ABSTRACT

COWLEY, JENNIFER ANTONIA. Towards a Theory of Mind Wandering in Relation to Task Type, Behavioral Responses, and Respective Adverse Consequences in Piloted Vehicles. (Under the direction of Douglas Gillan.)

In critical systems, human operators are often responsible for maintaining the safe operation of systems. However, cognitive factors (e.g., affect, attentional resources, fatigue, etc.) can negatively affect the operator's ability to stay focused on the task and make timely decisions. Although prior research has studied multiple factors that impact performance, such as distracted driving, situation awareness, multitasking, vigilance decrements, etc., mind wandering has yet to be integrated with these factors to yield a more comprehensive theory of human performance. To address this problem, we employed a mixed method, qualitative and quantitative, approach to study performance implications of mind wandering in the context of automobile driving and aviation piloting using undergraduate students and certified pilots, respectively. Only prominent results are reviewed.

In Part I, we assessed the degree of unrelatedness required between thoughts and the current task at hand before it was subjectively classified as mind wandering with respect to four types of relatedness (e.g., functional, conceptual, temporal relatedness and thought tense). Using a factor scenario survey approach, high and low levels of each type of relatedness were correlated with participant's perceptions of mind wandering. It was found that thoughts conceptually unrelated to the task were significantly more often identified as mind wanderings; the only main effect of these four types of relatedness. In addition, there were a few significant interactions with functional relatedness and the tense as well as temporal relatedness. Low functional, retrospective thoughts were more likely considered mind

wanderings than high functional, prospective thoughts. Also, low functional thoughts in the distant past or distant future were more likely considered mind wanderings (e.g., low temporal relatedness) than high functional thoughts in the near past or future (e.g., high temporal relatedness). In addition to these relatedness types studied in Part I, thoughts were also categorized on a continuum of task/thought decoupling with on-task thoughts at one end of the continuum (e.g., low decoupling) and mind wanderings that lack meta-awareness at the other (e.g., high decoupling) in Part IV of this research effort. Based on discussions with participants during these studies, the continuum of all thought categories is ordered from least to most decoupled as follows: on-task, off-task, mind wandering and unaware mind wandering.

In Part III, we explored whether the level of thought/task decoupling was related to performance decrements in the driving simulation environment. When participants were most frequently mind wandering without meta-awareness which is the most extreme decoupled state, the highest mean number of lane deviations (2.26 deviations) and the second highest mean seconds of speeding (5.53 seconds) were recorded. When participants were most frequently on-task, the lowest mean seconds of speeding (1.17 seconds) and second lowest number of mean lane deviations (1.00 deviations) also resulted. A similar relationship was found in the flight simulation with general aviation pilots even though the frequencies of off-task thinking were lower. When all on-task and off-task thought contents were examined, it was found that when pilots believed their thoughts were on-task, thought decoupling occurred as the pilot moved from high task load situations (e.g., take-off and climb) to lower task load situations (e.g., cruise). During take-off and climb, pilots' thinking involved actions/behaviors previously executed as well as thoughts about navigation. If poor performance occurred in the take-off phase of flight, thoughts about their performance often appeared during the climb phase of flight. During cruise, thinking became more decoupled from the task (e.g., thoughts about planning and strategizing) even though these thoughts were participant classified as on-task. Thus, as the pilots executed tasks from take-off to cruise, the thoughts gradually became more decoupled. Future work needs to assess performance implications of decoupling in the flight simulations.

Correlates to mind wandering were also assessed to understand when and why mind wandering may occur. Similar to findings in the prior literature, negative mood was found to be significantly and positively related to higher frequencies of off-task thinking but not mind wandering. Yet, task load, subjective mental workload, task type and working memory capacity did not significantly correlate with frequencies of off-task thinking and mind wandering.

Lastly, this work investigated the efficacy of three mind wandering metrics used in prior literature (e.g., all probe, self-caught and probe, and clicker). Efficacy was operationally defined as a metric that from the participant's viewpoint, facilitated quick task re-engagement, accurately captured the frequency of mind wanderings and did not unnaturally engender excessive meta-awareness. The results suggest that the clicker method was the most efficacious metric followed by the self-caught/probe and the all probe metrics.

This mixed method approach sought to build theory about what mind wandering is, when it arises and for what reasons. It is not clear how mind wandering fits with other research on related phenomenon like vigilance, multitasking, distracted driving, automatic/controlled processing, and novice/expert performance. Thus, a framework is needed to understand how the results reported herein overlap with other theory generated about these related phenomena. Future work should also evaluate how decoupling of the thinking from the task is facilitated or not facilitated by parameters of the task environment. In addition, a potential theory of mind wandering with respect to homeostatic levels of arousal is presented.

Towards a Theory of Mind Wandering in Relation to Task Type, Behavioral Responses, and Respective Adverse Consequences in Piloted Vehicles

by
Jennifer Antonia Cowley

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Psychology

Raleigh, North Carolina

2012

APPROVED BY:

Douglas J. Gillan Committee Chair	Anne C. McLaughlin
Christopher B. Mayhorn	Stan Dicks

Key Dismukes

DEDICATION

I dedicate this dissertation to everyone I love.

BIOGRAPHY

Jennifer Cowley received the Bachelors of Science in Animal Science from the University of Illinois at Urbana-Champaign (UIUC) in May of 1997. Her passion for research began at UIUC when she was employed at the UIUC Roy J. Carver Biotechnology Center to research transgenic livestock. During her tenure at NCSU, her research interests blossomed through her work experiences at the SAS Institute, Inc., the MITRE Corporation, and at the Software Engineering Institute at Carnegie Mellon University. Her interests include the study of human error (e.g., defining and measuring error and identifying consequences of error), warning systems design (including interface and universal design) to reduce human error, and research methods and metrics.

.

ACKNOWLEDGMENTS

I give thanks to my father, Greg Cowley for teaching me the art of exploring this fascinating world. In addition, I'd like to thank my mother, Joanne Cowley, for teaching me persistence and attention to detail. Also, I'd like to thank my husband, Travis Breaux, for his patience and unbounded belief in my abilities.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. Problem

Expert operators engaging in repetitive tasks commit a surprising number of errors, failures and mistakes (Loukopoulos, Dismukes & Barshi, 2009) despite extensive operator training and ample motivation to maintain system safety. Recent studies of well-trained commercial flight deck operators identified *human errors* as the most common contributor to commercial aviation accidents (Wiegman & Shappell, 2003). The resultant accidents are correlated with pilot deviations from normal, well-practiced flight task routines, including the circumstance when pilots must defer, or concurrently execute, one task with a second task (Loukopoulos, Dismukes, & Barshi, 2009). It is not entirely clear why deviations occur; the study of human behavior (e.g., the deviations from normal routines) cannot be devoid of the study of the thoughts that often precede the behavior. Mind wandering may be one potential cause of certain behaviors that can potentially lead to failures. Mind wandering, also called task-unrelated thinking, is when the operator is thinking about something other than the current task. Despite multiple studies on the effects of mind wandering in educational settings (Schooler, Reichel, & Halpern, 2005), mind wandering has not been well studied as a potential cause of human error in the transportation domain. In part, this may be because mind wandering is difficult to measure. To address this problem, researchers need to evaluate the efficacy of existing mind wandering metrics used in other domains for its application to the transportation domain. With validated metrics, researchers can more precisely pinpoint causal mechanisms of mind wandering and their contribution to human error. The research reported herein measures the relationship among mind wandering, the tasks juxtaposed to the mind wandering episode, and resulting human error, if any, while driving and flying simulated vehicles. To achieve this goal, the following aims have been identified and will be reported:

Aim #1: To identify existing metrics for measuring mind wandering, this document surveys the literature and applies the three most favorable metrics to driving and flight simulation studies.

Aim #2: To test the efficacy of previously identified cognitive mind wandering correlates (e.g., working memory capacity and negative affect) in the transportation domain, the research reported herein measured these potential correlates in a driving simulation experiment.

Aim #3: To measure the relationship between mind wandering, task performance and error, the present research examined automobile drivers and aviation pilots in simulation.

Together, these aims were selected to generate new data necessary for theory development to understand the potential implications of mind wandering in the transportation domain.

2. Literature Review

We now review the literature for relevant findings in the human error and mind wandering before summarizing the gaps in both areas of literature.

2.1.Human Error Defined

Multiple human error studies offer different definitions of human error as well as different viewpoints on causal mechanisms. For example, separate definitions define human error as failure of action (Gertman, Hallbert, Parris, Sattision, Brownson & Tortorelli, 2001), failure of cognition (Pounds & Isaac, 2002) or failure to achieve certain goals (Reason, 1982, 1995). Table 1 presents a list of definitions of human error. Because these investigators viewed the problem differently, the definitions are not unified into a single theory of human error. Similarly, human error has been defined through taxonomic classifications that express what factors led to such failures (Norman, 1984; Reason, 1990; Shappell, Detwiler, Holcomb, Hackworth, Boquet, Wiegmann, 2006). For example, Reason (1995) defines an error of omission as one that occurs when an intended behavior is not executed. Finally, several investigators study human error but never provide a definition (Edmondson, 2004; Leiden, Laughery, Keller, French, Warwick, & Wood, 2001; Macwan & Mosleh, 1994)

Human error may be difficult to define because caveats exist in the source of error. For example, the English Oxford Dictionary (EOD) defines error as, “the state or condition of being wrong in conduct or judgment.” The caveat with this definition is the implication that someone (an operator, an accident investigator) can identify “wrong” conduct or judgments. In regards to definitions of human error based on the failure to achieve some goal (Isaac & Ruitenberg, 1999) or failure to carry out a planned action (Reason, 1982, 1995), a caveat appears in dynamic environments, where goals may not be realized or goals may be shifting or planned actions are not appropriate in certain circumstances. Consider a scenario in which a set of planned actions are being executed to achieve a goal, when an unexpected event causes the planned actions to be modified or aborted. From the vantage point of an outside observer, this scenario may be considered an error, if the goal can no longer be achieved. However, from the vantage point of the operator, this is not an error (Dekker, 2005). Another caveat is that errors are often identified by the adverse consequences of an inaccurate behavior or judgment (e.g., incidents and accidents). However, adverse consequences do not always consistently and reliably occur following a behavior, or group of behaviors. Adverse consequences may be slow to manifest, in which case they are difficult to correlate with system changes or other factors in dynamic environments. Therefore, because no single error definition is comprehensive (Dekker, 2005), it may be prudent to evaluate error in a manner that is as comprehensive as possible.

