbard, Grunwald, Swift, & Salimi, 2011). TBI severity is typically classified using the labels mild, moderate and severe through scores on the Glasgow Coma Scale (GCS; Sternbach, 2000), Length of Post-Traumatic Amnesia (PTA; Marshman, Jakabek, Hennessy, Quirk & Guazzo, 2013) or Loss of Consciousness (LOC; Tindall, 1990) at time of injury. These labels are used as broad indicators to describe the severity of and global impact of brain functioning after TBI, but injury size and location are stronger indicators of specific neurocognitive impacts after injury. Individuals with TBI predominantly receive damage to the frontal and temporal regions of the brain, which can result in deficits in executive functioning and planning (Stuss, et al., 1985; Morris, et al., 2005). The following sections will discuss the epidemiology, executive dysfunction, impact on functional planning and measurement strategies with this population.

2.2.1.1 Prevalence & Epidemiology

According the Center for Disease Control (CDC), every year approximately 1.7 million individuals receive a TBI either as isolated injuries or comorbid with other injuries (Faul et al., 2010). Epidemiological data shows that a majority of reported cases of TBI are rated as mild or moderate at time of medical center admission, and the incidence of mild and moderate TBI cases is substantial (NCIPC, 2013). It is estimated that between 70-90 percent of all TBIs that receive acute medical attention are labeled as mild Traumatic Brain Injuries (mTBI) (Cassidy et al., 2004). Several factors could be skewing the rate of mTBI observed in medical settings. Not all mTBI recipients receive medical attention, which is partially attributed to unclear defining criteria of mTBI on medical center admission and mTBI occurrence without seeking medical treatment (Cassidey et al., 2004; Sosin, Sniezek & Thurman, 1996). In particular, systematic reviews of research in sports-injury related concussions show the reported rate of mTBI may be affected by heterogeneous or unreported criteria for mTBI in study methods, reducing the amount of mTBI identified (Cassidey et al., 2004; Langlois, Rutland-Brown, & Wald, 2006).

TBIs are not evenly spread out across all age groups, and increased rates of TBI are observed at early and late developmental stages. As compared to adults, children aged 0-4, and older adolescents aged 15-19, are more likely to experience a TBI (Faul et al., 2010). Rates of TBI are found to be highest for young children aged 0-4 with 1,256 per 100,000 of this population having TBI-related emergency department visits, followed by young adults aged 15-19 with 757 per 100,000 (Faul et al., 2010). Differences in TBI rates by race have been observed, showing increased TBI rates in black children age 0-9, as compared to white children (Langlois, Rutland-Brown & Thomas, 2005). Longitudinal studies of incidence in infants, children and adolescents show that the most common source of injury for individuals 0-14 years old are falls, and motor vehicle crashes and contact sport injuries for individuals 15-25 years old (McKinlay, Grace, Horwood, Fergesson, Riddler & MacFarlane, 2008). Rates of TBI decrease and plateau between the ages of 25-65. Older adults aged 65 and higher also have increased risk for TBI, and have the highest rates of TBI-related hospitalization and death (Faul et al., 2010). The leading cause for TBI in this age group is falling, and is estimated to account for half of all TBI incidents followed by motor vehicle accidents which account for almost one in ten TBI (Thompson, McCormick & Kagan, 2006). This increased TBI prevalence in older adults is attributed to their increasingly fragile health status at those ages and decreased neurological and cardiovascular protective factors (Thompson et al., 2006). In all age groups, TBI rates are higher for men than women (Faul et al., 2010).

2.2.2 Executive Dysfunction in Brain Injury

Generally, executive functioning is associated with higher-order cognitive functions and believed to be mediated primarily by the frontal lobes (Stuss, 2011). However, other research suggests that executive functions associated with different regions of the brain including several regions of the frontal lobes, subcortical structures and thalamic pathways (Stuss & Alexander, 2000; Lewis et al., 2003). Individuals with TBI predominantly receive damage to the frontal and temporal regions of the brain, which can result in deficits in executive functioning (Stuss, Ely, Hugenholtz, Richard, LaRochelle, Poirier & Bell, 1985). Individuals with damage to the frontal lobes or associated systems can experience impairment in executive functions often labeled as executive dysfunction or dysexecutive syndrome (Baddeley & Wilson, 1988). This results in a complex constellation of behavioral, emotional and cognitive sequelae (Busch, McBride, Curtiss & Vanderploeg, 2005; Rabinowitz & Levin, 2014).

Executive dysfunction has a wide impact on the cognitive abilities required for goal-directed behavior, and is one of the most frequent results of acquired and traumatic brain injuries (McDonald, Flashman & Saykin, 2002). The effects of a mild Traumatic Brain Injury (mTBI) are fewer and potentially shorter lasting, and often retain cognitive domains that are affected by more severe TBI (Binder, 1997). However, individuals that experience an mTBI are still at risk for persistent executive functioning deficits that contribute to restrictions in participation (Erez, Rothschild, Kaz, Tuchner & Hartman-Maeir, 2009). In mTBI, clinical predictors of injury severity are less accurate as predictors of outcomes (Stulemeijer, van der Werf, Borm & Vos, 2008). As many as 65% of moderate to severe TBI patients report long-term problems with cognitive functioning including impairment in executive functioning (Evans, Wilson, Needham & Brentnall, 2003; Whiteneck, Gerhart & Cusick, 2004).

2.2.2.1 Functional Limitations Resulting from Executive Dysfunction

The impact of executive dysfunction on everyday functioning and participation is difficult to predict in TBI, several factors should be considered including: the location and severity of the injury, comorbid medical complications, individual characteristics and presence of rehabilitation services (Rabinowitz & Levin, 2014). Executive functioning is a critical component to every-day functioning, and has been associated with as high as 37% of the variance in every-day functional status (Mcalister, Schmitter-Edgecombe & Lamb, 2016).TBI can have a profound effect on a person's ability to return to work, complete ADL and IADL, or live independently (Ponsford, Draper & Schönberger, 2008; Dikmen et al., 2009). Executive dysfunction can also impact social relationships and contribute to difficulties in community functioning (Godfrey & Shum, 2000; Ownsworth & Flemming, 2005; McCabe et al., 2007). Executive dysfunction has a strong influence on the successful completion of tasks that are novel, complex, fluid and require problem solving, can reduce the capacity for independent living and community functioning (Busch et al., 2005; Cahn-Weiner et al., 2002; Jeffersonet al., 2006).

2.2.3 Brain Injury and Planning

Individuals with TBI often demonstrate difficulties in one or more executive functioning domains, which can impact planning (Morris, et al., 2005). Individuals with Traumatic Brain Injuries (TBI) often experience difficulties in planning on both laboratory tests of planning ability and in planning goal-oriented actions required for successful independent living (Benett, Ong & Ponsford, 2005). On well-structured planning tests such as the Tower of Hanoi, individuals with focal frontal lesions take more moves to reach a goal and have more difficulty with complex planning problems (Morris, Ahmed, Syed & Toone, 1993; Morris, Miotto, Feigenbaum, Bullock, & Polkey, 1997). Additionally, individuals with TBI also experience difficulty on ill-defined planning tests, such as the Multiple Errands Test (Shallice & Burgess, 1991b), Party Planning Task (Shanahan, McAllister & Curtin, 2011; Pentland Todd & Anderson., 1998) and Six Elements Test (Wilson, Alderman, Burgess, Elmsie, & Evans, 1996). The poor performance on ill-defined planning instruments could be impacted from both plan generation and plan execution perspectives. After TBI, individuals can have difficulty creating plans or coming up with strategies how to accomplish IADL, and create goals that do not address the problems that they face (Levine et al., 1998; Boelen, Allain, Spikman & Fasotti, 2011). Individuals with brain injuries can also have difficulty anticipating upcoming steps in their plans, and remembering steps in their plans during execution (Duncan et al., 1996; Kliegel et al., 2007).

2.3 Naturalistic Instruments

2.3.1 Naturalistic Instrument Definition

Naturalistic instruments are performance-based measures that assess executive functioning through performance on tasks that resemble those encountered in everyday life. The tasks that are introduced are often complex, novel and present several rules that must be followed. Executive functions are evaluated through performance on the tasks, assessing behavior related to the efficiency, effectiveness, rule breaks and speed of task completion. As opposed to in-vivo behavioral observation, naturalistic instruments do not need to be conducted in the individual's home, work or community, and use standardized tasks and scoring to measure executive functions. Naturalistic instruments can be administered in the real-world, in clinical simulations, on computers or in virtual reality. Scoring for these tests are standardized to account for the individual's approach to the test demands, and the environmental and contextual factors that could contribute to their performance. For these reasons, it is considered that naturalistic simulated assessments are able to more accurately capture the impact of cognitive impairment on everyday performance, therefore have an inherent high degree of ecological validity (Burgess et al., 2006; Robertson & Schmitter-Edgecombe, 2017).

2.3.1.1 Naturalistic Instrument Administration Context

Naturalistic instruments are ecologically-oriented to reflect the environments and demands that are present in an individual's everyday lives. Naturalistic instruments can be administered in several different environments including: real-world/community environments, clinical labs designed to simulate an environment, computer systems or in virtual-reality (Parsons, 2016).

Naturalistic instruments conducted in the real-world have the strongest face validity to the real-world, however their usability is limited to one specific location in testing (Robertson & Schmitter-Edgecombe, 2017). Community-based instruments also involve additional time, financial and resource burden for administering due to the level of involvement required by clinicians and clients. Another limiting factor of real-world contexts is the real world is an environment in flux, and naturalistic community-based instruments are only valid in the setting where their psychometric properties were developed in, reducing the instrument's usability in clinical practice (Chaytor, Schmitter-Edgecombe & Burr, 2006). Variability in real-world testing conditions can affect performance in ways that are difficult to operationalize by instrument administrators, which complicates categorizing errors in task completion to specific domains of executive functioning or functional planning (Burgess et al., 2006). As opposed to clinical-setting instruments, working in the community requires increased time and financial resources involved in transportation and test administration (Nalder et al., 2017). In contrast to community-based instruments, naturalistic-simulated instruments are conducted in a simulated clinical space that artificially incorporates task and environmental demands similar to those encountered in the real-world. Many simulated-naturalistic tests are designed to be portable, and use materials that can be set up in more than one clinical space (Smith et al., 2014; Shanahan et al., 2011). Using virtual environments such as computer-based and virtual/augmented reality systems are also being developed to test a variety of cognitive abilities, including executive functioning and planning (Parsons, 2015). There is not a standardized method of measuring verisimilitude in naturalistic-simulated instruments conducted in clinical spaces nor virtual environments, which could affect their ecological validity (Parsons, 2016).

2.3.2 Naturalistic Instruments Strengths

A promising approach to address the limited ecological validity of executive functioning measurement is through the use of naturalistic instruments. Naturalistic instruments use tasks designed to resemble the environments and demands that are present in a person's daily life to address questions about their cognitive and functional abilities (Robertson & Schmitter-Edgecombe, 2017). There are several advantages to naturalistic instruments measures, such as their stronger emphasis on assessing functional cognitive abilities, demonstration of sound psychometric properties, and their face validity to clients and clinicians that work in every-day environments (Burgess et al., 2006, Parsons, 2016). Real-world demands are often complex, dynamic and require overlapping task demands, which may increase the sensitivity of naturalistic instruments to elicit executive functioning (Marcotte, Scott, Kamat, & Heato, 2010). an inherent high degree of ecological validity (Burgess et al., 2006; Chan et al., 2008).

2.3.3 Naturalistic Instrument Use with Brain Injury Populations

As executive functioning assessment is a critical aspect of TBI rehabilitation services, there are several research studies investigating the use of naturalistic instruments with this population. The most ubiquitous naturalistic instrument to date is the Multiple Errands Test (MET), which was originally designed to measure executive dysfunction after TBI (Shallice & Burgess, 1991b). Individuals with traumatic and acquired brain injuries have been involved with many adapted versions of the MET including: The MET – Hospital Version (MET-HV; Knight, Alderman & Burgess, 2002), MET Simplified Version (MET-SV; Alderman, Burgess, Knight & Henman, 2003), Baycrest MET (BMET; Dawson, Anderson, Burgess, Cooper, Krpan & Stuss, 2009), and

American Multiple Errands Test (AMET; Aitken, Chase, McCue & Ratcliff, 1993). Brown and Hux conducted two studies utilizing naturalistic instruments to assess every-day planning abilities of individuals with TBI (Brown & Hux, 2016; Brown & Hux, 2017). Both of these naturalistic instruments instructed participants to create a written plan about how to accomplish a series of tasks, and then instructed them to execute these tasks over a span of a few hours to multiple days (Brown & Hux, 2017). There are several other examples of naturalistic instruments that specifically measure planning with brain injury populations using both community-based and simulated environments (Bottari, Dassa, Rainville & Dutil, 2009; Chevignard et al., 2000; McGeorge et al., 2011; Novakovic-Agopian et al., 2014; O'Neil-Pirozzi et al., 2010; Shanahan et al., 2011). These studies have several psychometric limitations, including low sample size, no reported ecological validity or limited assessment of plan generation (Knutson & McCue, 2018). Robertson & Schmitter-Edgecombe (2017) provide a comprehensive review of "tasks" completed in naturalistic environments, which highlights the high occurrence of acquired brain injury and TBI in naturalistic assessment research.

2.4 Summary

Rehabilitation specialists are putting more emphasis in ecologically valid instruments that measure executive functioning and functional planning (Nalder et al., 2017; Lewis et al., 2011; Chan et al., 2008). Clinicians view naturalistic instruments as reflective of real-world functioning and valuable to community rehabilitation services (Nalder et al., 2017). There are gaps in the literature surrounding functional planning abilities of individuals with traumatic brain injuries, and a limited pool of research and clinical tools to investigate this topic. There are several advantages to ecologically-valid measures that support continued research efforts in the investigation of functional planning abilities of individuals with traumatic brain injuries.

Naturalistic instruments place stronger emphasis on assessing functional cognitive abilities, demonstrate of sound psychometric properties, and have inherent face validity to clients and clinicians that work in every-day environments (Burgess et al., 2006). Naturalistic instruments also show stronger ecological validity than traditional executive functioning tests (Chan et al., 2008; Burgess et al., 2006). Naturalistic instruments can be easier to integrate into treatment plans for community rehabilitation programs, because the testing environment is analogous to the environment where many rehabilitation interventions are implemented (Chaytor, Schmitter-Edgecombe & Burr, 2006). There is a growing literature investigating naturalistic-simulated instruments, however few published studies report relationships between test performance and real-world functioning or report psychometric properties (Robertson & Schmitter-Edgecombe, 2017). Grounding naturalistic instrument demands and environments in the "real-world" aligns the assessment of cognitive abilities from the performance domain of the International Classification of Functioning, Disability and Health (ICF; WHO, 2001) and represent a stronger rehabilitation focus to test results (Lewis et al., 2011). Researchers and clinicians have vested interest in developing and using instruments with high ecological validity, and reflect real world demands that elicit performance more representative of every-day functioning (Lewis et al., 2011; Nalder et al., 2017). Naturalistic instruments with stimuli and demands that reflect contexts relevant to community rehabilitation could simulate performance that closely resembles real-world capacity, and facilitate the selection of appropriate rehabilitation interventions for functional planning limitations.

This dissertation will conduct the development, creation, validation and piloting of a new naturalistic instrument measuring functional planning: The Planning in Life and Adapting to Novel Situations (PLANS). The PLANS is anticipated to help clinicians objectively assess client ability to generate effective plans of action, and execute those actions as they would be required to do in their daily lives. Improving the ability to assess functional planning limitations that can go unnoticed during traditional neuropsychological testing can increase the efficacy of rehabilitation interventions by more precisely identifying the functional planning limitations that can reduce effective independent living and community functioning. Naturalistic instruments enhance community rehabilitation clinician's ability to assess clients' functional planning to manage complex situations required in independent living and community functioning. Based on PLANS performance, new rehabilitation goals and environmental accommodations could be identified and guide clinical decisions in rehabilitation planning. Results from the PLANS could guide specific cognitive rehabilitation interventions to improve functional planning. Improving the ability to measure client functional planning could reduce the resource demands associated with community rehabilitation programs by decreasing the time and money spent assessing functional planning. Successful independent living for individuals participating in community rehabilitation programs is a result of a coordination of rehabilitation services, and requires intensive attention and resources for the individual. Clinicians using naturalistic simulated instruments will have testing results more germane to the cognitive and behavioral health services that are present in community rehabilitation, streamlining interpretation and implementation of testing results. Finally, this dissertation stands to contribute innovation to the field of measurement by producing a new ecologically-valid instrument designed to measure functional planning abilities.

3.0 The Development of the Planning in Life and Adapting to Novel Situations (PLANS)

3.1 Introduction

Individuals with traumatic brain injuries often experience impairment in executive functioning abilities, which can limit the capacity to complete functional planning tasks required for independent living and community functioning. Functional planning abilities are necessary for successfully completing goal-directed behavior, and facilitate adapting to change and solving novel problems required for independent living and community functioning. Community rehabilitation programs have a limited number of testing options for identifying how an individual's functional planning abilities impact functioning in the context of everyday living. A promising research development to address this issue is the use of naturalistic instruments. Naturalistic instruments are measurement instruments that assess cognitive functioning through an individual's performance on a "real-world" task. Naturalistic instruments demonstrate stronger ecological validity than traditional neuropsychological tests, and provide functionally-relevant information for assessment and goal planning in community rehabilitation. In order to address the deficit of ecologically-valid measures for functional planning, a multi-phase instrument development study was conducted to inform the development of a new naturalistic instrument to measure functional planning abilities in individuals with traumatic brain injuries, the Planning in Life and Adapting to Novel Situations (PLANS).

To identify the tasks, environment and scoring components of the PLANS, a systematic instrument development study was conducted. Two components of this development project are outlined: First, a scoping literature review of naturalistic planning instruments and second, semi-structured interviews with community rehabilitation clinicians that work with individuals with brain injuries. A third component of the development project, a retrospective analysis of clinical data from the Community Multiple Errands Test (CoMET; (Knutson, McCue, Terhorst & Kulzer, 2016; Knutson et al, 2019) is covered in the following chapter (as portrayed in Figure 1).

3.2 Research Aims

This study has several specific aims that converge to inform the development of a new instrument that measures functional planning abilities through performance on a naturalistic planning task, the Planning in Life and Adapting to Novel Situations (PLANS). A scoping literature review and semi-structured interviews with clinicians working in community rehabilitation programs with individuals with brain injuries are used to inform and guide the development of the PLANS scoring domains, tasks and environment.

Specific aims for this instrument development study include:

1. Assess previously published naturalistic simulated instruments and identify features relevant to a naturalistic functional planning instrument
    a. Identify administration settings used by published naturalistic planning instruments
    b. Identify scoring criteria used by published naturalistic planning instruments
    c. Identify tasks used by published naturalistic planning instruments
    d. Identify environments used by published naturalistic planning instruments
    e. Identify psychometric properties reported in studies of naturalistic planning instruments

2. Investigate the experiences of community rehabilitation clinicians that work with individuals with traumatic brain injuries and identify clinical factors relevant to functional planning assessment.
    a. Identify functional planning limitations encountered in community rehabilitation of individuals with traumatic brain injuries
    b. Identify the every-day tasks that are significantly impacted by planning limitations in community rehabilitation of individuals with traumatic brain injuries
    c. Identify the every-day environments that are significantly impacted by functional planning limitations in community rehabilitation of individuals with traumatic brain injuries

3.3 Materials and Methods

3.3.1 Scoping Literature Review of Naturalistic Planning Instruments

A scoping literature review was conducted to identify existing naturalistic instruments that measure planning using a real-world task. PubMed, PsychInfo & Cumulative Index to Nursing and Allied Health Literature (CINHAL) databases were accessed to identify articles featuring naturalistic instruments that measure functional planning. The search was conducted from July 2017 through February 2018. The primary investigator of this study conducted the search and selected articles based on the inclusion & exclusion criteria. The following keywords were used to conduct the search: 'naturalistic', 'simulated', 'community', 'instrument', ‘ecologically valid’, ‘planning’ and ‘errand’. Additional articles were identified using reference lists of identified articles. Studies were included if they met the following inclusion criteria: 1) the study design is a development, feasibility or psychometric study of an instrument 2) the study broadly mentions that the instrument uses every day, real world, simulated, naturalistic or ecologically valid testing environments 3) the article was published in English language 4) the instrument measures functional planning abilities. For the purposes of this instrument development study, naturalistic instruments were considered to measure planning if the task demands, scoring, or instrument description included details about plan formulation and/or plan execution. Exclusion criteria included: 1) article provided low detail on naturalistic instrument administration. If an included article did not provide adequate information about the naturalistic instrument’s administration or scoring, further searches were conducted using the article’s reference list and searches were conducted using Google and Google Scholar search engines for articles with additional detail.

Once the search was completed, each article was reviewed and its naturalistic instrument was classified according to the administration setting, scoring components, naturalistic environments, naturalistic tasks and psychometric properties reported. This data was extracted and classified using two independent raters (EK & JK) trained on using a designated article extraction form. When extraction was finished, results were shared between raters and conflicts in extraction were resolved by discussion involving a third party (MM).

3.3.2 Semi-Structured Interviews with Community Rehabilitation Clinicians

Interviews with community rehabilitation clinicians were conducted to investigate three pertinent factors in the planning limitations that individuals with TBI experience in community rehabilitation: 1) What functional planning limitations are encountered in community rehabilitation of individuals with brain injuries 2) what are some of the every-day tasks that are significantly impacted by planning limitations and 3) what are the every-day environments that are significantly impacted by functional planning limitations These factors were selected to identify the cognitive domains, tasks, and environments that would be of most importance to functional planning for community rehabilitation, and in turn would inform the scoring, task demands, and environmental demands of the instrument in development.

Interviews followed a specific script that was developed a priori beginning any recruiting efforts. Interviews were piloted with 2 rehabilitation clinicians who had 30+ years of experience working with individuals with brain injuries and working in a community setting or every-day cognition. The interview was adjusted based on feedback from the pilot work to increase the precision of questions about the tasks that are most impacted by functional planning limitations, and remove questions that were unnecessary to the instrument development project. The semi-structured interview includes questions with both qualitative and quantitative responses. Several questions with Likert-scale responses were used to capture quantitative responses from clinicians. In addition, questions with open-ended responses were used to capture qualitative information when additional depth and breadth was beneficial. In total, each interview asked 86 questions, with 65 Likert-scale questions and 21 open-ended questions. Each interview asked clinicians to list 3 every-day tasks impacted by functional planning impairment. For each task, they were asked questions related to the rehabilitation processes related to the task and the task demands associated with completion of the task. Interviewees were also asked to list 3 home environments where rehabilitation services are provided that are impacted by functional planning limitations and 3 community environments where rehabilitation services are provided that are impacted by functional planning limitations. For each environment interviewees were asked to rate the environment's importance to independent living or community functioning rehabilitation goals, and the environmental barriers and facilitators present that impact functional planning. The full interview can be referenced in Appendix B.

Local community rehabilitation organizations were contacted by the study PI with emails and flyers to recruit subjects for the semi-structured interviews. The recruitment efforts were directed toward experienced clinicians that worked with individuals with brain injuries in the community. Inclusion criteria included: 1) at minimum bachelor's degree education 2) minimum 6 months working in current job position 3) majority of case load includes individuals with brain injuries and 4) works in a community setting or in every-day functioning. 15 clinicians were contacted for interest in participating, but 5 did not express interest in participating in the study, 2 did not meet criteria for working in the community or every-day functioning. A total of 8 clinicians were enrolled in the study. Each clinician completed a verbal consent protocol before participating in the semi-structured interviews. The clinicians came from a variety of professional disciplines, and included Rehabilitation Counseling, Occupational Therapy, Certified Brain Injury Specialists, Speech Language Pathologists, Psychology and Physical Therapy.

Eight 50-90 minute semi-structured interviews were conducted with community rehabilitation clinicians. Interviews were audio-recorded, and responses were written in short-hand by the interviewer. Clinicians were provided a Likert scale with 5 responses representing '5 - extremely', '4 - very', '3 - moderately', '2 - somewhat' or '1 - not at all' to answer the quantitative questions. These scales can be referenced in Appendix B. Open-ended questions were transcribed by a third-party commercial company, but are not included in the instrument development study. Institutional Review Board approval for the protocol involving semi-structured interviews with community rehabilitation clinicians was obtained through the University of Pittsburgh's IRB board.

3.4 Analyses

Scoping literature review data was reviewed and used to summarize the task demands, environments, and scoring criteria for naturalistic planning instruments. Frequency tables were used to explore extracted data from included articles with categories having higher frequencies were evaluated as having more 'weight' in the synthesis.

Descriptive exploration of the semi-structured interview data set was completed using IBM SPSS (version 25.0). Quantitative data collected from the semi-structured interviews data investigated through descriptive statistics, including mean, median, standard deviation, range and frequency tables were produced for unique responses for tasks and environments. Percent agreement between the eight interviews was calculated for each quantitative question about cognition. Percent agreement was calculated for each task, home environment and community environment with greater than 2 shared responses.

3.5 Results

3.5.1 Scoping Literature Review of Naturalistic Planning Instruments

565 articles were identified during the literature search using the established terms and hand searching. After duplicates were removed (n=78) the primary investigator read through (n=487) titles and abstracts and excluded 395 articles that were irrelevant to the literature search topic.

92 full-text articles were assessed for eligibility, (n=3) were reviews, (n=5) excluded the study design in their articles, (n=35) articles featured instruments that did not measure planning, (n=23) articles featured instruments that partially measured planning, but focused on global executive functioning, and (n=5) articles did not adequately describe their instrument. A total of 21 articles that met inclusion criteria were included in the review and featured 18 different naturalistic instruments. The Party Planning Task (Shanahan et al., 2011; Chalmers et al., 1993; Pentland et al., 1998) and the Cooking Task (Craik et al., 1996; Doherty, et al., 2015) had more than one study that used the instrument with different administration protocol and scoring, and were included in the review. A full presentation of the 21 articles included in this review can be referenced in Appendix A, which reviews the instrument tasks, environment, scoring components and psychometric properties. There were 6 articles that used community-based environments for their instruments, 10 articles that used simulated environments, 3 articles that used computer environments and 2 articles that used virtual-reality environments. A flow diagram describing the literature review process can be referenced in Figure 2.

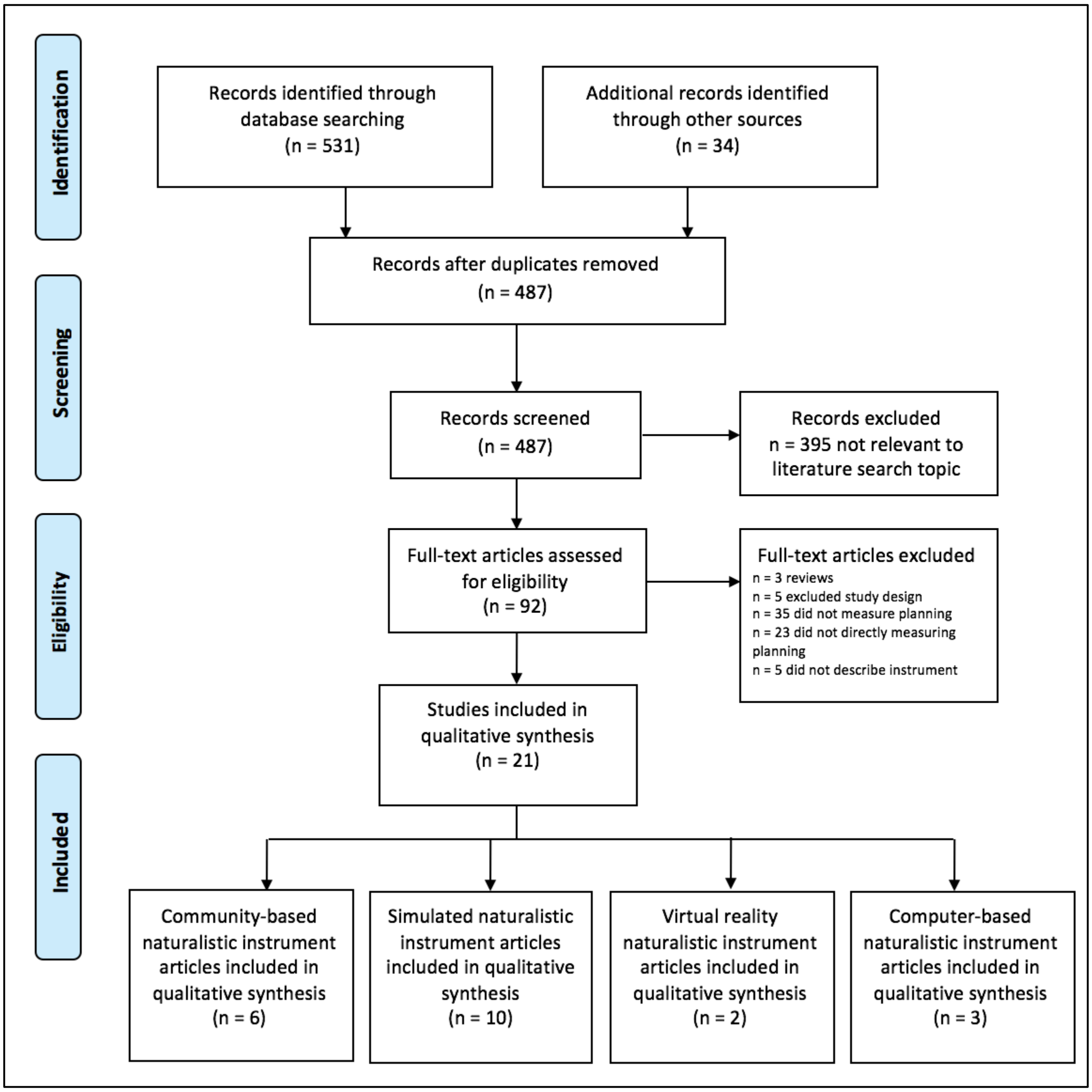


Figure 2 Scoping Literature Review Flow Diagram

3.5.1.1 Testing Environment & Administration

There is a wide spectrum of testing environments that are used by the naturalistic instruments in this scoping literature review. These differences are observed between testing administration environments but also within each category, as each instrument was unique. The testing environment is a critical factor to the development of a new naturalistic instrument because it can affect the complexity and tasks of the naturalistic planning instrument and the time and cost required to set up the environment. In the following discussion, the instruments are discussed with respect to the environments that the tests are administered in, and the time and effort required to administer the tests. A summary of these findings is presented in Appendix A.

The literature review identified 8 different naturalistic instruments that measured planning and were administered in a lab-based testing setting that simulated aspects of real-world or independent living tasks. Four of these instruments: The Party Planning Task (PPT; Shanahan et al., 2011; Chalmers et al., 1993; Pentland et al., 1998), The Classroom Chore Scheduling Task (CCST; Pea et al., 1987), the Errand Task (ET; Kliegel et al., 2007) and the Virtual Planning Test (VIP; O'neil-Pirozz et al., 2010) instructed participants to complete a task requiring them to schedule errands or plan out a weekly schedule in a clinical space. The stimuli for these instruments were portable, and could be contained in a folder or portfolio that could be carried to more than one clinical lab environment. The Naturalistic Action Test (NAT; Seter et al., 2011) was administered in a clinical space, however it used testing stimuli that were more characteristic of the real world, including food products, packaging, and ample real-world distractor objects. The Goal Processing Scale (GPS; Novakovic-Agopian et al., 2014) was unique among the naturalistic simulated instruments because it was administered in a clinical lab space using a computer web browser connected to the internet. The GPS provided participants an open-ended task requiring them to plan for a real-world trip and compare and contrast 3 different options they selected. The internet and web browsing falls somewhere between the real world and an artificial space, however it does reflect a meaningful space for those planning an event in the community. The Apartment Map Test (AMAP; Sanders et al., 2014) and the Day-Out-Test (DoT; Schmitter-Edgecombe et al., 2012) both used designated apartments owned by the research group to administer their naturalistic-simulated instruments. These fully furnished apartments were modeled after real-world apartments were and equipped with cameras and smart-home technologies for data collection purposes. Despite the resource cost, the AMAP and DoT reflect the highest degree of verisimilitude of the naturalistic simulated instruments identified in this scoping literature review.

The literature review identified 6 different naturalistic instruments that measured planning and were administered in the real-word. The Instrumental Activities of Daily Living Profile (IADL Profile, Bottari et al., 2009) and the Prospective Task Planning & Prospective Task Execution Test (PTTPTE; Brown & Hux, 2017) used the participant's home and local community as testing environments. The IADL profile testing took almost one day of activity to execute the full list of tasks, and the PTPPTE monitored participants over one week to determine if they had completed the assigned tasks. The Script Generation & Execution Task (SG&E; Chivengard et al., 2000) was conducted in a multiple community settings to complete a multiphase plan to prepare a meal. Planning a meal is a task that involved multiple environments including a kitchen within a hospital, an office within a hospital, a post office within a hospital and a local grocery store. The Multiple Errands Test – Contextualized Version (MET-CV, Valls-Seranno et al., 2017) and the Modified Multiple Errands Test Planning & Execution (MMET-P&E; Brown & Hux, 2016) are planning tasks that were administered on a rehabilitation or college campus. Finally, the Vocational Oriented Errand Planning Task (VOEPT; 2001) was conducted in two floors of a real-world office building used by the research team.

The literature review identified 2 different naturalistic instruments that measured planning that were administered using computer systems and computer software to display and administer naturalistic tasks. The Cooking Task (Craik et al., 1996; Doherty, et al., 2015) and the

Plan-A-Day Task (PAD; Holt et al., 2006) used computer software to administer a planning task. Both of these computer-based instruments did not require independent raters to score performances, and used the software to automatically record performance variables and calculate performance metrics after the test was completed.

The literature review identified 2 different naturalistic instruments that measured planning using virtual reality to simulate real-world contexts. The Virtual Action Planning Supermarket (VAP-S; Werner et al., 2009) and the Virtual Supermarket (VS; Klinger et al., 2007) used a computer system to administer their virtual-reality planning instruments. The VAPS and VS programs created a virtual environment modeled to represent a typical grocery store similar to one that could be encountered in the real-world. Participants interacted with these virtual environments from a first-person perspective, and were able to "travel" in between aisles and inspect items and environmental features of the grocery store. Similar to the computer-based instruments, the VAP-S and VS do not need independent raters to score performances, and use the software to automatically record performance variables and calculate performance metrics after the test was completed.

3.5.1.2 Scoring Domains of Planning

The naturalistic instruments identified in this scoping review measure an individual's ability to plan in the context of daily life by assessing performance on tasks that reflect the complex demands and structure of every-day life. The performance variables measured by the naturalistic planning instruments included in this review varied widely and many different behavioral and performance outcomes were used to measure planning. There were few overlaps between naturalistic instrument scoring systems, which can be referred to in Appendix A.

Planning is not a uniformly conceptualized construct, and is often obscured as a sub-component of executive functioning (Burgess, Simons, Coates & Channon, 2005). To synthesize the various scoring features of naturalistic planning instruments, we define *functional planning* as the ability to cognitively look ahead and identify goals, organize and sequence steps, and execute steps, required for completing tasks in daily life (Morris, et al., 2005). The following narrative highlights some of the scoring components that were observed in the naturalistic instruments included in this review that are most representative of functional planning.

The first important scoring component of functional planning identified in this review is the distinction between plan formulation and plan execution. Almost half of the naturalistic planning instruments in this review used scoring systems that measure the efficacy or accuracy of plan formulation, while the rest only score performance once a task is started. Planning is a process that first requires generating a mental representation of the steps required to reach a goal, and plan formulation is measured using multiple performance variables collected before an individual started working on the task. The amount of time taken to create a plan is the most common measurement used in planning formulation, and reflects the amount of time individuals used to plan before beginning the task. Planning formulation time does not necessarily reflect the efficiency of plan formulation, as shorter planning times may indicate low effort and poor plan representation. Plan efficiency is characterized by the number of steps required to reach a goal or the strategies that were considered to facilitate completing the plan. Plans that were rated as more efficient used the fewest steps to reach a goal, used an appropriate amount of detail to describe the steps, and selected strategies that would require less effort or time to complete the plan. In addition to the efficacy of a plan, the level of detail provided in a plan was also measured by some articles in this review. The level of detail provided in plan formulation was also included in the VOEP, PTP&PTE and the MMET-P&E naturalistic planning instruments. This score was calculated by counting the amount of information provided in a written or spoken plan created before beginning a task. The level of detail provided in a plan can provide insight into the comprehension of every-day demands, and the amount of effort put into plan generation. Difficulties in plan formulation demonstrate mixed relationships with the ability to successfully complete planning tasks in previous literature on naturalistic planning, and warrants further attention in naturalistic planning measurement.

Another important scoring component of functional planning identified in this review is rating an individual's ability to self-generate a realistic goal. Every-day planning tasks often have unstructured and open-ended goals, and can be completed in more than one way. Despite the importance that this plays in daily life, very few instruments that measure functional planning take this into account. One instrument, the Goal Processing Scale (GPS), used the open-ended and unstructured task of researching an event using an internet browser and required participants to come up with their own goals for the online research. The GPS used a subjective rating system to score the participant on the appropriateness of the goals they selected for their plan, and their comprehension of the task. The ability to measure these aspects of plan formulation could not be re-created in every naturalistic instrument in this review, because these planning factors rely on the task having an open-ended goal. The ability to create a goal in the face of open-ended and unstructured demands is a critical aspect of planning required in daily life, and are especially important for adapting to novel situations. Therefore, naturalistic planning instruments should include tasks with open-ended structure to allow for the measurement of realistic goal generation in plan formulation.

Another important scoring component of functional planning identified in this review is the deviation between an individual's plan formulation and their execution of that plan. Plan deviation was determined to be an important part of functional planning because of research findings from the Apartment Map Test that demonstrated individuals with executive functioning difficulties can deviate from their plan formulations (Sanders et al., 2014). The AMAP study found that their study sample with Mild Cognitive Impairment deviated from their formulated plans to correct for unanticipated challenges, or use strategies that they did not account for in their formulation. This process is referred to as online planning, and can increase the efficiency of every-day action (Ormerod, 2005). The IADL Profile also assessed plan deviation by measuring an individual's level of independence required to verify that their performance matches a self-generated goal (Bottari et al., 2009). Online planning can allow flexibility in the steps to achieve a goal, and facilitate adaption of a plan in the face of environmental and strategic factors. In addition, plans that are overly rigid or structured may be less generalizable to novel situations than flexible plans, and less advantageous in daily life. Flexibility and strategic deviation from a formulated plan is a cognitive skill with little attention in everyday planning measurement that could have clinical significance, and warrants further attention in naturalistic instrument development.