Table 1. *Definitions of Human Error*

Author, Year, Page	**Definition**
Reason, 1982, pp.14	When a planned action fails to achieve its desired outcome, and when this failure cannot be attributed to the intervention of some chance occurrence
Reason, 1995, pp.81	A failure of planned actions to achieve a desired goal
Isaac & Ruitenberg, 1999, pp. 11	Intended actions which are not correctly executed
Gertman, Hallbert, Parris, Sattision, Brownson & Tortorelli, 2001	Human errors are failures, which are the inability a human or component to perform its function as required.
Pounds & Isaac, 2002	Cognitive failures that lead to incorrectly executed actions
Woods, Johannesen, Cook, & Sarter, 1994	Human error is a specific variety of human performance that is so clearly and significantly substandard and flawed when viewed in retrospect that there is no doubt that it should have been viewed by the practitioner as substandard *at the time the act was committed or omitted.*
American Oxford English Dictionary, 8th Ed.	The state or condition of being wrong in conduct or judgment

One potentially comprehensive method of defining a human error is to assign probabilities to the expression of adverse consequences for each human error factor (i.e., contributing element) or combination of factors. Herein, an *error* is operationally defined as a violation of an explicit rule governing operator performance; e.g., a driver who deviates from a restricted lane. The identification of each human error factor might begin by observing and documenting the variability of normal, daily behaviors and resultant performance. Afterward, probabilities between certain behaviors and adverse consequences may emerge.

2.2. Mind wandering

Mind wandering may be one contributing factor to task-unrelated behavior correlated with error. In this section, we present a definition of mind wandering, following by background information such as who is most susceptible to mind wandering, when is mind wandering most likely to occur, and a few documented consequences of mind wandering.

2.2.1. What is mind wandering?

The synonyms to the term "mind wandering," including task-unrelated thinking, stimulus-independent thought, zone outs, task disengagement, self-interruptions and mind pops, are not well-defined or consistently defined across prior research. For example, *mind pops* (Kvavilashvili & Mandler, 2004) are thoughts that spontaneously arise unconsciously or semi-consciously while stimulus-independent thought could be consciously identified. *Self-interruptions*, also called *discretionary task interleaving* and which is part of the multitasking domain, implies deliberate distracted thinking before a task is complete, yet the boundaries between the multi-tasking domain and the mind wandering domain is unclear. Perhaps the boundary is a function of the degree of deliberate distraction. *Task-unrelated thinking*, another term involving task execution, is an experience where the content of a person's current thought is not associated with the primary task (or task-related thinking) but rather is composed of other unrelated thoughts, fantasies, feelings and musings (Smallwood, 2006). Alternatively, *mind wandering* is not always defined with respect to its relationship to a task. Smallwood and Schooler (2006) explained that mind wandering is the process by which our attention is decoupled from the immediate task context, which makes us become absent-minded. On the contrary, William James (1890) suggests that mind wandering occurs when the individual experiences a meandering string of thoughts, some of which may not be influenced the exogenous world of stimuli at a given moment in time. Thus, an individual can mind-wander through a series of thoughts that may or may not be related to events occurring in the task environment. For the purposes of this paper, we use the terms *mind wandering* and *task unrelated thinking* interchangeably to represent a subjective experience of thinking off-task thoughts, regardless of whether the thinking was deliberate or not, conscious or not, spontaneous or not.

A few properties of task-unrelated thinking have been identified in the literature; some of which are inherent in the definitions above. First, the change in thought content has been characterized as either abrupt, wherein unrelated thoughts spontaneously "pop" into the mind (Kvavilashvili & Mandler, 2004), or more gradual, wherein thoughts gradually meander off-task (James, 1890; Smallwood, Baracaia, et al., 2003). Also, task-unrelated thinking can be either voluntarily- or involuntarily-initiated (Forster Lavie 1999; Smallwood & Schooler, 2006) and at the time that the unrelated thought is initiated, the person can be aware or unaware that their thinking has deviated off-task (Schooler, 2002). When the individual is unaware of off-task thinking, a phenomenon known as temporal dissociation occurs, wherein the person is said to lack meta-awareness (Schooler, 2002) or the awareness of one's cognitive state. When meta-awareness diminishes, consciousness of the current task's status may be diminished as well as knowledge of task performance (Schooler, 2002).

2.2.2. When is mind-wandering likely to occur?

A few explanations about why people mind wander have been provided in the literature. Some explanations are task-specific, in which the explanation is related to the task type, while others are related to individual differences of the operator, such as mood or working memory capacity. Research on reading and problem solving tasks has provided some insight into why people mind wander. For example during reading tasks, those texts that were interesting to the readers facilitate "zone-outs" because the text is thought provoking while uninteresting texts facilitate attentional shifts to more interesting topics (Schooler, Reichel, & Halpern, 2005). During problem solving tasks, mind wandering may serve as an incubation period for solutions to be brought into conscious awareness (Anderson, 1975 and 1985; Bastik, 1982; Cattell, 1971; Giambra, 1995; Helmholtz, 1896; Smith, 1995).

Some of the explanations related to an operator's individual differences involve negative mood, attentional disorders and working memory. Individuals with mood and attentional disorders are more likely to experience episodes of off-task thinking (Giambra, Grodsky, Belongie, & Rosenberg, 1994–1995; Kane, Brown, McVay, Silvia, Myin-Germeys, and Kwapi 2007; Shaw & Giambra, 1993). In particular, those suffering from dysphoria, obsessive compulsiveness, and neuroticism tend to engage more frequently with task-unrelated thinking than individuals with non-negative affect (Eysenck & Byrne, 1992; Eysenck & Graydon, 1989; Smallwood, Davies, Heim, Finnigan, Sudberry, O'Connor, and Obonsawin, 2004). Those individuals with a negative mood may mind wander to more pleasant topics to alleviate negative feelings and to serve as a coping mechanism (Binder et al., 1999; Greenwald & Harder, 1995, 1997). In addition, an individual's working memory capacity may also be related to the frequency of mind wandering. In one dual task study, wherein two tasks are performed simultaneously, the investigators found that if the primary task demanded high levels of cognitive resources, then few resources were available for mind wandering; thus, the frequency of mind wandering was reduced (Teasdale, Dritschel, Taylor, Proctor, Lloyd, Nimmo-Smith, 1995). These findings paralleled other work that identified an inverse relationship between cognitive resource demand and the frequency of task-unrelated thinking (Forster & Lavie, 2009; Giambra, 1995; Teasdale, Dritschel, Taylor, Proctor, Lloyd, Nimmo-Smith, 1995; Teasdale, Proctor, Lloyd, & Baddeley, 1993; Smallwood & Schooler, 2006). This relationship may be mitigated by an individual's working memory capacity such that when an individual with a high working memory capacity is faced with a large task load that consumes cognitive resources, more task-unrelated thinking occurs compared to low working memory capacity individuals (Kane, Brown, McVay, Silvia, Myin-Germeys, & Kwapi, 2007). This may be because high working memory capacity individuals may have extra cognitive resources available to mind wander when the task load is high, as compared to low working memory individuals with the same task load. However, it is unclear whether we would observe any differences in mind wandering frequencies among high and low working memory capacity people given the same task load.

2.2.3. Task-unrelated thinking and task type

There are some indications that mind wandering occurs when the individual is performing certain task types. First, mind wandering may occur more often when performing tasks which require high levels of engagement or attention, called vigilance, which has been observed during reading tasks (Smallwood, Fishman, & Schooler, 2007). During vigilance tasks, the frequency of task-unrelated thinking ranges from 30-55% of thought content probes (Giambra, 1995; Smallwood, Davies, et al., 2004; Smallwood, O'Connor, Sudberry, Haskell,

& Ballantyne, 2004; McVay & Kane, 2009, Schooler, Reichle, & Halpern, 2005). In studies where individuals were presented with targets for their attention, this effect was pronounced when the target presentation rate was comparatively slow (Antrobus, 1968; Grodsky & Giambra, 1991; Teasdale et al., 1995). These findings involved those individuals whose thinking involves meta-awareness (Schooler & Schreiber, 2004) which is awareness of the contents of conscious thoughts. Other studies found that during reading tasks, mind wandering occurred on average of 1.6 probes out of a total of 6 probes per study and during 13% of those 1.6 probes, the participant lacked meta-awareness (Grodsky & Giambra, 1991). In addition, mind wandering occurrences may be related to certain task characteristics. For example, repetitive, familiar, long-duration, unvarying, somewhat cognitively undemanding tasks tend to be related to higher frequencies of mind wandering (Smallwood & Schooler, 2006; Antrobus, 1968; Grodsky & Giambra, 1991; Teasdale et al., 1995). Finally, automated tasks, which require little conscious effort and attention to execute, have also been correlated with higher frequencies of task-unrelated thinking compared to tasks that are not automated (Klinger, 1977; Smallwood, Obonsawin, Reid, 2003).

2.2.4. Consequences of Task-unrelated thinking

Cognitive processing and behavioral responses may be adversely affected by task-unrelated thinking. In a variety of studies, mind wandering may adversely affect cognitive processes such as attentional focus, stimuli encoding (i.e., converting the stimuli information into a memory), memory retrieval and mental model building (i.e., building a mental representation of a concept). During task-unrelated thinking, one theory is that less attention is available to dedicate to the contents of the exogenous environment. Thus, attention to the exogenous world becomes decoupled from attention to conscious thought (Klinger, 1999; Smallwood, Baraciaia, Lowe, & Obonsawin, 2003; Smallwood, Davies, Heim, Finnigan, Sudberry, O'Connor, & Obonsawin, 2003; Smallwood, Fishman et al., 2007). This may be particularly true in those individuals with low working memory capacity (McVay & Kane, 2009). As a consequence of this decoupled state, the individual may develop a superficial representation of the environment (Smallwood & Schooler, 2006). For example, mind wandering seems to reduce the encoding accuracy of task relevant information (Seibert & Ellis, 1991; Smallwood et al., 2003; Smallwood, Fishman, et al., 2007; Smallwood, McSpadden, et al., 2008; Smallwood, Obonsawin, & Heim, 2003). As a result, task-relevant memories are inaccurate (Smallwood, Baraciaia, Lowe, & Obonsawin, 2003), which can compromise mental model building (Smallwood, Fishman, et al., 2007).

Task-unrelated thinking may also adversely affect the execution of task-relevant actions. Off-task thinking has been related to response delays (Cheyne, Carriere, & Smilek, 2006; Smallwood, McSpadden et al., 2007), as well as higher rates of incorrect response execution (Smallwood, McSpadden et al., 2007). Finally, during mind wandering, an increase in the number of action slips is observed, which is the execution of the wrong behavior when performing an automated task (Broadbent & Cooper, 1982), as well as an increase in the number of errors of commission, which are the execution of the wrong behavior when the task requires controlled processing (Cheyne, Carriere, Smilek, 2006; Helton, Kern, & Walker, 2009).