Another important scoring component of functional planning identified in this review is scoring of errors during planning task execution. Assessing the errors present in plan execution is a common way to evaluate the effectiveness of a plan and the ability to successfully complete every-day tasks. Nine out of the twenty-one articles in this review included a process of differentiating task errors while completing a naturalistic planning task, and reflects a commonly used manner of scoring in naturalistic instruments. These "errors" in task accuracy shared amongst these instruments include errors of: omission, addition/commission, inversion and perseveration. These error scoring labels overlap with those used by the Multiple Errands Test, considered to be the genesis of naturalistic instruments, and are often encountered in errand-like naturalistic instruments (Shallice & Burgess, 1991b; Robertson & Schmitter-Edgecombe, 2017). It was more common for naturalistic instruments to quantify the number of tasks that were accurately complete, 17 of the articles in this review included a measure indicating how many tasks were accurately completed during the naturalistic instrument administration. However, relying solely on a count of accurately completed tasks may not provide the level of detail necessary to generate functional inferences describing failed task performance, which limits the ability to understand functional performance in naturalistic action.

Finally, an important scoring component of functional planning identified in this review is strategy use during task execution. A strategy is defined as a cognitive or task heuristic intended to complete a task more efficiently or effectively. Nine of the 21 articles in this review used a scoring system to operationalize strategic behavior that was used during plan execution. A good example of strategy measurement is found in the Party Planning Task, which identified strategies such as chunking, (grouping like tasks together), purpose (verbalizing how to accomplish tasks), monitoring (stopping and checking written plan) and error correction (spontaneous recognition and correction of an error during plan execution). These strategies were observed during the plan execution by trained raters using a behavioral rating scale designed specifically for the PPT. The type of tasks that participants are required to complete affect which strategies are feasible to use. In example, during the Cooking Task (CT; Doherty et al., 2015), participants must execute their plans very quickly, and manage 'cooking' several virtual food items at once making it difficult to observe strategy use during task execution. In this case, the researchers did not measure strategy use, instead using accurate task completion as a metric of the strategic level of the plan: a planner that cooked all items so that they were done at the same time used a strategic plan to accomplish this. The Apartment Map Test (AMAP; Sanders et al., 2014) measured strategy use based on a participant's ability to use a particular strategy to follow an arbitrary rule. During the AMAP, participants completed tasks in an apartment with several specific rooms that should only be visited once during the test. The strategy of searching (visiting incorrect rooms looking for information/task materials) was used to characterize inefficient (rule breaking) actions that were attempting to complete the tasks. In functional assessment, it is important to recognize when strategies are effective and complete steps efficiently and move the plan forward, and when strategies are used inefficiently or fail to complete the task at hand. This is an especially important concept when considering individuals with executive functioning impairments, who can perseverate on an ineffective strategy and fail to recognize that the strategy ineffective.

3.5.1.3 Tasks used by Naturalistic Planning Instruments

In order to measure planning in the context of day-to-day life, the naturalistic instruments in this literature review require participants to complete tasks that resemble tasks encountered in the real-world. Naturalistic instruments can use one or more real-world tasks such as cooking, shopping or planning an event. The type of tasks, complexity of tasks and the number of different tasks varied significantly between the instruments identified in this review.

10 of the articles in this review used errand tasks in their naturalistic instrument design. Errand tasks require participants to collect, deliver, or journey to a specific place and do something. Errand tasks can be quite diverse, and errands can be varied and adapted to fit different environments, including hospitals, apartments, campuses, shopping malls and post

offices. The Day Out Task (DoT; Schmitter-Edgecombe, 2012), Apartment Map Test (AMAP; Sanders et al., 2011), Script Generation & Execution (SG&E; Chevignard et al., 2000), Independent Activities of Daily Living Profile (IADL-Profile; Bottari et al., 2009), Multiple Errands Test – Hospital Version (MET-CV; Valls-Serranno et al., 2017), Vocational-Oriented Errand-Planning Task (VoEP; McGeorge et al., 2001), Prospective Task Planning & Prospective Task Execution (PTP&PTE; Brown & Hux, 2017) and the Modified Multiple Errands Test – Planning & Execution (M-MET P&E; Brown & Hux, 2016) required participants to complete errands tasks in simulated or community environments. The types of errands that are included in a naturalistic instrument can vary, and could include other tasks such as information gathering, shopping, and visiting different locations. Completing errands requires participants visit multiple locations or areas, and unless the instrument is conducted in virtual-reality errand tasks require lots of space. Examples of simulated errand tasks include the AMAP and the DoT, which required participants to carry out errands in a model apartment specially designed for research purposes. Their errands were crafted around errands that could be completed in an apartment, like watering plants, retrieving items, tidying rooms, and finding specific items. The remainder of the naturalistic instruments that use errand tasks were conducted in real-world environments. These errands were based in the environments that were available to the researchers, and were often based in areas local to the research lab. An exception to this is the SG&E, which required participants travel between a hospital-based kitchen, a grocery store and a hospital-based post office to complete multiple errands required for planning a dinner party. Errand tasks were the most common task types in this scoping literature review, which is expected due to the high amount of research based around the Multiple Errands Test (MET; Shallice & Burgess, 1991b). The MET was the first major effort in developing an ecologically valid instrument, and was the first to use errand tasks. Shallice & Burgess found that performance on errands was found to be sensitive to planning dysfunction (Shallice & Burgess, 1991b). Another benefit to using errand tasks in naturalistic instruments is they can be adapted and standardized to many clinical, real-world, and hospital contexts, and are one of the most popular tasks to use for assessment of executive functions. However, adapting errand tasks to a new environment removes the standardization and norming that were established, and further research is required to re-establish those standards.

7 of the articles included in this review used organizing an event as an analogue of a real-world task that requires functional planning in order to complete: The Party Planning Task (PPT; Shanahan et al., 2011; Chalmers et al., 1993; Pentland et al., 1998), the Virtual Planning Test (VIP; O'neil-Pirozz et al., 2010), the Instrumental Activities of Daily Living Profile (IADL Profile, Bottari et al., 2009), and the Script Generation & Execution Task (SG&E; Chivengard et al., 2000). These instruments provide participants with a large goal, such as planning a party or a dinner, and instructs participants to generate a plan that will sequence multiple steps required to reach that goal. In the PPT and the VIP, participants use informational cards or worksheets to plan and organize steps required to throw a party. These instruments use rules and constraints to make the planning for the party more difficult, and requires planning and problem solving in order to effectively manage the schedule and organize the tasks. The PPT and VIP are structured in their demands, goals, steps, and constraints in their planning tasks, and there is a "correct" or best way to organize the tasks. The IADL-Profile and the SG&E requires that participants not only plan an event, but also execute their plan. Participants were required to follow through on their plans, and complete tasks required for the event such as cooking, shopping, errands and information gathering.

In this review errand tasks and scheduling tasks were the most commonly selected tasks to frame around functional planning instruments, but there were several other tasks that are worth mentioning. A commonly used task for naturalistic instruments in this scoping literature review was cooking a meal. Cooking is a cognitively demanding task, that requires multi-tasking, organizing, planning, and time management in order to successfully and safely prepare a meal (Doherty et al., 2015). The IADL-Profile (IADL Profile; Bottari et al., 2009), Script Generation & Execution (SG&E; Chevignard et al., 2000), and the Cooking Task (Doherty et al., 2015; Craik et al., 2006) required participants to cook one or more food items as a task included in their instrument. Another task used by instruments in this review was using a map to plot a course. The Plan-A-Day Task (PAD; Holt t al., 2011), Classroom Chore Scheduling Task (Pea et al., 1987) and the Version A&B Errand Task (Kliegel et al., 2007) required participants to plan a route to complete several tasks at different hypothetical locations. Participants were provided a map based on a town or city, with highlighted locations such as a bank, town hall or hospital. A list of tasks was also provided, which required participants to be at locations at the map at a certain time in order to complete a hypothetical task (be at the bank at 12:00 am to cash a check). These tasks were designed with time constraints and logical rules that created complexities in planning the route, and reflect the sequencing required to complete tasks in daily life. The process of plotting the route required generating an efficient plan that sequenced and interweaved visiting the locations on the map. Other IADLs that were used by naturalistic planning instruments identified in this scoping literature review include using a phone to make a call, medication management, solving puzzles, and being at a place at a certain time.

3.5.1.4 Psychometric Properties of Naturalistic Planning Instruments

Psychometric properties of instruments used in health and rehabilitation practice are of critical importance in determining the value of an instrument for clinical decision making. However, a common critique of naturalistic instruments is their lack of psychometric properties reported and the low power generated from their psychometric studies (Robertson & Schmitter-Edgecombe, 2017). In this review, the current literature on naturalistic instruments that measure functional planning will be examined, and recommendations for instrument development will be proposed based on the findings of this review.

Clinicians and researchers are interested in the reliability of an instrument, or the degree to which the instrument is free from measurement error (Portney & Watkins, 2000). Typical reliability measures from performance-based instruments used by rehabilitation researchers include inter-rater reliability, test-re-test reliability and internal consistency (Mokkink et al., 2012). Inter-rater reliability is reported by 10 of the 21 articles included in this review, as observed in Appendix A. Instruments conducted in virtual reality or that use a computer to administer the instrument use automatic scoring systems within their software, and do not need to calculate an inter-rater reliability. Only instruments that are conducted in the community or in simulated environments report inter-rater reliability, and 63% of these articles in this review report inter-rater reliability. Very few studies report other types of reliability. The Naturalistic Action Test (Seter et al., 2011) reports internal consistency of their instrument. The Virtual Planning Test (VIP; O'Neil-Pirozz et al., 2010) is the only instrument to report test re-test reliability. This may be related to the desire for naturalistic instruments and instruments that measure aspects of executive functioning to have novel demands, which would limit the ability to establish test re-test due to learning effects (Burgess et al., 2006). The Multiple Errands Test

Contextualized version (MET-CV; Valls-Seranno et al., 2017) was the only article to include context reliability. Context reliability was established by comparing MET-CV performance between three different rehabilitation campus settings, examining if the MET-CV was affected by environment of testing. The MET-CV performance was not significantly different between the three testing environments indicating, which is a surprising finding in contrast to the difficulty of establishing equivalent alternate versions of another version of the MET in a hospital setting (Clark, Anderson, Nalder, Arshad & Dawson, 2017).

The validity of a naturalistic instrument is an important factor to consider when determining the use and purpose of the instrument. Researchers and clinicians are interested in the internal validity of the instrument, or examining what constructs the instrument measures, and the external validity of the instrument, or how generalizable the results are to the real-world (Waltz, Strickland & Lenz, 2010). Internal validity is represented by the psychometric properties of construct validity and **known-groups** validity, and external validity through ecological validity. Construct validity is a measure of the degree to which an instrument measures what it intends to be measuring. Out of the 21 articles included in this review, 11 investigated the construct or concurrent validity of their naturalistic planning instrument. Construct validity was established by investigating relationships between naturalistic planning instruments and one or more standardized tests that measure cognitive constructs such as prospective memory, executive functioning, working memory and planning. Known-groups validity is a measure of an instruments ability to discriminate between disability diagnoses by differences in performance between two or more groups. 12 of the articles included in this review investigated the known-groups validity of their instrument. Known-groups validity is typically investigated by measuring differences in performance between a neurologic group and a healthy control group. Ecological validity refers to the extent to which performance on a cognitive instrument can be generalized to cognitive performance in real-life settings. Ecological validity is typically established through comparisons between the naturalistic instrument performance and self or informant responses on clinical rating scales that measure executive dysfunction in daily life. Similar to the findings from a review conducted by Robertson & Schmitter-Edgecombe (2017), this review found that few articles investigate the ecological validity of their naturalistic instruments. Only 5 of the 21 articles in this review report ecological validity for their naturalistic instruments.

There are several implications for the development of naturalistic instruments that can be drawn from the findings of this scoping literature review. First, it is important to note that the psychometric strength of naturalistic instruments is very important, but the available literature demonstrates a weak focus on psychometrics. A limitation of this review is not all articles are instrument development focused, which may have limited the search scope and artificially reduced the amount of psychometric information present. Despite this sentiment, we can see that the Party Planning Task with multiple articles focusing on the same instrument only reports the ecological validity once, despite the goals of the articles and the study design being similar (Shanahan et al., 2011; Chalmers et al., 1993; Pentland et al., 1998).

3.5.2 Semi-Structured Interviews with Community Rehabilitation Clinicians

In total, eight community rehabilitation clinicians participated in the semi-structured interviews. The average time spent in an interview was 75.6 (*sd*= 13.9) minutes. Six of the interviews were conducted in person in clinical offices environment, and 2 were conducted over-the-phone. Demographic information about the interview subjects can be referenced in Table 1.

Table 1 Semi-Structured Interview Participant Demographics

	Mean (SD)	Range
# of years working in position	7.14 (6.73)	1.5-21
Highest education (#)		
Doctor	1	
Masters	6	
Bachelors	1	
TBI Caseload Size	20 (9.61)	3-35
Distribution of customer's TBI severity (%)		
Mild	23.75 (23.01)	0-70
Moderate	49.38 (26.98)	0-100
Severe	25.63 (30.46)	0-100
Distribution of customer's functional limitation (%)		
Mild	17.88 (13.61)	0-45
Moderate	61.00 (26.80)	15-100
Severe	21.00 (24.33)	0-75

There was no unaccounted missing data in the interview responses, and all clinicians completed the interview. The results of the semi-structured interviews will be reported for each section of the interview: 1) What functional planning limitations are encountered in community rehabilitation of individuals with brain injuries 2) what are some of the every-day tasks that are significantly impacted by planning limitations and 3) what are the every-day environments that are significantly impacted by functional planning limitations.

3.5.3 Cognitive Dreams

During the semi-structured interviews, the community rehabilitation clinicians responded to questions asking about planning limitations that individuals with brain injuries of moderate severity experience during community rehabilitation programing. The interview data for this section is presented in Table 2 and Table 3. Community rehabilitation clinicians rated that functional planning was an extremely important aspect to their services with individuals with brain injuries, and that it was a frequent goal that was established in community rehabilitation services (*m* = 4.75). They rated that their clients frequently experienced challenges planning in daily life, and that planning challenges were very limiting to independent living and community functioning goals. Clinicians were in perfect agreement that their clients were only somewhat aware of planning challenges that they experience. Additionally, clinicians reported to what degree individuals with brain injuries of moderate severity have challenges with several cognitive domains relevant to functional planning.

Several domains that received high ratings are related to plan formulation including: the ability to set realistic goals, choosing appropriate goals, logical reasoning, understanding abstract demands, and anticipating appropriate strategies to use. These domains reflect the ability to accurately understand the task demands of a potential activity, and creating a plan that will reach a realistic and relevant goal. Another group of cognitive domains that received high ratings are related to plan execution including: initiating a plan of action, detecting errors, compensating for changes in task demands, encountering a problem that does not match anticipated goals, adjusting a plan to compensate for obstacles and independently working toward a self-generated goal. These domains are related to an individual's ability to initiate and follow through with their plan, modify the plan in the face of new or unanticipated demands. The six cognitive domains that received the lowest ratings are: Following (simple) rules, recognizing when a goal has been completed, understanding (simple) directions, sequencing steps, asking effective questions and using (trained) strategies to help achieve a goal. The low ratings observed for these items may represent a minimization of cognitive challenges that could occur in routine, concrete tasks that individuals with brain injuries more easily complete.

Table 2 Semi-Structured Interview Cognitive Factors Responses

Planning Questions	% Agreement (n=8)	Average Rating
How important is functional planning to your services with this population?	57%	4.75*
In general, how significant are planning problems with individuals with brain injury and moderate impairment?	43%	4.5
On average, how often do your customers encounter problems in related to functional planning in their day to day life?	43%	4.5
To what degree can your customers independently handle functional planning demands?	32%	2.375
Is your client aware that functional planning is a problem for them?	00%	2*
How much do challenges in functional planning limit independent living goals?	39%	4.125
How much do challenges in functional planning limit community functioning?	32%	4.25
To what degree do you direct resources towards functional planning impairment?	39%	4.5
How often do you work with individuals with moderate severity brain injuries to create goals directed around functional planning required for independent living or community functioning?	54%	4.625*
How often do you work with individuals with moderate severity brain injuries to create goals directed around functional planning required for independent living or community functioning?	25%	4.125
** = % agreement > 5%*		

Table 3 Semi-Structured Interview Cognitive Domains Responses

Cognitive Domains	% Agreement (n=8)	Average Rating
Setting Realistic Goals	25%	4.125
Choose appropriate goals	32%	4.25
Appraising a problem	39%	3.875
Understanding directions	39%	3.5
Organization of thoughts	32%	3.375
Sequencing steps	39%	3.5
Logical reasoning	28%	4
Understand abstract demands	32%	4.25
Initiate plans of action	39%	4
Volition	32%	3.375
Follow specific directions	18%	3.75
Following rules	39%	3.125
Monitor their behavior	25%	3.875
Detecting errors	54%	4
Compensating for changes in task demands	32%	4.25
Anticipating appropriate strategies to use for a problem	43%	4.5
Encountering a problem that does not match anticipated goals	32%	4.375
Respond to feedback	32%	3.125
Independently work toward self-generated goals	32%	4.375
Adjust plans to compensate for obstacles	29%	4
Ask effective questions	21%	3.5
Use strategies to help achieve goals	39%	3.5
Recognize when they have completed a goal	17%	3.375
** = % agreement > 5%*		

3.5.3.1 Every-day Tasks

During the semi-structured interviews, the community rehabilitation clinicians were instructed to name three every-day tasks that are significantly impacted by functional planning limitations and were asked questions about the task in the context of community rehabilitation with individuals with brain injuries of moderate severity. The tasks and their quantitative interview data is presented in Table 4. The most frequently reported task was 'scheduling' of one's time, such as scheduling daily tasks, scheduling recreational time, and creating a weekly schedule. Clinicians were in high agreement that this task was rated as related to both independent living and community functioning contexts, and that their clients are aware that this is a challenge for them. The second most frequent task reported by the clinicians was planning for 'chores' that need to be completed around the house, such as laundry, house-keeping, and other instrumental activities of daily living. Clinicians were in high agreement that successfully completing chores are important to positive rehabilitation goals that this is an important task to clients and that clients have only moderate success independently completing chores. The third most frequent task reported was 'planning a recreational activity' such as planning going out with a friend or planning to travel in the community. Clinicians had high agreement that their clients had difficulty independently completing this task.

Table 4 Semi-Structured Interview Task Responses

	Home or Community Setting	How important is successful completion of this task to attaining positive community rehabilitation outcomes?	Are your clients aware that this task is a challenge for them?	Are your clients motivated to work on goals related to this task?	How important is successful completion of this task to your client?	How significant is this problem with individuals with moderate brain injury?	How often do your customers encounter these problems?	To what degree can your consumer independently complete this task?	Count (n)
Scheduling	IL & CF*	4.4	3*	2.6	3.4	4.6	4.4	2.2	5
Completing Chores	IL	4.25*	3.25	3.5	3.75*	4.25	3	2.75*	4
Planning a Recreational Activity	CF	3.33	2.67	3	2.33	4.67	4.67	2*	3
Daily Grooming	IL & CF	4.5	3	3.5	3	4.5	4.5	2.5	2
Grocery Shopping	IL & CF	4	2.5	3	3.5	4	4	3.5	2
Meal Planning	IL	3.5	2	3	3.5	4	4.5	2	2
Medication Management	IL & CF	5	3	3.5	4.5	5	4.5	1.5	2
Money Management	IL & CF	4.5	1.5	2	3	4	3	2	2
Managing Sleep Schedule	IL & CF	5	2	2	2	4	5	2	1
Planning Homework	CF	4	4	3	4	3	5	1	1

IL= Independent Living; CF= Community Functioning
** = % agreement > 50%*

3.5.3.2 Every-day Environments

During the semi-structured interviews, community rehabilitation clinicians were instructed to name three community environments and three home environments that are significantly impacted by functional planning limitations. For each environment, clinicians were asked about the importance of the environment to independent living and community functioning rehabilitation goals. One of the interviewees worked primarily in the community and only provided one home environment. Another interviewee only worked in the community, and did not provide any home environments. These omissions from the interviews were treated as missing data. A summary of the every-day environments is presented in Table 5.

The interviews provided a robust set of every-day environments, 12 different community settings and 7 different home settings identified throughout the semi-structured interviews. The difference in number of different environments is understandable as home environments have fewer distinct environments. The three community environments that were provided most frequently in the semi-structured interviews include: Grocery store, restaurant and coffee shop. Of these 3 items, the grocery store was rated as the most important to independent living and community functioning goals. The three home environments that were provided most frequently in the semi-structured interviews include: Kitchen, bathroom and bedroom. Of these 3 environments, the bathroom was identified as the most important to independent living and community functioning goals.

Table 5 Semi-Structured Interview Every-day Environments Responses

		How important is this setting to independent living goals?	How important is this setting to community functioning goals?	Count (n)
Community Setting	Grocery Store	4	4	5
	Restaurant	3.67	3.33	3
	Coffee Shop	3	2.5	2
	Department Store	4	3.5	2
	Volunteer Setting	3.5	4	2
	Academic Setting	4	3	1
	Bank	3	4	1
	Doctor's Office	5	5	1
	Gym	2	2	1
	Library	3	3	1
	Transportation	5	5	1
	Mall	3	3	1
Home Setting	Kitchen	3.42	4.57	7
	Bathroom	4.6	5*	5
	Bedroom	3.2	4.4	5
	Exercise Room	1	2	1
	Home Office	3	3	1
	Living Room	3	3	1
	Yard	1	3	1
** = % agreement > 50%*				

3.6 Summary

In summation, this chapter describes the results of two components of the PLANS instrument development study: the scoping literature review and the semi-structured interviews. These methodologies produced empirical evidence that will be used to inform the definition of functional planning, scoring components, tasks, environments and administration context of the

PLANS. For the sake of brevity, a detailed description of the conclusions drawn from all three instrument development components are discussed in detail in chapter 5, in the context of the synthesis of the instrument development study.

4.0 An Analysis of Performance on the Community Multiple Errands Test (CoMET) by Individuals with Cognitive Disabilities

4.1 Introduction

Individuals with cognitive disabilities can experience executive dysfunction that contributes to difficulty completing tasks in daily life and reduces their capacity for independent living and community functioning. There are limited assessment options to identify the impact of executive functioning impairment in the community for this population. The Community Multiple Errands Test (CoMET) is a community-based naturalistic instrument that assesses executive functioning in the context of community functioning. The CoMET involves completing several errands in a grocery store while following rules. This study aims to investigate CoMET performance and preliminary psychometric properties of the CoMET with individuals with cognitive disabilities. This analysis includes CoMET, neuropsychological and self-report questionnaire data from 53 individuals with cognitive disabilities that were participating in cognitive rehabilitation.

This study accomplishes two goals: 1) investigates the feasibility, reliability, construct validity and ecological validity of the CoMET and 2) contributes to the Planning in Life and Adapting to Novel Situations (PLANS) development study by providing valuable insight into issues in practicality and feasibility of community-based naturalistic instruments.

4.2 Research Aims

This study used a cross-sectional cohort design. The data set includes CoMET, neuropsychological and self-report questionnaire data collected from individuals with cognitive disabilities that participated in a cognitive rehabilitation program.

Specific research questions were generated to investigate the CoMET's utility and psychometric properties with individuals with cognitive disabilities:

1) How do individuals with cognitive disabilities perform on the CoMET?

2) Does the CoMET demonstrate construct validity?

3) Does the CoMET demonstrate ecological validity?

4) Does the CoMET demonstrate internal consistency?

It is hypothesized that the CoMET will demonstrate construct validity as a cognitive measure of executive functioning through statistically significant correlations between CoMET performance and neuropsychological tests measuring executive functioning. It is hypothesized that the CoMET will demonstrate ecological validity as a measure of every-day executive functioning through statistically significant correlations between CoMET performance and neuropsychological tests measuring executive functioning. It is hypothesized that the CoMET will demonstrate internal consistency and its items will be inter-correlated.

4.3 Materials & Methods

4.3.1 Community Multiple Errands Test (CoMET)

The CoMET is a naturalistic community-based instrument that measures cognitive abilities through the completion of 13 errand tasks in a local grocery store. The CoMET is adapted from protocol used in the American Multiple Errands Test (AMET; Aitken, Chase, McCue & Ratcliff, 1993). Tasks and rules from the AMET were altered to fit the environment and resources available, a commonly used strategy by MET adaptions. The tasks included in the CoMET vary in complexity, and some require the participant to plan ahead and problem-solve in order to accurately complete the tasks. The CoMET is particularly designed so that some task demands are not explicitly stated to the participant, and require the ability to comprehend the implicit demands into the problem solving required to complete these tasks. These tasks are 1) acquiring something to write with and 2) finding a stamp to mail a letter.

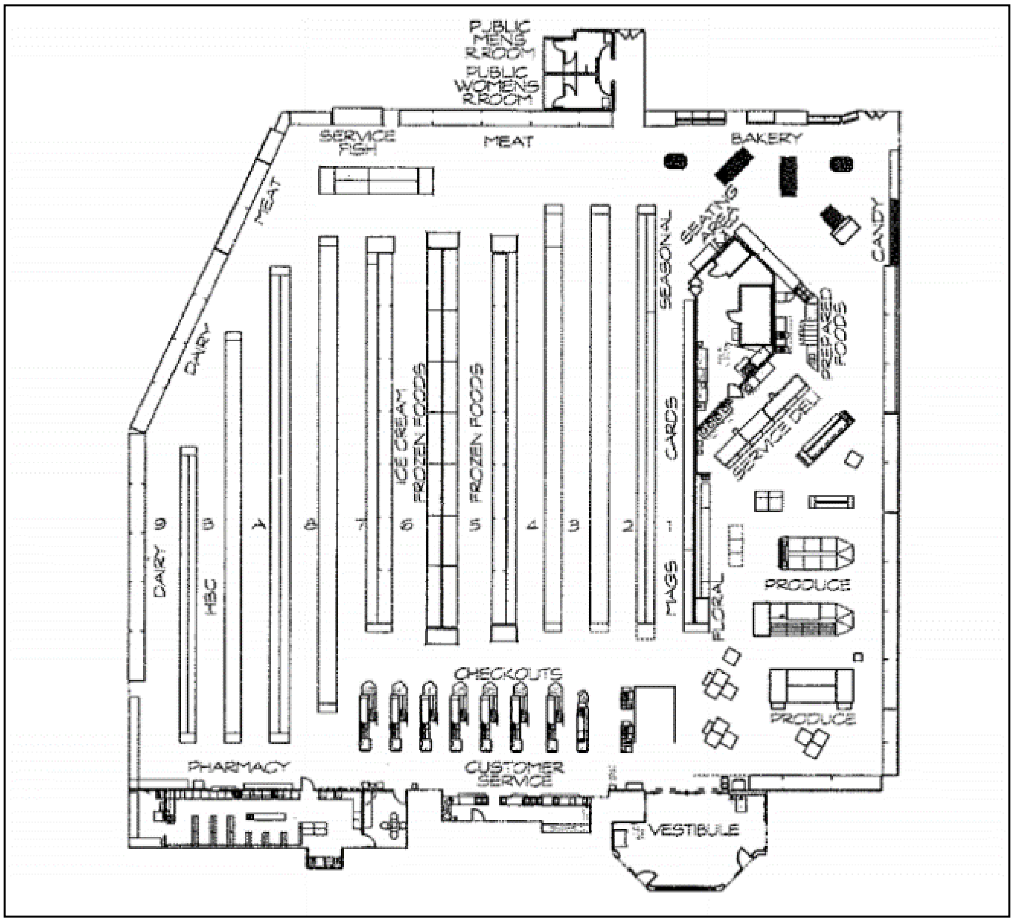


Figure 3 Map of CoMET Grocery Store

-Task Sheet-

1. Buy an instant lottery ticket.
2. You must go to the deli 15 minutes after the test begins and request a sample
3. You must visit the water fountain twice during the assessment and take a drink each time
4. Mail this list to Michael McCue, whose address is provided below

Information to collect while at grocery store

a. What is the price of a five pound bag of potatoes?

   Answer:____________________

b. Using the RedBox kiosk screen: search for and report how many movies titles are available that begin with the letter R?

   Answer:____________________

c. What was the winning number for the Pick Three Pennsylvania Lottery AM drawing yesterday?

   Answer:____________________

d. How much does a 40 oz. jar of Jif peanut butter cost?

   Answer:____________________

e. What aisle does the supermarket carry flour?

   Answer:____________________

f. What is the fewest amount of 13 gallon garbage bags you can buy in a box?

   Answer:____________________

g. What aisle does the supermarket carry rice?

   Answer:____________________

h. How much does one dozen bagels from the bakery cost?

   Answer:____________________

i. Which type of Musselman's brand applesauce has the least amount of sugar per serving?

   Answer:____________________

Test Rules

- Spend as little money as possible (within reason).
- Take as little time as possible (without rushing excessively).
- With the exception of your watch, ANY ITEMS NEEDED TO COMPLETE THE TASKS MUST BE PURCHASED.

Michael McCue
200 Atwood Street
Pittsburgh, PA 15213

Tell the examiner when you have completed all the tasks

Figure 4 Task List Provided to CoMET Participants

After traveling to the grocery store and receiving the instructions to the CoMET, participants are asked to repeat back the demands of the CoMET in their own words to assess their comprehension of the tasks. Clarification is provided on incorrect responses and subjects are encouraged to try again until successfully stating the CoMET demands. Participants are provided a map of the grocery store (as seen in Figure 3), the CoMET task sheet (as seen in

Figure 4), $10.00 and an envelope in which everything was contained. The money is used to purchase items during the CoMET, and participants are instructed to spend as little money as possible. Participants are given time to review the test materials before beginning and are encouraged to take time to plan. Participants begin the CoMET in the lobby of the grocery store, and navigate the store while attempting to complete the tasks. During the CoMET, participants are followed by the administrator who silently observes participant behavior from a distance and monitors their performance. The administrator scores the CoMET while participants complete the tasks, using a clipboard and scoring sheets to record participant performance. If an error in task performance is observed, the administrator records the error and marks the task as inaccurately completed. If participants broke one of the CoMET rules while completing the task, the task is marked as incorrectly completed. Administrator intervention occurs under specific conditions during the CoMET to facilitate participants that stop completing tasks or become stuck and cannot continue. At 10 minutes into the assessment, if participants had not acquired something to write with, they were informed that they should purchase something to write with in order to complete the CoMET tasks. If by 15 minutes they had not acquired something to write with, they were provided concrete and direct instructions to purchase a pen or pencil, and pointed to the specific aisle where they could be found. The same prompting structure is used when participants are completing the letter mailing task. The CoMET ends when the participant tells the administrator that they finished all of the tasks. Tasks that are not completed or not attempted at the end of the CoMET are scored as an omitted task. Participant task sheets were reviewed together by the examiner and participant, and performance errors and recorded answers were reviewed and discussed.

Scoring for the COMET includes several performance variables that were collected while participants completed the tasks. There are 8 performance variables collected for each CoMET administration, as observed in Table 6. Each of the 15 CoMET tasks are rated based on participant performance. Tasks are either counted as accurately completed, inaccurately competed, or omitted during CoMET performance. CoMET task errors were retrospectively coded by the type of error committed; attributed as a judgement, poor searching, rule break, comprehension, self-regulation or poor attention to detail or 'other' error. Retrospective error coding was conducted in several rounds by the PI, and each round was followed by review and consultation from a clinical supervisor until an agreement was reached. The final set of error codes, definitions and sample statistics can be referenced in Table 7. CoMET total time is a measure of total number of minutes elapsed between start of CoMET and end of CoMET. The level of participant prompting reflects the amount of prompting delivered by the CoMET administrator to complete implicit demand tasks: no subject prompting, indirect subject prompting and direct subject prompting respectively. The frequency of aisles entered represents how many times participants travelled down one of the 11 aisles of the grocery store. Number of aisles travelled was considered as a measure of task efficiency, as participants were encouraged to spend their time efficiently. A tally was kept recording each time participants asked questions to others in the grocery store. Participants were not provided instruction nor feedback about asking questions to other individuals in the store, and could do so without penalty. Finally, the total amount of money that participants spent was recorded after participants finished the CoMET. This was also considered a measure of efficiency, as participants were encouraged to spend as little money as possible to complete the tasks.

Table 6 Community Multiple Errands Test (CoMET) Performance Variables

1. CoMET Tasks Accuracy Completed	5. Completion Time
2. CoMET Task Errors	6. Frequency of Aisles Entered
3. CoMET Tasks Omitted	7. Frequency of Questions Asked
4. Level of Prompting	8. Money Spent

Table 7 Community Multiple Errands Test (CoMET) Error Codes

Error Codes	Definition
1. Judgement	Participant uses ineffective strategy or poor judgement to complete task
2. Poor Searching	Participant does not carefully search a store aisle or display and provides an incorrect answer
3. Rule Break	Participant breaks a CoMET rule while completing a task
4. Comprehension	Participant misinterprets the demands of the COMET and completes a task incorrectly
5. Self-Regulation	Participant completes a task impulsively or at the wrong time
6. Poor Attention to Detail	Participant does not finish a task all the way, or does not double-check written answers
7. Other	Participant fails to complete task due to low motivation

4.3.2 Neuropsychological Testing

Select neuropsychological tests are routinely completed with all participants of the cognitive rehabilitation program as part of their clinical assessment. The PI acted as the psychometrist that administered testing; they had experience and training on administering the selected tests and was supervised by a licensed clinical neuropsychologist. This testing was administered in a standardized manner and on average took 2 hours to complete. The neuropsychological tests included in the correlational analyses in this retrospective analysis include: Cognitive Estimates Test (Axelrod & Millis, 1994), Controlled Oral Word Association Test-FAS (Benton, Hamsher & Sivan, 1983), the Behavioural Assessment of Dysexecutive Syndrome Subtest Modified 6-Elements Test (Wilson et al., 1996), Alternate Uses test (Wilson, Christensen, Merrifield & Guilford, 1960) and Trails A (Reitan & Wolfson, 1985). These neuropsychological tests are designed to isolate cognitive functioning in several different cognitive domains as seen in Table 8. All scores are converted to normalized Z-scores using the appropriate norming references, except for the Alternate Uses test which uses a raw score ranging from 0-36.

Table 8 Neuropsychological Tests and Cognitive Domains Measured

Neuropsychological Test	Cognitive Domain
Cognitive Estimates	Logical reasoning, strategy development, divergent thinking
Six-Elements	Rule-following, planning, self-monitoring
COWAT	Verbal fluency
Alternate Uses	Divergent thinking, creativity
Trails A	Processing speed, psychomotor speed

4.3.3 Dysexecutive Questionnaire

To assess everyday executive abilities, the Dysexecutive (DEX) Questionnaire was administered to participants while they were involved in the cognitive rehabilitation program as part of their routine clinical assessment. The DEX Questionnaire was developed by Burgess et al. (1996) and has been validated in a larger study of the Behavioral Assessment of Dysexecutive Syndrome (BADS) by Wilson et al. (1998). This questionnaire is comprised of 20 items that assess everyday executive functioning from motivational behavior and cognitive perspectives. The questionnaire's items use a 5-point Likert-scale to rate the frequency of every-day executive problems, which produces a total score by summing all ratings. A higher total score indicates more executive dysfunction individuals experience in their daily lives, total scores can range from 0-80. The DEX Questionnaire is often used as a measure of every-day executive functioning in studies that examine the ecological validity of executive functioning instruments (Wood & Liossi, 2006; Sanders, Lowe, Schmitter-Edgecombe, 2014). DEX subscales were calculated using subscales proposed and reviewed by Simblett & Bateman (2011) as observed in Table 9.

Table 9 Dysexecutive Questionnaire Subscales

Study	Subscale	DEX Questionnaire Items
Wilson, Alderman, Burgess, Emslie & Evans, 1996	Behavior	2, 7, 9, 12, 13, 15, 16 and 20
	Cognition	3, 6, 14, 18 and 19
	Emotion	5, 8 and 11
Burgess, Alderman, Evans, Emslie & Wilson, 1998	Inhibition	1, 2, 9, 13, 15, 16 and 20
	Intentionality	4, 7, 17, 18 and 19
	Executive Memory	3, 6 and 14
Chaytor, Schmitter-Edgecombe & Burr, 2006	Behavioral Inhibition	2, 9, 11, 12, 13 and 20
	Goal Directed Behavior	4, 5, 8, 10 and 19
	Executive Memory/ Cognition	1, 6, 14 and 18
	Lack of Awareness	3, 7 and 17
Simblett & Bateman, 2011	Behavioral and Emotional Self-Regulation	3, 7, 8, 10, 13, 14, 15 and 17
	Metacognition	2, 5, 12, 16 and 20
	Executive Cognition	1, 4, 6 and 18

4.3.4 Participant Characteristics

The data set includes CoMET, neuropsychological and self-report questionnaire data from 53 individuals with cognitive disabilities that participated in a cognitive rehabilitation program. Primary cognitive diagnoses in this sample include: Autism Spectrum Disorder (ASD), Attention-Deficit Hyperactivity Disorder (ADHD), Specific Learning Disability (SLD) – Math, Reading and/or Writing, and Borderline Intellectual Functioning.

Inclusion criteria for the cognitive rehabilitation program include: age 18-50, native English speaker, verbally fluent for conversation (able to understand speech and understandable verbal communication), full scale and verbal IQ $\geq$ 70 according to individually administered IQ testing within the past 5 years (Wechsler Scale of Intelligence (WAIS) IV or WAIS III), and functional limitations related to one or more shared problems including, awareness, social skills, difficulty generalizing skill set, executive dysfunction or vocational skills. Informed consent was included in the consent process at the beginning of the cognitive rehabilitation program, and included language explaining that the data collected by the cognitive rehabilitation program would be used in future research studies, including retrospective analysis. Institutional Review Board approval for the analysis of this clinical data was obtained through the University of Pittsburgh's IRB board.