2.2.5. **Mind wandering and the operation of vehicles**

Mind wandering has the potential to affect operators of various types of manned and unmanned vehicles, in which the operator must control the vehicle within the bounds of established rules of operation. For example, drivers must obey the laws of driving in their jurisdiction, and aircraft pilots need to obey national laws governing air traffic. The occurrence of mind wandering, to my knowledge, has never been reported in research publications about task engagement in the cockpit flight deck; however, there is evidence that pilots do mind-wander. For example, Reason and Mycielska (1982) identified 16 different aviation accidents between the years 1955 and 1972 where the causal contributor may have been "absent-minded behaviors." The most common absent-minded behavior, called a slip, occurs when the unintentional execution of habitual steps in a task or omissions of steps in a task that is committed to habit. Because these 16 incidents were 3rd party retrospective accounts of the accidents, the occurrence of mind wandering cannot conclusively be ascertained. In addition, as of July 18, 2010, the Aviation Safety Reporting System (ASRS) database contained 7 reports of pilots who mind wandered in the cockpit (See Appendix A). Therefore, it is probable that mind wandering occurs in the cockpit flight deck during normal flight operations, but without a standard metric by which to measure mind wandering, we may find this phenomenon is underreported in the case of human error.

Much of the literature related to mind wandering while driving an automobile is categorized as "distracted driving" which commonly involves the use of an electronic device as a mechanism for distracted thinking (Hanowski, Perez & Dingus, 2005; Heck & Carlos, 2008; Jacobson & Gostin, 2010; Young & Regan, 2007). It has been shown that distracted driving, with or without an electronic device, results in performance decrements (Lee, Caven, Haake & Brown, 2001; Horrey & Wickens, 2006; Ishigami & Klein, 2009; Strayer & Drews, 2007; Strayer, Drews & Johnston, 2003). However, recently, researchers began making the distinction between distracted driving and mind wandering when driving and respective performance implications (Glaze & Ellis, 2003; He, Becic, Lee, & McCarley, 2011; Stutts, Reinfurt, Staplin, & Rodgman, 2001).

3. Research Overview

This research uses a grounded approach to studying mind wandering where hypotheses formed in one domain are applied to another. This research begins with the study of automobile drivers to understand how to best measure mind wandering and ends with the application of those identified measures to aviation pilots. The target domain is aviation because, compared to driving, flight task sequences are more standardized than driving task sequences.

The research goal is to study the relationship between mind wandering, task execution and errors for operator tasks involving manned vehicles; specifically flying and driving. This research effort is divided into four parts: each subsequent part builds upon the previous part. Part I is a factor scenario survey that serves to determine how aircraft pilots define mind wandering and where in a flight sequence they feel mind wandering is most likely to occur. These findings provide information needed for a study involving mind wandering pilots described in Part IV. Part II is a task analysis of a driving simulation that serves to identify workload fluctuations of automobile drivers in a standardized driving simulation route. Interviewing probes were established at physical locations in this driving simulation that target potential subjective experiences of varying workload. These probes are used in the driving simulation study in Part III; this study's purpose is to identify the most effective metric to measure mind wandering and to study the relationship between task type, task load, operator dysphoria, operator working memory capacity, mind wandering and adverse consequences. Part IV is a flight simulation study that employs the metric identified in Part III to study mind wandering with general aviation pilots.

The remainder of this chapter begins with an overview of the research domain, followed by the four parts described above. Each part introduces the research questions and/or hypotheses addressed, experimental design, results and discussion. An overall discussion of the study results follows the presentation of these four parts.

3.1.Research domain

Mind wandering can affect expert operators in a number of domains, such as transportation, energy and health care. Herein, the focus is predominantly on-tasks performed by aviation pilots and automobile drivers. In aviation, commercial pilots have a very tightly controlled environment in which policies and training aim to minimize variance in the flight task to which mind wandering could be a potential cause of variance. However, significant recruiting resources are needed to hire commercial aviation pilots, so general aviation pilots were studied instead. Whereas commercial pilots fly scheduled routes for an airline, general aviation pilots typically fly smaller planes on unscheduled routes. In addition, we are interested in the how well the findings from one sub-domain (e.g., aviation) generalize to another sub-domain (e.g., driving). Thus, we chose to study the driving simulation environment to evaluate existing research findings and calibrate our methodology. Automobile driving was chosen because the task is well-practiced among U.S. adults, participants are plentiful and easily accessible on a University campus, and participants are relatively inexpensive to recruit compared to pilots.

Our ambition was to study a job whose task performance variability is not due to lack of training and novelties in the task environment, whose task load varied, and whose task sequence was fairly standard such that, if a mind wandering event occurred, then the task performance variability could potentially be mapped to that mind wandering event. The aviation domain was the target domain to study mind wandering because the task set across flights is standardized, well-trained and repeatable in relatively the same task sequence. Because the frequency of mind wandering can potentially be mitigated with novel task situations, we chose a more stable task sequence so mind wandering would be impacted by the task parameters and not task novelty. Unlike some operators in other domains, pilots are highly-trained to fly in normal and abnormal conditions. Therefore, the variability of performance is likely not due to a lack of training. In addition, these pilots must follow policies issued by the Federal Aviation Administration, called Flight Aviation Regulations, and issued by their respective airlines in the flight operations manuals, regarding how to fly the aircraft. These policies potentially restrict the variability of task performance, which is easier to study than systems with highly variable task performance. Unlike some domains, the cockpit flight deck has a rich variety of tasks such as monitoring instrumentation, searching for aircraft conflicts in the airspace, performing quick mental calculations, and coordination with air traffic control, ground crew, co-pilot, etc. In addition, pilots perform the same sets of tasks each flight; however, the task sequencing may not always be repeated. One caveat is that, because the system is dynamic and susceptible to perturbations, repeating task sequencing may not always be possible and, to compensate, pilots may multi-task (a potential increase in task load) to complete all tasks in the time allotted (Loukopoulos, Dismukes, & Barshi, 2009). Therefore, this domain offers a rich variety of well-known, repeated tasks to not only study incidents of mind wandering, but to also observe the variance of pilot behaviors.

3.2.Research Questions and Experimental Hypotheses

A mixed method research approach, which was exploratory and experimental, was used to create theory and to test prior research findings from other domains in the context of the aviation domain. Exploratory methods were used to develop the theoretical construct of mind wandering, to categorize domain-specific tasks in a taxonomy, to identify when mind wandering occurs and to examine potential relationships at a given point in time between the content of task-unrelated thoughts, the task type, and the proximal resultant consequences of mind wandering. Experimental methods were used to evaluate the efficacy of different mind wandering metrics for the cockpit environment and the relationship between working memory capacity, dysphoria, task type, task load, and mind wandering. The exploratory research questions will be discussed in detail first, followed by the experimental hypotheses.

3.2.1. **Exploratory Research Questions**

The exploratory research sought to investigate the following research objectives: 1) how is task unrelated thinking defined; 2) what is the relationship amongst task-unrelated thinking, human error and adverse consequences; and 3) what is the efficacy of mind wandering metrics? We now review important supporting literature before discussing the exploratory study objectives and specific research questions that we defined to investigate these objectives.

Task-unrelated thinking is defined as an experience where the content of a person's current thought is not task-related but rather is composed of other unrelated thoughts, fantasies, feelings and musings (Smallwood, 2006). However, we lack any measurable definition of relatedness between a two topics of thought. Thoughts can be related on many dimensions, including:

- *Temporal relatedness* occurs when two thoughts are about two events that occur relatively close in time. *Tense* is the directionality of the temporal quality of a thought (e.g., retrospective or prospective).
- *Functional relatedness* occurs when one thought is built from another.
- *Conceptual relatedness* occurs when two thoughts belong to the same construct.

For example, suppose a driver is thinking about how long it will take to get home from the next intersection. Suddenly, the driver's thinking is interrupted by a memory to pick up his/her dry cleaning at the shopping center at the next intersection. The thought about how long it takes to get home and the thought related to the dry cleaning have some degree of temporal relatedness because the person is thinking about where they are in space and time and where they need to be in the next prospective (i.e., tense) moment. In contrast, these two thoughts have minimal functional or conceptual relatedness. An example of conceptual and functional relatedness may occur when a pilot is thinking about the before-start checklist that he/she is executing when thoughts about requesting take-off clearance from air traffic control (ATC) "pop" into consciousness. Thoughts regarding both the checklist and obtaining ATC clearance in the future are less temporally related than the previous example because there is a larger time interval between these two events. However, the thoughts about the checklist and ATC clearance are more conceptually related than the previous example, because they both involve the task of flying the aircraft. In addition, these two thoughts have some degree of functional relatedness, because the checklist must be done before requesting clearance for takeoff, so requesting clearances is functionally dependent upon the completion of the checklist.

Another dimension of interest is the *tense* (e.g., prospective, retrospective and no tense) of the off-task thought compared to the present moment when the thought occurred. For example, if a person is experiencing an off-task thought about an ‘event’ in the future (i.e., an episodic memory), the tense is prospective. However, some abstractions (e.g., social justice, college, etc.) are not about a particular point in time but may give rise to the subjective ‘feeling’ of tense. For example, a person may see a sunset that reminds them of the generic concept of the beach. When probed further, this individual has no associations of episodic memories with the beach but the beach has an association with the past or future. Other individuals may have a different experience. When probed further about the beach, these individuals may start naming a series of episodic memory associations, triggered by the sunset that occurred at a variety of time points in the past or future. By association, the abstraction may automatically acquire a tense due to its close proximity to an array of other episodic memories. To be precise, other individuals may explain that the abstraction recalled has no temporal quality at all.

The number of ways two thoughts could be related is not well studied in the literature. Only one study classified relatedness on multiple dimensions (temporal, purpose, object, movement and location), but this relatedness was between an intended action and an actual action rather than between two thoughts (Reason & Mycielska, 1982).

Before any research can be executed on exactly when mind wandering occurs, we need terminology to define mind wandering. A universally agreed-upon definition of mind wandering may not exist amongst study participants; however, we may approximate a definition by developing and testing terminological criteria (e.g., degree of relatedness between thoughts and task) for including and excluding thought experiences as types of mind wandering. These thought experiences arise while the current task is being executed and some degree of thought unrelatedness to the task must be experienced before it is perceived as mind wandering. Three types of relatedness will be studied: temporal, functional and conceptual. To give directionality to temporal relatedness, we also studied the notion of thought tense, i.e., prospective, retrospective or no tense. Thus, the first research objective is to study degree of temporal, functional and conceptual relatedness required before it is classified as mind wandering.