4.4 Analyses

Analysis of this data set was completed using IBM SPSS (version 25.0) with a significance level of .05 unless otherwise specified. All assumptions for statistical testing were checked prior to analyses. The dataset includes (n=53) individuals with cognitive disabilities participating in cognitive rehabilitation.

Descriptive statistics were computed for demographic variables, CoMET performance variables, neuropsychological test scores and DEX Questionnaire scores. Frequency tables of CoMET individual item performance were generated to inspect COMET performance and identify items with floor or ceiling effects. Criteria for CoMET item floor and ceiling effects are when 10% of the sample accurately completed a task or when over 90% of the sample accurately completed a task respectively.

Spearman-Rho correlations were used to investigate the construct validity, ecological validity and to investigate the effect of task familiarity and self-estimations on CoMET performance of the CoMET. The interpretation criteria used for correlation coefficients are: weak $r < .35$, moderate $r = .36-.67$ and strong $r > .68-1.0$ (Taylor, 1990). Construct validity and ecological validity correlational analyses set alpha=.01 to reduce chance of type 1 error. Heteroscedascity of variables included in correlation analyses was visually inspected from histograms created in descriptive exploration. Cronbach's Alpha was used to determine the internal consistency of the CoMET. CoMET item scores were calculated dichotomously for this analysis, reflecting "accurate" or "inaccurate" performance for each item. Cronbach's Alpha is typically interpreted using cut-off points of .8 as adequate for research instruments, however alpha levels of .7 are considered adequate for instruments in early development (Nunnally & Bernstein, 1994).

4.5 Results

4.5.1 Sample Demographics

The demographic characteristics of the sample are presented in Table 10. The sample included 44 men and 9 women, and the mean age was 19.9 (*sd*=2.0) years old. All participants were high school graduates, the mean years of education was 12.2 (*sd*=0.7) and 90.6% of the sample had no formal education after high school. The sample's mean full scale IQ was 88.49 (*sd*=11.59) and ranged from 73-133. The participants were predominantly Caucasian (86.2%), followed by Black (7.5%) and Asian (3.8%). A majority of the sample had a primary diagnosis of Autism Spectrum Disorder (66.0%), and the remainder had primary diagnoses of Attention Deficit Hyperactivity Disorder, Specific Learning Disorder, or an undefined Cognitive Disability (34.0%). One-way ANOVA and Chi-Square tests show there were no significant differences on any demographic characteristic between cognitive disability diagnoses (p= .086 - p= .904).

Table 10 CoMET Sample Demographic Characteristics

	Retrospective Analysis Data Sample (n=53)	
	Mean	Standard Deviation
Age (years)	19.9	2.0
Education (years)	12.2	0.7
Full Scale IQ (standard score)	88.5	11.59
Race n (%)		
White	46 (86.8)	
Black	4 (7.5)	
Asian	1 (1.9)	
Other	2 (3.8)	
Gender n (%)		
Male	44 (83)	
Female	9 (17)	
Primary Diagnosis		
Autism Spectrum Disorder	35 (66.0)	
Attention Deficit Hyperactivity Disorder	7 (13.2)	
Specific Learning Disorder	7 (13.2)	
Undefined Cognitive Disability	4 (7.5)	

4.5.2 CoMET Performance

Fifty-three individuals with cognitive disabilities completed the CoMET as part of their comprehensive clinical assessment during cognitive rehabilitation. Spearman Rho correlations between demographic variables and CoMET performance indicate that the sample's demographic factors were not significantly related to CoMET performance (r= -.250 - .146, p>0.05). One-way ANOVA found no significant differences on any CoMET performance variable between ASD, ADHD, SLD, or undefined diagnoses (p= .153 - .842). The sample's overall performance on the CoMET is presented in Table 11. The mean time that it took to complete the CoMET was 42.58 (sd=13.56) minutes, and completion times ranged from 25 - 105 minutes. The mean number of CoMET items accurately completed was 10.15 (sd=2.60) out of 15. The mean number of tasks inaccurately completed was 4.06 (sd=2.05) and omitted 0.77 (sd=1.71). Participants spent 3.06 dollars while completing the CoMET, and spending ranged from zero dollars to 8.14 dollars. CoMET subjects walked down 10.34 (sd=5.32) aisles of the grocery store on average, and ranged from as few as 3 aisles travelled to 31. Subjects asked an average of 2 (sd=2.4) questions to store employees. Twenty-two subjects (sd=41.5%) did not ask any questions to store employees during the CoMET. A significant moderate-strength negative relationship was found between the number of questions asked during the CoMET and the number of tasks omitted (r=-.414, p=.002).

The mean number of implicit tasks accurately completed during the CoMET was .60 (sd=.72) out of 2. The level of prompting delivered to the sample was recorded for the two implicit CoMET Tasks: did the participant acquire a pen and did the participant mail the letter. 60.4% of the sample require no prompting to acquire something to write with, while 17% required an indirect prompt and 22.6% required a concrete prompt. 86.8% of the sample did not require any prompting to purchase a stamp in order to mail the letter, while 11.3% required an implicit prompt and 1.9% required a direct prompt.

Table 11 Performance on the CoMET

		Retrospective Analysis Data Sample (n=53)		
		Mean	Standard Deviation	Range
1	How many tasks were completed with full accuracy (out of 15)	10.15	2.60	3-14
2	How many tasks were completed inaccurately	4.06	2.05	1-10
3	How many tasks were not attempted	0.77	1.71	0-5
4	How many implicit tasks were accurately completed by participant (out of 2)	0.60	0.72	0-2
5	How many explicit tasks were completed with full accuracy (out of 13)	9.55	2.26	3-13
6	How many times a participant travelled down a store aisle searching for information	10.34	5.32	3-31
7	How many times participants asked questions to individuals in the store	2	2.4	0-10
8	How much money was spent by the participant during the assessment	$3.06	$1.89	$0.00-$8.16
9	Total minutes spent in CoMET	42.58	13.56	25-105
1	Judgement Errors	1.23	1.25	0-7
2	Poor Searching Errors	0.43	0.66	0-2
3	Rule Break Errors	0.55	0.57	0-2
4	Comprehension Errors	0.53	0.72	0-3
5	Self-Regulation Errors	0.66	0.78	0-3
6	Poor Attention to Detail Errors	0.62	0.71	0-2
7	Other Errors	0.04	0.04	0-1

The CoMET individual item performance is presented in Table 12. All participants completed at least one task inaccurately or omitted a task, and there was not a perfect performance among this sample. Three of the information gathering items demonstrate ceiling effects: what is the price of a 5-pound bag of potatoes, what aisle does the supermarket carry flour, and what aisle does the supermarket carry rice. Performance on these items did not demonstrate adequate variance, and these items are poor fit for the CoMET as a consequence. All of these items require participants to find a specific part of the grocery store and identify a clearly displayed answer to the question, and may not be demanding enough to discriminate effective performance by individuals with cognitive disabilities participating in cognitive rehabilitation. A majority of participants inaccurately completed or omitted tasks with implicit demands (13 & 14). These tasks are not concretely described to the participants during the instructions, and require abstract thinking and divergent thinking to comprehend the demands. The sample's poor performance on these items reflect a limitation in their ability to comprehend the implicit demands that were presented in the CoMET.

Each CoMET item was associated with a unique distribution of error categories, which is presented in Table 12. Tasks 2, 3 and 15 required participants to complete a task at a certain time or a certain number of times, and are the only tasks that were coded with self-regulation error categories. Tasks 1 and 13 required participants to make a purchase, and are the only tasks that were coded with rule break errors. Judgement errors were the most common type of error on the CoMET tasks, partially because judgement errors were the broadest error category and could be applied to almost every task type. Poor searching errors were only observed in tasks that required careful information gathering, where participants did not search an aisle completely or did not search for information in the correct location. Poor attention to detail errors were almost as broadly distributed as judgment errors, as participants in general had difficulties with maintaining the quality of their task completion. Poor attention to detail errors were recorded when participants did not record their information correctly, or when they made a mistake completing a task that was not related to the other error types. Comprehension errors occurred most frequently in tasks 7 and 9, which required participants to recognize that a specific item or a specific set of information was required. Comprehension errors are observed in most CoMET items, and are attributed to misreading the question. Other errors were recorded when participants did not invest fully in the task, and were related to low motivation in completing the task. These errors were extremely rare, which speaks to the level of investment of the sample to the CoMET.

Table 12 Frequencies of CoMET Item Accuracy

		CoMET Analysis Sample (n=53)								
		AC (%)	Inaccurately Completed (%)							Om (%)
			J	PS	RB	C	I	AtD	O	
1	Buy an instant lottery ticket	40 (75.4)	1 (1.9)	--	10 (18.9)	--	--	--	1 (1.9)	1 (1.9)
2	Go to the deli 15 minutes after the test begins and request a sample	26 (49.1)	3 (5.7)	--	--	--	19 (35.8)	--	--	5 (9.4)
3	Visit the water fountain twice during the assessment and take a drink each time	40 (75.4)	--	--	--	--	3 (5.7)	--	--	10 (18.9)
4	What is the price of a 5-pound bag of potatoes *	48 (90.5)	2 (3.8)	--	--	1 (1.9)	--	2 (3.8)	--	0 (0.0)
5	How many movies begin with the letter A	36 (67.9)	11 (20.7)	1 (1.9)	--	--	--	2 (3.8)	--	3 (5.7)
6	What is the winning lottery number	30 (56.7)	4 (7.5)	1 (1.9)	--	7 (13.2)	--	4 (7.5)	--	7 (13.2)
7	How much does a jar of peanut butter cost	46 (86.7)	1 (1.9)	2 (3.8)	1 (1.9)	--	1 (1.9)	--	--	2 (3.8)
8	What aisle does the supermarket carry flour *	50 (94.3)	--	--	--	--	--	3 (5.7)	--	0 (0.0)
9	What is the fewest amount of 13-gallon garbage bags you can buy in a box	20 (37.7)	2 (3.8)	12 (22.6)	--	10 (18.9)	--	8 (15.1)	--	1 (1.9)
10	What aisle does the supermarket carry rice *	49 (92.4)	1 (1.9)	1 (1.9)	--	1 (1.9)	--	--	--	1 (1.9)
11	How much does one dozen bagels from the bakery cost	46 (86.7)	2 (3.8)	2 (3.8)	--	1 (1.9)	--	--	--	2 (3.8)
12	Which type of applesauce has the least amount of sugar per serving	36 (67.9)	2 (3.8)	4 (7.5)	--	7 (13.2)	--	2 (3.8)	--	2 (3.8)
13	Did the participant get a pen	17 (32.1)	17 (32.1)	--	19 (35.8)	--	--	--	--	0 (0.0)
14	Did the participant mail the letter	15 (28.4)	19 (35.8)	--	--	--	--	11 (20.7)	--	8 (15.1)
15	Did the participant tell examiner when they finished the CoMET	39 (73.6)	--	--	--	--	13 (24.5)	--	1 (1.9)	0 (0.0)

** Item with ceiling effects*
AC = Accurately Completed, J = Judgement Error, PS = Poor Searching Error, RB = Rule Break Error, C = Comprehension Error, I = Impulsivity Error, AtD = Attention to Detail Error, O = Other Error, Om = Omitted

4.5.3 Construct Validity

Results for the CoMET construct validity analyses can be observed in Table 13. A moderate-strength positive correlation was found between the Cognitive Estimation Test (CET) and the number of implicit tasks accurately completed (r=.533, p=.000). Weak negative relationships were found between the CET and the CoMET poor searching (r=-.356, p=.009) and rule break errors (r=-.375. p=.006). COWAT scores were significantly correlated with task accuracy on the CoMET. There was a negative moderate-strength correlation between the COWAT and the total number of tasks inaccurately completed (r=-.453, p=.001) and a moderate-strength positive correlation between the COWAT and the number of explicit tasks that were accurately completed (r=.388, p=.004) and total tasks accurately completed (r=.466, p=.007). A negative moderate-strength relationship was found between scores of Trails A and the number of judgement errors made during the CoMET (r=-.369, p=.007). Both the Alternate Uses and the Six-Elements Test were not significantly related to performance on the CoMET. These results support the hypothesis that the CoMET is a measure of executive functions through significant moderate strength correlations between the CoMET and the Six-Elements and CoWAT tests. These findings also suggest that CoMET performance is related to logical reasoning and psychomotor speed, through significant moderate correlations with Trails A and Cognitive Estimates Test.

Table 13 Correlations between CoMET and Neuropsychological Testing

	Cognitive Estimates	Six-Elements	Alternate Uses	COWAT	Trails A
How many tasks were completed with full accuracy (out of 15)	.318*	.240	.180	.366**	.166
How many tasks were completed inaccurately	-.295*	-.159	-.253	-.453**	-.109
How many tasks were not attempted	-.160	-.203	-.016	-.103	-.084
How many implicit tasks were accurately completed by participant (out of 2)	.533**	.286*	.203	.192	.208
How many explicit tasks were completed with full accuracy (out of 13)	.322*	.195	.168	.388**	.054
How many times a participant traveled down a store aisle searching for information	-233	-.201	-.176	.050	-.178
How many times participants asked questions to individuals in the store	.178	.099	.095	.156	.055
How much money was spent by the participant during the assessment	.000	-.002	.122	-.032	-.304*
Total minutes spent in CoMET	-.121	-.124	-.152	-.007	-2.60
Judgement Errors	-.169	-.266	-.097	-.206	-.369**
Poor Searching Errors	-.365**	.020	-.018	-.350*	.002
Rule Break Errors	-.375**	-.023	-.063	-.163	-.218
Comprehension Errors	-.071	-.228	-.228	.168	.159
Self-Regulation Errors	-.147	-.201	-.201	-.165	.070
Poor Attention to Detail Errors	.108	.016	.016	-.190	.001
Other Errors	-.123	-.123	-.123	-.091	.075
Correlation is significant at the 0.05 level (2-tailed). * *Correlation is significant at the 0.01 level (2-tailed).***					

4.5.4 Ecological Validity

Ecological validity of the CoMET was evaluated through comparisons of CoMET performance variables with concurrently administered Dysexecutive (DEX) Questionnaire self-report total score and validated subscales (Simblett & Bateman, 2011). For clarity, only subscales with significant relations to CoMET performance are displayed. Results for the CoMET ecological validity analyses can be observed in Table 14.

The DEX total score was not significantly correlated to any performance variable of the CoMET. A moderate-strength positive correlation was found between the DEX subscale Executive Memory (r=.376, p=.005) and how much money participants spent during the CoMET. The DEX Subscales Intentionality (r= .377, p=.005) and Goal Directed Behavior (r=.390, p=.004), had moderate-strength positive correlations with how much time participants took to complete the CoMET.

These findings suggest that the CoMET demonstrates ecological validity through significant weal and moderate strength correlations between CoMET speed measures (time and money) and self-report measures of every-day executive functioning and executive-related concepts.

Table 14 Correlations between CoMET and DEX Questionnaire Subscales

	DEX Subscales					
	Cog	Int	EM	GDB	EMC	EC
How many tasks were completed with full accuracy (out of 15)	-.160	-.009	-.119	-.077	-.186	-.126
How many tasks were completed inaccurately	.083	-.016	.049	.048	.118	.089
How many tasks were not attempted	.199	.098	-.054	.135	.197	.157
How many implicit tasks were accurately completed by participant (out of 2)	-.199	-.128	.144	-.125	-.167	-.158
How many explicit tasks were completed with full accuracy (out of 13)	-.106	.046	-.022	-.031	-.147	-.091
How many times a participant traveled down a store aisle searching for information	.067	.047	-.032	-.070	.124	.129
How many times participants asked questions to individuals in the store	-.244	-.170	-.133	-.006	-.244	-.254
How much money was spent by the participant during the assessment	.305*	.200	.376**	.066	.281*	.289*
Total minutes spent in CoMET	.193	.377**	.078	.390**	.168	.286
Judgement Errors	.130	.027	-.211	-.016	.140	.094
Poor Searching Errors	-.054	-.149	-.147	-.023	-.028	-.099
Rule Break Errors	.018	-.005	-.037	.128	-.046	.014
Comprehension Errors	-.147	-.036	-.222	-.031	-.172	-.047
Self-Regulation Errors	.069	-.046	.091	-.167	.146	.069
Poor Attention to Detail Errors	.138	.080	.136	.133	.099	.112

Notes: Cog=DEX Subscale Cognition, Int=DEX Subscale Intentionality, Exec Mem= DEX Subscale Executive Memory, GDB=DEX Subscale Goal Directed Behavior, EMC=DEX Subscale Executive Memory and Cognition, EC=Executive Cognition

*Correlation is significant at the 0.05 level (2-tailed).**
*Correlation is significant at the 0.01 level (2-tailed).***

4.5.5 Internal Consistency

CoMET item performance was recoded to facilitate investigating internal consistency, item's coding was reduced to two categories: accurate performance and inaccurate performance. Based on the dichotomized scoring of the 15 CoMET items internal consistency of the CoMET almost reached an adequate level, with Cronbach's Alpha = .652 and Cronbach's Alpha based on standardized items =.648. Item-total statistics found no single item if removed would significantly increase Cronbach's Alpha.

These findings do not support the hypothesis that the CoMET would demonstrate internal consistency. This finding suggests that the CoMET items are not all measuring the same concept equally.

4.6 Discussion

In summary, this study was conducted to assess the performance of individuals with cognitive disabilities on an open-ended community based naturalistic instrument: The Community Multiple Errands Test (CoMET). In our examination of the CoMET, we determined that the measure is feasible for use with individuals with cognitive disabilities and shows promise for measuring executive functions as indicated by relationships with executive functioning tests and self-report measures. This implies that this naturalistic instrument may be a useful tool for assessing executive functions in the real-world environment.

The CoMET represents a complex set of tasks designed to reflect the demands present in independent community functioning. Accurate performance on these tasks requires a variety of cognitive and executive domains. The CoMET uses a method of investigating cognitive abilities that is very salient to the goals and life-roles for transition-age individuals with cognitive disabilities. These participants were participating in a cognitive rehabilitation training program and all had individual goals related to success in their future vocational or independent living situations. The sample demonstrated high investment in completing the CoMET tasks, 95% of the tasks were at least attempted by the participants and the 5% of tasks that were omitted were typically related to cognitive challenges as opposed to effort or motivational issues.

Individuals with cognitive disabilities demonstrated high accuracy on tasks that were simple and straightforward, and were challenged by tasks that were not clearly stated or required problem solving. These 'implicit tasks' had very low accuracy ratings as compared to other CoMET tasks, and reflect a unique aspect to the test. Logical reasoning abilities involve applying a strategy to a situation where a readily available answer is not present, and this study found medium strength correlations between Cognitive Estimates Test scores and completion of implicit tasks (r=.533, p< .001). The participants were given time to plan before the CoMET, however many elected to begin the CoMET right away without taking time to consider the most efficient sequence to complete the tasks or which tasks were of higher priority than others.

Tasks that had high rates of inaccuracy were primarily attributed to errors of judgement and comprehension. This finding may be telling of the types of challenges that individuals with cognitive disabilities experience in daily life. Errors of comprehension could be related to the sample's difficulties in abstract and divergent thinking, complicating their interpretation of the demands of the CoMET. Judgement errors appear related to difficulties selecting effective problem-solving strategies to accomplish a task. It is possible that in the face of novel tasks, or tasks that are unfamiliar to the individuals, they did not have a schema to guide their performance. In some cases, bizarre and confusing strategies were selected, such as using a scale to weigh individual potatoes to identify the price of 5 pounds of potatoes or using a poster next to the Redbox machine to identify how many movies start with the letter 'A'. These findings highlight the importance of carefully defining error types that clearly reflect the problem behaviors that might be observed in community settings.

Overall, there was wide variation in scores on the CoMET performance variables, which resulted in high standard deviations. Despite the variability in performance, the CoMET demonstrated low internal consistency for an instrument in early development. To promote internal consistency of naturalistic instruments, future instrument development projects may choose to group tasks and items by the cognitive domains they intend to measure, and sum the scores on these scales as an internally consistent measure of cognitive functioning.

The CoMET demonstrated convergent construct validity with several neuropsychological tests related to logical reasoning, processing speed and category fluency. The Cognitive Estimates Test was strongly significantly correlated with how many implicit tasks were accurately completed, poor searching errors and rule break errors. This may suggest that the CoMET's implicit tasks and the arbitrary rule require abstract thinking and logical reasoning. The relationship between searching errors and logical reasoning may be explained by the logical reasoning involved with scanning aisles and products for information gathering tasks, where the information that is being searched for must be identified among various distractors, and could be obscured by other items. Logical reasoning could be necessary for deciding where to search for information in a grocery store, and could explain poor search errors related to searching in incorrect parts of the grocery store. COWAT-FAS was correlated with how many tasks were accurately completed, how many explicit tasks were accurately completed and how many tasks were incorrectly completed. This could be related to the demand for fluency in the CoMET, where participants were required to generate multiple strategies to accomplish the tasks, and a dearth of strategies could result in tasks being completed incorrectly when a first strategy does not result in the anticipated solution. An interesting finding from the construct validity analyses was the significant correlation between TMT A and judgement errors. While psychomotor processing speed may seem unrelated to judgement, the fact that the CoMET was administered in a community setting may emphasize the speed at which plans and strategies need to be formulated, and could contribute to impulsive or ineffective strategy selection at the cost of speed.

The CoMET also demonstrated ecological validity, as investigated through comparisons on the CoMET performance measures and self-report Dysexecutive (DEX) Questionnaire forms. Longer CoMET completion times were associated with higher scores on the Intentionality and Goal Directed Behavior DEX questionnaire subscales. Where higher scores on these DEX questionnaire subscales indicate more difficulty, this may indicate that longer completion times on the CoMET are a result of difficulties in intent and goal-persistence. In addition, a moderate positive relationship was found between how much money was spent on the CoMET and the executive memory subscale. This may indicate that those that spent more money on the CoMET had difficulty remembering the rule to spend as little money as possible. The correlations between cognition, executive memory and cognition, and the executive cognition DEX questionnaire subscales almost reached significance with how much money was spent during the CoMET, which may indicate that making purchases is an activity in the CoMET that has strong ties to every-day functioning.

4.7 Summary

In summation, this chapter describes the results of the third component of the PLANS instrument development study: the analysis of the CoMET dataset. This methodology produced empirical evidence that will be used to inform the definition of functional planning, scoring components, tasks, environments and administration context of the PLANS. For the sake of brevity, a detailed description of the conclusions drawn from all three instrument development components are discussed in detail in chapter 5, in the context of the synthesis of the instrument development study.

5.0 The Finalization and Validation of the Planning in Life and Adapting to Novel Situations (PLANS)

5.1 Introduction

Individuals with traumatic brain injuries often experience impairment in executive functioning, including functional planning. Functional planning involves the generation and execution of a plan of action to complete a series of sub-tasks required to reach a larger goal, in the context of real-world demands. Functional planning is necessary in goal-directed behavior, and facilitates adapting to change and solving novel problems as seen in tasks like cooking, shopping and planning an event. For individuals with brain injuries, functional planning limitations can impact the ability to control cognition and behavior necessary for coordination of day to day life, and can limit their ability to live independently and function in the community and are often targets of assessment and rehabilitation intervention. However, there are few psychometrically sound and ecologically-valid instruments available that measure functional planning.

Chapters three and four describe three components to an instrument development study designed to create a new naturalistic instrument that measures functional planning. This chapter describes the systematic process through which synthesis, construction and content validity study of the PLANS was conducted. The three components of the instrument development project were analyzed separately and their results synthesized together to develop the task demands, scoring criteria and the scenario framing the PLANS (as referenced in Figures 5 & 6). The PLANS was reviewed by content experts in instrument psychometrics, brain-behavior relationships and every-day cognitive functioning to establish content validity for the instrument. Once validated, the PLANS is intended for a pilot study to investigate the feasibility of use with individuals with TBI and establish preliminary psychometric properties.

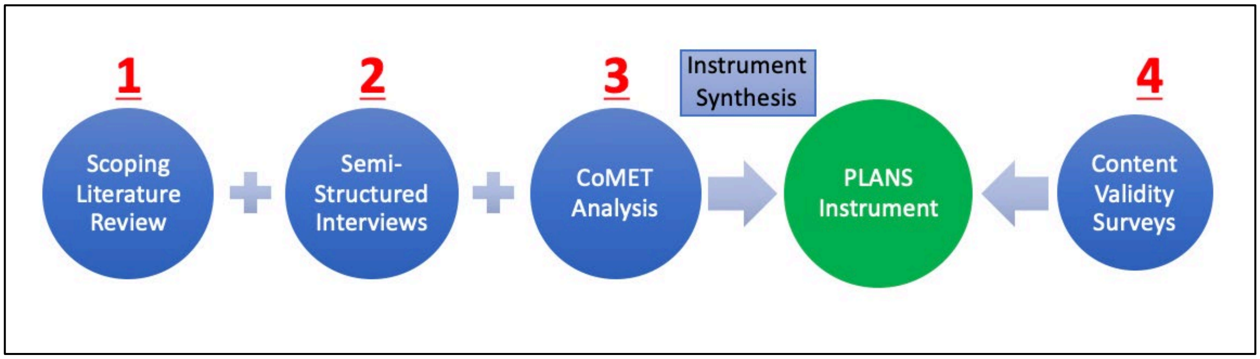


Figure 5 Instrument Development Methodology

5.2 PLANS Finalization Aims

The PLANS development process involves three systematic steps that will produce a novel naturalistic simulated instrument, the PLANS, and establish its content validity. The steps that were completed include:

1. Conduct synthesis of the 3 instrument development components to identify the task demands, environment and scoring components of the PLANS
2. Construct a draft of the PLANS, and have PLANS reviewed for face validity
3. Establish preliminary content validity for PLANS through content expert surveys

5.3 Methods and Procedure

This instrument development project used scoping literature review, semi-structured interview data, and performance data from a community-based naturalistic instrument, the Community Multiple Errands Test (CoMET; Knutson, McCue, Terhorst & Kulzer, 2015) to inform the testing environment, context, tasks and scoring of a new instrument designed to measure functional planning. This synthesis was supported by quantitative data provided from these three sources. This data was explored through two scopes: First, quantitative data was weighed against each other to make design choices with the most empirical support (findings with the highest frequency or mean indicate stronger support). Second, empirically supported findings were balanced with respect to design choices that are feasible, practical and made sense as a whole-instrument. The decision to frame the PLANS as a simulated-naturalistic instrument was committed early in the development, and feasibility decisions were made for administering the PLANS in a clinical space while sitting at a table.

The synthesis was conducted between April 2018 and June 2018, the construction of the instrument occurred over a 2-month period between May 2018 and July 2018.

Following the creation of the first PLANS draft, a content validity survey was developed. This survey was developed following the specific guidelines and criteria for testing content validity as outlined by Waltz, Strickland & Lenz (2010). The full content validity survey can be referenced in Appendix C. For the PLANS tasks, stimuli and scoring domains, experts were asked to rate the item's relevance to functional planning, the clarity of the item's definition and the sufficiency of the item to represent the span and scope of its stated purpose. Surveys were administered through Qualtrics, a web-based program. This survey platform follows University of Pittsburgh Medical Center and University of Pittsburgh guidelines for data security, survey fidelity and HIPPA compliance.

Experts in brain behavior relationships and instrument development were recruited to participate in the content validity surveys. Experts were carefully selected by the PI of the study from professional references by research team. Experts were provided digital copies of the PLANS administration manual, scoring rubrics and stimuli and were also provided a video demonstration of a PLANS performance by a mock participant. Recruitment for content validity surveys was active between July 2018 and October 2018. 17 content experts were contacted to take part in the content validity surveys.

5.4 Analyses

Percent agreement was calculated between the (n=4) content validity experts, with agreement > 70% indicating adequacy. Content Validity Index - Scale (CVI-S) is calculated for the PLANS by the number of experts giving a rating 3 or 4 to the relevancy, clarity and sufficiency of each item, divided by the total number of experts (Zamanzadeh et al., 2015). CVI are presented for each PLANS item, and the CVI-S is calculated for relevancy, sufficiency and clarity.

5.5 Synthesis

The synthesis of the PLANS followed an a priori protocol that compared the results of the three instrument development components and weighed them against each other to provide the best support for the PLANS administration context, scoring components, tasks and environment. A graphic representing this process can be referenced in Figure 6.

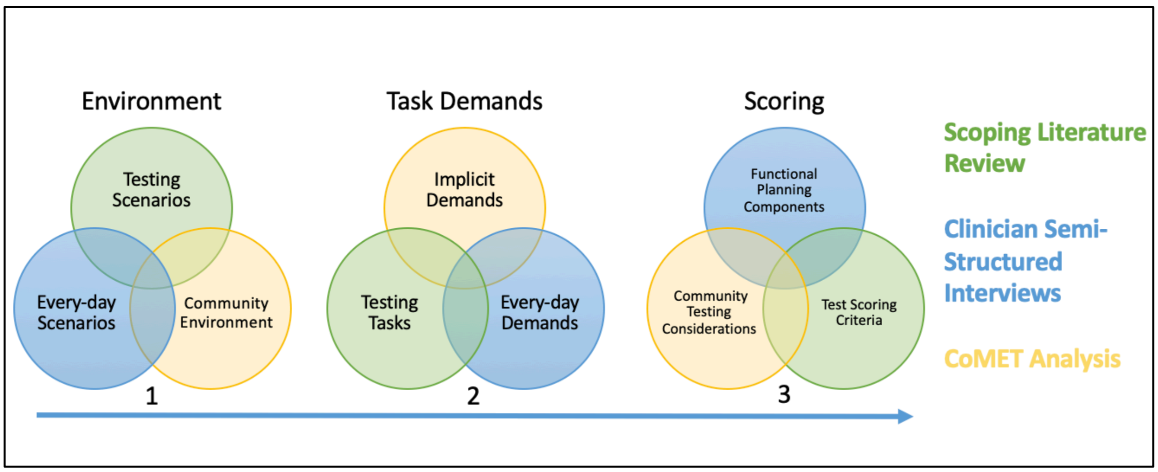


Figure 6 Synthesis Flow Diagram

The findings from the synthesis supported that the new instrument should be designed as a naturalistic simulated instrument. In the scoping literature review, naturalistic simulated testing environments had the most amount of article published (n=11). These instruments also showed promising psychometric properties such as inter rater reliability, construct validity and ecological validity, which are of significant importance to a new clinical or research instrument. A distinct benefit of simulated naturalistic instruments is their portability, which would enhance the standardization of testing protocol. Portable instruments are also more practical for the clinicians, as they are not limited to one site for testing and can be used by multiple individuals and organizations simultaneously. Naturalistic simulated instruments also required fewer time and monetary resources than real-world testing, and took fewer resources to develop as opposed to computer-based or virtual reality environments. Simulated instruments varied in their verisimilitude to the real-world, and ranged from pencil and paper stimuli to full-scale model apartments. Despite this wide range, naturalistic simulated instruments were significantly related to measures of real-world functioning, indicating that they are clinically useful measures for detecting every-day cognitive functioning independent to degree of verisimilitude. It is hypothesized that naturalistic simulated instruments do not need to be exact reflections of real-world conditions, but need to possess cognitive demands similar to those experienced in the real world. Semi-structured interviews found that novelty, time limits, tasks that require anticipation, unstructured tasks and sensory distractions are examples of real-world demands that can be incorporated into a naturalistic simulated instrument. The semi-structured interviews and CoMET performance data corroborated these findings, as clinicians agreed that assessing cognition in the real-world was a resource-heavy activity and the CoMET required significant time and financial resources as opposed to naturalistic simulated instruments.

The context for the new naturalistic instrument is a planning task that has individuals prepare a menu and plan a trip to a grocery store. The clinicians reported that community environments were of more importance to rehabilitation goals and the most impacted by limitations in functional planning with their clients. The decision to use a simulated testing environment was balanced with the importance of community environments by creating a task that required individuals to prepare and plan for a task that would involve functioning in the community. The grocery store was selected as the travel-location for several reasons. First, it was the most commonly reported community environment in the semi-structured interviews (n=5) and the environment of highest importance to independent living goals. Clinicians identified that most individuals that seek to live independently and function in the community, and would need to be able to plan meals and make purchases at a grocery store. Grocery stores are a common context for naturalistic instruments, as the CoMET and both community-based naturalistic instruments (n=2) and virtual-reality naturalistic instruments (n=2) from the scoping literature review had individuals complete tasks in grocery stores. Grocery stores have not been used as a context for naturalistic simulated instruments, and the new naturalistic instrument will employ a novel design by being the first naturalistic simulated instrument to involve the grocery store context. Since it is not feasible to simulate many of the environmental aspects of the grocery store into a clinical space, the new instrument will use specially designed cookbooks, grocery store catalogues and grocery store maps that reflect the visually distractions present in grocery stores.

The new naturalistic simulated instrument utilizes several different types of tasks that individuals need to complete during a functional planning activity. The scoping literature review identified several instruments that shared commonly used tasks as part of their naturalistic instrument: planning an event (n=7), budgeting/shopping (n=7), information gathering (n=7), and using a map to plot a course (n=3). The semi-structured interviews also found that scheduling/planning (n=5), planning a recreational activity (n=3), grocery shopping (n=2) and money management (n=2) were tasks significantly impacted by functional planning limitations and of importance to independent living and community functioning. Aspects of each of these task types were included in the new naturalistic simulated instrument, as they could feasibly be interweaved into a functional planning activity. The CoMET performance data found that individuals had difficulty with information gathering tasks and made mistakes related to poor searching strategies. Additionally, the sample's performance showed low accuracy on tasks that had implicit or unstated demands. These low-structure task demands are often present in the demands of daily life. Thus, are valuable to include in the functional planning activity.

The naturalistic simulated instrument includes aspects of scoring related to plan generation and plan execution. The scoping literature review identified several instruments created distinction between plan generation and plan execution in their scoring systems (n=10). There are several factors that comprise both stages of planning, including measures of speed, efficiency and accuracy of the plan. The semi-structured interviews found several highly important cognitive domains that overlapped with the findings of the scoping literature review. Highly rated cognitive domains were associated with both plan generation and plan execution, supporting their inclusion in the instrument scoring. Scoring for the new naturalistic simulated instrument does not involve subjective ratings of "errors". This decision was made in part because of the difficulties encountered during CoMET scoring, and the conceptual muddiness observed in the "error" coding of community-based and simulated naturalistic instruments from the scoping literature review. Inspired by automatic score calculations observed in computer and virtual-reality naturalistic instruments, the new naturalistic simulated instrument bases its scoring of objective performance measures. This includes using objective behavioral observations (e.g. # seconds on task), products of performance (e.g. written plan, grocery lists, drawn paths through grocery store map) frequency counts (e.g. # of times participant switches task) and responses to Likert questions (e.g. how efficient are you at completing planning tasks in every-day situations?). In the effort to ease scoring burden on clinicians and facilitate calculation of molecular aspects of plan generation and execution, a programmed excel can be used to enter observations and automatically process and compute values representing the above-mentioned planning variables.

5.5.1 Construction of the PLANS

The products of the synthesis informed the first draft of the PLANS. The first draft of the PLANS was designed to take place in a simulated clinical environment, and uses every-day tasks involved with meal preparation and planning a trip to the grocery store. Specially-designed materials were developed based off real-life items required to plan a meal and a trip to the grocery store, including: a cookbook, a grocery store catalogue and a grocery store map. These items were carefully crafted to standardize the information available during the task. The ingredients, prices, and grocery store aisle locations were set with specific values that limited the range of correct responses, reducing the chance that participants could accidently pass the PLANS rules. PLANS materials were designed to be as visually-complex as those encountered in grocery stores, to match the amount of visual scanning and information gathering required during grocery shopping. The first draft of the PLANS was designed with several rules that complicate the tasks and introduce constraints restricting the amount of correct selections. This increases the amount of anticipation, organization, sequencing and efficiency required during planning. The first draft of the PLANS was designed to be portable and manualized to increase its utility as a clinical tool. The first draft of the PLANS was designed to take between 30 and 45 minutes to complete, balancing the time-demands of real-world tasks with the average amount of time necessary to complete simulated instruments.

5.6 PLANS Content Validity

Four experts completed the content validity surveys. The Content Validity Index – Scale (CVI-S) was calculated for PLANS item relevancy, clarity and sufficiency by calculating the average of item CVI for each domain. The results form the surveys resulted in the CVI-S for PLANS Relevancy = 1.0, PLANS Clarity = 0.83 and PLANS Sufficiency = 0.83. Results for individual categories of the PLANS content validity survey can be observed in Table 15. In 10 of the 18 survey items, experts were in perfect agreement that the relevancy, clarity & sufficiency were very sufficient or mostly sufficient and adequate for piloting.

Table 15 PLANS Content Validity

	Relevancy		Clarity		Sufficiency	
	Score	% Agreement	Score	% Agreement	Score	% Agreement
Tasks	3.25	100%	3.5	75%	3.75	100%
Stimuli	3.75	100%	4	100%	4	100%
Planning and Anticipation	3.75	100%	3.5	100%	3.75	100%
Plan Generation	3.75	100%	3.75	100%	3.5	100%
Plan Execution	4	100%	3.5	100%	3.75	100%
Plan Detail	3.5	100%	2.25	25%	2.75	75%
Plan Accuracy	4	100%	3.75	100%	3.5	75%
Plan Anticipation	4	100%	3.25	75%	4	100%
Plan Generation Speed	4	100%	4	100%	4	100%
Plan Efficiency	4	100%	4	100%	3.75	100%
Task Familiarity	4	100%	4	100%	3.25	75%
Task Estimations	4	100%	3.5	100%	3.75	100%
Plan Execution Accuracy	4	100%	3.75	100%	3.75	100%
Adherence to Plan	4	100%	3.25	100%	4	100%
Strategic Behavior	4	100%	3.75	100%	3.75	100%
Rule Following	3.75	100%	4	100%	4	100%
Plan Execution Speed	3.75	100%	4	100%	4	100%
Accuracy	4	100%	3.75	100%	4	100%

Content validity survey results prompted several revisions to the PLANS before piloting. Raters provided feedback that several of the scoring domains appeared time-consuming and effort-intense to score from PLANS performance, which limits the utility of the instrument. The Plan Accuracy scoring domain did not reach adequate CVI, with raters providing feedback that categorizing 'errors' in performance on the meal preparation task seemed difficult. This was responded to by revising scoring domains to remove subjective error coding from plan execution accuracy, and to develop a computerized scoring system for the PLANS scoring. This computerized scoring system is able to automatically calculate complex scores such as: how much time is spent using each PLANS material, plan omission & sequencing errors, # of steps matched in plan generation and plan execution, # of ingredients in the menu selected, etc. This removed the subjectivity associated with the scoring domains, leaving the only metric to score by PLANS administrators in what order PLANS steps are completed and how much time participants spend on each step. PLAN Detail scoring was removed from the PLANS due to low ratings from content experts in clarity and sufficiency of the scoring metric. Items that received low clarity ratings were revised to clarify instructions or wording of tasks. The PLANS tasks and the scoring domain plan efficiency were reviewed and edited to increase the clarity of the instructions and scoring criteria, respectively. Following these revisions, a final version of the PLANS was prepared for piloting.

6.0 Feasibility and Preliminary Psychometrics of the Planning in Life and Adapting to Novel Situations (PLANS)

6.1 Introduction

Individuals with traumatic brain injuries can experience impairment in executive functioning, which limits the capacity to complete planning tasks required for independent living and community functioning, including the ability to plan. Functional planning abilities are necessary for successfully completing goal-directed behavior, and facilitate adapting to change and solving novel problems required for independent living and community functioning. Functional planning is an important factor in TBI rehabilitation, however there are few instruments available that provide information generalizable to community rehabilitation intervention. Naturalistic instruments demonstrate stronger ecological validity than traditional planning tests, and provide functionally-relevant information for assessment and goal planning in community rehabilitation.

The Planning in Life and Adapting to Novel Situations (PLANS) is a new naturalistic simulated instrument designed to fill the vacant niche of testing options for identifying how an individual's functional planning abilities impact functioning in the context of everyday living. The result of a complex instrument development project, the PLANS is intended as an option for community rehabilitation clinicians to measure functional planning abilities of their clients with moderately severe brain injuries. The PLANS uses naturalistic simulated tasks administered in a clinical space to assess functional planning abilities and reflects the planning demands of community and independent living. The PLANS is designed to guide rehabilitation interventions with stronger ecological validity than current instruments, and incorporates functional assessment of plan formulation and execution in an open-ended planning task using novel stimuli and design.

Before a newly developed instrument can be used in research or clinical contexts, a series of standardized studies should be conducted to establish the feasibility, reliability and validity of the instrument with its target population. In order to investigate the feasibility of the PLANS for measuring functional planning, a pilot study was conducted with a sample of individuals with traumatic brain injuries. This study investigated if the PLANS can be completed by the sample, and investigate preliminary performance norms. Other preliminary psychometric properties of the PLANS will be investigated including its inter-rater reliability, construct validity and ecological validity.

6.2 Research Aims

The primary aim of this study is to establish feasibility of administering the PLANS with individuals with traumatic brain injuries. In addition, PLANS preliminary psychometric properties are investigated. It is hypothesized that the PLANS will demonstrate convergent construct validity with planning measures and discriminant construct validity with non-planning measures. It is hypothesized that the PLANS will demonstrate ecological validity through significant moderate-strength relationships with an every-day executive functioning measure. Preliminary inter-rater reliability, construct validity, and preliminary ecological validity will be investigated through analyses of PLANS performance, neuropsychological testing and clinical rating scales.

1. The first aim for this study is to investigate the feasibility of the PLANS for use with individuals with TBI.

a. Pilot performance will be assessed though descriptive statistics of PLANS performance variables by the pilot samples

b. PLANS feasibility will be determined if subjects complete >90% of the tasks involved with the CoMET

2. The second aim for this study is to investigate reliability of the PLANS scoring and items.

a. Inter-rater reliability will be investigated through analyzing agreement between 2 trained raters

3. The third aim for this study is to investigate validity of the PLANS.

a. Convergent validity of the PLANS will be investigated through relationships between the PLANS scores and scores of other cognitive tests measuring planning

b. Ecological validity of the PLANS will be investigated through relationships between PLANS performance and the scores from a clinical rating scale of everyday cognitive functioning

6.3 Materials & Methods

6.3.1 The PLANS

The Planning in Life and Adapting to Novel Situations (PLANS) is a naturalistic simulation designed to measure functional planning. The PLANS is framed to reflect the real-world task of planning to go out in the community. The task that is presented is to plan a dinner for two and prepare for a trip to the grocery store. This activity involves several different subtasks including: identifying a budget, selecting recipes out of a cookbook, identifying the prices and ingredients of the recipes, and planning a route through a grocery store to shop for those ingredients. The subject also needs to listen to a voicemail that describes additional items that need to be purchased while at the grocery store. The participants are given a series of rules that need to be followed while completing the tasks, and need to follow several meal-specific constraints while selecting their dinner menu. Participants have 5 minutes to generate a plan of how they will accomplish the preparation activity, and are told they have 30 minutes to execute their plan and complete the preparation tasks.

The full administration manual for the PLANS is available in Appendix D. The PLANS requires between 30-45 minutes administer. The introduction to the PLANS takes around 10 minutes, with half that time spent going through the instructions and the other half spent asking the participant questions about their familiarity with shopping tasks. Participants have up to 5 minutes to generate a written plan of how they will conduct the preparation tasks, and are given as much time as necessary to execute the plan. While participants generate their plan, the examiner is observing their behavior and noting the number of times a step in the plan is revised once it has been written down. After the participant finished their plan or after time runs out, the examiner instructs the participant to carry out their plan. Participants are instructed that the plan execution should take 30 minutes, and a timer is set to go of 30 minutes after they begin executing the plan, but they can continue working after the timer ends. While the participant executes their plan, the examiner records the order that tasks are completed in; including the start time and end time for the task. Several PLANS scoring measures are calculated after the PLANS is completed. These scoring measures use data from the examiners score sheet and the participants written materials to determine scores. An excel calculator is available to input examiner records and participant records and output a summary of performance. The assessment flow for the plans can be referenced in Table 16.

Table 16 PLANS Administration Flow

Stage	Time	Examiner actions	PLANS Scoring
Instructions	10 minutes	Read through instructions	N/a
Plan Generation	5 minutes	Start timer Mark frequency that participant revises a step in plan. Check plan to identify if the PLANS rules are included	**Plan Generation** Plan Generation Time Plan Generation Steps Plan Generation Omission Errors Plan Generation Sequencing Errors Plan Generation Rules Anticipated Plan Generation Plan Alterations
Plan Execution	30 – 45 minutes	Start timer Record the order that tasks are completed in, how much time is spent on each task, and monitor performance recording frequency of strategic behaviors. Record the final menu selected, the total cost of items and the order of grocery aisles visited. Collect the written plan, the scrap paper, the grocery list, and the grocery map.	**Plan Execution** Plan Execution Time Plan Execution Steps Plan Execution Omission Errors Plan Execution Sequencing Errors Plan Execution Rules Followed **Meal Accuracy** Grocery Budget Disparity Grocery # of errors (prices & aisles) Grocery Incorrect Aisles Visited

		Post-administration: Input responses into Excel Worksheet to calculate Plan Alteration & Execution Accuracy scores.	

The PLANS is designed so that it can be administered sitting at a table in a clinical office. The examiner and participant sit at opposite ends of the table, and the examiner has the scoring materials obscured with a clipboard or document stand. Participants have several objects facing them at the front of the table. From left to right they have: a binder containing the PLANS cookbook, grocery store map, and grocery store catalogue, a scrap piece of paper to write on, and a sheet of paper with the PLANS rules. On the right side of the table, participants are provided a calculator and a pen that they use to complete the tasks. A digital timer is set up so that participants can keep track of time, and should be slightly angled so that the examiner can note the time as well. Behind the other testing materials, an envelope with the budget for the PLANS and an audio player with a pre-recorded voicemail are set up within reach of the participant. After participants complete the plan generation stage, they are provided a worksheet they can use to record the ingredients, prices and aisle locations for their menu. The layout of the materials can be referenced in Figure 7.

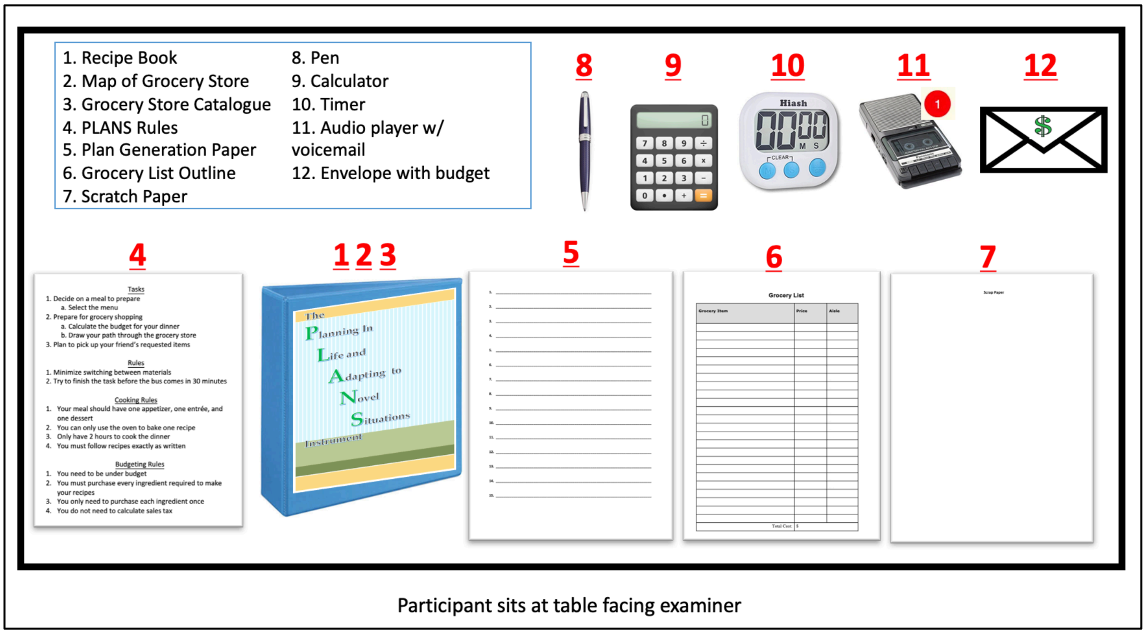


Figure 7 PLANS Materials Layout

6.3.2 Neuropsychological Testing

6.3.2.1 BADS Key Search

The Key Search is a subtest of the Behavioral Assessment of Dysexecutive Syndrome (BADS) assessment battery developed by Wilson, Evans, Emslie, Alderman & Burgess (1998). This test is used as a measure of an individual's ability to generate an efficient and effective plan of action, and a measure of ability to monitor one's behavior. This test also includes elements of "ill-defined" demands, as the individual needs to take into account factors that are not "explicitly stated". BADS subtests profile scores between 0-4 that reflect effectiveness of performance. In this study, the BADS Key Search has been selected to test convergent validity due to its relations to planning (Norris & Tate, 2000).

6.3.2.2 BADS Zoo Map

The Zoo Map is a subtest of the BADS assessment battery (Wilson et al.1998). This test is used as a measure of planning, rule-following, and on-line monitoring. This test requires individuals to plan a route through a map of a zoo that visits a certain number of locations while following arbitrary rules. This test has a high-demand and a low-demand sub-tests; zoo map version A and zoo map version B, respectively. While the goal of the test stays the same, the zoo map version A involves more rules and less direct instructions, which requires further planning ahead before action in order to produce an errorless performance. Zoo map version B involves following a set of concrete instructions listing the zoo locations in a particular sequence. Zoo map version A and B produce raw scores between 0-8. These raw scores are summed and matched with a BADS subtests profile score between 0-4 that reflect effectiveness of performance on both versions. In this study, the BADS Zoo Map A has been selected to test convergent validity due to its relations to planning (Norris & Tate, 2000).

6.3.2.3 BADS Modified Six-Elements

The Modified Six-Elements Test is a subtest of the BADS assessment battery (Wilson et al.1998). This test is a simplified version of the original Six-Elements Test that was used by Shallice and Burgess (1991b), paired with the Multiple Errands Test in order to study ecologically valid measures of executive functioning. This task requires individuals to organize their time and attempt some of the six separate tasks in a certain amount of time while following arbitrary rules. This test is used to measure an individual’s ability to plan, organize and monitor behavior. This test is an appropriate choice for a concurrent validity for this study specifically because it emphasizes plan execution through initiation of the plan, monitoring their progress and regulate their behavior towards completion while attending to the test rules. BADS subtests profile scores between 0-4 that reflect effectiveness of performance. In this study, the BADS Six-Elements Test has been selected to test convergent validity due to its relations to planning (Norris & Tate, 2000).

6.3.2.4 Delis-Kaplan Executive Functions System (DKEFS) Tower Test

The D-KEFS Tower Test is an example of a well-defined test that measures problem-solving and planning (Delis, Kaplan & Kramer, 2001). The test involves a series of 9 problems of increasing difficulty where an individual must generate a sequence of moves that will achieve a specified end state while making as few moves as possible and while following rules. The Tower Test is inspired by the Tower of London test, originally developed by Shallice (1982) as a tool to study Supervisory Attentional System theory of executive control in individuals with frontal lobe damage. In this study, the D-KEFS Tower Test has been selected to test convergent validity because it is a measure of planning (Shallice, 1982, Chan et al., 2007).

6.3.2.5 Cognitive Estimations Test

The Cognitive Estimations Test (CET) was first established by Shallice and Evans (1978) as a measure of cognitive estimation. Cognitive Estimation was proposed to be an aspect of executive functioning, and included the ability to reason, problem solve and use deductive reasoning to come to a logical conclusion. This is considered a separate cognitive domain from semantic knowledge that the individual possesses, however fund of knowledge is related to cognitive estimation (Della Salle et al., 2004). Individuals with brain injuries to their frontal lobes experienced difficulties in providing answers to questions estimating time, weight, height and other quantitative measurements (e.g. how tall is the empire state building in feet?) (Shallice & Evans, 1978). This test is included as a convergent measure to the construct validity of the

PLANS. The CET is not a direct measure of planning, however the ability to use cognitive estimation, logical reasoning and deduction are important in the completion of everyday tasks, and are hypothesized to be related to functional planning. Functional planning that occurs in the context of everyday demands requires individuals use reasoning and logic to set realistic goals and anticipate challenges that would impede completion of functional planning tasks. The CET will produce a raw score from 0-27 reflecting the accuracy of estimation. While being used as a convergent validity measure for an instrument measuring functional planning, the CET is limited in as it demonstrates inconsistent psychometric properties, including low internal validity and low divergent validity (Spencer, Johnson-Greene, 2009).

6.3.2.6 Boston Naming Test

The Boston Naming Test (BNT) is a commonly used neuropsychological test for measuring word retrieval (Kaplan, 1983). The BNT has 60 items of line drawings that participants need to name. As items continue, the drawings become more difficult to name because they represent less frequent words or more complex words. BNT performance is associated with language abilities, and is included in the Boston Diagnostic Aphasia Examination to identify aphasia-related disorders (Kaplan, 1983). The Boston Naming Test will produce a score from 0-60 indicating the number of line drawings correctly names. The use of BNT as a measure of divergent validity for the PLANS is justified, as there are few relations between word finding and planning (Obler et al., 2010).

6.3.2.7 Cognistat 5 Pencil & Paper Test

The Cognistat 5 pencil and paper edition) was used as a brief assessment of cognitive functioning (Cognistat, 2013. The Cognistat 5 includes 4 tests that require memory registration, orientation to place, time and person, visual construction and memory recall. The Cognistat is a brief 5-minute abbreviated version of the full Cognistat test. The full Cognistat test demonstrates the ability to detect and classify neurocognitive status in individuals with brain injuries (Doninger et al., 2006). In this pilot study, the Cognistat-5 pencil and paper test is used as a measure of global cognitive impairment.

6.3.3 Dysexecutive Questionnaire

The Dysexecutive (DEX) Questionnaire will be used to assess everyday executive functioning. The DEX Questionnaire was developed by Burgess et al. (1996) and has been validated in a larger study of the Behavioral Assessment of Dysexecutive Syndrome (BADS) by Wilson et al. (1998). This questionnaire is comprised of 20 items that assess everyday executive functioning from motivational behavior and cognitive perspectives. The questionnaire's items use a 5-point Likert-scale to rate the frequency of every-day executive problems, which produces a total score by summing all ratings. A higher total score indicates more executive dysfunction individuals experience in their daily lives, total scores can range from 0-80. In order to measure every-day executive functioning abilities, both subjects and an informant will fill out the DEX questionnaire. Informant DEX questionnaires were administered through Qualtrics, a web-based program. Informants could be a respective guardian, family member, spouse, clinician or close friend. A total score is calculated from subject and informant DEX ratings. The DEX Questionnaire is often used as a measure of ecological validity in studies that examine the ecological validity of executive functioning and naturalistic instruments (Wood & Liossi, 2006, Sanders, Lowe, Schmitter-Edgecombe, 2014). DEX Questionnaire subscales behavioral and emotional self-regulation; metacognition and executive cognition were calculated using scales proposed by Simblett & Bateman (2010) (see Table 9).

6.4 Research Subjects

6.4.1 Enrollment

The subjects included in this pilot study consisted of a sample of individuals with traumatic brain injuries. Subjects were recruited from local community rehabilitation programs, through Pitt+Me online subject registries, and through flyers located in authorized public spaces.

Screening procedures were utilized to identify individuals with a history of a TBI that experience chronic cognitive symptoms. A screening interview was used to identify eligible participants. The screening interview included a demographic survey to assess inclusion criteria and the HELPS screening tool to screen for history of brain injury (Picard, 1991). The HELPS is a brief 5 minute interview that asks about an individual's history assessing for events and symptoms that indicate a high likelihood of TBI. The HELPS establishes criteria for a positive brain injury screen as affirming statements in three question categories: 1.) An event that could have caused a brain injury (trauma or event) 2.) A period of loss of consciousness or altered consciousness after the injury or another indication that the injury was severe, and 3.) The presence of two or more chronic problems that were not present before the injury.

Inclusion criteria for participation in this pilot study included: 1) Positive HELPS screen with chronic cognitive symptoms 2) At least 6-months post-injury 3) Age 18-59 years old 4) Native English speaker. Exclusion criteria will be determined through screening questions and subject ability to complete over-the-phone or in-person recruitment communications. Exclusion criteria included: 1) Severely impaired mental status 2) Presence of significantly impaired expressive or receptive language abilities 3) Sensory impairment that would prevent completing the PLANS and other cognitive tests 4) History of psychiatric illness (bipolar disorder, schizophrenia, personality disorders). Subjects that had legally authorized guardians were not excluded from the study, and legally authorized representatives were contacted for proxy consent in addition to subject assent.

Individuals with brain injuries were provided a $50 WePay debit card as compensation for their time and effort.

6.4.2 Participants

Participants were enrolled between September 2018 and December 2018. 40 individuals were screened for eligibility. 18 participants were excluded: 6 had negative HELPS screenings (n=5 no chronic cognitive problems and n=1 no altered consciousness after trauma), 4 were over age limit, 3 had a history of psychiatric conditions, 1 was unable to complete screening due to significant language or behavioral impairment, 1 was unable to complete screening due to significant behavioral impairment, 1 had substance abuse concerns, 1 had a non-TBI (Multiple Sclerosis), and 1 was a non-native English speaker. After being enrolled, 2 participants cancelled more than 2 appointments and were removed from the study. In total, 20 subjects completed study procedures after being enrolled.

6.5 Study Procedure

The research design for this study is a cohort design. It was non-experimental and therefore, did not utilize randomization, blinding, or allocation processes. Participants completed all research activities over one session that took no longer than 3.5 hours in total. The study involved administering the PLANS along with a series of neuropsychological tests and completing a questionnaire. Subjects were asked to identify someone close to them (such as a spouse, clinician, friend or family member) to complete a questionnaire about their day-to-day functioning. The PI emailed a web link to the subjects containing a web link for the DEX Questionnaire administered through the Qualtrics© online format. Informants provided the name of the subject they were filling out the survey for, and their relationship to the subject, and surveys were prefaced with information about the study and confidentiality.

Subjects met the primary investigator at a designated clinical office (University of Pittsburgh clinical offices, ReMed Brain Injury Rehabilitation clinical offices or the Hiram G Andrews Center) to complete the study protocol. These offices are in quiet and secluded spaces, and a sign was posted at the front of the door to inform others not to disturb the testing session. The research tasks were administered all at one time point, and in a face-to-face format.

The study flow for the PLANS pilot can be observed in Figure 8. The PI and subjects read through and signed an informed consent document before beginning study protocol. Demographic information including age, time since injury, marital status, employment status, race and education was collected. Study protocol included administering several measures of cognitive functioning. First, subjects completed the Cognistat 5 (Cognistat, 2013). After completing the Cognistat 5, subjects were introduced to the PLANS. Participants were given the instructions and rules necessary for completion of the PLANS. Each testing space was equipped with a Go Pro Session-5© camera on a tri-pod mount to capture and record subject performance and behaviors during testing. Subjects were informed about camera placement before being introduced to the PLANS instructions. After completing the PLANS, participants completed short cognitive interviews with the primary investigator assessing their comprehension of the instructions.

Afterward, subjects completed a testing battery with the primary investigator of this study. The PI had experience and training in administering the selected tests and was supervised by a licensed clinical neuropsychologist. These tests were administered in the following order: Behavioral Assessment of Dysexecutive Syndrome (BADS) Key Search Test, BADS Zoo Map Test, BADS Six-Elements Test, Cognitive Estimation Test, Boston Naming Test and Delis-Kaplan Executive Function System (D-KEFS) Tower Test. At the end of this testing battery, subjects completed the Dysexecutive (DEX) Questionnaire Self form.

Institutional Review Board approval for this pilot study was obtained through the University of Pittsburgh's IRB board.

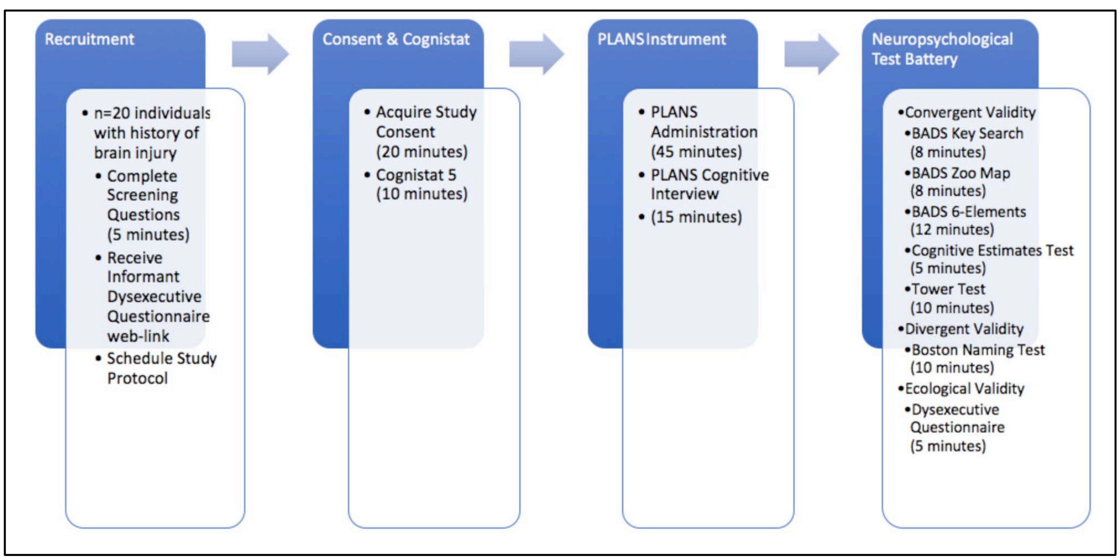


Figure 8 PLANS Pilot Study Flow

6.5.1 PLANS Rater Training & Procedure

Two Master's level students from the University of Pittsburgh's Clinical Mental Health and Rehabilitation Counseling program were recruited to act as clinician raters of PLANS video performances. These students received 1 hour of formal training on how to administer and score the PLANS, along with training how to use the video observation software. Raters worked independently of each other, did not have access to study data or subject personal identifiers and were never in any contact with research subjects.

Raters were monitored over 4 weeks as video ratings were completed, and independently coached on PLANS scoring classifications. In the first week of coding, rater conflicts were assessed and resolved through discussion and training involving the primary investigator.

The software package NOLDUS Observer-XT© version 12.5 was used to code video performance on the PLANS. The Observer-XT is a widely-used software package developed by NOLDUS that allows for the coding and analysis of behavior data from video files. Video files were coded with PLANS steps, marking what step participants were on at any given time, allowing for tracking the duration and sequence of steps completed during the PLANS.

6.6 Analyses

Analysis of this data set was completed using IBM SPSS (version 25.0) with a significance level of .05 unless otherwise specified and Noldus Observer-NT software (version 12.5). All assumptions for statistical testing were checked prior to analyses. This data set includes (n=20) individuals with brain injuries. There is no missing data in this sample. A significant outlier is present in the sample, which scored ≥ 2 standard deviations below the sample mean on the PLANS and neuropsychological testing. This case's data is retained for sample demographics, inferential statistics including those investigating preliminary construct validity and preliminary ecological validity and inter-rater reliability analyses but removed from the descriptive statistics of performance on the PLANS by the sample. Inferential statistics are conducted with and without the outlier data, and findings that are present in both analyses are reported.

Descriptive statistics were generated for subject demographics and all study variables to derive the mean (with standard error), median, standard deviation, range, skewness and kurtosis (with standard errors) and identify any major outliers. Simple histograms were plotted to allow visual inspection of frequency distributions. Additionally, heteroscedascity of variables included in correlation analyses were visually assessed from histograms created in descriptive exploration. Non-parametric statistical tests were utilized for all inferential tests, as study variables were not normally distributed.

Feasibility of the PLANS was investigated in this study through descriptive statistics and frequency tables. Descriptive statistics were inspected from PLANS performance variables to identify the distribution of performance of the pilot sample. Feasibility of the PLANS is determined if a majority of the sample was able to complete the PLANS and attempt 90% of the tasks.

Inter-rater reliability of PLANS scoring is evaluated using Kappa test statistics. Kappa test statistic ranges from 0 (agreement equal to chance) to 1 (perfect agreement) and values above .60 indicate substantial agreement between raters (Sim, & Wright, 2005; Portney & Watkins, 2009, p. 598-600). Inter-rater reliability is only calculated for the "Plan Execution" scoring component, as this is the only “clinician scored” process in the PLANS administration. Therefore, it is the only PLANS scoring component that requires inter-rater reliability as every other performance variable is recorded via participant written products or computed by the programmed excel score sheet.

The construct validity and ecological validity of the PLANS was investigated by conducting 2-tailed correlations between PLANS scores and neuropsychological testing or DEX Questionnaire scores, respectively. Spearman-Rho tests were used with alpha = .05. The interpretation criteria used for correlation coefficients are: weak $r < .35$, moderate $r = .36$-.67 and strong $r > .68$-1.0 (Taylor, 1990).

6.7 Results

The results reported for this pilot study include performance on the PLANS neuropsychological testing and DEX Questionnaire data. The inter-rater reliability, preliminary construct validity and preliminary ecological validity of the PLANS are also reported.

6.7.1 Subject Demographics

The demographic characteristics of the sample are presented in Table 17. The sample included 6 men, 13 women, and one non-binary individual. The mean age was 44 years old, and the samples’ age ranged from 22 years old to 59 years old. The median time since injury was 7.5 and 30% of the sample experienced their injury less than 5 years ago and 40% less than 10 years ago. 80% of the sample experienced a TBI from either a fall (40%) or a motor vehicle accident

(40%), and the remainder experienced multiple traumatic events resulting in a TBI (falls, motor vehicle accidents and other trauma) with the mean number of events = 2.5. All participants were high school graduates, the mean years of education was 14.25 (*sd*=2.25). 85% of the sample was not gainfully employed and 75% were not employed in any capacity.

The only demographic characteristic that was significantly correlated with PLANS performance was # of years of education, which was significantly ($p < .05$) correlated with Time Estimation Disparity (r=-.454, p=.044), Plan Generation Steps (r=.50, p=.025), Plan Generation Omission Errors (r=-.528, p=.017), Plan Generation Plan Alterations (r=.583, p=.007) and Plan Execution Steps (r=.541, p=.014).

Table 17 PLANS Pilot Demographics

	PLANS TBI Group (n=20)		
	Mean	Standard Deviation	Range
Age (years)	44.0	13.2	20-59
Years Since Injury (years)	11.7	12.1	1-38
TBI Etiology			
Motor Vehicle Accident	8 (40)		
Fall	8 (40)		
Multiple Concussion	4 (20)		
Race n (%)			
White	16 (80)		
Black	2 (10)		
Asian	1 (5)		
Multiracial	1 (5)		
Gender n (%)			
Male	6 (30)		
Female	13(65)		
Non-Binary	1 (5)		
Highest Education Achieved			
High school degree or equivalent (E.g. GED)	4 (20)		
Some college but no degree	8 (40)		
Associate's degree	1 (5)		
Bachelor's degree	0 (0)		
Graduate degree	7 (35)		
Employment Status			
Employed, working 40 or more hours per week	1 (5)		
Employed, working 1-39 hours per week	4 (20)		
not employed, not looking for work	4 (20)		
Retired	2 (10)		
Disabled, not able to work	7 (35)		
Not employed, volunteering part time	2 (10)		
Cognistat MCI (MCI Score /6)	0.65	1.22	0-4
BADS Zoo Map (Z-Score)	-0.48	1.37	-2.16-1.38
BADS Six Elements (Z-Score)	-0.28	1.15	-3.15-0.60
BADS Key Search (Z-Score)	-0.11	0.86	-1.97-1.06
Cognitive Estimates Test (Z-Score)	-0.85	1.48	-3.23-2.17
D-KEFS Towers Achievement (Z-Score)	-0.05	1.1	-0.66-2.00
Boston Naming Test	-0.38	2.34	-9.27-1.20

6.7.2 PLANS Performance

Overall, the PLANS shows promising feasibility for use with individuals that have experienced a TBI. Overall, 95% of the pilot sample completed more than 90% of the tasks presented to them in the PLANS, indicating that the PLANS instructions and demands are not producing a significant floor effect. The outlier subject only completed one of the PLANS tasks, and it is suggested that the demands of the PLANS were too complex and open-ended for them to complete, as this individual needed several attempts to repeat back the instructions correctly. It is possible that the PLANS may not be a suitable instrument for use with individuals that would not be able to comprehend the basic demands and instructions. Failure to grasp the basic demands of the PLANS could be clinically meaningful as an indicator for this individual's ability to function independently in the community or independent living contexts.

For the 19 individuals that effectively completed the PLANS, there were several interesting findings related to their ability to generate and execute plans of action. The sample's performance on plan generation, plan execution and meal accuracy can be observed in Table 18.

Table 18 PLANS Pilot Performance

	PLANS Pilot Sample (n=19)	
	Mean	Range
Plan Generation		
Plan Generation Time (minutes)	3.54 (1.03)	1.5-5.0
Plan Generation Steps	5.26 (2.38)	2-12
Plan Generation Omission Errors	3.42 (2.14)	0-7
Plan Generation Sequencing Errors	3.21 (1.51)	1-7
Plan Generation Rules Anticipated (/4)	0.89 (1.1)	0-3
Plan Execution		
Plan Execution Time (minutes)	29.22 (7.66)	15.95-41.12
Plan Execution Steps	13.42 (5.51)	6-24
Plan Execution Omission Errors	1.16 (0.69)	0-2
Plan Execution Sequencing Errors	10.26 (5.52)	2-22
Plan Execution Rules Followed (/5)	3.58 (0.77)	2-5
Meal Accuracy		
Grocery Budget Disparity ($)	$-11.14 ($22.72)	$-80.00-$17.00
Grocery # of errors (prices & aisles)	6.47 (7.70)	0-25
Grocery Incorrect Aisles	3.11 (2.92)	0-9

The written plans generated by the subjects before completing the PLANS on average omitted 3.42 (*sd* = 2.14) of the tasks that were necessary to complete the PLANS and had 3.21 (*sd* = 1.51) tasks in an inefficient order. The written plans ranged from having as many as 12 individual steps to a few as 2 and on average only listed 5.26 (*sd* = 2.38) steps. Although not scored, 68% of written plans contained one or more steps outside the scope of the meal preparation activity, in example: "get on the bus to go to the grocery store", "purchase all the items on my list" or "set the table for the dinner". In these cases, these steps were not scored or interpreted as fulfilling another required task (go to store and purchase items was scored as Draw path through grocery store map). These inappropriate steps in generated plans may reflect participants inappropriately investing themselves in the task at hand or difficulty comprehending the scope of the requirements for the PLANS. Very rarely were the PLANS meal rules included in the written plan, and 58% of the sample did not include any rules in their written plan. The types of errors observed in plan generation suggest that individuals with TBI may have difficulty anticipating challenges and comprehending the demands of complex, novel, and open-ended planning problems, and have difficulty creating efficient steps toward their goals.

The average amount of time required to complete the plan execution was 29.22 (*sd* = 7.66) minutes. In contrast to the planning generation, the sample's plan executions were very different in their efficiency and effectiveness. The plans executed by the subjects to complete the PLANS on average omitted 1.16 (*sd* = 0.69) of the tasks that were necessary to complete the PLANS and had 10.26 (*sd* = 5.52) tasks in an inefficient order. The reduced number of tasks omitted is an improvement compared to the plan generation omission errors, and indicates that the sample did not anticipate steps (nor rules) they would be required to complete in the task execution. Overall, their task executions did not follow the plans that they generated, and consisted on average of 13.42 (*sd* = 5.51) steps. Subjects had difficulty following the PLANS efficiency rule to "minimize switching between PLANS materials", and often went back and forth between the cookbook, grocery store catalogue, and grocery store map; details about the individual PLANS steps can be observed in Table 19. In the context of functional planning, this performance indicates that the sample had significant challenges monitoring their behavior, and needed to visit each PLANS material more than once in order to gather all the information necessary before moving on to the next step.

Three PLANS tasks took the majority of the completion time: selecting the dinner menu from the cookbook, recording the ingredient information from the grocery catalogue and calculating the dinner total cost. These tasks required significant visual scanning, visual attention, and concentration to complete, and reflect the demands on these areas required to complete shopping tasks in a real-world grocery store. Overall, the subjects committed several types of errors in completing these tasks, including recording incorrect aisle prices and locations, missing ingredients necessary to prepare their selected recipes, and incorrectly calculating the total price of their shopping lists. While these errors may be considered independent from the measurement of functional planning, they are clinically meaningful in the context of an individual's ability to complete everyday planning tasks that require precision and accuracy. An individual may be able to generate an efficient plan, and could execute steps in a plan in an effective sequence but if they commit many errors in information gathering they can still have significant difficulty reaching their anticipated goals.

Table 19 PLANS Individual Step Performance

		PLANS Pilot Sample (n=19)	
	Task	Mean Duration of Each Step (seconds) (SD)	Mean Frequency of each Step (SD)
1	Check envelope	39.25 (39)	0.95 (0.22)
2	Check tape recorder	66.09 (57)	1.16 (0.77)
3	Calculate remaining budget	66.46 (82)	0.16 (0.38)
4	Select menu from cookbook	206.26 (201)	2.63 (1.9)
5	Use catalogue to record ingredient info	170.74 (192)	3 (2.4)
6	Calculate menu total cost	130.56 (113)	1.95 (1.1)
7	Identify necessary store aisles	67.25 (80)	1.58 (1.2)
8	Draw path through store map	105.76 (106)	1.58 (1.5)

6.7.3 Inter-Rater Reliability

Inter-rater reliability was calculated using Noldus Observer Software for the duration and sequence that subjects completed steps during the plan execution phase of the PLANS. Percent agreement and Kappa were calculated between two clinician raters from the proportion of time PLANS steps were in agreement over total PLANS steps rated, where the two rater scores were contrasted and was calculated dichotomously as raters in agreement (1) or raters in disagreement (0). From the time subjects started their plan execution, each second of the PLANS was coded as one of 8 steps: 1 = Check the envelope; 2 = Checking the friend's voicemail; 3 = Checking the remaining budget for dinner after finding price of friend's requested items; 4 = Selecting recipes from cookbook; 5 = Recording ingredients information from grocery store catalogue; 6 = Calculating total cost of dinner; 7 = Identifying which grocery store aisles need to be visited; 8 = Drawing path through grocery store. Table 20 shows that both raters were in agreement with each other 89% of the time, with only 7477.2 seconds of disagreement across all PLANS videos indicating excellent reliability. Cohen's Kappa was calculated to measure the rate of agreement not due to random chance. Cohen's Kappa for the combined rater results Kappa = .86 indicates excellent inter-rater reliability, which was significant at the .05 level.

To evaluate if there was variability in rater agreement between the individual PLANS steps, the percent agreement was calculated for agreement within each step, which is reported in Table 21. Steps 3 and 7 have low percent agreement as compared to other steps, indicating that these steps were less likely to be accurately coded by the raters. Rationale for the low percent agreement for step 3 is that this step was rarely attempted, which reduces the number of ratable instances as compared to other steps. Rationale for the low percent agreement for step 7 is that the raters had difficulty identifying when participants were completing this step, as it required the map to be open and the participant to not have started drawing their path through the store.