The second research objective is to determine what task types and cognitive requirements (e.g., cognitive skills, abilities, mental processes) are significantly correlated with different types of mind wandering. Task analytic approaches (e.g., job task analysis, cognitive task analysis, hierarchical task analysis) to decomposing tasks into subtasks and their respective cognitive requirements are widely-used techniques within the social sciences (Brannick, Levine & Morgeson, 2007; Morgeson & Dierdorff, 2011; Sanchez & Levine 1999). However, we believe one of the problems with published task analytic approaches is the lack of standardized terminology needed for naming task types and their respective cognitive requirements. To circumvent this problem, we used the US Department of Labor's O*Net™ (http://www.onetonline.org/) site which publishes a standardized list of jobs and their respective task types, task names, task cognitive requirements, etc. for multifarious job roles in the United States, including automobile drivers and aircraft pilots. In addition, O*Net™ established a taxonomy used to decompose jobs into a hierarchy of *activities, sub-activities* and *tasks* (Ammerman, Becker, Jones, & Tobey, 1987). In this taxonomy, multiple tasks comprise a sub-activity and 5-9 sub-activities comprise an activity. While definitions of activities and sub-activities are not provided, a task is defined as a "…'concise, specific statement of one purposeful job action of a worker, generally performed by one individual within some limited period of time… tasks are intermediate in specificity between the high level activities/sub-activities and the procedural steps and actions by which they are accomplished" (Ammerman, Becker, Jones, & Tobey, 1987, pp. 3-12).

O*Net™ is a venue for publishing the results of a series of standardized job task analyses (JTAs) conducted by researchers on behalf of the Department of Labor. The O*Net™ website provides not only the results of JTAs conducted on a wide variety of job roles, but it also provides a standardized list of defined knowledge, skills, mental processes and cognitive ability (KSMAs) requirements for each job role. During each JTA conducted by O*Net™ researchers, participants were asked to rate the importance level of identified KSMAs with respect to the job role, and researchers further tabulated the mean importance ratings (listed as a percentage) for all KSMAs. Only those KSMAs with a mean importance rating of 50% or higher were by default listed on the website (http://www.onetcenter.org/dataCollection.html); however, a viewer can adjust the site's filter settings to view lower mean-rated KSMAs. The participant sampling is not explicit, but the website does mention that ratings were conducted repeatedly on samples of job operators to build reliability and validity into importance ratings. In order to post the results of a job role on this site, a minimum of 15 returned questionnaires was required, but sample sizes by job roles are not provided (http://www.onetcenter.org/dl_files/omb2011/AppendixF.pdf).

Research Question 1 focuses on the relation between the type of mind wandering experienced (e.g., temporal, functional, conceptual and tense) and the task being performed,: Are task type with a particular set of characteristics and mind wandering related? If so, are task type and the type of mind wandering related?

From the distracted driving literature, we may glean that adverse consequences result from device-dependent or device-independent distracted driving (Strayer & Drews, 2007; Strayer, Drews & Johnston, 2003). This result leads us to believe that adverse consequences may also occur with mind wandering. One explanation based on human information processing models is that task-unrelated thoughts can be a type of endogenously generated stimuli (e.g., the recall of a fantasy, musing or memory) that competes in the person's sensory system with other representations of exogenous stimuli (e.g., visual information about task performance) for processing resources. If the endogenous stimuli or thoughts interfere with the processing of the exogenous stimuli in working memory, then behavioral errors in response to exogenous stimuli may occur and adverse consequences may result. For example, an operator might be thinking about a change in a corporate scheduling policy when a visual warning appears. The operator may be attending to their own thoughts about this policy leaving little attentional resources left to notice the warning and the system experiences an adverse consequence or failure. There is some support for this phenomenon in the distracted driving literature; for example, distractions can reduce the detection and recognition of critical events (Strayer & Drews, 2007; Strayer, Drews & Johnston, 2003) and they are correlated with inappropriately delayed responses to critical events (Lee, Caven, Haake, & Brown, 2001; Strayer & Johnston, 2001). Thus, **Research Question 2** concerns the relationship between mind wandering, behavioral responses or non-responses, and adverse consequences within close temporal proximity to the wandering: Are certain types of mind wandering related to certain behaviors and adverse consequences?

The third exploratory research objective is to identify a mind wandering metric that is efficacious, which means the metric that participants believed was least disruptive and most accurate at capturing the frequency of experienced mind wandering occurrences. A mind wandering metric is a technique used to measure the occurrence, characteristics and contents of a mind-wandering event. One metric published in mind wandering research is called the *probe technique*, also called the probe-caught method, in which the interviewer interrupts, or *probes*, the participant's task execution to solicit the content of that person's thoughts (Antrobus, 1968; Giambra, 1995; Schooler, Reichle, & Halpern, 2005; Smallwood, Baracaia, et al., 2003; Smallwood, Davies, et al., 2004; Smallwood, Obonsawin, & Heim, 2003; Smallwood, O'Connor, et al., 2004; Teasdale, Dritschell, et al., 1995). A probe is not a random system-generated interruption; it's an event whose timing is predetermined and executed by the investigator. Probe techniques are typically used to learn of the participant's thought content at a given point in task execution as well as used to assess the level of conscious awareness of that participant's thinking. People who mind wander with little or no conscious awareness of the wandering are said to have mind wanderings that lack *meta-*

awareness. There are two types of probe techniques in the literature: the self-classification probe method (Teasdale, Dritschell, et al., 1995; Teasdale et al., 1993) and the experimenter-classification probe method (Smallwood, Obonsawin, & Reid, 2003). The *self-classification probe method* trains participants to self-categorize their own thought content according to a classification provided by the investigator. In contrast, the *experimenter-classification probe method* occurs when the investigator exclusively categorizes the participant's thought content of the probe.

Non-probe mind wandering metrics are also used to identify the occurrence and thought content of mind wandering episodes. One such metric, called the *self-caught method* (Cunningham, Scerbo, & Freeman, 2000; Giambra, 1993), allows participants to self-interrupt the task to report when he/she experiences mind wandering. Another non-probe metric involves the *retrospective recall* of mind wandering episodes after task execution is complete (Seibert & Ellis, 1991); no task interruption occurs when using this metric. A small number of studies used a combined probe and self-caught metric to capture mind wanderings of individuals with and without meta-awareness (Schooler, Reichle, & Halpern, 2005). These non-probe mind wandering techniques rarely capture wanderings that lack meta-awareness.

Metrics for any research study are typically selected to minimize measurement error such that the metric is sensitive enough to capture the phenomenon accurately and the occurrence of the phenomenon is not altered by its measurement. For example, following the first probe, the participant may consciously increase awareness of thought content more and, consequently, unnatural frequencies of mind wandering may occur. Thus, the mind wandering metric itself inflates measurement error. Furthermore, when the participant is asked to self-classify mind wandering episodes when using the self-classification probe method, this person may consequently be extra conscious of mind wandering episodes (inflating meta-awareness) than when not asked. Consequently, mind wandering frequencies as well as the occurrence of wanderings with meta-awareness, may be artificially inflated (Smallwood & Schooler, 2006).

Beyond the potential increase of measurement error given certain metrics used, another concern is that these metrics have not been validated in the driving and flight simulators. Therefore, **Research Question 3** aims to assess the most efficacious metric for the cockpit domain given the types of tasks inherent to this environment: What metric or metrics exist for measuring mind wandering in the cockpit that also minimize disruption of a person's train of thought and enhance the ability to recall the probe content?

3.2.2. **Experimental Hypotheses**

We now present relevant terminology before reviewing the study objectives and experimental hypotheses.

Working memory capacity is loosely defined as the maximum amount of information moved into and out of the brief working memory store (Baddely & Hitch, 1974; Miller, Galanter & Pribram, 1960). Previous work in other domains has identified a relationship between the task load or cognitive demands that the task requires, the operator's working memory capacity, and occurrence of mind wandering (McVay & Kane, 2010). Without assessing working memory capacity of participants, prior research shows an inverse relationship between task load and the frequency of mind wandering (Smallwood & Schooler, 2006;

Antrobus, 1968; Grodsky & Giambra, 1991; Teasdale et al., 1995). However, there is some indication that the participant's working memory capacity may influence the relationship between task load and mind wandering frequency (McVay & Kane, 2010). One explanation is that in light of a single pool of cognitive resources (Kahneman, 1973) needed to perform a task, individuals with more resources (i.e., high working memory capacity individuals) also have more resources on reserve during task execution to use for mind wandering as compared to those individuals with less resources (i.e., low working memory capacity). One would expect that higher task loads during task execution would increase the likelihood that low working memory capacity individuals will not mind wander, because residual resources for mind wandering are not available. It might be expected that that for low task loads, no differences in mind wandering frequency would exist, yet as task load increases, mind wandering frequency differences should be observed. Hypothesis 1 aims to assess the expected results in high task load conditions.

Hypothesis 1: Operators with low working memory capacity will mind wander more frequently during high task loads compared to other individuals with normal or high working memory capacity.

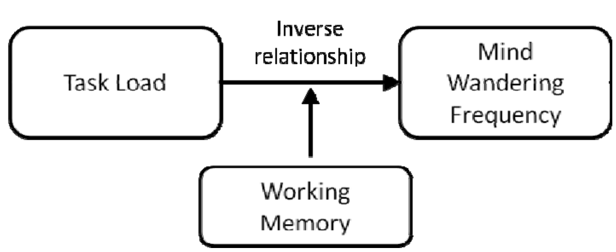


Figure 1. The inverse relationship between task load and mind wandering frequency that is effected by working memory capacity

Another factor influencing the frequency of mind wanderings from prior literature is negative mood, also called dysphoria; however, the reasons for this relationship can only be speculated. Negative mood and low working memory capacity could potentially be correlated. The operator with low working memory capacity may historically be aware of their own performance limitations in cognitively demanding tasks and, thus, may suffer from negative mood. Since melancholy individuals may mind wander to alleviate negative mood (Eysenck & Byrne, 1992; Eysenck & Graydon, 1989; Smallwood, Davies, Heim, Finnigan, Sudberry, O'Connor, and Obonsawin, 2004), individuals with low working memory capacity may also mind wander to alleviate the negative feelings attributed to mental strain in high task load situations. Therefore, we are interested in comparing the mind wandering frequency of low working memory capacity individuals with the mind wandering frequency of depressed individuals. Hypothesis 2 aims to reaffirm these findings.

Hypothesis 2: Dysphoric operators will mind wander more frequently than those who have normal or positive affect.

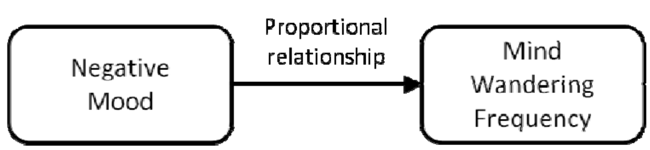


Figure 2. A proportional relationship between negative mood and mind wandering frequency

4. Part I. Factor Scenario Survey

We now present Part I, which describes the factor scenario survey and includes a discussion of the survey purpose, method, procedure, results and discussion respectively.