Table 20 Inter-Rater Reliability of the PLANS

			Statistic	Confidence Interval
Observer A and Observer B Combined Results	(n=20)	Agreement (sec)	60422.3	
		Disagreement (sec)	7477.2	
		% Agreement (/100)	89.0	
		Kappa (/1.0)	0.86*	(.84-.87)
*Indicates significance				

Table 21 PLANS Individual Step Rater Agreement

	(n=20)	Percent Agreement (/100)
Check Envelope Step	17	76.0
Check Voicemail Step	17	81.2
Calculate Remaining Budget Step	5	56.1
Select Menu Step	20	92.7
Record Ingredient Information Step	19	84.1
Calculate Remaining Budget Step	19	89.2
Identify Store Aisles Step	17	60
Draw Path Through Store Map Step	19	88.4

6.7.4 Construct Validity

Spearman-Rho correlations were conducted between the PLANS scoring domains (Plan Generation, Plan Execution & Meal Accuracy), convergent validity neuropsychological tests (BADS Zoo Map, BADS 6-Elements, BADS Key Search, Cognitive Estimation Test, DKEFS-Tower Test) and divergent validity neuropsychological tests (Boston Naming Test). Results for the PLANS construct validity analyses can be observed in Table 22. The BADS 6-Elements Test was demonstrated significant moderate and strong strength correlations with plan generation (total # steps (r=.591), omission errors(r=-.722), sequencing errors(r=.450) and rules anticipated (*r*=.545)), plan execution (total # steps (*r*=.493)) and meal accuracy (grocery budget disparity (*r*=.518) and number of incorrect aisles visited (*r*=-.553)). The Cognitive Estimate Test was significantly correlated with plan generation (total # steps (*r*=-.531), omission errors (*r*=.579) and sequencing errors (*r*=-.574)). The DKEFS Towers Test demonstrated significant moderate strength correlations with plan execution (total # of steps (*r*=.523) , plan execution time (*r*=-.485)), and meal accuracy (grocery budget disparity (*r*=.599) and # of incorrect aisles visited (*r*=.564)). The Boston Naming test was not correlated with either plan generation or plan execution scoring measures, but demonstrated significant moderate strength correlations with measures of meal accuracy (grocery budget disparity (*r*=.488) and # of incorrect aisles visited (*r*=-.510)).

Overall, these findings support preliminary evidence of construct validity for the PLANS as a measure of functional planning. Neuropsychological tests of planning, strategy formulation and logical reasoning are predominantly related to the efficiency and effectiveness of a written plan of action. By requiring participants to write down a step by step plan on how to complete their meal preparation task, the PLANS measure the covert cognitive operations involved with generating plans required to complete every-day tasks, which is not often included in neuropsychological nor naturalistic planning instruments. More attention may be warranted toward how logical reasoning is necessary to generate plans in the context of daily life. Plan execution was significantly correlated with the DKEFS- Tower Test, which is sensible because the Tower Test involves completing a “puzzle-like” planning task in an efficient manner while following a simple rule. Between the plan generation and plan execution stages, there are differences in the amount and strength of correlations between the planning neuropsychological tests. This may suggest that the two stages of planning proposed in the PLANS are distinct from one another, and represent two different aspects of planning. Plan generation demonstrated preliminary construct validity with neuropsychological tests of logical reasoning and planning, while plan generation demonstrated preliminary construct validity with neuropsychological tests of planning, and meal accuracy demonstrated preliminary construct validity with neuropsychological tests of planning, logical reasoning and word finding.

It is reassuring that the Boston Naming Test was not correlated with any planning scores, but very interesting that it did correlate with accuracy of information gathering and budget calculation. The Boston Naming Test is used as a divergent measure because it is not related to planning, but it is more an indicator of crystalized intelligence. Every-day planning requires information gathering, and this finding may support that the PLANS is not only measuring planning, but also what challenges subjects have in detail-oriented tasks and accurate task completion. It is noteworthy that no neuropsychological test was correlated with errors or rules followed in plan execution. One explanation could be the relative low variability of the sample on omission errors and rules followed, which might be an issue with the sample size being lower than necessary to establish the power required for these non-parametric correlations.

Table 22 Correlations between PLANS and Neuropsychological Testing

N=20 non-normed np data	BADS Zoo Map Test A	BADS 6 Elements Test	BADS Key Search Test	Cognitive Estimation Test	DKEFS Tower Test Achievement	DKEFS Tower Test Mean Move Time	DKEFS Tower Test Rule Break Ratio	Boston Naming Test
Plan Generation								
Plan Generation Time (minutes)	0.211	0.376	0.334	-0.286	0.3	-0.134	-0.134	-0.32
Plan Generation Steps	0.288	.591**	0.091	-.531*	0.238	0.098	0.098	-0.192
Plan Generation Omission Errors	-0.264	-.722**	-0.14	.579**	-0.312	-0.183	-0.183	0.255
Plan Generation Sequencing Errors	0.01	.450*	0.209	-.574**	0.143	0.415	0.415	-0.255
Plan Generation Rules Anticipated (/4)	.456*	.545*	0.244	-0.368	0.257	-0.132	-0.132	-0.297
Plan Execution								
Plan Execution Time (minutes)	0.405	-0.095	-0.04	0.25	-0.076	-.485*	0.235	-0.038
Plan Execution Steps	.661**	.493*$	0.344	-0.306	.523*	0.186	-0.268	0.293
Plan Execution Omission Errors	-0.385	-0.392	-0.109	0.221	-0.304	0.121	0.251	-0.076
Plan Execution Sequencing Errors	.625**	0.2	0.209	-0.219	0.097	-0.01	0.013	0.048
Plan Execution Rules Followed (/5)	0.274	0.255	0.142	-0.206	0	-0.23	0.105	-0.163
Meal Accuracy								
Grocery Budget Disparity ($)	0.175	.518*$	0.268	-0.255	.599**	.554*	-0.386	.488*
Grocery # of errors (prices & aisles)	0.505*	0.053	-0.22	0.193	-0.013	0.317	0.387	-0.067
Grocery Incorrect Aisles	-0.182	-.553*	-0.359	.503*	-0.441	-0.257	.564**	-.510*

*Correlation is significant at the 0.05 level (2-tailed). **
*Correlation is significant at the 0.01 level (2-tailed).***
Significant at 0.05 level (2-tailed) if outlier present $

6.7.5 Ecological Validity

Spearman-Rho correlations were conducted between the PLANS scoring domains (Plan Generation, Plan Execution & Meal Accuracy) and the DEX Questionnaire (where a higher score indicates more severe everyday executive problems). The DEX Questionnaire Self & Informant ratings were included in this analysis, along with several scales as identified in previous factor analysis of the DEX-Questionnaire (Simblett & Bateman, 2011). Results for the PLANS ecological validity analyses can be observed in Table 23.

No significant correlations were found between the PLANS and DEX self-report measures. Significant moderate strength correlations were found between plan generation and plan execution time and DEX total score and DEX scales related to executive control and self-regulation. When the outlier is included in the analysis, statistically significant correlations are found between plan generation sequencing errors and DEX-Meta Cognitive factor and plan execution omission errors and DEX-Executive Cognition factor. These findings reflect low preliminary ecological validity for the PLANS, as the only PLANS score related to every-day executive function was time.

Table 23 Correlations between PLANS and DEX Measures

N=20 non-normed np data	DEX-Informant	DEX-BESR	DEX-MC	DEX-EC
Plan Generation				
Plan Generation Time (minutes)	-.551*	-.500*	-.590**	-.560*
Plan Generation Steps	-0.307	-0.25	-0.326	-0.319
Plan Generation Omission Errors	0.371	0.325	0.393	0.367
Plan Generation Sequencing Errors	-0.417	-0.386	-.486*$	-0.337
Plan Generation Rules Anticipated (/4)	-0.166	-0.158	-0.168	-0.29
Plan Execution				
Plan Execution Time (minutes)	.464*	.512*	.473*$	0.4
Plan Execution Steps	0.111	0.092	-0.104	0.041
Plan Execution Omission Errors	0.438	0.345	0.352	.505*$
Plan Execution Sequencing Errors	-0.007	0.003	-0.164	-0.015
Plan Execution Rules Followed (/5)	0.263	0.306	0.163	0.123
Meal Accuracy				
Grocery Budget Disparity ($)	0.183	0.042	0.013	0.155
Grocery # of errors (prices & aisles)	0.231	0.285	0.231	0.255
Grocery Incorrect Aisles	0.179	0.277	0.216	0.184

*Correlation is significant at the 0.05 level (2-tailed) **
*Correlation is significant at the 0.01 level (2-tailed) ***
Significant at 0.05 level (2-tailed) if outlier present $
DEX-BESR = DEX Factor "Behavioral and Emotional Self-Regulation";
DEX-MC = DEX Factor "Metacognition" ; DEX-EC = DEX Factor "Executive Cognition"

7.0 Discussion

This dissertation project aimed to develop and provide preliminary research support for a new naturalistic simulated instrument to measure functional planning of individuals with TBI: The Planning in Life and Adapting to Novel Situations (PLANS). A scoping literature review, semi-structured interviews and retrospective analysis of clinical data from a community-based instrument were conducted to provide guiding information for this development. This chapter will present three sections that discuss the implications, future directions and limitations of the studies included in this dissertation.

7.1 Implications

7.1.1 Novel Instrument Development Methodology

Naturalistic instruments are unique among instrument development projects because they must incorporate aspects of real-world scenarios in their administration. This requires new considerations in the development process that diverge from traditional instrument development such as: which tasks will be completed, what scenario will be selected, what environment will be modeled, what stimuli will be utilized and where the instrument will be conducted (real-world vs clinic vs computer-based). In order to comprehensively and empirically account for these novel factors, a unique and multi-pronged approach was developed for this dissertation project.

A notable outcome of this dissertation is the novel methodology used to produce the PLANS. The PLANS instrument development progress was a complex, rigorous and multi-faceted project, which is unique among naturalistic instruments. The goal of this methodology was to ground the PLANS in both clinical and research basis. Utilizing clinical data and stakeholder input were valuable to understanding the clinical importance of planning in every-day action and defining functional planning. The scoping literature review helped to identify the most successful aspects of validated naturalistic planning instruments. The importance of stakeholder and clinical input is especially salient when developing naturalistic instruments, as the task demands and environments are of critical importance to the face validity and verisimilitude to the independent living and community functioning contexts of meaning to target populations. There is a dearth of literature available that describes the processes used to develop new naturalistic instruments, and it is regarded that these instruments used low rigor in their development. One rare example was identified of a naturalistic instrument development that conducted scoping literature reviews and collected input from stakeholders in rehabilitation science (Radomski et al., 2013). However, the PLANS development included an additional stage incorporating findings from retrospective analysis of clinical data from another naturalistic instrument. There are very few examples of naturalistic instrument development projects, and even fewer provide a detailed rationale for their development process. The PLANS development project is unique in its multi-faceted, comprehensive development stages, and provided high empirical support for the construction on the PLANS.

Despite the success of this project, more effort is warranted to explore the most effective ways to integrate stakeholder and clinical involvement in naturalistic instrument development. There is currently not a standard for establishing "verisimilitude" of a naturalistic instrument, so researchers and clinicians can only make arbitrary claims as to how well their instrument represents the real world. Additionally, there are no standardized survey tools available to investigate the perceptions of stakeholders as to the salience of naturalistic tasks to the demands present in a client's daily functioning. These topics were investigated using newly created semi-structured interviews and content validity surveys (see Appendix B and Appendix C, respectively), however researchers would benefit from a unified approach to the development of naturalistic instruments. Despite these contemporary issues in naturalistic instrument development, this dissertation stands as a potential model for future development projects.

7.1.2 PLANS Utility

The rationale for undertaking this instrument development process was to address several limitations of planning measurement in community brain injury rehabilitation. Traditional planning instruments demonstrate low ecological validity, and are not sensitive to mild executive functioning impairment that can impact every-day cognitive functioning. Instruments that measure planning do not provide detailed information about an individual's ability to plan in day to day life, and are limited in their ability to guide prescriptive cognitive rehabilitation for functional planning limitations. Additionally, these instruments are poor determiners of the capacity for functional planning as it relates to independent living and community functioning rehabilitation goals. Naturalistic instrument scoring can be biased through the subjective application of 'error codes', and scoring performance can be a significant resource burden for clinicians. This dissertation provides preliminary evidence that the PLANS can be utilized in a manner that addresses these concerns, which is described below.

Preliminary evidence from the PLANS pilot study shows encouraging research and clinical support, and several inferences can be made about the utility of this tool. First, the PLANS can be used to measure functional planning, as moderate and strong significant correlations were found between PLANS and neuropsychological testing and measures of every-day executive functioning. PLANS scores could help establish whether an individual has the capacity to complete complex planning tasks that require anticipation, creating efficient plans and potential to follow through on those plans. A second utility for the PLANS is in guiding prescriptive cognitive rehabilitation interventions for individuals with functional planning limitations. Using novel performance metrics, the PLANS can measure the typically covert cognitive construct of plan generation. Conventional planning measures do not thoroughly assess plan generation, and solely focus on plan execution. The PLANS contributes a new tool that can give clinicians evidence supporting whether plan generation, plan execution or both are the root of functional limitations individuals experience in every-day planning. As observed in the pilot's performance on the PLANS, it is apparent that individuals can significantly deviate from plans generated and plans executed on every-day tasks, therefore this is especially intriguing and warrants further research attention. The third utility for the PLANS is perhaps as an additional tool for diagnostic purposes in clinical and research contexts. The PLANS demonstrated preliminary construct validity with measures of executive functioning, planning and reasoning, and could be useful assessing the nature and extent of executive dysfunction in individuals with TBI. The lack of a ceiling effect on PLANS performance may indicate that the PLANS could detect mild executive functioning impairment resulting from TBI that traditional performance-based planning test could not detect, which would be helpful for guiding rehabilitation intervention for those with injuries of mild and moderate severity. Results from the pilot study support that the PLANS may be especially appropriate in assessing executive functioning mild to moderately severe TBI, as the PLANS could be sensitive to even mild impairment in functional planning.

The PLANS addresses feasibility issues for naturalistic instruments by emphasizing resource efficency in its test administration and scoring. The PLANS is conducted in a clinical office setting, which is convenient and accessible for clients and clinicians. The time resources for administering the PLANS are minimal in comparison to community-based administration, but long enough that it reflects the cognitive demands of every-day planning tasks. In addition, the PLANS utilizes a novel scoring system that relies on clinicians observing objective behavioral patterns while the PLANS is completed, removing subjective error classification burden. Scoring for the PLANS is streamlined through relying on a client's behavioral observations and written products, as these scoring metrics are inserted into a programmed excel calculator that automatically generates planning errors, rule breaks, plan inefficiencies and errors in meal preparation. The PLANS incorporated these administration and scoring efficiencies as a means of increasing the feasibility of the PLANS for use, and the results from this pilot study supported the goal of increased feasibility for naturalistic instruments.

7.1.3 Preliminary PLANS Psychometrics

The PLANS pilot study achieved several of its intended goals related to investigating its psychometric properties. The PLANS was found to be feasible for use with individuals with traumatic brain injuries, with over 95% of the sample completing the meal preparation task. As identified by the poor outlier performance on the PLANS, the PLANS was found to be too complex for individuals with significant neuropsychological impairment, and may provide more useful assessment of functional planning of individuals with mild and moderate severity injury. The PLANS was not found to have any floor or ceiling effects in its scoring measures, which indicates that the scoring metrics and task demands were appropriate.

The PLANS demonstrated preliminary construct validity with measures associated with executive functioning and planning, which is an expected result of the pilot study. Interestingly, errors on the PLANS meal accuracy were associated with not only planning but also the discriminant validity measure of word retrieval. This may indicate that the errors that were committed on calculating, recording and information gathering may be associated with other cognitive domains than planning, such as word recognition attention, processing speed or crystalized intelligence. Another expected finding from the pilot study was that the PLANS was moderately correlated with informant ratings of every-day cognitive functioning. It is sensible that the PLANS was not related to self-appraisals of cognitive functioning because of anosognosia present after TBI. While correlations were not as strong as anticipated, this could be related to the appropriateness of the DEX questionnaire as a measure of 'everyday executive functioning', as the PLANS is measuring everyday planning ability. Future psychometric studies of the PLANS may utilize multiple measures of every-day cognitive functioning as a more comprehensive and complete investigation of ecological validity. Finally, the PLANS demonstrated excellent inter-rater-reliability, which is incredibly promising for usability by trained clinical professionals. The inter-rater reliability of the study raters increased dramatically after the first 4 subjects that they observed, which deflates the potential inter-rater reliability of the PLANS. This finding may indicate that more rater training before starting on the pilot subjects would have benefited rater agreement. High inter-rater reliability may also be attributed to the relatively low demands on raters, and the utilization of objective scoring metrics and computerized scoring computation, which reduces the degree to which measurement error can impact instrument reliability.

7.2 Future Directions

7.2.1 Psychometric Studies

The PLANS was found feasible for use with individuals with brain injuries and demonstrated exciting preliminary construct validity and ecological validity. Therefore, these results support that future work should be conducted to continue investigating the psychometric properties of the PLANS. This work would resemble the study design of the PLANS pilot study with several proposed alterations and additions. First a more comprehensive psychometric evaluation of the PLANS would recruit a larger, stratified sample of individuals with brain injuries. A larger sample (n>100) would be able to generate analyses of higher statistical power than the pilot, and investigate the construct and ecological validity of the PLANS with higher generalizability. Stratifying the sample by age, gender, and TBI severity would be used to establish a set of standardized norms for PLANS performance. These future psychometric evaluations of the PLANS should include increased screening procedures such as collecting medical records (Glasgow Coma Scale, Post-Traumatic Amnesia Loss of Consciousness or MRI scan) to establish severity of TBI. Psychometric evaluations of the PLANS would benefit from investigating its known-groups validity. This would help strengthen its utility as a screening tool for TBI by determining the degree to which PLANS performance differs between sample with and without history of TBI. This would be conducted by recruiting a control group selected to match the TBI sample on demographic variables such as age, race and gender. Sample stratification could also be used to investigate the and ecological validity of the PLANS, by comparing PLANS performance between community-dwelling and non-community-dwelling individuals with TBI.

One proposed addition to this future work would be the inclusion of additional executive functioning surveys used to investigate PLANS ecological validity. The pilot study's sole reliance on the DEX Questionnaire may have limited the study's ability to comprehensively evaluate the degree of executive functioning and functional planning limitations that individuals experience in their daily lives. Having informants complete the BRIEF, CEFI, D-REF and/or the CFQ in addition to the DEX Questionnaire would provide a more comprehensive evaluation of executive functioning limitations encountered in daily life.

7.2.2 Revision & Refinement

The PLANS was found to be feasible for use with individuals with mild and moderate traumatic brain injuries, and demonstrated preliminary reliability and validity, however future work with the PLANS may include revision and revalidation of the PLANS. Instrument development is an iterative process, and involves making adjustments to instructions, scoring and testing demands based on psychometric studies, test practicality and content expert feedback.

The PLANS could benefit from changes to increase clarity of the instructions to increase the approachability of the task by individuals with severe traumatic brain injuries. The PLANS pilot study found that the plan generation and plan execution scoring variables were correlated with neuropsychological tests of planning and executive functioning, while the meal accuracy scoring variables were correlated with both planning and non-planning measures. A proposed revision to the PLANS may involve increasing the amount of task demands that are similar to the plan generation and plan execution variables, which may result in a naturalistic instrument more focused on assessing planning in every-day life. Another route to take related to PLANS revision is development of an alternate version of the PLANS. This alternate version could change the task, the rules, the instructions or the level of complexity of the demands while retaining the scoring system already developed. Having alternate versions of cognitive tests is useful to combat the inherent learning effects associated with naturalistic instruments. Another value of an alternate version of the PLANS is the flexibility it would allow for clinicians to select a task that is salient to the independent living or community functioning goals of their client. If a client would not be expected to conduct meal preparation or independent trips to a grocery store, an alternate version of the PLANS modeled around another task or another scenario may have better face validity to the idiosyncrasies of the client.

Another potential route for PLANS revision could be the pursuit of virtual-reality or computer-based administration methods. The PLANS already integrates partial computerized administration through the specially developed excel calculator used for scoring of the PLANS. Full computer-administered testing could provide several benefits, including increased verisimilitude and standardization of simulated testing environments, removal of clinician burden for scoring, and advanced analysis of performance data using machine learning methods. The cost and resources of developing specialized computer programs to administer performance based instruments is more demanding than pencil & paper tests, and requires coordination with specialized professionals (Martínez-Pernía, et al., 2017). However, the costs of software development and the hardware to run these programs is decreasing, which increases the feasibility of developing new computer and virtual reality naturalistic instruments (Martínez-Pernía et al., 2017).

7.2.3 Potential Commercial Applications

Increased empirical evaluation and refinement of the PLANS would strengthen its utility as a clinical and research tool, which may warrant attention from psychometric testing companies. A publisher may value the clinical and commercial applications of the PLANS which is fit for mass production and dissemination to clinical assessment professionals, neuropsychologists, community rehabilitation organizations or research laboratories. The PLANS is designed with several features that would be valuable to clinicians and publishers, which would facilitate the translation to commercial applications. The PLANS is portable, time and resource low, and provides objective information about an individual's functional planning. As opposed to a majority of naturalistic instruments, publishers could take the stimuli, scoring and materials from the PLANS and mass-produce it for a wide dissemination.

7.3 Limitations

This dissertation had several limitations that are worth mentioning, as they could impact the validity of study outcomes and alter the progress with the PLANS moving forward. The limitations for each study are provided, along with a discussion of how they could impact the generalizability of study findings.

The PLANS development study involved both a scoping literature review and semi-structured interviews, and there are methodological issues associated with both. The scoping literature review was limited by the low amount of literature available on naturalistic instruments that measure planning. The use of naturalistic instruments has only been popularized in the literature in the past two decades, and there is a low amount of literature available. The literature that is available is often impacted by low sample sizes and low power in their psychometric evaluations of the instrument, which makes it difficult to evaluate the naturalistic instruments comprehensively. The use of semi-structured interviews was limited by the relatively low sample size (n=8). While a small sample size is appropriate for use in the instrument development study, as there are no inferential statistics or assumptions made, adding more interviews would increase the span of information that was acquired by the clinician stakeholders. There is not a standardized survey used to investigate stakeholder involvement in instrument development projects, so a brand-new survey was developed (Appendix B). While this survey was piloted (n=2) before being employed in the instrument development study, its reliability and validity is unknown.

There are several limitations to the CoMET analysis that limit the generalizability of its findings. The sample used for this study is heterogeneous in their diagnoses, and has a wide range of cognitive functioning. The sample's performance on the CoMET reflects a transition-age population that was participating in cognitive rehabilitation, however it may be unwise to attribute the sample's performance to one specific diagnosis. While the sample size is large enough to establish adequate power for the analyses chosen, if more data was available it may have increased the amount of correlations that reached the conservative p-value. The error coding for this study was conducted post-hoc by the research team based on the paper records collected from the CoMET administration, however it may have been beneficial to establish a specific coding system a priori. Coding of naturalistic action in a community setting can be complicated if done in the moment observing in vivo. It would benefit the objectivity of performance recording if the instrument is piloted and preliminary psychometric properties are established before using the instrument in a clinical capacity. Additionally, establishing inter-rater reliability for the instrument before using it clinically will increase the standardization of its scoring systems. Using self-report measures for the DEX questionnaires may have introduced bias into the ecological validity analyses, as individuals with cognitive disabilities could have limited insight into the challenges that they experience in daily life. In other studies that investigate the ecological validity of naturalistic instruments, a self-report and an informant-report DEX questionnaire are collected to get a more approximate measure of every-day cognitive functioning. Finally, administering the CoMET required excessive time, financial and work resources to administer, which is a challenge to implement in many cognitive rehabilitation programs. Future work may benefit from developing new instruments to measure executive functioning that use a clinical-environment or a computer system to reduce time and resource burden.

The PLANS pilot study had several limitations related to its screening process. First, the one outlier that was part of the sample was found to be eligible for the study despite being significantly lower cognitive functioning compared to the cohort. There were no screening processes to detect the cognitive challenges that this individual experienced. Despite this, the outlier's performance is valuable to the study in that it indicates that the PLANS may be more appropriate to assess functional planning of mild and moderate severity TBI. Another screening limitations was that a majority of the subjects were females (65%), which significantly contrasts with TBI epidemiology research that estimates that 1.5:1 to 2.7:1 of TBI cases are males (Bruns & Hauser, 2003). One explanation for this inverted gender ratio may be the exclusion criteria of this pilot study. Of the 18 participants that were excluded from the study, 14 (77.8%) were men. One explanation for this may be the consistency of community-dwelling TBI survivors willing to participate in research. It is rationale that the sample reflects the recruitment efforts, as there was an approximately even number of men (50%) and women (45%) that were screened for this study. Screening procedures may have been biased towards excluding men from the pilot study, as 70% of men met at least one of the ineligibility criteria (3 history of psychiatric disorder, 4 over age limit, 2 unable to communicate, 2 negative HELPS Screen, 1 non-native English speaker). One way to address these screening concerns is through a larger psychometric study of the PLANS, which would increase the sample size and could correct the non-representative gender ratio observed in this small pilot study. In addition, increased screening efforts would be added to create a ceiling for TBI severity, reducing the chances of enrolling outlier participants.

7.4 Conclusion

Functional planning is an important aspect of assessment in community rehabilitation of individuals with traumatic brain injuries, however, there is limited research investigating this cognitive construct. This dissertation project aimed to develop and provide preliminary research support for a new naturalistic simulated instrument to measure functional planning of individuals with TBI: The Planning in Life and Adapting to Novel Situations (PLANS). A scoping literature review, semi-structured interviews and retrospective analysis of clinical data from a community-based instrument were conducted to provide guiding information for this development. The

Planning in Life and Adapting to Novel Situations (PLANS) was developed throughout the synthesis of these studies, including creation of a full administration manual, stimuli and scoring materials.

The feasibility PLANS for use with individuals with TBI was supported by this project. The preliminary psychometric properties of the PLANS are promising and support further research of this naturalistic instrument.

This dissertation provides the foundations for new assessment methods for every-day planning through the PLANS. This naturalistic simulated instrument provides a comprehensive assessment of an individual's ability to generate and execute plans of action as they are required in the context of independent living and community functioning. This instrument shows promising psychometric properties, and warrants further research efforts to strengthen these preliminary findings.

Appendix A Scoping Literature Review Findings Table

Year & Authors	Instrument Title	Sample Size	Scoring	Tasks	Study Psychometrics
Virtual-Reality					
Werner et al. (2009)	Virtual Action Planning Supermarket (VAP-S)	Mild Cognitive Impairment n=30 Healthy Controls n=30	**Plan Generation:** Plan generation time **Plan Execution:** Plan execution time, task accuracy (correct, incorrect), plan efficacy (# of stops, stopping time, distance traveled)	Participants used a computer system with a virtual grocery store software and travel through the store purchasing several items listed on the screen. 10 minutes to administer	Validity **Known-Groups Validity:** Significant group differences were found on subtests of covered distance, trajectory duration, and total duration of pauses. VAP-S trajectory duration was able to correctly classify 93% of healthy participants and 80% of persons with MCI while MMSE+BADS correctly classified 87% of health adults and 80% of those with MCI **Ecological Validity:** Four out of the eight VAP-S outcome measures were correlated with BADS scores (ranging from $r = -0.40$ to $r = -0.63$, $p < 0.01$). Especially strong correlations were found between BADS and the trajectory duration and total number of stops ($r = -0.63$ and -0.58, respectively).
Klinger et al. (2006)	The Virtual Supermarket (VS)	Parkinson's Disease n=5	**Plan Generation:** Plan generation time (time until 1st action) **Plan Execution:** Plan	Participants used a computer system with a virtual grocery store software and travel	Not Reported

Year & Authors	Instrument Title	Sample Size	Scoring	Tasks	Study Psychometrics
			execution time, task accuracy (good actions, intrusions), plan efficacy (# of stops, stopping time, distance traveled)	through the store purchasing several items listed on the screen. 10-20 minutes to administer	
			Computer-Based		
Doherty et al. (2015)	Cooking Task	Healthy Controls n=46	**Plan Generation:** N/a **Plan Execution:** Task Accuracy (adjustment score, achievement score, accuracy ratio, residual time), Task Efficiency (range score, discrepancy score)	Participants used a computer system with a touch screen that displayed several food items with timers, and were instructed to prepare a meal by 'cooking' the items so that they were all ready at the same time while setting a virtual table. 20 minutes to administer	Not reported
Holt et al. (2011)	Plan-a-Day Task (PAD)	Schizophrenia n=80	**Plan Generation:** Plan Formulation Time **Plan Execution:** Plan Execution Time, Task Accuracy (# of problems solved w/o corrections), Plan Efficiency (planning ratio)	Participants used a computer system with a mouse and keyboard that displayed a town map and a series of tasks and were instructed to plan a route through the town to accomplish several errands 20 minutes to administer	Reliability **Internal Consistency:** Cronbach's alpha for PAD items indicate moderate to high internal consistency, total solution time=.78, planning ratio=.67, accuracy score=.47 **Test Re-Test:** Using a 4-week washout, test-retest reliability was high for solution time, but learning effects were observed for planning ratio & accuracy, solution time r=.82, planning ratio r=.39, and accuracy r=.33 Validity **Construct Validity:** PAD solution time was significantly correlated with Tower Task

Year & Authors	Instrument Title	Sample Size	Scoring	Tasks	Study Psychometrics
					time (r=.42, p<.01) and Zoo Map time (r=.37, p<.01). No correlations between planning tests and accuracy scores. **Ecological Validity:** Correlations between PAD and Global Assessment of Functioning was significant, r= -.32, p<.01
Craik et al. (2006)	The Breakfast Task	Healthy Younger Adults n=30 Healthy Older Adults n=30	**Plan Generation:** N/a **Plan Execution:** Task Accuracy (adjustment score, accuracy ratio), Plan Efficiency (range score, discrepancy score, # of times food checked)	Participants used a computer system with a mouse and keyboard that displayed several screens with food items and timers, and were instructed to prepare a meal by 'cooking' the items so that they were all ready at the same time while setting a virtual table 20 minutes to administer	Validity **Construct Validity:** Alpha Span test was correlated with discrepancy score (r= -.45) and Digit Span test was correlated with range score (r= -.36). **Known-Groups Validity:** Older adults performed significantly worse than young adults in the most complex and demanding trials of the cooking task, which required increased working memory and prospective memory **Ecological Validity:** Number of sequencing strategies in the real-life task correlated with deviation from ideal start times (r= -.47) and the range score (r= -.45), and the number of checks (r=.47)
Community-based					
McGeorge et al. (2001)	Vocational-Oriented Errand-Planning Task	Acquired Brain Injury n=5 Healthy Controls n=5	**Plan Generation:** Plan efficacy (1-6 rating of plan quality) **Plan Execution:** Task accuracy (# of errands completed)	Participants were brought inside an office building and instructed to complete as many errands as possible in 20 minutes.	Validity **Known-Groups Validity:** Participants with ABI completed fewer errands (F(1, 8) = 8.86, p < 0.05) and had lower plan ratings (F(1,8) = 5.66, p < 0.05) than healthy controls

Year & Authors	Instrument Title	Sample Size	Scoring	Tasks	Study Psychometrics
				20-30 minutes to administer	
Brown et al. (2017)	Prospective Task Planning & Prospective Task Execution	Severe Acquired Brain Injury n=9 Healthy Controls n=9	**Plan Generation:** Plan generation time, # of information units recorded, strategies (planning behaviors) **Plan Execution:** Task accuracy (# of tasks attempted, # of tasks accurately completed), rule breaks (# of rules violated)	From their homes, participants created a plan for execution of daily tasks in accordance with pre-established rules and over the subsequent 10 days, participants independently attempted to complete the tasks 11 days to administer	Reliability **Inter-Rater Reliability:** Percent agreement = 95.46% between raters on information units recorded in prospective planning phase Validity **Known-Groups Validity:** Significant group differences on # of tasks completed (U1=8.50, p=.004), # of rule violations (U1=24.00, p=.045 & frequency of rule violations (U1=24.00, p=.009)
Brown et al. (2016)	Modified MET Planning & Modified MET Execution	Severe Acquired Brain Injury n=9	**Plan Generation:** Plan generation time, # of information units recorded, strategies (planning behaviors), estimation of performance (how long it will take to complete tasks) **Plan Execution:** Plan execution time, task accuracy (# tasks accurately completed), rule breaks (# of rules violated, frequency of rule breaks),	Participants were brought to a college campus and instructed to complete a scavenger hunt with 12 errand tasks while following a set of rules. 60 minutes to administer	Not Reported
Valls-Serranno et al. (2017)	Multiple Errands Test - Contextualized Version (MET-CV)	Substance dependence n=60 Healthy Controls n=30	**Plan Generation:** Plan generation time **Plan Execution:** Plan execution time, task accuracy (# tasks accurately completed), error types (inefficiencies, interpretation failures, task failures),	Participants were brought to a rehabilitation campus and instructed to complete a series of errand tasks including shopping, information gathering and scheduling	Reliability **Testing Context Reliability:** No significant differences on MET-CV scores between 3 different testing environments, with the exception of total performance time between 2 rehabilitation campuses (p = .006). Validity **Construct Validity:** MET-CV task failures

Year & Authors	Instrument Title	Sample Size	Scoring	Tasks	Study Psychometrics
			Strategies (# times participant consults signs), rule breaks (# of rules broken, frequency of rule breaks)	an appointment. 20-30 minutes to administer	correlated with letters and numbers test (r= -.341, p<.001), rule breaks correlated with SOC (r=.273, p<.001) and interpretation failures correlated with the Zoo Map Test part 1 (r=-.309, p<.05) **Known-Groups Validity:** Significant differences were found between groups for task failures (d=.771), and inefficiencies (d=.879), as well as initial planning time (d=918), number of times signals were consulted(d=1.43) and all task failures (d=.533-.879)
Chevignard et al. (2000)	Script Generation & Execution	TB n=11 Healthy Controls n=10	**Plan Generation:** Task familiarity (0-2 rating of familiarity with hosting a dinner party), plan generation time, plan generation errors (task: omission, addition, inversion, estimation, commentary; neuropsychological: (context neglect, control errors, environmental adherence, distractibility, dependency, behavioral disorder) **Plan Execution:** Plan execution time, plan execution errors (task: omission, addition, inversion, estimation, commentary; neuropsychological: (context neglect, control errors, environmental adherence,	Participants were brought to a hospital kitchen and instructed to generate scripts (plans) for accomplishing three every-day tasks required to host a dinner party then go into the community (post office & grocery store) and execute the tasks. 2-3 hours to administer	Validity **Construct Validity:** No significant correlations between script execution and tests of executive functions. **Known-Groups Validity:** Individuals with Traumatic Brain Injuries made significantly more errors on all three tasks (p<.01), script execution time (p=0.0019). Individuals with Traumatic Brain Injuries had significantly higher number of errors on all three scripts (p<.01) and script generation time (p<.01) **Ecological Validity:** No significant correlations between script execution and clinical rating scales of executive functioning completed by caregivers and patients.