4.1. Goals

This survey has two goals: 1) to collect aircraft pilot opinions of what factors comprise the relatedness construct and how these opinions aid in operationally defining mind wandering; and 2) to collect pilot speculations about during which flight tasks their fellow pilots are most likely to experience mind wandering. The data collected for goal 2 will be used in Part IV to assist in determining probe locations in a flight simulation.

4.2. Method

We now discuss the method, including the participants, sampling, sample size, sample demographics, and the apparatus and stimuli used in the study.

4.2.1. Participants

Sampling. The target population for this research was fixed-winged aircraft pilots. A list of aviation magazines that this target population would most likely read was compiled using online search engines. Contacts for each magazine were solicited to include a study advertisement in an upcoming issue, but several contacts declined or did not respond. The following magazines advertised the study and our subjects came from their readership: Flight Training magazine, Aeroplane, AINalerts (a newsletter for corporate pilots), Aviation Consumer/AVweb and the Atlantic Flyer. The advertisement listed the study purpose, the risk and benefits to participation, the survey hyperlink and the investigator's contact information.

Sample Size. To encourage survey participation, an Apple iPod was raffled off to a single participant with a completed survey record; however, determining the total sample size was uncertain. A total of approximately 39 out of 114 total participants who attempted the survey provided near-complete surveys; all other participants provided incomplete surveys. When a participant entered the survey, a record was created in the dataset. However, a single record could not be attributed to a single participant's efforts due to the design of the survey software, called SurveyGizmo ™. The survey software addressed this issue by generating multiple record identifiers, such as IP addresses, geo-location data , participation time stamps, and record completion status (e.g., partial or complete), which may be used to infer similar and dissimilar response attempts. A partial record occurred when a subject did not view every page of the survey, whereas a complete record occurred when every page was viewed. However, a complete record does not indicate that every question was responded to. In fact, there were no 'complete' records in which all survey questions were answered. This may have resulted in part because participants were not forced to answer each question. In addition, multiple records from a single participant could not be detected in the dataset. One participant could technically use multiple IP addresses to participate in the survey and one IP address could also serve multiple participants. This is especially true with community computers serving multiple users in, for example, an airport lounge.

The study dataset contains eight records with redundant IP addresses and all IP redundancies had one partial and one complete record. Since participants were encouraged to re-enter the survey if they could not finish within their personal time constraints, partial records could come from early termination of the survey. The total number of responses to the survey invitation was 122, among which we removed eight duplicate IP addresses to yield 114 unique responses. Among the 114 responses, a total of 39 respondents visited all pages of the survey. The remaining 75 respondents were predominantly participants who read the consent form and left the survey. Among these 75 participants, there were participants who answered a few questionnaire items. Some of our planned analyses did not require participants to complete the entire survey; thus, for those analyses, we included responses from the 75 participants who answered the relevant questions.

Sample Demographics. The resultant demographics of this sample are provided in Appendix B but because the sample size was unclear, sample percentages were not calculated. The sample predominantly consists of male, native English speakers, who held private certificates but who were not students working on a certificate at the time. The 34 pilots who flew under Title 14 CFR Part 91 (for general aviation) had a mean total of 1181 flight hours (*SD*=6881 hours), while the 11 pilots flying under Title 14 FAR Part 121 (for commercial cargo operations) had a mean total of 6.8 flight hours (*SD*=23 hours). None of the participants in this survey self-reported any total flight hours under Part 135 (for commercial passenger operations). Other quantitative data, such as the total number of aircraft flown, the average number of days/month flown, etc., were collected, but the standard deviations were larger than the means, so data trends could not be deduced with any certainty; these results are listed in Appendix B. Based on the geographical location of the IP address, all but two participants were located in 42 different US states; one IP address was from Germany and one from Panama. California hosted 15 different IP addresses, which is the most frequent state of IP addresses of survey respondents. However, nationality should not be inferred from this geographical information because American pilots often fly international flights and pilots of different nationalities might be based in the United States.

4.2.2. Apparatus and Stimuli

A browser-based survey design and deployment tool, SurveyGizmo ™ was used to host the survey. There are three main parts to this survey: the factor scenario portion, the flight task portion and the demographics portion.

Factor Scenarios. Factor scenario surveys are a methodology used to measure participant's attitudes and beliefs with respect to a variety of construct factors (Converse& Presser, 1986). This type of survey was employed to collect opinions about the type (i.e., temporal, functional and conceptual relatedness) and level of relatedness between on-task thinking (what the participant should have been thinking about) and off-task thoughts (what the participant was actually thinking about). The benefit of using factor scenarios is that the researcher can create scenarios that reflect the factorial combinations representing a single construct and ask participants to make judgments about that construct based on factorial combinations. Three factors (temporal, functional and conceptual relatedness) with two levels (high and low) were Latin-square counterbalanced across a series of scenarios such that each scenario represents one factorial combination. In addition, the 'tense' of the off-task thought was also counterbalanced across scenarios to provide directionality (retrospective or prospective) for temporal relatedness. Hence, a 2 (temporal: high and low) X 2 (functional: high and low) X 2 (conceptual: high and low) X 2 (tense: retrospective and prospective) factorial design was implemented with one factorial combination per scenario, resulting in a total of 16 unique scenarios.

Each scenario was carefully constructed to clearly describe two thought experiences: one with respect to the driving task and one deviant thought. The first thought was always consistent across all 16 scenarios (e.g., "You are driving home from work when you are detoured onto another route that you are semi-familiar with. At this very moment, your thinking is focused on down shifting your manual transmission into third gear in order to make a detour-directed turn.") followed by a statement that reflected the factorial combination. Example statements for each of the following factor levels are in Table 2. Ratings on temporal, functional and conceptual relatedness were collected as well as an overall relatedness rating and a yes/no judgment on whether this factor scenario was an example of mind wandering. The response scale for each type of relatedness was an 11-point, Likert scale with "not at all related" as zero on the left-most pole and "extremely related" as a 10 on the right-most pole. Judgments of whether or not this scenario represented an example of mind wandering was captured with two "yes" / "no" radio buttons.

All scenarios were loaded into SurveyGizmo™, but to keep the survey length manageable, each participant viewed only four selected scenarios. The survey tool randomly selected one scenario from four different scenario clusters and each cluster was devised such that each participant had an example of high and low levels of at least three of the four factors (temporal, functional, conceptual and tense).

Care was taken to assess the validity of the scenarios' representation of the factorial combinations in a small pilot test before launching the survey publically. Ten graduate and undergraduate students (*n*=10) took the survey scenario and ratings were assessed for whether they reflected the high and low levels of each factor. Modifications were made based on the beta test feedback and the factor scenario portion was adjoined to the remainder of the survey and launched.

Table 2. Example Sentences Representing Each Factor Level in the Factor Scenario Survey

Factor	Level	Example Sentence
Tense	Prospective	"You are now thinking about a behavior you will doing in ______ minutes."
	Retrospective	"You are now thinking about a behavior you just completed ______ minutes."
Temporal	High	"You are now thinking about something that happened 10 minutes ago."
	Low	"You are now thinking about something that happened 1 week ago."
Functional	High	"You are now thinking about how you just grinded your gears." The action of gear shifting in the first thought cued you to think about gear shifting in the second thought.
	Low	"You are thinking about a conversation you had about your children." This has no functional relatedness with gear shifting.
Conceptual	High	"You are now thinking about the last time you drove to your parents house." The concept of driving in the present moment is highly related to the concept of driving in the second thought.
	Low	"You are now thinking about a cruise you plan to take next week." Driving and taking a cruise have no conceptual relatedness.

Flight Tasks. The initial flight task list portion of this survey was generated using a list provided in Loukopoulous, Dismukes & Barshi (2009). The list was reviewed by two retired cargo pilots and one passenger airline pilot and modifications were made to reflect a broader audience of pilots with a variety of certificates even though many of the tasks were appropriate for flying large aircraft with two or more pilots. To organize the list, these flight tasks were categorized into “Before Start Phase,” “Pushback,” Takeoff, Climb, and Cruise,” “Descent and Approach,” and “Landing, Taxi-in and Shutdown” phases. Participants were asked, “During which of these tasks would your colleagues be most susceptible to experience mind wandering?” Respondents selected one of the two response options, “Yes” or “N/A (not applicable to the flight certificate you hold).” A text box was provided at the end of the list to collect any comments or elaborations participants felt compelled to discuss.

Demographics. The demographics portion of this survey collected information about age and gender as well as details about their flight experience (e.g., accumulated flight hours, certificates held, aircraft types flown, etc.).

All portions of the survey were aggregated and the entire survey was beta-tested by one retired cargo pilot and revisions were made for semantic clarity. The survey was reviewed a second time by the investigator, launched and survey advertisements were then posted in the aviation magazines. Because different publications have different circulation dates, the survey was available for 6 months to accommodate publications that are circulated biannually.

4.2.3. **Procedure**

Participants entered the survey using their Internet browser, read the consent form and then completed the survey. If participants did not consent, they were blocked from completing the survey. The survey ended with a small debriefing paragraph, which also provided the investigator’s contact information.

4.3. Analyses and Results

No proposed hypotheses or research questions were addressed in this section. However, a brief review of the results from the factor scenario survey and the flight task checklist data are offered.

<u>*Factor Scenario Results*</u>. Appendix C includes the mean ratings for each type of relatedness for all 16 factor scenarios (*n*~15 participants per scenario). In theory, a well-written factor scenario will yield participant ratings that reflect the scenario factor combinations. For example, if a factor scenario was written to reflect an off-task thought that is retrospective and has a high temporal relatedness (e.g., an event that took place in the distant past), then participant ratings on tense and temporal relatedness should be retrospective and have high ratings of temporal relatedness. For scenario validation purposes, the participants' factor ratings could not be compared to the scenario's intended factor combinations because the total number of responses per scenario was too small for between-subjects assessments. In addition, parametric statistical analyses (e.g., logistic regression, structural equation modeling) could not be used to determine the relationship between factor, their respective levels and the occurrence of mind wandering. Instead, a contingency table logit analysis was used to analyze the relationship between researcher's factor levels for all 16 scenarios and participant's judgment about whether the scenario was an example of mind wandering. The factor levels were not the participant ratings of each factor, but the intended scenario factor levels. A 5-way (2^5) contingency table was created to analyze the relationships between the four factors (e.g., tense, temporal relatedness, functional relatedness and conceptual relatedness) and the dichotomous response to the question, "Was this an instance of mind wandering?" All factors except for tense had two levels; high and low. The two variable levels for tense were retrospective or prospective. See Table 3 for all factor combinations

that comprise the 16 scenarios and the number of participants who judged that scenario as an example of mind wandering.