Year & Authors	Instrument Title	Sample Size	Scoring	Tasks	Study Psychometrics
			distractibility, dependency, behavioral disorder), plan efficacy (amount of money spent), task comprehension (ability to grasp open-nature tasks) task familiarity (0-2 rating for each task)		
Bottari et al. (2009)	Instrumental Activities of Daily Living Profile (IADL-Profile)	TB n=96	**Plan Generation:** Selecting appropriate goals (level of independence in goal formulation) Plan efficacy (Level of independence in plan generation) **Plan Execution:** Task accuracy (level of independence carrying out a task), Plan adherence (level of independence in verifying task performance matches generated goal)	Participants were instructed that they have an overarching goal that requires planning, carrying out and completing several sub goals (planning a meal for guests). 3 hours to administer	<u>Reliability</u> **Inter-rater Reliability:** Varied between 100-64.3 percent agreement for IADL-R items and scoring domains. Some items with very little score variability produced high IRR but low Kappa.
			<u>Simulated</u>		
Novakovic-Agopian et al. (2014)	Goal Processing Scale (GPS)	Acquired Brain Injury, n=19	**Plan Generation:** Plan formation time, selecting appropriate goals (planning domain 0-8 rating), task comprehension (planning domain 0-8 rating) **Plan Execution:** Plan efficacy (initiation domain 0-8 rating, flexible problem solving domain 0-8 rating), task accuracy (task execution domain 0-8), plan adherence	Participants were instructed to gather and compare information about 3 different activities (or products/services—as designated on alternate forms) of their choice, using the available means while following specified rules in a limited time (30	<u>Reliability</u> **Inter-rater Reliability:** ICC (2,1) ranged from 0.75 to 0.98 for the GPS overall composite score and the subdomain scores <u>Validity</u> **Construct Validity:** GPS scores were correlated with learning and memory tests: overall (r=.61, p=.01), self-monitoring (r=.59, p=.02), learning and memory (r=.46, p=.02), flexible problem solving (r=.55, p=.04)

Year & Authors	Instrument Title	Sample Size	Scoring	Tasks	Study Psychometrics
			(maintenance of attention domain 0-8 rating, sequencing and switching of attention domain 0-8 rating), insight (learning and memory domain 0-8 rating)	minutes). 30-40 minutes to administer	**Ecological Validity:** GPS scores were correlated with performance on a multiple errands test. Significant correlations found between overall performances on GPS and MET: (-0.59, p=.012) and between GPS and MET rule breaks (-0.60, p=.01)
Shanahan et al. (2011)	Party Planning Task (PPT)	Pediatric Severe TB n=2	**Plan Generation:** Plan formulation time **Plan Execution:** Plan execution time, task accuracy (# of tasks scheduled accurately), task errors (omission, time, allocation), strategies (chunking, purpose, error correction, monitoring)	Participants were brought to a clinical space and provided a list of tasks and a schedule and instructed to organize an imaginary party whilst meeting a number of scheduling constraints. 20-40 minutes to administer	Reliability **Inter-rater Reliability:** Percent agreement was 100% between two raters
Kliegel et al. (2007)	Version A (Real World) & Version B (Space Theme) Errand Task	Healthy Young Adults n=52 Healthy Older Adults n=52	**Plan Generation:** Not Reported **Plan Execution:** Task accuracy (# of correctly scheduled tasks), task errors (omission, perseveration, intrusion), rule breaks (# of rules broken), plan efficiency (optimized route score), strategies (planning behaviors before beginning, minimizing distances)	Participants were brought to a clinical space with a map and a set of tasks and instructed to plan how they would carry out a series of 6 errand tasks as many of the errands as possible using the shortest possible route. 10 minutes to administer	Validity **Known-Groups Validity:** Significant differences were found between groups on planning performance (t(50) = 5.83, p<.01)
Sanders et al. (2014)	The Apartment Map Task (Amap)	Mild Cognitive Impairment n=37 Healthy	**Plan Generation:** Plan Formation Accuracy (# of tasks included in plan), Plan Formation Efficacy (# of tasks	Participants were brought to a model apartment set up in a college research center,	Validity **Construct Validity:** For adults with MCI, plan execution accuracy was correlated with Memory Assessment Scale list learning

Year & Authors	Instrument Title	Sample Size	Scoring	Tasks	Study Psychometrics
		Controls n=37	interweaved), **Plan Execution:** Task Accuracy (# of tasks accurately completed), Plan Execution Efficacy (# of tasks interweaved), Strategy Use (# of times individual searched for information in the wrong part of the testing site), Plan Adherence (proportion of tasks executed that were included in the initial plan)	and were provided a map and a list of everyday tasks to complete. Participants were instructed to develop and write out a strategy to successfully complete a list of tasks. Subsequently, participants carried out the tasks in the apartment with the aid of their formulated plan 15 minutes to administer	(r=.44, p < .01) **Known-Groups Validity:** Significant differences were found between groups on plan formation accuracy (t(72) = 1.14, p < .001, d = 0.96), plan formulation efficacy (t(72) = –4.08, p < .001, d = 0.95), execution accuracy (t(72) = 3.42, p = .001, d = 0.80), execution efficacy (t(72) = –3.85, p < .001, d = 0.89), plan execution time (t(71) = 2.17, p < .05, d = 51) & plan adherence (t(72) = –5.92, p < .001, d = 1.38) **Ecological Validity:** For adults with MCI, plan formulation accuracy was correlated with DEX-informant ratings (-.51, p <.01)
Schmitter-Edgecombe et al. (2012)	The Day Out Task (DOT)	Mild Cognitive Impairment n=38 Healthy Controls n=38	**Plan Generation:** Plan Formulation Time **Plan Execution:** Task Accuracy (# of tasks completed accurately, # of tasks completed), Plan Execution Efficacy (# of tasks completed in efficient order, # of tasks interweaved)	Participants were brought to a model apartment set up in a college research center, and were provided a map and a list of tasks to complete for a day out in the community. 15 minutes to administer	Reliability **Inter-rater reliability:** Percent agreement was 96.2% for task accuracy scores and 99.7% for task sequencing Validity **Construct Validity:** Memory Assessment Scale prose recall test was correlated with plan execution accuracy score (r=-.39, p<.05) and task incomplete score (r=-.46, p<.005); the Activity Paradigm test was correlated with plan execution accuracy score (r=-.32, p<.05), task incomplete score (r=-.36, p<.05), and plan efficacy score, (r=.6, p<.005); BADS Zoo Map test correlated with plan execution accuracy

Year & Authors	Instrument Title	Sample Size	Scoring	Tasks	Study Psychometrics
					score (r=-.4, p<.005) and task incomplete score (r=-.34, p<.05) **Known-Groups Validity:** Significant differences were found between groups on plan execution time (p=.01), Plan execution accuracy (p<.01), and subtask completion (p=.01) **Ecological Validity:** For the MCI group, DOT sequencing score was the only unique predictor of IADL performance, B = −.45, t = −1.99, p = .05
Seter et al. (2011)	Naturalistic Action Test (NAT)	Schizophrenia n=48 Healthy Controls n=26	**Plan Generation:** Plan Formulation Time (seconds spent planning, proportion of time spent planning and completing task), Strategies (planning behaviors observed before starting each task) **Plan Execution:** Plan Execution Time, Task Accuracy (Accuracy Score, % Accomplishment Score), Error Type (Omission, Commission, Addition, Substitution, Perseveration, Quality, Gesture Substitution, Spatial Misorientation, Spatial Misestimation, Tool Omission)	Participants were brought to a clinical space with many objects resting on a table and instructed to complete three everyday tasks (Preparing a slice of toast, wrapping a gift and packing a lunchbox and a schoolbag) 15 minutes to administer	Reliability **Inter-Rater Reliability:** Raters demonstrated high reliability for coding NAT planning behaviors, percent agreement % 88.90, Kappa=.72 **Internal Reliability:** Coding for planning behavior score and task accuracy demonstrated good internal consistency, planning behavior = Cronbach's Alpha: .84 and tasks accuracy = Cronbach's Alpha: .84-.71 Validity **Known-Groups Validity:** Participants with Schizophrenia demonstrated worse performance than healthy controls on all NAT variables.

Year & Authors	Instrument Title	Sample Size	Scoring	Tasks	Study Psychometrics
Chalmers et al. (1993)	Party Planning Task (PPT)	Healthy Adolescents n=35 Healthy Adults n=35	**Plan Generation:** Selecting Appropriate Goals (participant transformation of PPT goals) **Plan Execution:** Plan Execution Time, Task Accuracy (# of tasks scheduled accurately), Error Types (omission, time, allocation), Strategies (use of planning aid), Awareness/insight (1-4 rating how satisfied they are with the plan, 1-4 rating how well they thought they organized the party)	Participants were brought to a clinical space and instructed to organize an imaginary party whilst meeting a number of scheduling constraints 30-45 minutes to administer	Reliability **Inter-Rater Reliability:** Agreement between two trained raters ranged from 91%-95% for PPT scoring criteria Validity **Known-Groups Validity:** Significant differences were found between groups on plan execution time and omission errors (p<.05)
Pea et al. (1987)	Classroom Chore Scheduling Task	Healthy Children n=32	**Plan Generation:** Plan Formulation Efficacy (# of plans generated, distance traveled in planned path) **Plan Execution:** N/a	Participants were brought to a clinical space and provided a map of a classroom and a list of chores and were instructed to create an efficient plan to accomplish several classroom chores	Validity **Construct Validity:** No significant relationships between CCST and digit span test nor Wechsler Intelligence Scale for Children
O'Neil-Pirozz et al. (2010)	Virtual Planning Test (VIP)	TB n=75	**Plan Generation:** Plan Formulation Time **Plan Execution:** Plan Execution Time, Task Accuracy (# of accurate task choices, # of omitted tasks), Error Types (sequencing error, trip-distractor error, week-related-distractor error)	Participants were brought to a clinical space and provided a schedule and a series of playing cards reflecting tasks and were instructed to create a schedule of tasks necessary to prepare for a trip	Reliability **Test Re-Test Reliability:** VIP test-re-test reliability correlation coefficients ranged from .341 to 8.55 for all performance variables, and all were significant at p<.05
Pentland et	Party Planning	Adolescents	**Plan Generation:** Plan	Participants were	Reliability

Year & Authors	Instrument Title	Sample Size	Scoring	Tasks	Study Psychometrics
al. (1998)	Task (PPT)	w/ Severe Head Injury n=16 Adolescents w/ Mild/Moderate Head Injury n=17 Healthy Controls n=17	Formulation Time **Plan Execution:** Plan Execution Time, Task Accuracy (# of tasks scheduled accurately), Error Types (omission, time, allocation), Strategies (tasks chunked)	brought to a clinical space and instructed to organize an imaginary party whilst meeting a number of scheduling constraints 20-30 minutes to administer	**Inter-Rater Reliability:** Percent agreement between two trained raters on the PPT achieved 88% for time requirements, and 82% for the allocation criteria <u>Validity</u> **Known-Groups Validity:** The group with severe HI had significantly worse performance than the other two groups on measures of time, allocation criteria, overall accuracy, and efficiency measures

Appendix B Clinician Semi-Structured Interview

Introductory Script: Good morning/afternoon, and thank you for consenting to participate in this research study. During the next 60 to 90 minutes, I would like to ask you some questions about functional planning impairments that you encounter in individuals with brain injuries that participate in your clinical practice. Functional planning abilities control thought and action required in day to day living. They include being able to cognitively look ahead and identify goals, organize and sequence steps, and execute steps, required for completing tasks in daily life. Functional planning abilities are necessary to set goals, initiate behavior, monitor action and modify plans in the face of obstacles until goals are completed. Some examples of tasks requiring functional planning may include: coordinating a schedule, planning and carrying out daily tasks, or planning and carrying out an activity. When answering these questions, you should think about some of the more recent clients that you have worked with that had moderately severe brain injuries and experienced deficits in their ability to plan and organize their daily lives, and not just one individual in particular. Do not provide any identifiable information about your clients. Completing this interview is a voluntary process, and you can choose to end the interview at any time. I can repeat questions, but I cannot provide clarification about the questions in this interview. Do you have any questions?

Topic 1: What planning impairments do clinicians encounter in community rehabilitation settings?

"First, I would like to ask you some questions to help better understand the population of individuals with brain injuries that you work with…"

What is the distribution of your customers with brain injuries that have mild, moderate or severe injury severities?

Mild-Moderate- Severe-

What is the distribution of your customers with brain injuries that experience mild, moderate or severe impairment in functioning?

Mild-Moderate- Severe-

"Do your clients experience problems with Functional Planning abilities? As a reminder, Functional Planning abilities control thought and action required in day to day living. They include being able to cognitively look ahead and identify goals, organize and sequence steps, and execute steps, required for completing tasks in daily life. Functional planning abilities are necessary to set goals, initiate behavior, monitor action and modify plans in the face of obstacles until goals are completed. Please consider your clients with moderately severe brain injuries when answering these questions"

How important is functional planning to your services with this population?

5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important

In general, how significant are planning problems with individuals with brain injury and moderate impairment?

5- extremely significant | 4- very significant | 3-moderately significant | 2- slightly significant | 1- not at all significant

Oh average, how often do your customers encounter problems in related to functional planning in their day to day life?

5- Daily | 4- 4-6 times a week | 3-2-3 times a week | 2- once a week | 1- never

To what degree can your customers independently handle functional planning demands?

5- independent | 4- somewhat independent | 3-moderately independent | 2- somewhat independent | 1- dependent

Is your client aware that functional planning is a problem for them?

5- extremely aware | 4- very aware | 3-moderately aware | 2- somewhat aware | 1- not at all aware

How much do challenges in functional planning limit independent living goals?

5- extremely limiting | 4- very limiting | 3-moderately limiting | 2- somewhat limiting | 1- not at all limiting

How much do challenges in functional planning limit community functioning?

5- extremely limiting | 4- very limiting | 3-moderately limiting | 2- somewhat limiting | 1- not at all limiting

To what degree do you direct resources towards functional planning impairment?

5- always | 4-most of the time | 3- about half the time | 2- sometimes | 1- never

What factors determine if you would direct resources towards intervention in functional planning impairment?

How often do you work with individuals with moderate severity brain injuries to create goals directed around functional planning required for independent living or community functioning?

5- always | 4-most of the time | 3- about half the time | 2- sometimes | 1- never

"For individuals with moderate severity brain injuries and impairment in functional planning abilities, to what degree to they experience challenges with…"

1 Setting realistic goals

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

1 Choose appropriate goals

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

1 Appraising a problem

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

1 Understanding directions

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

2 Organization of thoughts

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

2 Sequencing steps

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

2 Logical reasoning

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

2 Understand abstract demands

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

3 Initiate plans of action

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

3 Volition

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

3 Follow specific directions

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

3 Following rules

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

4 monitor their behavior

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

4 detecting errors

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

4 compensating for changes in task demands

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

4 anticipating appropriate strategies to use for a problem

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

4 encountering a problem that does not match their anticipated goals

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

4 respond to feedback

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

4 independently work toward self-generated goals (necessary for daily life)

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

4 Adjust plans to compensate for obstacles

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

4 Ask effective questions

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

5 Use strategies to help achieve goals

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

5 Recognize when they have completed a goal

5- extremely challenging | 4- very challenging | 3-moderately challenging | 2- somewhat challenging | 1- not at all challenging

Topic 2: What are some of the every-day tasks that are limited by functional planning limitations that are incorporated in rehabilitation intervention?

Can you give me up to three examples of everyday tasks that are most significantly impacted by functional planning limitations?
Is task 1 related to independent living goals, community functioning goals, or other?
Is task 2 related to independent living goals, community functioning goals, or other?
Is task 3 related to independent living goals, community functioning goals, or other?

Task 1:___________________________

How important is successful completion of this task to attaining positive community rehabilitation outcomes?
5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important
Are your clients aware that this task is a challenge for them?
5- extremely aware | 4- very aware | 3-moderately aware | 2- somewhat aware | 1- not at all aware
Are your clients motivated to work on goals related to this task?
5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important
How important is successful completion of this task to your client?
5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important
How significant is this problem with individuals with moderate brain injury?
5- extremely significant | 4- very significant | 3-moderately significant | 2- slightly significant | 1- not at all significant
How often do your customers encounter these problems?
5- Daily | 4- 4-6 times a week | 3-2-3 times a week | 2- once a week | 1- never
To what degree can your consumer independently complete this task?
5- independent | 4- somewhat independent | 3-moderately independent | 2- somewhat independent | 1- dependent

What cognitive skills are required for successful completion of this task?

Is there an environment when your clients are most likely to encounter problems completing this task?
What kinds of things make it more likely that your customers will experience challenges completing this task? (difficult tasks, transition, distracting environment, etc.)

Task 2:___________________________

How important is this task to community rehabilitation?
5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important
Are your clients aware that this task is a challenge for them?
5- extremely aware | 4- very aware | 3-moderately aware | 2- somewhat aware | 1- not at all aware
Are your clients motivated to work on goals related to this task?
5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important
How important is successful completion of this task to your client?
5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important

How significant is this problem with individuals with moderate brain injury?
5- extremely significant | 4- very significant | 3-moderately significant | 2- slightly significant | 1- not at all significant
How often do your customers encounter these problems?
5- Daily | 4- 4-6 times a week | 3-2-3 times a week | 2- once a week | 1- never

To what degree can your consumer independently complete this task?
5- independent | 4- somewhat independent | 3-moderately independent | 2- somewhat dependent | 1- dependent
What cognitive skills are required for successful completion of this task?

Is there an environment when your clients are most likely to encounter problems completing this task?

What kinds of things make it more likely that your customers will experience challenges completing this task? (difficult tasks, transition, distracting environment, etc.)

Task 3:___________________________

How important is this task to community rehabilitation?
5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important
Are your clients aware that this task is a challenge for them?
5- extremely aware | 4- very aware | 3-moderately aware | 2- somewhat aware | 1- not at all aware
Are your clients motivated to work on goals related to this task?
5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important
How important is successful completion of this task to your client?
5- extremely significant | 4- very significant | 3-moderately significant | 2- slightly significant | 1- not at all significant
How significant is this problem with individuals with moderate brain injury?
5- extremely significant | 4- very significant | 3-moderately significant | 2- slightly significant | 1- not at all significant
How often do your customers encounter these problems?
5- Daily | 4- 4-6 times a week | 3-2-3 times a week | 2- once a week | 1- never
To what degree can your consumer independently complete this task?
5- independent | 4- somewhat independent | 3-moderately independent | 2- somewhat dependent | 1- dependent

What cognitive skills are required for successful completion of this task?

Is there an environment when your clients are most likely to encounter problems completing this task?

What factors make it more likely that your customers will experience challenges completing this task? (difficult tasks, transition, distracting environment, etc.)

Topic 3: What are some of the every-day environments that are limited by functional planning impairment that are incorporated in rehabilitation intervention?
Could you name three settings in the home that you provide rehabilitation services with individuals with brain injuries?
(For each environment ask)

Setting 1:________________________________
How important is this setting to community rehabilitation?
5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important
How important is this setting to independent living goals?

5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important
What environmental factors are present that limit functional planning abilities?
What environmental factors are present that make it more likely that your customers would successfully use functional planning abilities?

Setting 2:________________________________
How important is this setting to community rehabilitation?
5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important
How important is this setting to independent living goals?
5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important
What environmental factors are present that limit functional planning abilities?
What environmental factors are present that make it more likely that your customers would successfully use functional planning abilities?

Setting 3:________________________________
How important is this setting to community rehabilitation?
5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important
How important is this setting to independent living goals?
5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important
What environmental factors are present that limit functional planning abilities?
What environmental factors are present that make it more likely that your customers would successfully use functional planning abilities?

Could you name three settings in the community that you provide rehabilitation services with individuals with brain injuries?
(For each setting ask)

Setting 1:________________________________
How important is this setting to community rehabilitation?
5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important
How important is this setting to independent living goals?
5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important
What environmental factors are present that limit functional planning abilities?
What environmental factors are present that make it more likely that your customers would successfully use functional planning abilities?

Setting 2:________________________________
How important is this setting to community rehabilitation?
5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important
How important is this setting to independent living goals?
5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important
What environmental factors are present that limit functional planning abilities?
What environmental factors are present that make it more likely that your customers would successfully use functional planning abilities?

Setting 3:________________________________
How important is this setting to community rehabilitation?

5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important

How important is this setting to independent living goals?

5- extremely important | 4- very important | 3-moderately important | 2- slightly important | 1- not at all important

What environmental factors are present that limit functional planning abilities?

What environmental factors are present that make it more likely that your customers would successfully use functional planning abilities?

Appendix C Content Validity Survey

You are invited to participate in a web-based online survey investigating the content validity of the Planning in Life and Adapting to Novel Situations (PLANS) Instrument. The PLANS is a naturalistic simulated instrument designed to measure functional planning abilities through performance on a simulated real-world task. This instrument is the product of a mixed-methods instrument development project. You will be will be provided with a list of behavioral objectives that guided the construction of the PLANS, a definition of these terms, and a list of items designed to specifically test these objectives. You will be asked to

l) review the PLANS materials including: Instructions, stimuli & materials, scoring rubrics and administration materials (provided as PDFs) and a pilot video performance of the PLANS

2) assess the relevancy, clarity & sufficiency of the content for assessment of functional planning abilities of individuals with brain injuries of moderate severity.

This is a research study being conducted by Evan Knutson, a PhD candidate in Rehabilitation Science at the University of Pittsburgh- School of Health and Rehabilitation Sciences. It should take approximately 30 minutes to complete this survey.

PARTICIPATION
Your participation in this survey is voluntary. You may refuse to take part in the research or exit the survey at any time without penalty. You are free to decline to answer any particular question for any reason.

BENEFITS
You will receive no direct benefits from participating in this research study. However, your responses may help us learn more about the ways in which impairments in planning abilities impact day-to-day functioning, in order to guide the development of a new instrument to measure functional planning abilities

RISKS
There are no foreseeable risks involved in participating in this study other than those encountered in day-to-day life.

CONFIDENTIALITY
Your survey answers will be sent to a link at Qualtrics.com where data will be stored in a password protected electronic format.

CONTACT
If you have questions at any time about the study or the procedures, you may contact the Primary Investigator of this study, Evan Knutson, at the email address eck26@pitt.edu, or by phone at (610) 597-1449.

ELECTRONIC CONSENT: Please select your choice below. You may print a copy of this consent form

for your records. Clicking on the “Agree” button indicates that · You have read the above information· You voluntarily agree to participate
Agree (1)
Disagree (2)
End of Block: Consent

Start of Block: Introduction

The PLANS materials can be accessed at the following cloud storage location: https://pitt.box.com/s/lgxa73v2j0kqab95yesjox28au5

The password used to access this folder is: **PLANS**

The folder contains all the testing materials for the PLANS, including stimuli (Cookbook, Map & Catalogue), participant materials (Grocery Worksheet, Scrap Paper & Rules) and examiner materials (Script & Scoring Sheets).

The Administration Manual for the PLANS is also provided, with instructions how to administer & score the PLANS.

In addition, a mock-video performance is uploaded to the folder. This video shows a shortened PLANS performance that gives a detailed perspective of how the PLANS is administered and how a model participant attempts the tasks.

What is your name?
____

Introduction

The questions in this content validity survey will first ask you to review and rate the PLANS task demands and rules. Then, you will be asked to review each of the PLANS scoring domains. The PLANS scoring sections are: plan generation, task familiarity & estimation, and plan execution.

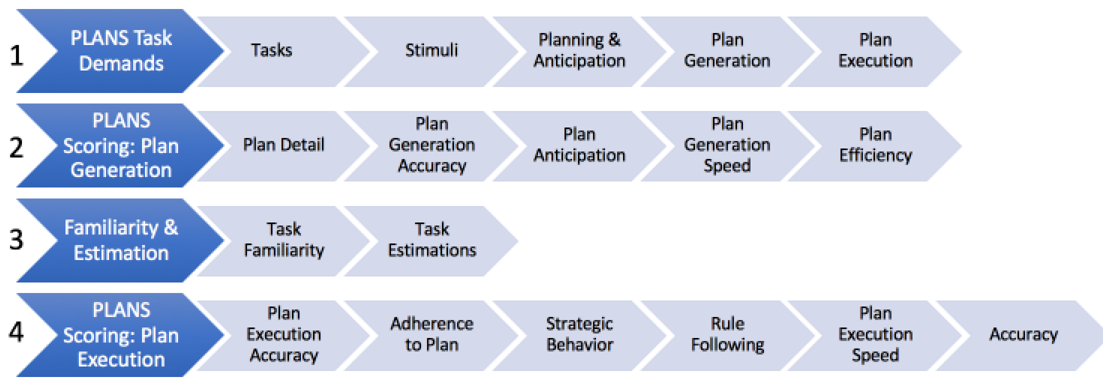
For each aspect of this content validity survey, please rate the presented material using the following categories:

Relevancy*: The degree to which the presented information is relevant to the assessment of functional planning with individuals with brain injuries of moderate severity*

Clarity*: The degree to which the presented information is clear and easy to understand.*

Sufficiency*: The degree to which the presented information is sufficiently covers the span and scope of its stated purpose*

You should have the PLANS materials (instructions, stimuli, scoring, video performance) available for reference when completing the content validity survey.



End of Block: Introduction

Start of Block: Tasks

PLANS Tasks
For each of the tasks that are listed, please use the following categories (relevancy, clarity & sufficiency) to rate the tasks on their ability to simulate tasks that one may encounter in independent living and community functioning contexts.

Money management - identifying the budget and calculating expenses
Listening to a voicemail and collecting important information - recording important information from the pharmacy and identifying the price and name of prescription
Select a menu from a cookbook - identifying menu items that meet requirements of PLANS
Creating a grocery list - identifying ingredients necessary for purchase, and recording their store location and price
Calculating a budget - calculating the total amount of money needed for purchases
Identifying a path through a grocery store map - drawing a path through the grocery store indicating where they would need to walk in order to complete task

Relevancy
Not relevant (1)
Item needs some revision (2)
Relevant, but needs minor revision (3)
Very relevant (4)

Please explain what aspect is irrelevant, and provide a suggestion for revision:
____

Clarity
Not clear (1)
Item needs some revision (2)
Clear, but needs minor revision (3)
Very clear (4)

Please explain what aspect is unclear, and provide a suggestion for revision:
____

Sufficiency
Not Sufficient (1)
Item needs some revision (2)
Sufficient, but needs minor revision (3)
Very sufficient (4)

Please explain what aspect is insufficient, and provide a suggestion for revision:

End of Block: Tasks

Start of Block: Stimuli

PLANS Stimuli *Please refer to the PLANS materials: PLANS Cookbook, PLANS Grocery Store Catalogue, PLANS Grocery Store Map & Pharmacy Voicemail. Using the following categories (relevancy, clarity and sufficiency), please rate these items on their ability to replicate materials that one may encounter in independent living and community functioning contexts.*

Relevancy
Not relevant (1)
Item needs some revision (2)
Relevant, but needs minor revision (3)
Very relevant (4)

Please explain what aspect is irrelevant, and provide a suggestion for revision:

Clarity
Not clear (1)
Item needs some revision (2)
Clear, but needs minor revision (3)
Very clear (4)

Please explain what aspect is unclear, and provide a suggestion for revision:

Sufficiency
Not Sufficient (1)
Item needs some revision (2)
Sufficient, but needs minor revision (3)
Very sufficient (4)

Please explain what aspect is insufficient, and provide a suggestion for revision:

End of Block: Stimuli

Start of Block: Planning & Anticipation

PLANS Planning & Anticipation

For each of the items that are listed, please use the following categories (relevancy, clarity & sufficiency) to rate the tasks on their ability to elicit planning and anticipatory thinking in PLANS participants.

Appetizer, Entrée, Dessert Restriction - Participants need to prepare a meal with one appetizer one entree and one dessert

Budgetary Restriction - Participants need to select a menu that will be under their budget

Oven Restriction - Participants need to select a menu that only uses the oven for one recipe

Cooking Time Restriction - Participants need to select a menu that can be prepared in 2 hours

Portion Restriction - participants need to select a recipe that serves 2 people Minimize switching between materials (cookbook, catalogue, map, money envelope, tape recorder) - Participants are informed that they should only use each PLANS item once

Relevancy

Not relevant (1)

Item needs some revision (2)

Relevant, but needs minor revision (3)

Very relevant (4)

Please explain what aspect is irrelevant, and provide a suggestion for revision:

Clarity

Not clear (1)

Item needs some revision (2)

Clear, but needs minor revision (3)

Very clear (4)

Please explain what aspect is unclear, and provide a suggestion for revision:

Sufficiency

Not Sufficient (1)

Item needs some revision (2)

Sufficient, but needs minor revision (3)

Very sufficient (4)

Please explain what aspect is insufficient, and provide a suggestion for revision:

End of Block: Planning & Anticipation

Start of Block: Plan Generation

PLANS Plan Generation
Functional planning is being investigated by the PLANS from two stages: Plan Formulation & Plan Execution. Considering the following instructions, please rate this demand of the PLANS on its relevance, clarity and sufficiency for measuring plan generation.
"Now, I would like you to create a plan how to accomplish this task in an efficient manner, following the important rules. Try to anticipate what step you need to complete first, second, third and so on. Provide as much detail as possible in your plan. You have up to 5 minutes to write out your plan, and I encourage you to talk-aloud while creating the plan if it would help you."

Relevancy
Not relevant (1)
Item needs some revision (2)
Relevant, but needs minor revision (3)
Very relevant (4)

Please explain what aspect is irrelevant, and provide a suggestion for revision:

____

Clarity
Not clear (1)
Item needs some revision (2)
Clear, but needs minor revision (3)
Very clear (4)

Please explain what aspect is unclear, and provide a suggestion for revision:

____

Sufficiency
Not Sufficient (1)
Item needs some revision (2)
Sufficient, but needs minor revision (3)
Very sufficient (4)

Please explain what aspect is insufficient, and provide a suggestion for revision:

____

End of Block: Plan Generation

Start of Block: Plan Execution

PLANS Plan Execution
Functional planning is being investigated by the PLANS from two stages: Plan Formulation & Plan Execution. Individuals first generate a plan how to accomplish the task, then the execute that plan and complete the task. Please consider the instructions, tasks, and scoring of the PLANS and rate the entirety of plan execution on its relevance, clarity and sufficiency in measuring plan execution.

Relevancy
Not relevant (1)
Item needs some revision (2)
Relevant, but needs minor revision (3)
Very relevant (4)

Please explain what aspect is irrelevant, and provide a suggestion for revision:

Clarity
Not clear (1)
Item needs some revision (2)
Clear, but needs minor revision (3)
Very clear (4)

Please explain what aspect is unclear, and provide a suggestion for revision:

Sufficiency
Not Sufficient (1)
Item needs some revision (2)
Sufficient, but needs minor revision (3)
Very sufficient (4)

Please explain what aspect is insufficient, and provide a suggestion for revision:

End of Block: Plan Execution

Start of Block: Scoring Description

The next questions in this content validity survey will ask you to review each of the PLANS scoring domains for plan generation. It may help you answer the questions if you have the PLANS score sheet available as a reference. For each aspect of this content validity survey, please rate the presented material using the following categories:
Relevancy: The degree to which the presented information is relevant to the assessment of functional planning with individuals with brain injuries
Clarity: The degree to which the presented information is clear and easy to understand.
Sufficiency: The degree to which the presented information is sufficiently covers the span and scope of its stated purpose

The following scoring criteria apply to the **Plan Generation**, which is the written plan that participants generate before attempting the tasks

End of Block: Scoring Description

Start of Block: Plan Detail

Plan Generation: Plan Detail
Definition*: The amount of information and level of detail contained in a plan.* ***Demands****: The PLANS instructs individuals to generate a plan how to approach the various tasks. The instructions state that the plan should be written in as much detail as necessary so that someone else could understand their plan. This implies that every step in the plan should be clearly written and include the important rules and constraints that need to be followed.* ***Measure****: Written plans will be assessed by the examiner for the level of detail provided. An information unit represented any critical piece of information within the task instructions that could stand independently (e.g., nouns, adjectives, verbs).* # of individual pieces of information in written plan

Relevancy
Not relevant (1)
Item needs some revision (2)
Relevant, but needs minor revision (3)
Very relevant (4)

Please explain what aspect is irrelevant, and provide a suggestion for revision:

____

Clarity
Not clear (1)
Item needs some revision (2)
Clear, but needs minor revision (3)
Very clear (4)

Please explain what aspect is unclear, and provide a suggestion for revision:

____

Sufficiency
Not Sufficient (1)
Item needs some revision (2)
Sufficient, but needs minor revision (3)
Very sufficient (4)

Please explain what aspect is insufficient, and provide a suggestion for revision:

____

End of Block: Plan Detail

Start of Block: Plan Accuracy

Plan Generation: Plan Accuracy
Definition*: The degree to which the generated plan accurately reflects the correct steps and the order required to reach the proposed goal.*
Demands*: The PLANS has been constructed so that there is only one* correct sequence of tasks. *The PLANS Instructs individuals to generate a plan how to approach the various tasks. The plan should include all of the steps that must be completed in order to accurately reach the goals outlined in the PLANS. This requires individuals to accurately appraise the demands of the PLANS, anticipate the obstacles to completing these tasks and project themselves in the future envisioning what they need to accomplish in order to fulfil these tasks.*
Measure*:* Written plans will be assessed by the examiner by comparing the individual's generated plan to an answer key containing the most correct sequence of tasks. The correct order is presented in the left-most column, and the order generated by the participant is recorded in the parentheses. The steps in the generated plan will be coded by the type of error that is committed (if an error is committed).

Relevancy
Not relevant (1)
Item needs some revision (2)
Relevant, but needs minor revision (3)
Very relevant (4)

Please explain what aspect is irrelevant, and provide a suggestion for revision:

Clarity
Not clear (1)
Item needs some revision (2)
Clear, but needs minor revision (3)
Very clear (4)

Please explain what aspect is unclear, and provide a suggestion for revision:

Sufficiency
Not Sufficient (1)
Item needs some revision (2)
Sufficient, but needs minor revision (3)
Very sufficient (4)

Please explain what aspect is insufficient, and provide a suggestion for revision:

End of Block: Plan Accuracy

Start of Block: Plan Anticipation

Plan Generation: Plan Anticipation
Definition: *The amount of information in a generated plan that reflects accurate anticipation of rules, constraints, and demands. Anticipation of demands involves correctly identifying the rule, constraint or demand and matching it to the step in the plan where it is relevant.*
Demands: *The PLANS has several rules, constraints and demands that need to be considered when generating a plan how to accomplish the task.*
Measure: *The following list represents the factors that should be anticipated for in the generated plan. The generated plans is checked by the examiner to identify if these factors are included to any degree in the written plan, and if they are included in a step where they apply. The item is correct only if it is included in the plan and it is included in a step where it is relevant.*

Relevancy
Not relevant (1)
Item needs some revision (2)
Relevant, but needs minor revision (3)
Very relevant (4)

Please explain what aspect is irrelevant, and provide a suggestion for revision:

____

Clarity
Not clear (1)
Item needs some revision (2)
Clear, but needs minor revision (3)
Very clear (4)

Please explain what aspect is unclear, and provide a suggestion for revision:

____

Sufficiency
Not Sufficient (1)
Item needs some revision (2)
Sufficient, but needs minor revision (3)
Very sufficient (4)

Please explain what aspect is insufficient, and provide a suggestion for revision:

____

Plan Anticipation

Start of Block: Plan Speed

Plan Generation: Plan Speed
Definition*: How much time was spent generating the plan. This does not directly reflect the amount of effort put into the plan or the speed at which one can formulate a plan, but is a measure of the total time spent working on the plan.*
Demands*: The PLANS instruct individuals to take 5 minutes before beginning to generate a series of steps they will need to complete to complete the tasks. They are instructed to create a plan that is detailed to a degree that could be clearly understood by someone else.*
Measure*: Number of seconds spent planning is recorded by the examiner, ranging from 0-300 (max time*

Relevancy
Not relevant (1)
Item needs some revision (2)
Relevant, but needs minor revision (3)
Very relevant (4)

Please explain what aspect is irrelevant, and provide a suggestion for revision:

Clarity
Not clear (1)
Item needs some revision (2)
Clear, but needs minor revision (3)
Very clear (4)

Please explain what aspect is unclear, and provide a suggestion for revision:

Sufficiency
Not Sufficient (1)
Item needs some revision (2)
Sufficient, but needs minor revision (3)
Very sufficient (4)

Please explain what aspect is insufficient, and provide a suggestion for revision:

End of Block: Plan Speed

Start of Block: Plan Efficacy

Plan Generation: Plan Efficacy
Definition*: The degree to which plan generation is efficiently conducted, and the efficiency of the plan generated.*
Demands*: The PLANS requires individuals generate a plan describing how they will complete the various tasks presented to them, while following rules and constraints. Plan generation as an activity should be efficient, and the rules, constraints and demands of the PLANS should be accurately incorporated into the plan steps as they are written. In order to be efficient, the plan that is generated should have as few steps as required in order to complete the tasks.*
Measure*: Plan generation efficacy will be measured by a frequency of times that an individual revises their plan. Any changes, additions, reorganizations removed tasks will be counted during the plan generation phase. The efficacy of the generated plan will be measured by counting the number of individual steps that are created.* # of times participant adjusts their written plan # of individual steps in the written plan

Relevancy
Not relevant (1)
Item needs some revision (2)
Relevant, but needs minor revision (3)
Very relevant (4)

Please explain what aspect is irrelevant, and provide a suggestion for revision:
____

Clarity
Not clear (1)
Item needs some revision (2)
Clear, but needs minor revision (3)
Very clear (4)

Please explain what aspect is unclear, and provide a suggestion for revision:
____

Sufficiency
Not Sufficient (1)
Item needs some revision (2)
Sufficient, but needs minor revision (3)
Very sufficient (4)

Please explain what aspect is insufficient, and provide a suggestion for revision:

End of Block: Plan Efficacy

Start of Block: Intro PLANS Familiarity & Estimations

Familiarity & Estimation
For each aspect of this content validity survey, please rate the presented material using the following categories: Relevancy: The degree to which the presented information is relevant to the assessment of functional planning with individuals with brain injuries Clarity: The degree to which the presented information is clear and easy to understand. Sufficiency: The degree to which the presented information is sufficiently covers the span and scope of its stated purpose
The following scoring criteria apply to the Task Familiarity & Estimation, which are conducted before participants receive instructions to the PLANS.

End of Block: Intro PLANS Familiarity & Estimations

Start of Block: Task Familiarity

Plan Generation: Task Familiarity
Definition*: How much experience an individual has at completing tasks with similar demands to the PLANS.*
Demands*: The PLANS is comprised of several different tasks (money management, meal preparation, information gathering) framed around a real-world scenario, which may be salient to an individual's current or previous life roles.*
Measure*: Before beginning the PLANS, individuals are given the following questions, and asked to rate their level of familiarity.*

Relevancy
Not relevant (1)
Item needs some revision (2)
Relevant, but needs minor revision (3)
Very relevant (4)

Please explain what aspect is irrelevant, and provide a suggestion for revision:

Clarity
Not clear (1)
Item needs some revision (2)
Clear, but needs minor revision (3)
Very clear (4)

Please explain what aspect is unclear, and provide a suggestion for revision:

Sufficiency
Not Sufficient (1)
Item needs some revision (2)
Sufficient, but needs minor revision (3)
Very sufficient (4)

Please explain what aspect is insufficient, and provide a suggestion for revision:

End of Block: Task Familiarity

Start of Block: Estimation

Task Estimation

Definition*: The degree to which individuals possess accurate anticipatory self-awareness in anticipation of completing a specific task and similar and intellectual awareness of their strengths and weaknesses in general.*
Demands*: Answering these questions requires individuals to have accurate insight and awareness of their own abilities, and how those abilities interact with the demands of the PLANS tasks.*
Measure*: Before beginning the PLANS, individuals are given the following questions, and asked to rate their abilities and make estimations about their performance. Questions 1 & 2 are related to intellectual awareness and questions 3 & 4 are related to anticipatory awareness.*

Relevancy
Not relevant (1)
Item needs some revision (2)
Relevant, but needs minor revision (3)
Very relevant (4)

Please explain what aspect is irrelevant, and provide a suggestion for revision:

____

Clarity
Not clear (1)
Item needs some revision (2)
Clear, but needs minor revision (3)
Very clear (4)

Please explain what aspect is unclear, and provide a suggestion for revision:

____

Sufficiency
Not Sufficient (1)
Item needs some revision (2)
Sufficient, but needs minor revision (3)
Very sufficient (4)

Please explain what aspect is insufficient, and provide a suggestion for revision:

____

End of Block: Estimation

Start of Block: Intro Plan Execution

For each aspect of this content validity survey, please rate the presented material using the following categories:
Relevancy: The degree to which the presented information is relevant to the assessment of functional planning with individuals with brain injuries
Clarity: The degree to which the presented information is clear and easy to understand.
Sufficiency: The degree to which the presented information is sufficiently covers the span and scope of its stated purpose The following scoring criteria apply to the Plan Execution, which is the participant's performance on the tasks.