Table 3. A 5-Way Contingency Table with Four 2-Level Factors (tense, temporal, functional, conceptual) and One Dichotomous Response Variable (total number of "yes" responses)

Tense	Temporal	Functional	Conceptual	Wandering= yes	Total N	%
Retrospective	High	High	High	8	13	62
Retrospective	High	Low	High	10	15	67
Retrospective	High	High	Low	5	15	33
Retrospective	High	Low	Low	11	15	73
Retrospective	Low	High	High	10	13	77
Retrospective	Low	Low	High	3	4	75
Retrospective	Low	High	Low	11	13	85
Retrospective	Low	Low	Low	13	15	87
Prospective	High	High	High	1	15	7
Prospective	High	Low	High	9	13	69
Prospective	High	High	Low	1	4	25
Prospective	High	Low	Low	13	13	100
Prospective	Low	High	High	10	15	67
Prospective	Low	Low	High	8	13	62
Prospective	Low	High	Low	2	4	50
Prospective	Low	Low	Low	12	12	100

All main effects and 2-way interactions were first loaded into the model to evaluate the model fit (i.e., the maximum likelihood estimates); 3- and 4-way interactions were not considered due to interpretation difficulties of results. Three 2-way interactions were non-significant (i.e., tense*temporal, tense*conceptual and temporal*conceptual) and removed from the model to improve model fit. The deviance model fit was $\chi^2(8)=11.89$, $p=.16$, indicating a good model fit. The resultant likelihood ratio chi-square test for model fit was $\chi^2(7)=49.63$ $p<.0001$ indicating that the model with all explanatory variables loaded was a better fitted model compared to a model without explanatory variables (intercept only) loaded. A significant likelihood ratio chi-square indicates that the model with explanatory variables is significantly better than a model with no explanatory variables at *explaining* the variance of the outcome variable. The Likelihood ratio Chi-Square statistics was chosen for model fit assessments, instead of the Wald Chi-Square statistics, because it is comparatively superior with small sample sizes and/or extreme data patterns (Jennings, 1986).

The estimates, standard errors, chi-square values and *p*-values for all main effects and statistically significant interactions of the resultant model are presented in Table 4. The null hypothesis for the estimates is that the coefficient is equal to zero; thus, a significant chi-square test indicates that the estimate is significantly different from zero. The main effect of conceptual relatedness as well as all three interaction terms were statistically significant; however, the results are often difficult to interpret. The significant main effect of conceptual relatedness ($p=.0069$) indicates that given a linear relationship between all the explanatory

Table 4. Main Effects and Interactions of 5-Way Contingency Table Analysis (n=17)

	Estimate	*SE*	χ^2	*p*-value
Intercept	1.8753	0.5629	11.0995	0.0009
tense	0.6348	0.5353	1.4066	0.2356
temporal	-0.0496	0.5385	0.0085	0.9266
functional	-0.3418	0.7629	0.2007	0.6541
conceptual	-1.5013	0.5556	7.301	0.0069
tense*functional	-2.1428	0.7643	7.8611	0.0051
temporal*functional	-1.9472	0.7404	6.9173	0.0085
functional*conceptual	1.8721	0.7627	6.0247	0.0141

variables and the outcome variable, for every 1 unit increase in conceptual relatedness, the predicted probability of participant rating it as *not* an example of mind wandering increases by 1.5 (β= -1.5). Stated another way, those scenarios with the low conceptual relatedness factor (i.e., the thought was not conceptually related to the task) were 50% higher odds to be classified as an example of mind wandering than scenarios with high conceptual relatedness. To assist in the interpretation of estimates for the significant interactions, the mean ratings of the outcome variable for all two-way combinations for these significant interactions are listed in Table 5. A score of 0 on the outcome variable indicates that the participant did not believe this scenario was an example of mind wandering and a score of 1 indicates that it was an example of mind wandering. The odds ratio estimates and the Wald confidence intervals of the three significant interaction terms (tense*functional, temporal*functional and functional*conceptual) are listed in Table 6. An odds ratio is statistically significant if the estimate has a confidence limit that does not include the value 1. Each statistically significant interaction term will be interpreted separately below.

*Tense*Functional Interaction.* The estimate for prospective (tense=1), high functional relatedness thoughts was 0.221 with Wald 95% Confidence Limits between 0.076 and 0.645. To interpret, the odds of retrospective thoughts with low functional relatedness to be examples of mind wandering were 4.5 times higher (=1/.221) than prospective thoughts with high functional relatedness of being considered examples of mind wandering.

*Functional*Conceptual Interaction.* The estimates for high functional, low conceptual relatedness were 0.093 (Wald 95% Confidence Limits were 0.029 to 0.299) while the estimate for high conceptual relatedness, low functional relatedness was 0.223 (Wald 95% Confidence Limits were 0.075 to 0.662). To interpret, the odds that low functional/high conceptual thoughts would be considered examples of mind wandering by participants were 10.7 times higher (=1/.093) than high functional/low conceptual thoughts. Also, the odds of low conceptual/ high functional relatedness thoughts as examples of mind wandering were 4.5 times higher (=1/.223) than thoughts involving high conceptual/low functional relatedness.

Table 5. Two-way aggregate factor combinations and their mean % of "yes this is an example of mind wandering" responses

Factor Combinations	Factor Levels	Mean % Mind wandering=Yes
Tense X Functional	Prospective X Low	0.83
	Retrospective X Low	0.76
	Retrospective X High	0.64
	Prospective X High	0.37
Functional X Conceptual	Low X Low	0.90
	Low X High	0.68
	High X High	0.53
	High X Low	0.48
Temporal X Functional	Low X Low	0.81
	High X Low	0.77
	Low X High	0.70
	High X High	0.32

*Functional * Temporal Interaction.* The estimate for high temporal/high functional thoughts was 0.136 with Wald 95% Confidence Limits between 0.050 and 0.368. To interpret, the odds that thoughts with low temporal/low functional relatedness are considered examples of mind wandering were 7.4 times (=1/.136) higher than thoughts with high temporal/high functional relatedness.

Table 6. Odds Ratio Estimates and Respective Wald Confidence Intervals of Logistic Regression Analysis Interaction Terms: Temporal, Functional, Conceptual and Tense

		95% Confidence Limits	
	Estimate	Low bound	High bound
Tense with low functional relatedness	1.89	0.66	5.39
Tense with high functional relatedness	0.22	0.08	0.65
Temporal with low functional relatedness	0.95	0.33	2.73
Temporal with high functional relatedness	0.14	0.05	0.37
Conceptual with low functional relatedness	0.22	0.08	0.66
Conceptual with high functional relatedness	1.45	0.52	4.04

Flight Task List Results. The flight tasks were clustered according to phases of flight and all frequencies of "yes, pilots experience mind wandering during this task sequence" are displayed in a series of bar charts in appendix D. Frequencies of yes responses for each flight task were rank ordered and Table 7 shows the top and bottom 10 ranked flight tasks and their respective frequencies. Most of the top 10 tasks involve more passive verbs like 'monitoring' and 'review' while those with the lowest frequencies involve action verbs like 'begin,' 'request,' 'obtain,' etc.

Table 7. Top and Bottom 10 Ranks of Flight Tasks During Which Pilots Experienced Mind Wandering

Rank	Phase of Flight	Flight Task	Count
Top 10	Takeoff, Climb and Cruise	Monitor ATC Communications	32
	Takeoff, Climb and Cruise	Monitor Weather/Update Weather	26
	Taxi Phase	Monitor ATC Ground	26
	Takeoff, Climb and Cruise	Monitor Dispatch Communications	23
	Before Start Phase	Complete Pre-Flight Inspection of Aircraft	22
	Descent and Approach Phase	Monitor ATC	21
	Before Start Phase	Review Paperwork	21
	Pushback Phase	Monitor the interphone	16
	Landing, Taxi-in and Shutdown	Monitor ATC Ground Communications	17
	Before Start Phase	Prepare/Review Chart	17
Bottom 10	Before Start Phase	Obtain clearance	2
	Pushback Phase	Request clearance from ATC-Ground or Clearance Delivery	3
	Pushback Phase	Obtain engine start clearance	3
	Pushback Phase	Obtain pushback clearance	3
	Takeoff, Climb and Cruise	Call out speeds/monitor speeds	3
	Takeoff, Climb and Cruise	Participate in the pre-takeoff checklist	4
	Takeoff, Climb and Cruise	Begin or monitor take off roll	4
	Takeoff, Climb and Cruise	Participate takeoff procedure	4
	Descent and Approach	Monitor speedback	4
	Descent and Approach	Monitor or execute landing procedure	4

4.4.Discussion

The results of this factor scenario survey serves two purposes: first, to understand how aircraft pilots classify mind wanderings on four different dimensions, and second, to identify when in the flight task sequences, mind wandering was most likely and least likely to occur. Three types of relatedness dimensions were assessed (i.e., temporal, functional and conceptual) as well as whether the off-task thought was about prospective or retrospective events (i.e., tense). The only statistically significant main effect was conceptual relatedness; low conceptual relatedness between on-task and off-task thoughts was significantly more likely to be judged as examples of mind wandering. For example, if the participant should be thinking about landing an aircraft and they were actually thinking about Christmas trees, landing and Christmas trees are not of the same concept.

Three significant interactions were also revealed that involve the "functional relatedness" dimension. First, low functional/retrospective thoughts were more likely to be considered "mind wanderings" than high functional/prospective thoughts. For example, if the participant was getting ready to downshift a manual transmission and thoughts arose about a vacation cruise taken a year ago (low functional, retrospective thought), this was more likely to be rated as mind wandering compared to thoughts about shifting gears at work on factory machines in the future (high functional, prospective thought). Second, low functional/ low temporal relatedness thoughts were considered mind wandering more often than high functional/high temporal relatedness. Thus, if participants were having thoughts about events in the distant past or distant future (low temporal relatedness) that were not functionally related to driving in the present moment (e.g., recollections of your first day of school while downshifting the manual transmission in the automobile), then these were more likely to be judged as examples of mind wandering than high functional thoughts about events closer in time (e.g., thinking about when you downshifted the gears on your bicycle workout that morning at the same time while downshift your automobile gears). Third, low conceptual, high functional thoughts were more likely to be classified as mind wandering than high conceptual, low functional For example, if the participant should be thinking about downshifting a manual transmission to make a turn ahead and the participant was actually thinking about something with low conceptual/high functional relatedness, such as shifting gears on a bicycle ride, then this type of thinking was more likely to be identified as examples of mind wandering than high conceptual/low functional thoughts, such as thinking about a conversation you had with your teen about the driver's education teacher.