End of Block: Intro Plan Execution

Start of Block: Task Accuracy

Plan Execution: Plan Execution Accuracy
Definition*: The degree to which an individual's actions accurately reflects the steps required to reach a proposed goal.*
Demands*: The PLANS instructs individuals to execute a plan and complete several tasks while following pre-defined rules and constraints. This requires individuals to accurately appraise the demands of the PLANS, anticipate the obstacles to completing these tasks and monitor their behavior while completing the tasks.*
Measure: *The examiner monitors the participant's behaviors while they complete the tasks, and records the order that tasks are completed in, how much time is spent on each task, and monitoring performance for any obvious errors or rule breaks. This information is transposed onto the following scoring tables that organize this information by the correct sequence of tasks and the amount of errors in each step (if any). The correct order is displayed in the left-most column, and the order generated by the participant is recorded in the parentheses. The steps in the generated plan is coded by the type of error that is committed (if an error is committed*

Relevancy
Not relevant (1)
Item needs some revision (2)
Relevant, but needs minor revision (3)
Very relevant (4)

Please explain what aspect is irrelevant, and provide a suggestion for revision:

____

Clarity
Not clear (1)
Item needs some revision (2)
Clear, but needs minor revision (3)
Very clear (4)

Please explain what aspect is unclear, and provide a suggestion for revision:

____

Sufficiency
Not Sufficient (1)
Item needs some revision (2)
Sufficient, but needs minor revision (3)
Very sufficient (4)

Please explain what aspect is insufficient, and provide a suggestion for revision:

____

End of Block: Task Accuracy

Start of Block: Adherence

Adherence to Generated Plan

***Definition**: The degree to which the steps and details of the generated plan accurately reflect the behaviors observed during plan execution.*

***Demands**: The PLANS instructs individuals to generate a written plan detailing the sequence of steps they will follow to complete the various tasks. This plan should be used as a structure to guide their behavior while executing the tasks, but deviations may occur between the plan generated and the plan executed. Plan deviations can be positive and help the individual complete the tasks, or negative and hinder performance or result in errors or rule breaks. As an example of positive deviation, individuals may encounter an unanticipated obstacle while completing the tasks, or notice an error in their generated plan. In these cases, individuals may adjust their behavior to deviate from the plan and take part in "opportunistic planning", which facilitates overcoming an obstacle in a manner that reduces errors. In a negative deviation, individuals may forget to follow aspects of their written plan or have difficulty monitoring their progress in executing the plan, and deviate from the plan in a way that breaks a rule or completes a task inaccurately. In either case, it is clinically relevant to note if plans are adhered to, and how the individual adjusts their plans.*

***Measure**: Examiners will use the coding from the plan generation and compare it to the plan execution, noting the number of deviations between the sequence of tasks, rules and anticipations. These deviations will be recorded using the following items, and labeled as positive or negative deviations.*

Relevancy
Not relevant (1)
Item needs some revision (2)
Relevant, but needs minor revision (3)
Very relevant (4)

Please explain what aspect is irrelevant, and provide a suggestion for revision:

____

Clarity
Not clear (1)
Item needs some revision (2)
Clear, but needs minor revision (3)
Very clear (4)

Please explain what aspect is unclear, and provide a suggestion for revision:

____

Sufficiency
Not Sufficient (1)
Item needs some revision (2)
Sufficient, but needs minor revision (3)
Very sufficient (4)

Please explain what aspect is insufficient, and provide a suggestion for revision:

End of Block: Adherence
Start of Block: Strategic Behavior

Plan Execution: Strategic Behavior
***Definition**: A strategy reflects intentional behavior or thought that facilitates completing a task. Strategies can be internal (metacognitive) or external (tools, checklists, organization, devices etc.).*
***Demands**: Individuals that partake in the PLANS may employ one or more strategies that could facilitate completing the tasks. Strategies are not explicitly mentioned in the instructions, but individuals may employ a strategy that they use to complete planning activities in day to day functioning during the PLANS, which is clinically valuable.*
***Measure**: During observation of PLANS performance, the type of strategy uses and the frequency of its use will be recorded by the examiner. There is a provided list of strategies that have been observed in piloting the PLANS, but there are blank spaces in the score sheet to write down strategic behavior that falls outside of predefined categories. Internal strategies are difficult to observe in performance, and can be investigated through asking the individuals about their strategy use after they finish the PLANS.*

Relevancy
Not relevant (1)
Item needs some revision (2)
Relevant, but needs minor revision (3)
Very relevant (4)

Please explain what aspect is irrelevant, and provide a suggestion for revision:

Clarity
Not clear (1)
Item needs some revision (2)
Clear, but needs minor revision (3)
Very clear (4)

Please explain what aspect is unclear, and provide a suggestion for revision:

Sufficiency
Not Sufficient (1)
Item needs some revision (2)
Sufficient, but needs minor revision (3)
Very sufficient (4)

Please explain what aspect is insufficient, and provide a suggestion for revision:

End of Block: Strategic Behavior

Start of Block: Rule Following

Rule Following
Definition*: The degree to which individuals are able to monitor their behavior to conform to rules and constraints outlined in the instructions.*
Demands*: The PLANS has several rules that are described in its instructions. These rules are designed to increase the complexity of the PLANS and encourage individuals to plan out their approach to the tasks in an efficient manner. There are also several constraints to the tasks that individuals need to keep in mind while completing the PLANS.*
Measure*: After the PLANS has been finished, the examiner will review the scoring for task accuracy and the information gathering domains and use the responses to determine if the rules and constraints of the PLANS were met*

Relevancy
Not relevant (1)
Item needs some revision (2)
Relevant, but needs minor revision (3)
Very relevant (4)

Please explain what aspect is irrelevant, and provide a suggestion for revision:
____

Clarity
Not clear (1)
Item needs some revision (2)
Clear, but needs minor revision (3)
Very clear (4)

Please explain what aspect is unclear, and provide a suggestion for revision:
____

Sufficiency
Not Sufficient (1)
Item needs some revision (2)
Sufficient, but needs minor revision (3)
Very sufficient (4)

Please explain what aspect is insufficient, and provide a suggestion for revision:
____

End of Block: Rule Following

Start of Block: Execution Speed

Plan Execution Speed
Definition*: How much time was spent executing the plan.*
Demands*: The PLANS instructions provide a specific time at which the individual should be finished with the tasks, and encourages working quickly while remaining accurate.*
Measure*: Number of seconds spent executing plan is recorded by the examiner.*

Relevancy
Not relevant (1)
Item needs some revision (2)
Relevant, but needs minor revision (3)
Very relevant (4)

Please explain what aspect is irrelevant, and provide a suggestion for revision:

Clarity
Not clear (1)
Item needs some revision (2)
Clear, but needs minor revision (3)
Very clear (4)

Please explain what aspect is unclear, and provide a suggestion for revision:

Sufficiency
Not Sufficient (1)
Item needs some revision (2)
Sufficient, but needs minor revision (3)
Very sufficient (4)

Please explain what aspect is insufficient, and provide a suggestion for revision:

End of Block: Execution Speed

Start of Block: Execution Accuracy

Plan Execution: Accuracy

Definition*: The degree to which an individual can search, process and report information.*

Demands*: The PLANS reflects the degree of reading, information gathering and computation that is typical of the daily activities of meal planning and grocery store shopping. This includes processing visually complex materials such as a cookbook, a catalogue, and a map. Some of the tasks that need to be completed in the PLANS require a significant amount of information processing and gathering, transposing.*

Measure*: The PLANS requires that individuals report several pieces of information: A list of ingredients necessary for the recipes selected, the prices and aisles of those ingredients, and a total price of the ingredients. In addition, participants must create a path through the grocery store indicating where they would walk to acquire the grocery items. These values are compared to a programmed score system developed in Excel. The degree of accuracy is recorded as the $ value deviation between the correct answers and the participant's grocery list. The grocery store path is scored on the amount of correct store locations visited in the correct sequence, and paired with a frequency measurement of the # backtracking in the grocery store path. Responses are compared to a computer-based scoring calculator.*

Relevancy
Not relevant (1)
Item needs some revision (2)
Relevant, but needs minor revision (3)
Very relevant (4)

Please explain what aspect is irrelevant, and provide a suggestion for revision:

Clarity
Not clear (1)
Item needs some revision (2)
Clear, but needs minor revision (3)
Very clear (4)

Please explain what aspect is unclear, and provide a suggestion for revision:

Sufficiency
Not Sufficient (1)
Item needs some revision (2)
Sufficient, but needs minor revision (3)
Very sufficient (4)
Please explain what aspect is unclear, and provide a suggestion for revision:

End of Block: Execution Accuracy

Start of Block: Additional Questions

Do you have any global recommendations that would increase the relevance of the PLANS instrument to assessment of functional planning?

____

Do you have any global recommendations that would increase the clarity of the PLANS instrument?

____

Do you have any global recommendations that would increase the sufficiency of the PLANS instrument to assessment of functional planning?

____

Is there any additional information you would like to provide regarding the content validity of the PLANS?

____

End of Block: Additional Questions

Appendix D PLANS Administration Manual

Planning in Life and Adapting to Novel Situations (PLANS)
Administration Manual

Introduction

The PLANS is a naturalistic-simulated instrument that provides detailed information about an individual's ability to plan in the context of real-world demands. The term functional planning is used to describe *the generation and execution of a plan of action to complete a series of sub-tasks required to reach a larger goal, in the context of real-world demands*. Functional planning resembles other measures of planning as both involve setting goals, cognitively looking ahead, organizing and sequencing steps and executing those steps to reach a goal. However, the model guiding functional planning investigates the ability to plan in real-world contexts, and includes factors that are relevant to the complex demands of the real world. Real-world planning demands often have unstructured goals and rules, which requires an individual to process complex information and create multiple sub-goals required to complete a task. Functional planning includes domains that account for problem anticipation and abstract thinking skills required to set realistic goals based on unstructured demands. Functional planning also emphasizes two distinct stages that comprise planning in the real world: plan formulation and plan execution. Plan formulation includes processing the task demands of a situation, envisioning a goal and creating a series of steps in order to achieve the goal(s). Plan execution is following through the plan that was generated, and monitoring one's behavior to follow the steps that were generated, and recognizing when the goal has been attained.

Background and Clinical Significance

According to the Center for Disease Control (CDC), every year approximately 1.7 million individuals receive a Traumatic Brain Injury (TBI) either as isolated injuries or comorbid with other injuries (Faul, Xu, Wald, & Coronado, 2010). Epidemiological data shows that a majority of reported cases of TBI are rated as mild or moderate severity at time of medical center admission, and the incidence of mild and moderate TBI cases is substantial. Individuals with TBI can experience short-term and long-term impairment related to cognitive and behavioral domains.

Individuals with brain injuries often experience impairment in executive functioning, including the ability to use functional planning. Functional planning involves *the generation and execution of a plan of action to complete a series of sub-tasks required to reach a larger goal, in the context of real-world demands*. Functional planning is necessary in goal-directed behavior, and facilitates adapting to change and solving novel problems as seen in tasks like cooking, shopping and planning an event. For individuals with brain injuries, functional planning limitations can impact the ability to control cognition and behavior necessary for coordination of day to day life, and can limit their ability to live independently and function in the community. Executive functioning impairment and functional planning limitations are commonly shared symptoms of both Acquired Brain Injuries (ABI) and TBI, and are often targets of assessment and rehabilitation intervention. However, there are few psychometrically sound and ecologically-valid instruments available that measure functional planning.

Naturalistic instruments are a promising assessment strategy to address the limitations of measuring functional planning. Naturalistic instruments have high degrees of ecological validity, meaning instruments designed to look and feel like real-life, and have testing demands that reflect every-day demands. Naturalistic instruments infer cognitive abilities from performance on tasks that resemble real-world demands in every-day environments, using behavioral observations and quality of performance as measures of neurocognitive functioning. Scoring criteria for these instruments are standardized and account for the individual's approach to the test demands, and the environmental factors that could contribute to performance. Real-world demands are often complex, dynamic and require overlapping task demands, which may increase the sensitivity of these instruments to elicit executive functioning and functional planning limitations. The clinical implications of performance on naturalistic instruments are also quite valuable, as the types of errors, rule breaks, initiation difficulties and inefficient behaviors observed in naturalistic instruments have a high degree of "representativeness" to the types of errors that could be observed in similarly demanding situations in day-to-day performance.

The Planning in Life and Adapting to Novel Situations (PLANS) instrument is a new clinical tool developed as an option for rehabilitation clinicians to measure functional planning. The PLANS uses naturalistic simulated tasks in a simulated clinical space to assess functional planning abilities and reflects the planning demands of community and independent living. The PLANS is designed to guide rehabilitation interventions with stronger ecological validity than current instruments, and incorporates functional assessment of plan formulation and execution in an open-ended planning task using a novel instrument design. The specific cognitive constructs underlying planning selected with input from rehabilitation clinicians to target cognitive and functional domains of the most importance to daily life and cognitive rehabilitation goals, which can be used to guide prescriptive rehabilitation plans to address planning limitations.

PLANS Development

The PLANS development used a novel mixed-methods design, and incorporated three empirical sources to inform PLANS development. The first data source is retrospective analysis of pilot data from clinical data collected from the administration of the Community Multiple Errands Test (CoMET) (Knutson, McCue, Terhorst & Kulzer, 2016). The CoMET was developed by the primary author of the PLANS, and consisted of over 2 years of design and piloting of a community-based naturalistic instrument that reflected aspects of problem solving and plan execution present in a community setting. Over 50 subjects are included in the pilot sample, and performance patterns raised clinical questions about the cognitive constructs relevant to the ability to plan and complete every-day planning tasks. The second data source is a scoping literature review of planning instruments that incorporate naturalistic demands and environment. The scoping literature review extracted data identified the simulated tasks, environments, and scoring systems of other published planning instruments, and identified the most critical elements for the PLANS instrument. The third data source uses structured interviews that were conducted with rehabilitation clinicians that work with individuals with traumatic brain injuries. The interviews directed rehabilitation clinicians to describe the cognitive demands, tasks, and environments most significant to functional planning in brain injury populations from the perspective of clinicians, which were used to support clinically meaningful factors of cognitive and community rehabilitation in the PLANS design.

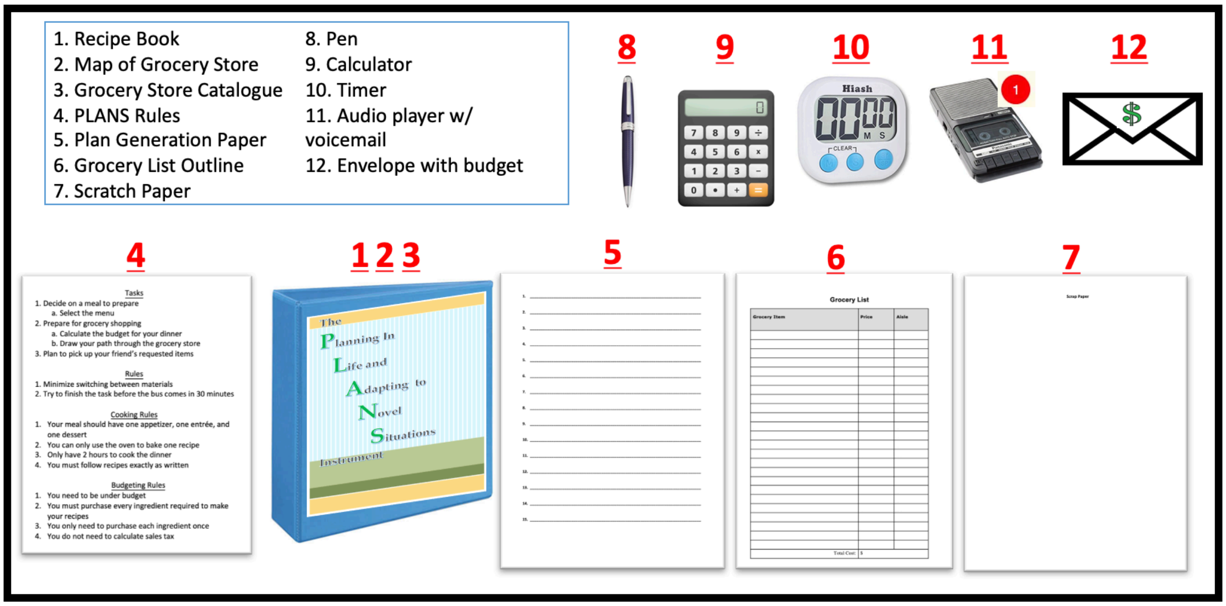
Description of PLANS

The PLANS is framed to reflect the real-world task of planning to go out in the community. The task that is presented is to plan for a trip to the grocery store. This activity involves several different subtasks including: identifying a budget, selecting recipes out of a cookbook, identifying the prices and ingredients of the recipes, and planning a route through a grocery store to shop for those items. The individual also needs to listen to a voicemail that describes another item that needs to be purchased at the grocery store. The participants are given a series of rules that need to be followed while completing the tasks. Participants have 5 minutes to generate a plan how they will accomplish these tasks, and are given 30 minutes to complete the tasks.

Logistics & Set-up
The PLANS requires between 30-45 minutes to administer. The introduction to the PLANS takes around 10 minutes, with half that time spent going through the instructions and the other half spent asking the task familiarity questions. Participants have up to 5 minutes to generate a plan, and are given as much time as necessary to execute the plan. While participants generate their plan, the examiner is observing their behavior and noting the number of times a step in the plan is revised once it has been written down. After the participant finished their plan or after time runs out, the examiner instructs the participant to carry out their plan. Participants are instructed that the plan execution should take 20 minutes, and a timer is set to go off 30 minutes after they begin executing the plan, but they can continue working after the timer ends. While the participant executes their plan, the examiner records the order that tasks are completed in including the start time and end time for the task. The examiner also observes for strategic behaviors that occur during task completion, and records the behavior, during what task it happens and how often it happens. Several of the PLANS scoring measures are calculated after the PLANS is completed, and use the examiners notes and the participants written plan, grocery list, scrap paper and drawn route through the grocery store to determine scores. The assessment flow for the plans can be referenced below.

Stage	Time	Examiner actions	PLANS Scoring
Instructions	10 minutes	Read through instructions	N/a
Plan Generation	5 minutes	Start timer Mark frequency that participant revises a step-in plan. Check plan to identify if the PLANS rules are included	**Plan Generation** Plan Generation Time Plan Generation Steps Plan Generation Omission Errors Plan Generation Sequencing Errors Plan Generation Rules Anticipated Plan Generation Plan Alterations
Plan Execution	30 – 45 minutes	Start timer Record the order that tasks are completed in, how much time is spent on each task, and monitor performance recording frequency of strategic behaviors. Record the final menu selected, the total cost of items and the order of grocery aisles visited. Collect the written plan, the scrap paper, the grocery list, and the grocery map. Post-administration: Input responses into Excel Worksheet to calculate Plan Alteration & Execution Accuracy scores.	**Plan Execution** Plan Execution Time Plan Execution Steps Plan Execution Omission Errors Plan Execution Sequencing Errors Plan Execution Rules Followed **Meal Accuracy** Grocery Budget Disparity Grocery # of errors (prices & aisles) Grocery Incorrect Aisles Visited

The PLANS is designed so that it can be administered sitting at a table in a clinical office. The examiner and participate sit at opposite ends of the table, and the examiner has their scoring materials obscured with a clipboard or document stand. Participants have several objects facing them at front of them at the table. From left to right they have a binder containing the PLANS cookbook, grocery store map, and grocery store catalogue, they have a piece of paper to write on, and they have a sheet of paper with the PLANS rules. On the right side of the table, participants are provided a calculator and pen that they use to complete the tasks. A digital timer is set up so that participants can keep track of time, and should be slightly angled so that the examiner can note the time as well. Behind the other testing materials, an envelope with the budget for the PLANS and an audio player with a pre-recorded voicemail are set up within reach of the participant. After participants complete the plan generation stage, they are provided a worksheet they can use to record the ingredients, prices and aisle locations for their menu. The layout of the PLANS materials can be referenced below.



Participant sits at table facing examiner

Tasks

Plan Generation

In order to prepare for their trip to the grocery store, participants are provided instructions that describe several tasks that need to be completed. The tasks are described in an arbitrary order and participants must determine the most efficient order to complete the tasks. Before starting, participants are given 5 minutes to write down a plan describing the steps they will take to complete the tasks. Instructions are given to make the plan as detailed as possible, so that someone else would be able to follow the plan. Participants are allowed to keep the written plan and reference it while they execute the tasks.

Plan Execution

Participants are instructed they have 20 minutes to execute their plan. The participants can attempt to complete the tasks the way that they are described in their plan, or they can deviate from their plan and attempt to plan in the moment. In either case, the tasks are developed so that they need to be completed in a specific order, as the participant needs information from some tasks to guide the rest of the tasks. The tasks can only be completed in one sequence that is considered “correct”, which is described in detail below:

First, participants must identify what their budget is to purchase necessary items from the grocery store. The participant needs to look in the envelope to identify the amount of money that their friend has provided them. Second, participants should listen to the voicemail that is left for them. The participants should press play and listen to the 30 second audio clip, and write down the item they need to purchase and the amount it costs. Third participants should open the PLANS binder to the cookbook. Here, participants need to select an appetizer, entrée and dessert for their meal. The participants should reference the PLANS rules page while selecting their menu, so that they select recipes that conform to the rules. The participants should write the ingredients of their selected recipes down on their worksheet.

Then, participants should flip to the binder tab with the grocery store catalogue. Scanning the catalogue, they should compare their written ingredients list to the catalogue items and record the price and aisle location of each item. Participants should calculate a total amount that their shopping trip will cost, and generate a list of aisles that they need to walk through in the grocery store. Finally, participants should then flip to the grocery store map, and use their aisle locations to draw an efficient path through the store. Participants should draw one continuous line that passes through every store location, and does not double back. The path should start at the grocery store entrance, and end at the grocery store exit. The path must pass through the pharmacy and the cash registers.

Instructions for Administering the PLANS & Script

Before introducing the participant to the clinical office where you are completing the PLANS, you will need to set up the testing space as it is described. Have ready the following documents: the PLANS scoring sheets and the grocery ingredient worksheet. Participants should be brought into the clinical office and sit in front of the PLANS materials, opposite of you. Introduce yourself and say the following:

"Today, you will be participating the Planning in Life and Adapting to Novel Situations (PLANS) instrument. The PLANS in an assessment that looks at how well you can plan and carry out every-day tasks. Some of these tasks involve anticipating challenges, information gathering, and problem solving. You will be given an overall goal to accomplish, several sub-tasks and a set of rules that you must follow. I will describe the tasks and provide you instructions, but I will be here just to observe, not to assist. We are doing this to assist you in identifying your strengths at everyday planning, and also to help us better understand any weaknesses you may have in planning and carrying out every-day tasks, so that we can work together to assist you in developing strategies for becoming more effective. We only ask that you try your best on this assessment. Any questions?"

If the participant has questions, repeat the instructions that they did not understand and clarify instructions until they report that they understand and move on. Say the following to the participants:

"Before I introduce the PLANS to you, I would like to ask you some questions about your experiences in every-day functioning. Using a scale of never – rarely – sometimes – often – or always, please answer the following questions: (1) when you eat at home, how often do you cook for yourself? (2) how often do you visit the grocery store by yourself? (3) how often do you follow a recipe when preparing a meal? (4) how often do you manage money to make purchases?"

Record the participant's answers for each question, and say the following:

"I have two more questions for you before we begin. First, using a scale of very effective – mostly effective – neither effective or ineffective – mostly ineffective – or very ineffective, how accurate are you at creating a plan to complete a series of tasks? Second, using a scale of very fast – mostly fast – neither fast or slow – mostly slow – or very slow, how fast are you at creating a plan to complete a series of tasks?"

Once the participant is sitting at the table and you have introduced yourself, you should say the following:

"For this activity, I would like you to consider this new situation: Your friend has just moved to a new home, and to celebrate you want to cook dinner for both of you today. Your task is to prepare for the dinner. Your friend gave you a set amount of money in this envelope, a cookbook, a catalogue of grocery store prices, a map of their local grocery store and a voicemail on this audio recorder. You will use these materials to help you prepare for your trip to the grocery store. The cookbook has 15 recipes in it, 5 appetizers, 5 entrees and 5 desserts, and it provides information on the ingredients, prep time & cook time, appliances needed, and serving size. To prepare for the dinner, you will need to create menu, a list

of what needs to be purchased and how much they will cost, and draw a path through the grocery store indicating where you will walk to purchase the items. You have a pen, calculator and budget sheet to record the grocery store list and prices, and there is a timer of when the bus will arrive. They also asked to purchase a medication for them from the pharmacy at the grocery store, but did not give you any information about the prescription. Using this voicemail from the pharmacy, record the most important information you need to pick up the medication. You can listen to this more than once."

Ask the participant if they understand the instructions so far. If they do not, repeat the instructions that they did not understand until they report that they understand. If participants do not understand the rules after 2 repetitions, move on. Point to the PLANS rules sheet that is in front of the participant, then say the following:

"There are some important rules that you should follow during while completing this task First, try to minimize switching between materials that your friend provided, so that you only use each item once"
You should point to recorder, envelope, recipe book, map & catalogue and clearly state that these 5 items should only be used once.
Point to the second PLANS rule and say:
"Now, you should try to finish before the bus comes in 30 minutes. If you go over your time, you can catch the next bus in an hour."

Then, you should move down the plans rules and continue to describe the rules that need to be followed. All 10 rules should be clearly stated as they are written on the page, and point to each rule as you say it out loud. Say the following:

"You also need to keep in mind the following rules about cooking dinner: Your meal should have one appetizer, one entrée, and one dessert, you can only use the oven to bake one recipe, you only have 2 hours to prepare the meal, you can multitask and prepare a recipe while others cook, you need to follow recipes exactly as written and your friend's kitchen has all necessary cookware and appliances to prepare these meals.
You should keep in mind the following rules about budgeting: You need to be under the budget your friend gave you, you must purchase every ingredient required to make your recipes, you only need to purchase an ingredient once and you do not need to calculate sales tax."

After reading these rules, you should ask the participant if they understand the rules. If they do not, repeat the instructions that they did not understand until they report that they understand. If participants do not understand the rules after 2 repetitions, move on. Say the following:

"Ok, good job. I have two additional questions for you before you begin: First, how long do you think it will take you to complete the PLANS tasks?"

Record their response and say the following:

"And second, what part of the PLANS do you think will be the most difficult for you"

Record their response and say the following:

"OK, thanks for answering those questions.
Now, I would like you to create a plan how to accomplish this task in an efficient manner, following the important rules. Try to anticipate what step you need to complete first, second, third and so on. Provide as much detail as possible in your plan. You have up to 5 minutes to write out your plan, and I encourage you to talk-aloud while creating the plan if it would help you."

When participants are ready, set the timer for 5 minutes and start the timer. After time timer ends or the participant indicates they are finished writing down their plan, say:

"Now you should begin this task. From this point on, I cannot answer any more questions or provide any input until your finish, and you should pretend that you are completing this activity by yourself. When you are finished, please let me know. Keep in mind that the bus that you will use to go to the grocery store comes in 20 minutes, so you should try to work quickly in order to catch the bus Once I start the timer, you may begin."

When the participants are ready, set the timer for 30 minutes and start the timer. When the participant indicates that they have finished the PLANS, say the following:

"Good job, and thank you for participating in this assessment!"

Scoring Guidelines
Task Familiarity & Experience

Task Familiarity

Definition: How much experience an individual has at completing tasks with similar demands to the PLANS.

Demands: The PLANS is comprised of several different tasks (money management, meal preparation, information gathering) framed around a real-world scenario, which may be salient to an individual's current or previous life roles.

Measure: Before beginning the PLANS, individuals are given the following questions, and asked to rate their level of familiarity.

Questions	0	1	2	3	4
1. When you eat at home, how often do you cook for yourself?	Never	Rarely	Sometimes	Often	Always
2. How often do you visit the grocery store by yourself?	Never	Rarely	Sometimes	Often	Always
3. How often do you follow a recipe when preparing a meal?	Never	Rarely	Sometimes	Often	Always
4. How often do you manage money to make purchases?	Never	Rarely	Sometimes	Often	Always

Task Estimation
Definition: The degree to which individuals possess accurate anticipatory self-awareness in anticipation of completing a specific task and similar and intellectual awareness of their strengths and weaknesses in general.
Demands: Answering these questions requires individuals to have accurate insight and awareness of their own abilities, and how those abilities interact with the demands of the PLANS tasks.
Measure: Before beginning the PLANS, individuals are given the following questions, and asked to rate their abilities and make estimations about their performance. Questions 1 & 2 are related to intellectual awareness and questions 3 & 4 are related to anticipatory awareness.

1.How accurate (effective) are you at creating a plan to complete a series of tasks?	5	4	3	2	1
	Very effective	Mostly effective	Neither effective or ineffective	Mostly ineffective	Very ineffective
2.How fast (efficient) are you at creating a plan to complete a series of tasks?	5	4	3	2	1
	Very fast	Mostly fast	Neither fast or slow	Mostly slow	Very slow

3. How long will it take you to finish these tasks?	_______:_____
4. What part of these tasks do you think will be the most difficult for you to complete?	

Plan Generation

Plan Generation Accuracy

Definition: The degree to which the generated plan accurately reflects the steps required to reach a proposed goal.

Demands: The PLANS Instructs individuals to generate a plan how to approach the various tasks. The plan should include all of the steps that must be completed in order to accurately reach the goals outlined in the PLANS. This requires individuals to accurately appraise the demands of the PLANS, anticipate the obstacles to completing these tasks and project themselves in the future envisioning what they need to accomplish in order to fulfil these tasks.

Measure: Written plans are assessed by the examiner by comparing the individual's generated plan to an answer key containing the most correct sequence of tasks. The correct order is presented in the left-most column, and the order generated by the participant is recorded in the parentheses. The steps in the generated plan will be coded and sequencing and omission errors will be calculated using excel score calculator sheet.

Step #	Ideal Order	Participant Written Plan
1 (__)	Check budget	
2 (__)	Collect pharmacy information	
3 (__)	Calculate remaining budget	
4 (__)	Select menu	
5 (__)	Record grocery list	
6 (__)	Calculate menu budget	
7 (__)	Identify aisle locations	
8 (__)	Draw path through store	
9 (__)	Tell Examiner when finished	

Plan Anticipation

Definition: The amount of information in a generated plan that reflects accurate anticipation of rules, constraints, and demands. Anticipation of demands involves correctly identifying the rule, constraint or demand and matching it to the step in the plan where it is relevant.

Demands: The PLANS has several rules, constraints and demands that need to be considered when generating a plan how to accomplish the task.

Measure: The following list represents the factors that should be anticipated for in the generated plan. The generated plans is checked by the examiner to identify if these factors are included to any degree in the written plan, and if they are included in a step where they apply. The item is only counted if it is included in the plan.

Anticipation:

Factor	Was the factor included in the plan?
Appetizer, Entrée, Dessert Restriction	Yes / No
Budgetary Restriction	Yes / No
Oven Restriction	Yes / No
Cooking Time Restriction	Yes / No
Portion Restriction	Yes / No

Plan Generation Speed

Definition: How much time was spent generating the plan. This does not directly reflect the amount of effort put into the plan or the speed at which one can formulate a plan, but is a measure of the total time spent working on the plan.

Demands: The PLANS instruct individuals to take 5 minutes before beginning to generate a series of steps they will need to complete to complete the tasks. They are instructed to create a plan that is detailed to a degree that could be clearly understood by someone else.

Measure: Number of seconds spent planning is recorded by the examiner, ranging from 0-300 (max time).

Time required to generate plan:	___ : ___

Plan Generation Efficacy
Definition: The degree to which plan generation is efficiently conducted, and the efficiency of the plan generated.
Demands: The PLANS requires individuals generate a plan describing how they will complete the various tasks presented to them, while following rules and constraints. Plan generation as an activity should be efficient, and the rules, constraints and demands of the PLANS should be accurately incorporated into the plan steps as they are written. In order to be efficient, the plan that is generated should have as few steps as required in order to complete the tasks.
Measure: Plan generation efficacy is measured by a frequency of times that the participant revises their plan. Any changes, additions, reorganizations or removed tasks are counted during the plan generation phase. The efficacy of the generated plan is measured by counting the number of individual steps.

# of times participant adjusts their written plan	
# of steps in the written plan	

Plan Execution

Plan Execution Accuracy

Definition: The degree to which an individual's actions accurately reflects the steps required to reach a proposed goal.
Demands: The PLANS instructs individuals to execute a plan and complete several tasks while following pre-defined rules and constraints. This requires individuals to accurately appraise the demands of the PLANS, anticipate the obstacles to completing these tasks and monitor their behavior while completing the tasks.
Measure: The examiner monitors the participant's behaviors while they complete the tasks, and records the order that tasks are completed in, how much time is spent on each task, and monitoring performance for any obvious errors or rule breaks. This information is transposed onto the following scoring tables that organize this information by the correct sequence of tasks and the amount of errors in each step (if any). The correct order is displayed in the left-most column, and the order generated by the participant is recorded in the parentheses. Plans are transposed into the excel score calculator, and plan omissions and sequencing errors are computed.

Task	Time Started	Time Stopped	Time on Task	Comments

Step #	
1 (__)	Check budget
2 (__)	Collect pharmacy information
3 (__)	Calculate remaining budget
4 (__)	Select menu
5 (__)	Record grocery list
6 (__)	Calculate menu budget
7 (__)	Identify aisle locations
8 (__)	Draw path through store
9 (__)	Tell Examiner when finished

Adherence to Generated Plan

Definition: The degree to which the steps and details of the generated plan accurately reflect the behaviors observed during plan execution.

Demands: The PLANS instructs individuals to generate a written plan detailing the sequence of steps they will follow to complete the various tasks. This plan should be used as a structure to guide their behavior while executing the tasks, but deviations may occur between the plan generated and the plan executed. Plan deviations can be positive and help the individual complete the tasks, or negative and hinder performance or result in errors or rule breaks. As an example of positive deviation, individuals may encounter an unanticipated obstacle while completing the tasks, or notice an error in their generated plan. In these cases, individuals may adjust their behavior to deviate from the plan and take part in "opportunistic planning", which facilitates overcoming an obstacle in a manner that reduces errors. In a negative deviation, individuals may forget to follow aspects of their written plan or have difficulty monitoring their progress in executing the plan, and deviate from the plan in a way that breaks a rule or completes a task inaccurately. In either case, it is clinically relevant to note if plans are adhered to, and how the individual adjusts their plans.

Measure: Examiners use the records from the plan generation and the recorded order of tasks from the plan execution. Plans are transposed into the excel score calculator, and the following measures are computed.

Plan Adherence:

	#
# of tasks added to plan execution not present in plan formulation	
# of tasks completed in a different order than plan formulation	
# of tasks omitted in execution that were not present in plan formulation	

Rule Following
Definition: The degree to which individuals are able to monitor their behavior to conform to rules and constraints outlined in instructions.
Demands: The PLANS has several rules that are described in its instructions. These rules are designed to increase the complexity of the PLANS and encourage individuals to plan out their approach to the tasks in an efficient manner. There are also several constraints to the tasks that individuals need to keep in mind while completing the PLANS.
Measure: After the PLANS has been finished, the examiner reviews the scoring for task accuracy and the information gathering domains and use the responses to determine if the rules and constraints of the PLANS were met. In addition, a programmed scoring sheet can automatically determine is the constraints were followed by inputting the menu selected.

Constraints and Rules Followed:

Appetizer, Entrée, Dessert Restriction	Yes	No
Budgetary Restriction	Yes	No
Oven Restriction	Yes	No
Cooking Time Restriction	Yes	No
Portion Restriction	Yes	No

Plan Execution Speed
Definition: How much time was spent executing the plan.
Demands: The PLANS instructions provide a specific time at which the individual should be finished with the tasks, and encourages working quickly while remaining accurate.
Measure: Number of seconds spent executing plan is recorded by the examiner.

Time required to execute plan:	

Plan Execution Menu Accuracy
Definition: The degree to which an individual can search, process and report information.
Demands: The PLANS reflects the degree of reading, information gathering and computation that is typical of the daily activities of meal planning and grocery store shopping. This includes processing visually complex materials such as a cookbook, a catalogue, and a map. Some of the tasks that need to be completed in the PLANS require a significant amount of information processing and gathering, transposing.
Measure: The PLANS requires that individuals report several pieces of information: A list of ingredients necessary for the recipes selected, the prices and aisles of those ingredients, and a total price of the ingredients. These values are compared to a programmed score system that displays the accurate values. The # of incorrect responses is recorded in the excel score sheet. The difference between the correct and provided total budget is calculated by the excel score sheet.

Menu:

#	Appetizer	Entrée	Dinner
1	$	$	$
2	$	$	$
3	$	$	$
4	$	$	$
5	$	$	$
6	$	$	$
7	$	$	$
8	$	$	$
9	$	$	$
10	$	$	$

Items Shared between Recipes		
Total Cost	$	

Grocery Store Areas Visited:

1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

Examiner Materials – PLANS Score Sheet

Pre-PLANS Questions

Questions	4	3	2	1	0
1. When you eat at home, how often do you cook for yourself?	Always	Often	Sometimes	Rarely	Never
2. How often do you visit the grocery store by yourself?	Always	Often	Sometimes	Rarely	Never
3. How often do you follow a recipe when preparing a meal?	Always	Often	Sometimes	Rarely	Never
4. How often do you manage money to make purchases?	Always	Often	Sometimes	Rarely	Never

5. How accurate (effective) are you at creating a plan to complete a series of tasks?	4	3	2	1	0
	Very effective	Mostly effective	Neither effective or ineffective	Mostly ineffective	Very ineffective
6. How fast (efficient) are you at creating a plan to complete a series of tasks?	4	3	2	1	0
	Very fast	Mostly fast	Neither fast or slow	Mostly slow	Very slow

7. How long will it take you to finish these tasks?	______:_____
8. What part of these tasks do you think will be the most difficult for you to complete?	

Was the participant able to repeat back the demands of the task on their first try?	Yes	How many times did the participant require?

PLANS Plan Formulation

Time required to generate plan:	______:_____
# of times participant revises their plan:	

Anticipation:

Factor	Was the factor included in the plan?	
Budgetary Restriction	Yes	No
Oven Restriction	Yes	No
Cooking Time Restriction	Yes	No
Portion Restriction	Yes	No

PLANS Plan Execution

Task	Time Started	Time Stopped	Time on Task
Time required to execute plan:			

Appetizer	Entrée	Dinner

Grocery Store Areas Visited:

1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
How many instances of backtracking are on the map?		

Bibliography

Aitken, S., Chase, S. McCue, M., & Ratcliff, G. (1993). An American adaptation of the Multiple Errands test: Assessment of executive abilities in everyday life (Abstract). *Archives of Clinical Neuropsychology, 8*(3), 212.

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.

Alvarez, J. A., & Emroy, E. (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review, 16*, 17-42.

Ardila, A., Pineda, D., & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Archives of clinical neuropsychology, 15*(1), 31-36.

Axelrod, B. N., & Millis, S. R. (1994). Preliminary standardization of the cognitive estimation test. *Assessment, 1*(3), 269-274.

Banich, M. T. (2009). Executive function: The search for an integrated account. *Current Directions in Psychological Science, 18*, 89-94.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychological bulletin, 121*(1), 65.

Barkley, R. A. (2001). The executive functions and self-regulation: An evolutionary neuropsychological perspective. *Neuropsychology review*, *11*(1), 1-29.

Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly, 71*(2), 287-311.

Beck, C. T., & Gable, R. K. (2001). Ensuring content validity: An illustration of the process. *Journal of Nursing Measurement, 9*(2), 201– 215.

Bennett, P. C., Ong, B., & Ponsford, J. (2005). Assessment of Executive dysfunction following traumatic brain injury: Comparison of the BADS with other clinical neuropsychological measures. *Journal of the International Neuropsychological Society*, *11*(05), 606-613.

Benton, A. L., Hamsher, K., & Sivan, A. B. (1983). Multilingual aphasia examination (3rd ed.). Iowa City, IA: AJA Associates.

Bernardin, H. J., & Buckley, M. R. (1981). Strategies in rater training. *Academy of Management Review*, *6*(2), 205-212.

Binder, L. M. 1997. A review of mild head trauma: Part II. Clinical implications. *Journal of Clinical and Experimental Neuropsychology, 19*: 432–457.

Boelen, D. H., Allain, P., Spikman, J. M., & Fasotti, L. (2011). Script generation and the dysexecutive syndrome in patients with brain injury. *Brain injury, 25*(11), 1091-1100.

Bogod, N. M., Mateer, C. A., & Macdonald, S. W. (2003). Self-awareness after traumatic brain injury: A comparison of measures and their relationship to executive functions. *Journal of the International Neuropsychological Society, 9*(3), 450-458.

Bottari, C. L., Dassa, C., Rainville, C. M., & Dutil, É. (2010). The IADL Profile: Development, content validity, intra-and interrater agreement. *Canadian Journal of Occupational Therapy, 77*(2), 90-100.

Bottari, C., Dassa, C., Rainville, C., & Dutil, E. (2009). The factorial validity and internal consistency of the Instrumental Activities of Daily Living Profile in individuals with a traumatic brain injury. *Neuropsychological Rehabilitation, 19*(2), 177-207.

Bottari, C., Iliopoulos, G., Shun, P. L. W., & Dawson, D. R. (2014). The clinical reasoning that guides therapists in interpreting errors in real-world performance. *The Journal of head trauma rehabilitation, 29*(6), 18-30.

Boyd, T. M., & Sautter, S. W. (1993). Route-finding: A measure of everyday executive functioning in the head-injured adult. *Applied Cognitive Psychology*, *7*(2), 171-181.

Broadbent, D. E., Cooper, P. F., FitzGerald, P., & Parkes, K. R. (1982). The cognitive failures questionnaire (CFQ) and its correlates. *British journal of clinical psychology, 21*(1), 1-16.

Brown, J. A., & Hux, K. (2016). Functional assessment of immediate task planning and execution by adults with acquired brain injury. *NeuroRehabilitation, 39*(2), 191-203.

Brown, J., & Hux, K. (2017). Ecologically Valid Assessment of Prospective Memory for Task Planning and Execution by Adults With Acquired Brain Injury. *American journal of speech-language pathology, 26*(3), 819-831.

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, 4*(6), 547-558.

Burgess, P. W., Alderman, N., Forbes, C., Costello, A., Laure, M. C., Dawson, D. R., Anderson, N, D., Gilbert, S. J., Dumontheil, I & Channon, S. (2006). The case for the development and use of "ecologically valid" measures of executive function in experimental and clinical neuropsychology. *Journal of the international neuropsychological society, 12*(2), 194-209.

Burgess, P. W., Alderman, N., Wilson, B. A., Evans, J. J., & Emslie, H. (1996). *The dysexecutive questionnaire. Behavioural assessment of the dysexecutive syndrome.* Bury St. Edmunds, UK: Thames Valley Test Company.

Burgess, P. W., Alderman, N., Wilson, B. A., Evans, J. J., & Emslie, H. (1996). The dysexecutive questionnaire. *Behavioural assessment of the dysexecutive syndrome*. Bury St. Edmunds, UK: Thames Valley Test Company.

Burgess, P., Simons, J. S., Coates, L. M. A., & Channon, S. (2005). The search for specific planning processes. In *The cognitive psychology of planning*, (pp. 199-227). Psychology Press.

Busch, R. M., McBride, A., Curtiss, G., & Vanderploeg, R. D. (2005). The components of executive functioning in traumatic brain injury. *Journal of clinical and experimental neuropsychology, 27*(8), 1022-1032.

Cahn-Weiner, D. A., Boyle, P. A., & Malloy, P. F. (2002). Tests of executive function predict instrumental activities of daily living in community-dwelling older individuals. *Applied neuropsychology, 9*(3), 187-191.

Cassidy, J. D., Carroll, L., Peloso, P., Borg, J., Von Holst, H., Holm, L., Krauss, J. & Coronado, V. (2004). Incidence, risk factors and prevention of mild traumatic brain injury: results of the WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury. *Journal of rehabilitation medicine, 36*(0), 28-60.

Centers for Disease Control and Prevention. (2011). Nonfatal traumatic brain injuries related to sports and recreation activities among persons aged≤ 19 years---United States, 2001--2009. MMWR: *Morbidity and mortality weekly report, 60*(39), 1337-1342.

Centers for Disease Control and Prevention. (2013). Traumatic Brain Injury. Retrieved from http://www.cdc.gov/traumaticbraininjury.

Chalmers, D., & Lawrence, J. A. (1993). Investigating the effects of planning aids on adults' and adolescents' organisation of a complex task. *International Journal of Behavioral Development, 16*(2), 191-214.

Chan, R. C., Shum, D., Toulopoulou, T., & Chen, E. Y. (2008). Assessment of executive functions: Review of instruments and identification of critical issues. *Archives of clinical neuropsychology, 23*(2), 201-216.

Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychological Review, 13*(4), 181–197.

Chaytor, N., Schmitter-Edgecombe, M., & Burr, R. (2006). Improving the ecological validity of executive functioning assessment. *Archives of clinical neuropsychology, 21*(3), 217-227.

Chevignard, M., Pillon, B., Pradat-Diehl, P., Taillefer, C., Rousseau, S., Le Bras, C., & Dubois, B. (2000). An ecological approach to planning dysfunction: script execution. *Cortex, 36*(5), 649-669.

Chua, K. S., Ng, Y. S., Yap, S. G., & Bok, C. W. (2007). A brief review of traumatic brain injury rehabilitation. *Annals-Academy of Medicine Singapore, 36*(1), 31.

Cicerone, K. D., & Tupper, D. E. (1991). Neuropsychological rehabilitation: Treatment of errors in everyday functioning. In *The neuropsychology of everyday life: Issues in development and rehabilitation* (pp. 271-292). Springer, Boston, MA.

Cicerone, K. D., Dahlberg, C., Kalmar, K., Langenbahn, D. M., Malec, J. F., Bergquist, T. F., Felicetti, T., Giacino, J. T., Harley, J. P., Harrington, D. E., Herzog, J., Kniepp, S., Laatsch, L., Morse, P. A. (2000). Evidence-based cognitive rehabilitation: recommendations for clinical practice. *Archives of physical medicine and rehabilitation, 81*(12), 1596-1615.

Cicerone, K., Levin, H., Malec, J., Stuss, D., & Whyte, J. (2006). Cognitive rehabilitation interventions for executive function: moving from bench to bedside in patients with traumatic brain injury. *Journal of cognitive neuroscience*, *18*(7), 1212-1222.

Clark, A. J., Anderson, N. D., Nalder, E., Arshad, S., & Dawson, D. R. (2017). Reliability and construct validity of a revised Baycrest Multiple Errands Test. *Neuropsychological rehabilitation, 27*(5), 667-684.

Craik, F. I., & Bialystok, E. (2006). Planning and task management in older adults: Cooking breakfast. *Memory & Cognition, 34*(6), 1236-1249.

Cronbach, L. J. (1990). Essentials of psychological testing. New York, happer and Row publishers.

Cuberos-Urbano, G., Caracuel, A., Vilar-López, R., Valls-Serrano, C., Bateman, A., & Verdejo-García, A. (2013). Ecological validity of the Multiple Errands Test using predictive models of dysexecutive problems in everyday life. *Journal of clinical and experimental neuropsychology*, *35*(3), 329-336.

Davies, S. P. (2004). Planning and problem solving in well-defined domains. In *The cognitive psychology of planning,* (pp. 45-48). Psychology Press.

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.

Delis, D. C. (2012). Delis rating of executive functions. *Bloomington, MN: Pearson*.

Delis, D. C., Kaplan, E., and Kramer, J. H. (2001). *Delis-Kaplan Executive Functioning System (D-KEFS)*. San Antonio: The Psychological Corporation.

Della Sala, S., MacPherson, S. E., Phillips, L. H., Sacco, L., & Spinnler, H. (2004). The role of semantic knowledge on the cognitive estimation task. *Journal of neurology, 251*(2), 156.

Dikmen, S. S., Corrigan, J. D., Levin, H. S., Machamer, J., Stiers, W., & Weisskopf, M. G. (2009). Cognitive outcome following traumatic brain injury. *The Journal of head trauma rehabilitation, 24*(6), 430-438.

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*.

Doninger, N. A., Ehde, D. M., Bode, R. K., Knight, K., Bombardier, C. H., & Heinemann, A. W. (2006). Measurement properties of the neurobehavioral cognitive status examination (Cognistat) in traumatic brain injury rehabilitation. *Rehabilitation Psychology*, *51*(4), 281.

Donovan, N. J., Kendall, D. L., Heaton, S. C., Kwon, S., Velozo, C. A., & Duncan, P. W. (2008). Conceptualizing functional cognition in stroke. *Neurorehabilitation and Neural Repair, 22*(2), 122-135.

Duncan, J. (1995). Attention, intelligence, and the frontal lobes. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 721-733). Cambridge, MA, US: The MIT Press.

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive psychology, 30*(3), 257-303.

Erez, A. B. H., Rothschild, E., Katz, N., Tuchner, M., & Hartman-Maeir, A. (2009). Executive functioning, awareness, and participation in daily life after mild traumatic brain injury: A preliminary study. *The American Journal of Occupational Therapy*, *63*(5), 634.

Evans, J. J., Wilson, B. A., Needham, P., & Brentnall, S. (2003). Who makes good use of memory aids? Results of a survey of people with acquired brain injury. Journal of the International *Neuropsychological Society, 9*(6), 925-935.

Faul, M., Xu, L., Wald, M. M., & Coronado, V. G. (2010). Traumatic Brain Injury in the United States. Emergency Department Visits, Hospitalizations and Deaths, 2006-2010.

Fleiss, J. L. (1986). Reliability of measurement. *The design and analysis of clinical experiments*, 1-32.

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological science, 17*(2), 172-179.

Gilhooly, K. J. (2005). Working memory and planning. *The cognitive psychology of planning*, (pp. 71-88). Psychology Press.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). Test review behavior rating inventory of executive function. *Child Neuropsychology*, *6*(3), 235-238.

Godfrey, H. P., & Shum, D. (2000). Executive functioning and the application of social skills following traumatic brain injury. *Aphasiology, 14*(4), 433-444.

Goldstein, G. (1996). Functional considerations in neuropsychology. In R. J. Sbordone & C. J. Long (Eds.), *Ecological validity of neuropsychological testing* (pp. 75–89). Delray Beach, Florida: GR Press/St. Lucie Press.

Harlow, J. M. (1848). Passage of an iron rod through the head. The Boston Medical and Surgical Journal, 39(20), 389-393.

Hart, T., Whyte, J., Kim, J., & Vaccaro, M. (2005). Executive function and self-awareness of "real-world" behavior and attention deficits following traumatic brain injury. *The Journal of head trauma rehabilitation, 20*(4), 333-347.

Hartman-Maeir, A., Harel, H., & Katz, N. (2009). Kettle Test—A brief measure of cognitive functional performance: Reliability and validity in stroke rehabilitation. *American Journal of Occupational Therapy, 63*(5), 592-599.

Hayes-Roth, B., & Hayes-Roth, F. (1979). A cognitive model of planning. *Cognitive science, 3*(4), 275-310.

Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in cognitive sciences, 16*(3), 174-180.

Holt, D. V., Rodewald, K., Rentrop, M., Funke, J., Weisbrod, M., & Kaiser, S. (2011). The Plan-a-Day approach to measuring planning ability in patients with schizophrenia. *Journal of the International Neuropsychological Society, 17*(2), 327-335.

Hommel, B., Ridderinkhof, R. K., & Theeuwes, J. (2002). Cognitive control of attention and action: Issues and trends.

Huck, S. W. (2012). Reading statistics and research (6th edition). *Boston: Pearson.*

Im, B., Hibbard, M., Grunwald, I., Swift, P. T., & Salimi, N. (2011). Medical aspects of disability: A handbook for the rehabilitation professional. Flanagan, S. R., Zaretsky, H., & Moroz, A. (Eds.). *New York, NY: Springer Publishing Company.*

Jefferson, A. L., Paul, R. H., Ozonoff, A. L., & Cohen, R. A. (2006). Evaluating elements of executive functioning as predictors of instrumental activities of daily living (IADLs). *Archives of clinical neuropsychology, 21*(4), 311-320.

Jovanovski, D., Zakzanis, K., Ruttan, L., Campbell, Z., Erb, S., & Nussbaum, D. (2012). Ecologically valid assessment of executive dysfunction using a novel virtual reality task in patients with acquired brain injury. *Applied Neuropsychology: Adult, 19*(3), 207-220.

Kaplan, E. (1983). *The assessment of aphasia and related disorders (Vol. 2).* Lippincott Williams & Wilkins.

Kliegel, M., Martin, M., McDaniel, M. A., & Phillips, L. H. (2007). Adult age differences in errand planning: The role of task familiarity and cognitive resources. *Experimental aging research, 33*(2), 145-161.

Klinger, E., Chemin, I., Lebreton, S., & Marié, R. M. (2006). Virtual action planning in parkinson's disease: A Control study. *Cyberpsychology & Behavior, 9*(3), 342-347.

Knight, C., Alderman, N., & Burgess, P. W. (2002). Development of a simplified version of the multiple errands test for use in hospital settings. *Neuropsychological Rehabilitation*, *12*(3), 231-255.

Knutson E., McCue M., Terhorst L. & Kulzer J. (2016) C-35 Relationships Between Community-Multiple Errands Test and Neuropsychological Measures. *Archives of clinical neuropsychology, 31(*6), 653-653.

Knutson, E., McCue, M. (2018, May). *The Application of Natural Instruments for Assessing Functional Cognition of Individuals with Brain Injuries.* Poster presentation at the 36th Annual Brain Injury Rehabilitation Conference. Williamsburg, VA.

Knutson, E., McCue, M., Skidmore, E., Arenth, P., Beck, K., Kulzer, J. & Terhorst, L. (2019). Performance on the Community Multiple Errands Test (CoMET) by Adults with Cognitive Disabilities. Manuscript submitted for publication.

Kovar, M. G., & Lawton, M. P. (1994). Functional disability: activities and instrumental activities of daily living. *Annual review of gerontology and geriatrics, 14*, 57-75.

Kruglanski, A. W., Shah, J. Y., Fishbach, A., & Friedman, R. (2018). A theory of goal systems. In *The Motivated Mind* (pp. 215-258). Routledge.

Langlois, J. A., Rutland-Brown, W., & Thomas, K. E. (2005). The incidence of traumatic brain injury among children in the United States: differences by race. *The Journal of head trauma rehabilitation, 20*(3), 229-238.

Langlois, J. A., Rutland-Brown, W., & Wald, M. M. (2006). The epidemiology and impact of traumatic brain injury: a brief overview. *The Journal of head trauma rehabilitation, 21*(5), 375-378.

Langlois, J.A., Rutland-Brown W., & Thomas K.E. (2004). Traumatic Brain Injury in the United States: Emergency Department Visits, Hospitalizations, and Deaths. Atlanta (GA): Centers for Disease Control and Prevention, National Center for Injury Prevention and Control.

Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: self-maintaining and instrumental activities of daily living. *The gerontologist, 9*(3), 179-186.

Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: self-maintaining and instrumental activities of daily living. *The gerontologist, 9*(3), 179-186.

Levine, B., Schweizer, T. A., O'Connor, C., Turner, G., Gillingham, S., Stuss, D. T., Manly, T. & Robertson, I. H. (2011). Rehabilitation of executive functioning in patients with frontal lobe brain damage with goal management training. Frontiers in human neuroscience, 5, 9.

Levine, B., Stuss, D. T., Milberg, W. P., Alexander, M. P., Schwartz, M., & MacDonald, R. (1998). The effects of focal and diffuse brain damage on strategy application: Evidence from focal lesions, traumatic brain injury and normal aging. Journal of the International Neuropsychological Society, 4(3), 247-264.

Lewis, M. W., Babbage, D. R., & Leathem, J. M. (2011). Assessing executive performance during cognitive rehabilitation. *Neuropsychological rehabilitation, 21*(2), 145-163.

Lewis, S. J., Dove, A., Robbins, T. W., Barker, R. A., & Owen, A. M. (2003). Cognitive impairments in early Parkinson's disease are accompanied by reductions in activity in frontostriatal neural circuitry. *Journal of Neuroscience, 23*(15), 6351-6356.

Lezak, M. D. (1982). The problem of assessing executive functions. *International journal of Psychology*, 17(1-4), 281-297.

Lezak, M. D. (2004). *Neuropsychological assessment*. Oxford University Press, USA.

Lhermitte, F., Pillon, B., & Serdaru, M. (1986). Human autonomy and the frontal lobes. Part I: Imitation and utilization behavior: a neuropsychological study of 75 patients. *Annals of neurology, 19*(4), 326-334.

Long, C. J. (1996). Neuropsychological tests: A look at our past and the impact that ecological issues may have on our future. *Archives of Clinical Neuropsychology, 22*(4), 327-337.

Luria, A. R. (1966). *Higher cortical functions in man*. New York: Basic Books.

Manly, T., Hawkins, K., Evans, J., Woldt, K., & Robertson, I. H. (2002). Rehabilitation of executive function: facilitation of effective goal management on complex tasks using periodic auditory alerts. Neuropsychologia, 40(3), 271-281.

Marcotte, T. D., Scott, J. C., Kamat, R., & Heaton, R. K. (2010). Neuropsychology and the prediction of everyday functioning. In T. D. Marcotte & I. Grant (Eds.), Neuropsychology of everyday functioning (pp. 5–38). New York, NY: The Guilford Press.

Marshman, L. A., Jakabek, D., Hennessy, M., Quirk, F., & Guazzo, E. P. (2013). Post-traumatic amnesia. *Journal of clinical neuroscience*, *20*(11), 1475-1481.

Martínez-Pernía, D., Núñez-Huasaf, J., Del Blanco, Á., Ruiz-Tagle, A., Velásquez, J., Gomez, M., Blesius, C. R., Ibanez, A., Fernandez-Manjon, B. & Slachevsky, A. (2017). Using game authoring platforms to develop screen-based simulated functional assessments in persons with executive dysfunction following traumatic brain injury. *Journal of biomedical informatics, 74*, 71-84.

Mcalister, C., Schmitter-Edgecombe, M., & Lamb, R. (2016). Examination of variables that may affect the relationship between cognition and functional status in individuals with mild cognitive impairment: A meta-analysis. *Archives of Clinical Neuropsychology, 31*(2), 123-147.

McCabe, P., Lippert, C., Weiser, M., Hilditch, M., Hartridge, C., & Villamere, J. (2007). Community reintegration following acquired brain injury. *Brain Injury, 21*(2), 231-257.

McCue, M., & Pramuka, M. (1998). Functional assessment. In *Rehabilitation* (pp. 113-129). Springer, Boston, MA.

McDonald, B. C., Flashman, L. A., & Saykin, A. J. (2002). Executive dysfunction following traumatic brain injury: neural substrates and treatment strategies. *NeuroRehabilitation, 17*(4), 333-344.

McGeorge, P., Phillips, L. H., Crawford, J. R., Garden, S. E., Sala, S. D., Milne, A. B., ... & Callender, J. S. (2001). Using virtual environments in the assessment of executive dysfunction. *Presence: Teleoperators & Virtual Environments, 10*(4), 375-383.

McKinlay, A., Grace, R. C., Horwood, L. J., Fergusson, D. M., Ridder, E. M., & MacFarlane, M. R. (2008). Prevalence of traumatic brain injury among children, adolescents and young adults: prospective evidence from a birth cohort. *Brain injury, 22*(2), 175-181.

Miotto, E. C., & Morris, R. G. (1998). Virtual planning in patients with frontal lobe lesions. *Cortex, 34*(5), 639-657.

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.

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.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L, M & de Vet, H. C. (2012). *COSMIN checklist manual*. Amsterdam: COSMIN.

Moore, D. J., Palmer, B. W., Patterson, T. L., & Jeste, D. V. (2007). A review of performance-based measures of functional living skills. *Journal of psychiatric research, 41*(1), 97-118.

Moore, D. J., Palmer, B. W., Patterson, T. L., & Jeste, D. V. (2007). A review of performance-based measures of functional living skills. *Journal of psychiatric research, 41*(1), 97-118.

Morris R., Kotitsa M., Bramham J. (2005) Planning in patients with focal brain damage: From simple to complex task performance. *The cognitive psychology of planning*, (pp. 164-191). Psychology Press.

Morris, R. G., Ahmed, S., Syed, G. M., & Toone, B. K. (1993). Neural correlates of planning ability: Frontal lobe activation during the Tower of London test. *Neuropsychologia, 31,* 1367–1378.

Morris, R. G., Miotto, E. C., Feigenbaum, J. D., Bullock, P., & Polkey, C. E. (1997). The effect of goal–subgoal conflict on planning ability after frontal- and temporal-lobe lesions in humans. *Neuropsychologia, 35*, 1147–1157.

Morris, R., & Ward, G. (2004). Introduction to the psychology of planning. In *The cognitive psychology of planning* (pp. 11-13). Psychology Press.

Mosenthal, A. C., Livingston, D. H., Lavery, R. F., Knudson, M. M., Lee, S., Morabito, D., Manley G.T., Nathens A., Jurkovich G.J., Hoyt D. B., & Coimbra, R. (2004). The effect of age on functional outcome in mild traumatic brain injury: 6-month report of a prospective multicenter trial. *Journal of Trauma and Acute Care Surgery, 56*(5), 1042-1048.

Mueller, J. A., & Dollaghan, C. (2013). A systematic review of assessments for identifying Executive function impairment in adults with acquired brain injury. *Journal of Speech, Language, and Hearing Research*, *56*(3), 1051-1064.

Murphy, K. R., & Balzer, W. K. (1989). Rater errors and rating accuracy. *Journal of Applied Psychology*, *74*(4), 619.

Naglieri, J. A., & Goldstein, S. (2013). *Comprehensive executive functioning index*. Toronto: Multi Health Systems.

Nalder, E. J., Clark, A. J., Anderson, N. D., & Dawson, D. R. (2017). Clinicians' perceptions of the clinical utility of the Multiple Errands Test for adults with neurological conditions. *Neuropsychological rehabilitation, 27*(5), 685-706.

National Center for Injury Prevention and Control (US). (2003). Report to Congress on mild traumatic brain injury in the United States: Steps to prevent a serious public health problem. Centers for Disease Control and Prevention.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.

Norman, D. A., & Shallice, T. (1986). Attention to action. In *Consciousness and self-regulation* (pp. 1-18). Springer, Boston, MA.

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.

Novakovic-Agopian, T., Chen, A. J. W., Rome, S., Rossi, A., Abrams, G., D'Esposito, M., ... & Kennedy, C. (2014). Assessment of subcomponents of executive functioning in ecologically valid settings: the goal processing scale. *The Journal of head trauma rehabilitation, 29*(2), 136-146.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory (McGraw-Hill Series in Psychology)* (Vol. 3). New York: McGraw-Hill.

O'Cathain, A., & Thomas, K. J. (2004). " Any other comments?" Open questions on questionnaires–a bane or a bonus to research?. *BMC medical research methodology, 4*(1), 25.

O'keeffe, F., Dockree, P., Moloney, P., Carton, S., & Robertson, I. H. (2007). Awareness of deficits in traumatic brain injury: A multidimensional approach to assessing metacognitive knowledge and online-awareness. *Journal of the International Neuropsychological Society, 13*(1), 38-49.

O'Neil-Pirozzi, T. M., Goldstein, R., Strangman, G. E., Katz, D. I., & Glenn, M. B. (2010). Test–re-test reliability of the virtual planning test in individuals with traumatic brain injury. *Brain injury, 24*(3), 509-516.

Obler, L. K., Rykhlevskaia, E., Schnyer, D., Clark-Cotton, M. R., Spiro, A., Hyun, J., Kim, D., Goral, M. & Albert, M. L. (2010). Bilateral brain regions associated with naming in older adults. *Brain and language, 113*(3), 113-123.

Obonsawin, M. C., Crawford, J. R., Page, J., Chalmers, P., Cochrane, R., & Low, G. (2002). Performance on tests of frontal lobe function reflect general intellectual ability. *Neuropsychologia, 40*, 970-977

Ormerod, T. C. (2004). Planning and ill-defined problems. In *The cognitive psychology of planning* (pp. 63-66). Psychology Press.

Ownsworth, T. L., McFarland, K., & Mc Young, R. (2000). Self-awareness and psychosocial functioning following acquired brain injury: An evaluation of a group support programme. *Neuropsychological Rehabilitation, 10*(5), 465-484.

Ownsworth, T., & Clare, L. (2006). The association between awareness deficits and rehabilitation outcome following acquired brain injury. *Clinical psychology review, 26*(6), 783-795.

Ownsworth, T., & Fleming, J. (2005). The relative importance of metacognitive skills, emotional status, and executive function in psychosocial adjustment following acquired brain injury. *The Journal of head trauma rehabilitation, 20*(4), 315-332.

Parsons, T. D. (2015). Ecological validity in virtual reality-based neuropsychological assessment. In *Encyclopedia of Information Science and Technology, Third Edition* (pp. 1006-1015). IGI Global.

Parsons, T. D. (2016). Ecological Validity. In *Clinical Neuropsychology and Technology* (pp. 11-27). Springer International Publishing.

Pea, R. D., & Hawkins, J. (1987). Planning in a chore-scheduling task. In *Blueprints for thinking: The role of planning in cognitive development*, 273-302.

Pentland, L., Todd, J. A., & Anderson, V. (1998). The impact of head injury severity on planning ability in adolescence: A functional analysis. *Neuropsychological Rehabilitation, 8*(3), 301-317.

Ponsford, J., Draper, K., & Schönberger, M. (2008). Functional outcome 10 years after traumatic brain injury: its relationship with demographic, injury severity, and cognitive and emotional status. Journal of the International *Neuropsychological Society, 14*(2), 233-242.

Portney, L. G., & Watkins, M. P. (2009). *Foundations of clinical research: applications to practice* (Vol. 3). Upper Saddle River, NJ: Prentice Hall.

Prigatano, G. P. (2005). Disturbances of Self-awareness and Rehabilitation of Patients With Traumatic Brain Injury: A 20-Year Perspective. *The Journal of head trauma rehabilitation, 20*(1), 19-29.

Prigatano, G. P., & Schacter, D. L. (Eds.). (1991). *Awareness of deficit after brain injury: Clinical and theoretical issues.* Oxford University Press.

Rabin, L. A., Burton, L. A., & Barr, W. B. (2007). Utilization rates of ecologically oriented instruments among clinical neuropsychologists. *The Clinical Neuropsychologist, 21*(5), 727-743.

Rabinowitz, A. R., & Levin, H. S. (2014). Cognitive sequelae of traumatic brain injury. *Psychiatric Clinics of North America, 37*(1), 1-11.

Rabinowitz, A. R., & Levin, H. S. (2014). Cognitive sequelae of traumatic brain injury. *The Psychiatric Clinics of North America, 37*(1), 1.

Radomski, M. V., Weightman, M. M., Davidson, L. F., Finkelstein, M., Goldman, S., McCulloch, K., Roy, T., C., Scherer, M. & Stern, E. B. (2013). Development of a measure to inform return-to-duty decision making after mild traumatic brain injury. *Military medicine, 178*(3), 246-253.

Reitan, R. M., & Wolfson, D. (1985). *The Halstead-Reitan neuropsychological test battery: Theory and clinical interpretation (Vol. 4).* Reitan Neuropsychology.

Robertson, K., & Schmitter-Edgecombe, M. (2015). Self-awareness and traumatic brain injury outcome. *Brain injury, 29*(7-8), 848-858.

Robertson, K., & Schmitter-Edgecombe, M. (2017). Naturalistic tasks performed in realistic environments: a review with implications for neuropsychological assessment. *The Clinical Neuropsychologist, 31*(1), 16-42.

Roca, M., Parr, A., Thompson, R., Woolgar, A., Torralva, T., Antoun, N., Manes F. & Duncan, J. (2009). Executive function and fluid intelligence after frontal lobe lesions. *Brain, 133*(1), 234-247.

Saatman, K. E., Duhaime, A. C., Bullock, R., Maas, A. I., Valadka, A., & Manley, G. T. (2008). Classification of traumatic brain injury for targeted therapies. *Journal of neurotrauma, 25*(7), 719-738.

Salthouse, T. A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology, 19*(4), 532.

Salvucci, D. D., & Taatgen, N. A. (2008). Threaded cognition: An integrated theory of concurrent multitasking. *Psychological review*, *115*(1), 101.

Sanders, C., Low, C., & Schmitter-Edgecombe, M. (2014). Assessment of planning abilities in individuals with mild cognitive impairment using an open-ended problem-solving task. *Journal of clinical and experimental neuropsychology, 36*(10), 1084-1097.

Sanders, C., Low, C., & Schmitter-Edgecombe, M. (2014). Assessment of planning abilities in individuals with mild cognitive impairment using an open-ended problem-solving task. *Journal of clinical and experimental neuropsychology, 36*(10), 1084-1097.

Sarazin, M., Pillon, B., Giannakopoulos, P., Rancurel, G., Samson, Y., & Dubois, B. (1998). Clinicometabolic dissociation of cognitive functions and social behavior in frontal lobe lesions. *Neurology, 51*(1), 142-148.

Sawilowsky, S. S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods, 8(*2), 597 – 599.

Schmeichel, B. J., & Tang, D. (2015). Individual differences in executive functioning and their relationship to emotional processes and responses. *Current Directions in Psychological Science, 24*(2), 93-98.

Schmitter–Edgecombe, M., McAlister, C., & Weakley, A. (2012). Naturalistic assessment of everyday functioning in individuals with mild cognitive impairment: The day-out task. *Neuropsychology, 26*(5), 631.

Seter, C., Giovannetti, T., Kessler, R. K., & Worth, S. (2011). Everyday action planning in schizophrenia. *Neuropsychological Rehabilitation, 21*(2), 224-249.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society, London, D, 298*(1089), 199-209.

Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge University Press.

Shallice, T., & Burgess, P. W. (1991a). Higher-Order Cognitive Impairments and Frontal lobe lesion in man. In H. S. Levin, H. M. Eisenberg, & A. L. Benton (Eds.), *Frontal lobe function and dysfunction* (pp. 125-138). Oxford: Oxford University Press.

Shallice, T., & Burgess, P. W. (1991b). Deficits in strategy application following frontal lobe damage in man. *Brain, 114*(2), 727-741.

Shallice, T., & Evans, M. E. (1978). The involvement of the frontal lobes in cognitive estimation. *Cortex, 14*, 294–303.

Shanahan, L., McAllister, L., & Curtin, M. (2011). The party planning task: a useful tool in the functional assessment of planning skills in adolescents with TBI. *Brain injury*, *25*(11), 1080-1090.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin, 86*(2), 420.

Shum, D., Gill, H., Banks, M., Maujean, A., Griffin, J., & Ward, H. (2009). Planning ability following moderate to severe traumatic brain injury: Performance on a 4-disk version of the Tower of London. *Brain Impairment, 10*(3), 320-324.

Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical therapy*, *85*(3), 257-268.

Simblett, S. K., & Bateman, A. (2011). Dimensions of the Dysexecutive Questionnaire (DEX) examined using Rasch analysis. *Neuropsychological Rehabilitation, 21*(1), 1-25.

Smith, L. B., Radomski, M. V., Davidson, L. F., Finkelstein, M., Weightman, M. M., McCulloch, K. L., & Scherer, M. R. (2014). Development and preliminary reliability of a multitasking assessment for executive functioning after concussion. *American Journal of Occupational Therapy*, *68*(4), 439-443.

Sosin, D. M., Sniezek, J. E., & Thurman, D. J. (1996). Incidence of mild and moderate brain injury in the United States, 1991. *Brain injury, 10*(1), 47-54.

Spencer, R. J., & Johnson-Greene, D. (2009). The Cognitive Estimation Test (CET): psychometric limitations in neurorehabilitation populations. *Journal of Clinical and Experimental Neuropsychology, 31*(3), 373-377.

Sternbach, G. L. (2000). The Glasgow coma scale. *The Journal of emergency medicine*, *19*(1), 67-71.

Stulemeijer, M., van der Werf, S., Borm, G. F. and Vos, P. E. 2008. Early prediction of favourable recovery 6 months after mild traumatic brain injury. Journal of Neurology, Neurosurgery, and Psychiatry, 79(8): 936–942.

Stuss, D. T. (2011a). Functions of the frontal lobes: relation to executive functions. *Journal of the international neuropsychological Society, 17*(5), 759-765.

Stuss, D. T. (2011b). Traumatic brain injury: relation to executive dysfunction and the frontal lobes. *Current opinion in neurology, 24*(6), 584-589.

Stuss, D. T., & Alexander, M. P. (2000). Executive functions and the frontal lobes: a conceptual view. *Psychological research, 63*(3-4), 289-298.

Stuss, D. T., Ely, P., Hugenholtz, H., Richard, M. T., LaRochelle, S., Poirier, C. A., & Bell, I. (1985). Subtle neuropsychological deficits in patients with good recovery after closed head injury. *Neurosurgery, 17*(1), 41-47.

Stuss, D.T. & Benson, D.F. (1986). *The frontal lobes*. New York: Raven Press.

Tanguay, A. N., Davidson, P. S., Nuñez, K. V. G., & Ferland, M. B. (2014). Cooking breakfast after a brain injury. *Frontiers in behavioral neuroscience*, *8*.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International journal of medical education, 2*, 53.

Taylor, R. (1990). Interpretation of the correlation coefficient: a basic review. *Journal of diagnostic medical sonography, 6*(1), 35-39.

Thompson, H. J., McCormick, W. C., & Kagan, S. H. (2006). Traumatic brain injury in older adults: epidemiology, outcomes, and future implications. *Journal of the American Geriatrics Society, 54*(10), 1590-1595.

Thompson, T. E., Felce, D. E., & Symons, F. J. (2000). *Behavioral observation: Technology and applications in developmental disabilities*. Paul H Brookes Publishing.

Tindall, S.C. (1990) Level of consciousness. In: Walker HK, Hall WD, Hurst JW, editors. Clinical Methods: The History, Physical, and Laboratory Examinations. Boston: Butterworths. pp. 296–299.

Toglia, J., & Kirk, U. (2000). Understanding awareness deficits following brain injury. *NeuroRehabilitation, 15*(1), 57-70.

Tombaugh, T. N., & McIntyre, N. J. (1992). The mini-mental state examination: a comprehensive review. *Journal of the American Geriatrics Society*, *40*(9), 922-935.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner Review: Do performance-based measures and ratings of executive function assess the same construct?. *Journal of Child Psychology and Psychiatry, 54*(2), 131-143.

Tranel, D., Hathaway-Nepple, J., & Anderson, S. W. (2007). Impaired behavior on real-world tasks following damage to the ventromedial prefrontal cortex. *Journal of Clinical and Experimental Neuropsychology, 29*(3), 319-332.

Valls-Serrano, C., Verdejo-García, A., Noël, X., & Caracuel, A. (2017). Development of a Contextualized Version of the Multiple Errands Test for People with Substance Dependence. *Journal of the International Neuropsychological Society*, 1-13.

Waltz, C. F., Strickland, O. L., & Lenz, E. R. (Eds.). (2010). Validity of Measures with Karen L. Soeken (163-201). In *Measurement in nursing and health research*. Springer publishing company.

Ward, G. (2005). Planning and the executive control of thought and action. *The cognitive psychology of planning*, (pp. 89-110). Psychology Press.

Weightman, M. M., McCulloch, K. L., Radomski, M. V., Finkelstein, M., Cecchini, A. S., Davidson, L. F., ... & Scherer, M. R. (2017). Further development of the Assessment of Military Multitasking Performance: iterative reliability testing. *PLoS one, 12*(1), e0169104.

Werner, P., Rabinowitz, S., Klinger, E., Korczyn, A. D., & Josman, N. (2009). Use of the virtual action planning supermarket for the diagnosis of mild cognitive impairment. *Dementia and geriatric cognitive disorders, 27*(4), 301-309.

Whiteneck, G. G., Gerhart, K. A., & Cusick, C. P. (2004). Identifying environmental factors that influence the outcomes of people with traumatic brain injury. The Journal of head trauma rehabilitation, 19(3), 191-204.

Wildman, B. G., Erickson, M. T., & Kent, R. N. (1975). The effect of two training procedures on observer agreement and variability of behavior ratings. *Child Development*, 520-524.

Wilson, B. A., Alderman, N., Burgess, P. W., Emslie, H., & Evans, J. J. (1996). *Behavioural assessment of the dysexecutive syndrome.* Bury St Edmunds: Thames Valley Test Company.

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.

Wilson, R. C., Christensen, P. R., Merrifield, P. R., & Guilford, J. P. (1960). Alternate uses test manual. *Beverly Hills, Calif.: Sheridan Supply*.

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.

Zamanzadeh, V., Rassouli, M., Abbaszadeh, A., Majd, H. A., Nikanfar, A., & Ghahramanian, A. (2015). Details of content validity and objectifying it in instrument development. *Nursing Practice Today, 1*(3), 163-171.

Zelazo, P. D., & Müller, U. (2002). Executive function in typical and atypical development. *Blackwell handbook of childhood cognitive development*, 445-469.

Zook, N. A., Davalos, D. B., DeLosh, E. L., & Davis, H. P. (2004). Working memory, inhibition, and fluid intelligence as predictors of performance on Tower of Hanoi and London tasks. *Brain and cognition, 56*(3), 286-292.