In summary, conceptual relatedness has a consistent impact on ratings. Whenever it is low conceptual, it is consistently classified as mind wandering. However, functional relatedness in two-way interactions also has an impact, but it is not as prominent as conceptual relatedness. Low functional relatedness paired with tense and temporal relatedness, is typically classified as a mind wandering event but when paired with conceptual, low conceptual trumps functional relatedness such that high functional, low conceptual is more often classified as mind wandering than low functional, high conceptual relatedness.

In addition, most of the fixed-winged pilots mentioned that monitoring-types of flight tasks are more likely to host mind wandering events compared to tasks that are more participatory. Tasks which require active participation (e.g., participating in checklists, obtaining clearances, etc.) tend not to produce mind wandering events in pilots. This finding coincides with existing literature that sustained attention-tasks that involve monitoring and reviewing are more likely to experience vigilance decrements (Davies & Parasuraman, 1982; Parasuraman, 1987). However, it is unclear whether mind wandering correlates with vigilance decrements. Some literature suggests that decrements result from task underloading and participant under-arousal, which are forms of mindlessness (Manly et al., 1999; Nachreiner & Hanecke, 1992; Robertson et al., 1997). On the contrary, other studies indicate that the subjective mental workload is high during these types of monitoring tasks and decrements result from a depletion of attentional resources (Hitchcock, Dember, Warm, Moroney & See, 1999; Matthews, Davies, Westerman & Stammers, 2000). Unknown is what level of cognitive workload is highly correlated with mind wandering events and whether mind wandering alters that workload level. Part II will begin to address these unknowns.

5. Part II. Task Analysis

We now present Part II, which describes the task analysis study, including the study purpose, method (i.e., participants, apparatus and stimuli), procedure, results and discussion.

5.1. Goal

The goal of Part II was to assess how task load fluctuates during simulated driving tasks; tasks that will be used in Part III with participant drivers. One way to approximate task load experienced by participant drivers is to tally the cognitive requirements for each task. Task sequences and their respective cognitive requirements for a simulated driving route were identified using a hybrid task analysis of a traditional Cognitive Task Analysis (CTA) and a Job Task Analysis (JTA). The CTA method provides a serial, hierarchically arranged, step-by-step processes of executing cognitions (e.g., memory downloads, decision making, etc.) and behaviors that comprise work tasks. However, a CTA does not yield a list of cognitive requirements (i.e., cognitive abilities and processes) that map onto these cognitions and behaviors. While some JTA methods do provide similar types of task sequencing information as the CTA, JTAs also provide the required knowledge, skills, mental processes and abilities (KSMAs) for each step in a task sequence. Therefore, the hybrid task analysis offers what cognitions are being used on the human operator-side and what cognitive requirements are required to complete the task. The KSMAs used in this study were borrowed from a standardized list available in the O*Net™ website (www.onetonline.org), a product of the U.S. Department of Labor, which lists tasks and respective KSMAs for each job. In the absence of this O*Net list, an extensive literature search would have to be conducted to locate the definitions of various KSMAs and then a large effort to aggregate these KSMAs under a unifying set of terminology would have needed to be done.

5.2.Method

The method section is divided into three subsections: the participants, apparatus and stimuli, and procedure.

5.2.1. Participants

Traditional task analyses are based on data collected from subject matter experts (SMEs). Identifying SMEs for the driving task is relatively easy because U.S. adults often have substantial driving experience. Evidence suggests that task analyses created from a small sample of SME's lack inter-rater reliability (Brannick & Levine, 2002); consequently, the task analysis was engendered by the principle investigator to scope the work appropriately.

5.2.2. Apparatus and Stimuli

Stimuli. A custom, computer simulated driving route was purchased from the simulation manufacturer to fit the research requirements. The research study in Part III requires route segments that would vary driver mental workload by manipulating the number of steering inputs, speed changes and events per unit of time across time. An *event* is operationally defined as an unexpected obstruction to the lane of travel, such as an unexpected pedestrian crossing or unexpected car entering into the lane of travel. For example, one route segment includes a long straightaway with no steering inputs, no traffic signals, no visual clutter, no events and one speed change which was intended to minimize mental workload. Another example segment designed to elevate mental workload involves dense traffic scenes with visual clutter (e.g., parked cars, oncoming traffic, surrounding buildings, etc.) that require lane changes, and attention to traffic signals and unexpected events. Figure 3 is a diagram of the driving route used to build the task analysis. This is the same driving route that will be applied to the study in Part III.

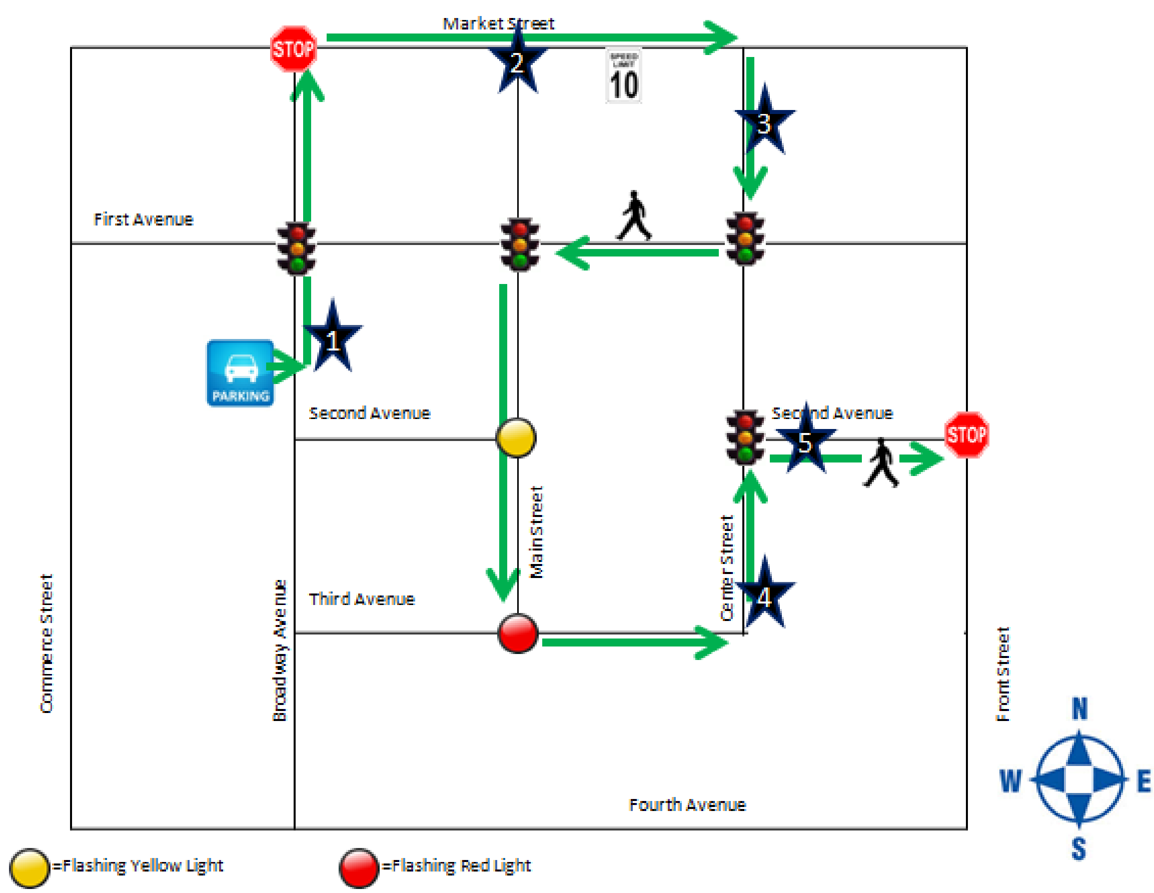


Figure 3. The Driving Route with Traffic Signals, Traffic Signs, Pedestrian Events and Probe Points (Black Stars)

Apparatus. The driving simulation included software and hardware. The software (BRDrivingSimEXM) and browser-based questionnaires were hosted on a Dell M4500 laptop with a Windows 7 Pro 32-bit operating system and a 1024 x 768 screen resolution. The hardware included an automatic and manual transmission gear shifting set, a steering wheel, and foot pedal set. All participants used the automatic transmission. The turn-signaling hardware was deactivated for this task analysis because it does not behave like typical automobile signaling.

5.2.3. **Procedure**

The procedure for building the task analysis is a two-step process beginning with the cognitive task analysis (CTA) created by the investigator with respect to the driving route in Figure 3. Next, the O*Net™ list of KSMAs from various job task analyses (JTA) were reviewed to determine which KSMAs might be required for each CTA task. We now discuss the CTA and the JTA procedures as well as the selection of the probe points for analysis.

CTA Procedure. When developing the CTA, the investigator drove the driving test route used in Figure 3 and repeatedly paused the driving to document tasks being executed as well as the required human information processing elements (e.g., perception, judgment, decision making, motor response, etc.). Next, a task list was devised and the route was re-driven without pausing to review the efficacy of the task list. The tasks were organized under two competing goals the investigator was experiencing termed, “maintain system safety” and “complete the driving route accurately.” Some tasks were grouped under one or both terms. For example, “check blind spots” was grouped under “maintain system safety” while “determine physical location where deceleration should begin” appears under both goals because deceleration is needed to make the next turn and to make the turn safely. Then using the guidelines offered in Ammerman, Becker, Jones and Tobey (1987), the complete task list was grouped into sub-activities which were in turn grouped into activities which are defined in section 3.2.1, above.

JTA Procedure. After the CTA was created, the standardized list of O*Net™'s knowledge, skills, mental processes and cognitive abilities (KSMAs) of automobile driving job roles (e.g., cab drivers, light truck delivery) were then matched to CTA tasks. Knowledge and skill requirements were extraneous to the study goals and hence, not included in the analysis.

O*Net™ displays KSMAs in categories presented in Appendix E, consisting of: cognitive abilities (CA), physical ability (PA), psychomotor ability (PM), sensory ability (SA) or mental processes (MP). While some of these categories seem orthogonal, incisive distinctions between cognitive abilities and mental processes are not articulated in O*Net™. Herein, we interpret cognitive abilities to be the abilities that a job performer uses to implement a mental process: for example, mathematical reasoning is a cognitive ability, whereas the act of planning and solving a math problem is a mental process that relies on cognitive and other abilities. In addition, while psychomotor abilities refer to physical responses that involve mental coordination and calculation, physical abilities are more broadly construed to include physical responses that are automatic and do not involve cognitive control.

Probe Point Selection. Once the CTA and JTA were complete for the entire driving route, the task analytic information was used to determine probe point locations to be used in Part III. The black stars in Figure 3 represent all probe points selected. A probe point is a pre-determined physical location in the driving scenario where the driving simulation is manually paused to solicit participant's thoughts. Each probe includes the exact probe location spanning plus and minus two seconds at the freeze point. This period was devised because the "freeze" is manually generated so a window of time was necessary to accommodate timing variance. Because driver velocities in Part III may vary across participants, those traveling at higher velocities covered more physical distance than those at lower velocities. Thus, probes had to be selected such that no matter what velocity each participant was traveling at (within reason), the content viewed (e.g., the scenery) and the actions required (e.g., steering inputs) within that 4 second window had to be relatively consistent across all participants. Also, probes could not overlap. Roughly seven candidate probe points were initially developed, but only five resultant probes were chosen based on the criteria above as well as cognitive load assessments. We operationally defined cognitive load at each probe point as the total number of abilities and mental process requirements listed in the task analysis. Five probes with highest and/or lowest cognitive load scores were selected. See Appendix F for the types of abilities and mental process requirements at each of the five probe points and Table 8 for aggregate frequencies across probes.

5.3.Analyses and Results

To assess significant differences of task load fluctuations across selected probes, a chi-square test was conducted on cognitive load totals gleaned from the task analysis. A complete listing of abilities and mental process requirements for each of the five selected probe points are provided in Appendix F and aggregate frequency data is displayed in Table 8. Probes #1 and #5 had the highest total frequencies for both mental processes and probes but it is unclear whether there are any statistically significant differences across probes. Therefore, two chi-square tests were conducted (probe X total abilities and probe X total mental processes) to assess whether significant differences existed. Neither chi-square test resulted in statistically significant differences across probes for the frequencies of abilities (p=.71) or the frequencies of mental processes (p=.64).

Table 8. Total Number of Abilities and Mental Processes (e.g., Cognitive Load) Required for Each Probe

Probe	Total Abilities	Total Mental Processes
#1	31	24
#2	20	25
#3	12	13
#4	19	19
#5	29	24

* p<.05

5.4.Discussion

The purpose of the task analysis was to provide data to help set the physical location of the probes such that the probes would provide a variety of cognitive load fluctuations. The chosen probes will then be applied to the driving simulation study outlined in Part III. Since cognitive load is a factor studied in Part III, we ran a chi-square test to determine the level of cognitive load. The steps in the driving sequence were decomposed into activities, sub-activities and tasks (Ammerman, Becker, Jones & Tobey, 1987) and were further decomposed into their abilities and mental process requirements offered by O*Net^{TM}. The total number of abilities and mental processes required across probes differed but not significantly in a statistical sense. It would be expected that probes #1, #4 and #5 would have the highest frequencies because the participant is behaviorally transitioning (e.g., executing steering inputs and decelerating) at the time of the probe. These three probes occurred just seconds after a turn was completed; however, probe #4 appeared to take place in a less visually cluttered environment compared to probes #1 and #5 where street cars were parked, traffic lights were changing and pedestrians were in the vicinity. Probes 2 and 3 occurred in a less visually cluttered environment following a long stretch of uninterrupted straightaway driving so the frequencies of abilities and mental processes would be expected to be lower comparatively. See Figures 4-8 for pictures of what participants visualize at each probe.

Figure 4. The driving scene observed by the participant during Probe #1, which consists of a city street with cars parked along the right-hand side and an approaching traffic light

Figure 5. The driving scene observed by the participant during Probe #2*, which consists of an open highway with no buildings or other cars*

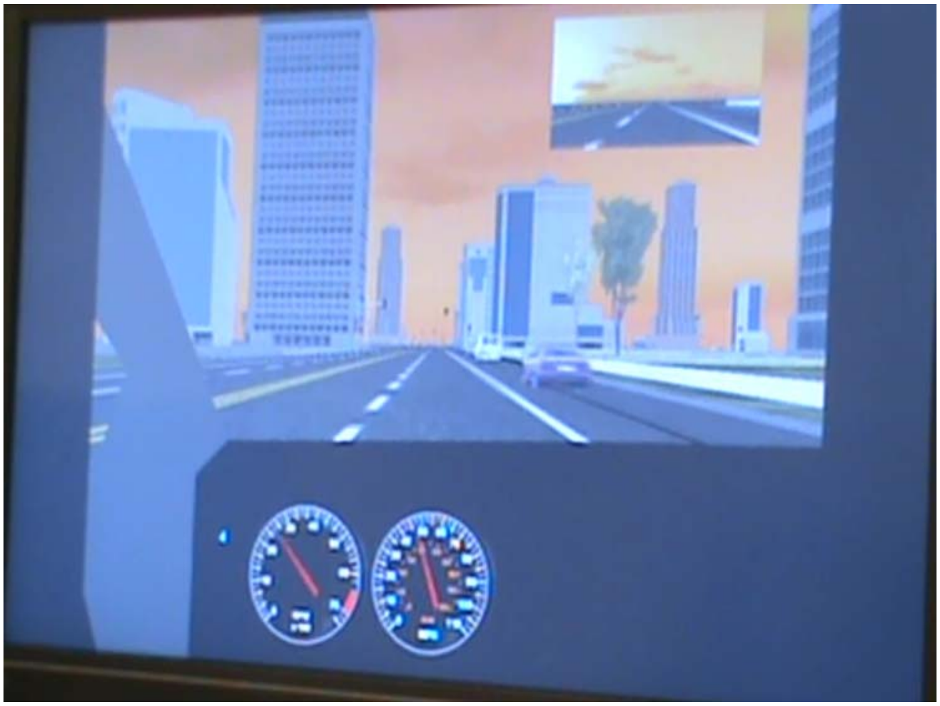


Figure 6. The driving scene observed by the participant during Probe #3, which consists of a city street with cars driving in the lane to the right and an approaching traffic light

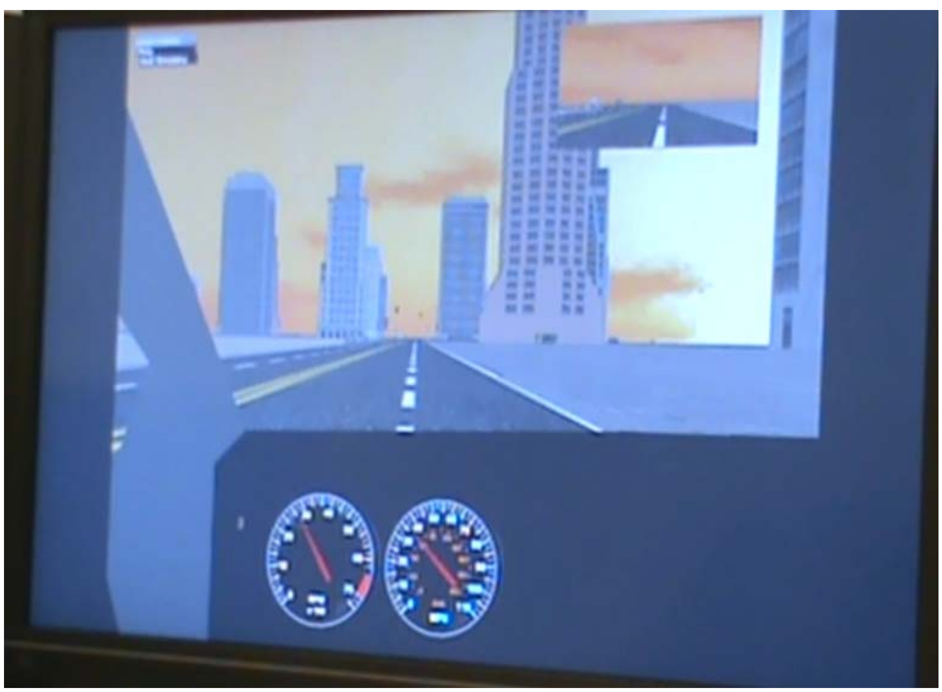


Figure 7. The driving scene observed by the participant during Probe #4, which consists of a city street with no cars

Figure 8. The driving scene observed by the participant during Probe #5, which consists of an open highway with light traffic in the right lane

There are efficacy concerns with using this approach to computing cognitive load. Non-simulated driving in the natural world is somewhat mindless and often involves automatic cognitive processing (Schneider & Shiffrin, 1977), a mentally effortless association between stimulus and response that has little active control and attention by the human. Hence, our assessments of cognitive load based on visual elements, steering inputs, etc. may not accurately reflect cognitive load in the natural world because automaticity precludes active attention to these elements. Ideally, drivers who use predominantly controlled processing, a form of active thinking of the process involved in executing a task (Schneider & Shiffrin, 1977), in novel situations may experience cognitive loads similar to what we computed. However, our intent was to reduce the novelty of simulated driving to approximate the cognitive processes used in natural driving, which would most likely entail more automatic processing. Remember there is also a temporal quality to cognitive processing. As the novelty of the simulation wears off throughout time, the driver may be more likely to shift from controlled to automatic processing used in naturalistic driving tasks. Thus, our computed cognitive load may be less accurate as time advances. To address this concern, we measured subjective mental workload at each probe in Part III, section 6.3.1.

6. Part III. The Driving Simulation Study

This section discusses the driving simulation study, which includes the study purpose, method, results and discussion.

6.1. Purpose

Part III serves two purposes: first, to determine the most efficacious metric for measuring mind wandering in the cockpit; and second, to determine the relationship between working memory capacity, task load, mind wandering and the adverse consequences of mind wandering. All hypotheses and research questions will be addressed in this section.

6.2. Method

This section describes the participants, including sampling, sample size and demographics, and the apparatus and stimuli.

6.2.1. **Participants**

Sampling. Participants were recruited from a pool of undergraduate psychology students at North Carolina State University. Participants who agreed to participate in the study received experimental credit in an Introduction to Psychology course and $10 cash compensation. The study advertisement was intentionally kept vague to minimize sampling bias towards those individuals who prefer driving related experimental studies, but risks and benefits were included in the advertisement. Each participant who arrived at the experimental appointment was assigned to one of the nine experimental conditions, based on a counterbalanced order of assessments administered (see Appendix N).

Sample Size and Demographics. All means and standard deviations for the sample demographics are listed in Appendix G for all 120 participants. The sample demographics were divided into basic demographics (age, gender, native language), driving demographics (e.g., licensing, amount and type of driving, etc.) and distractibility demographics (more or less distractible than your peers?). These demographics were collected in the post-driving questionnaire. To summarize, the sample consisted of 64% male and 36% female undergraduate students; 91% of the participants self-declared that English was their native language. The driving demographics include a mean number of 2.38 driving years of experience *(SD*=1.9 yrs.). Less than half (47%) of the sample self-reported that at the time of the study, they currently did not drive or drove less than 1 round-trip per week on average. Respondents reported 47% (*SD*=37%) of their driving time was in high density areas, while 24% (*SD*=26%) of their driving time was in low density areas. Given that this was a study about distracted thinking, participants were asked whether they believed themselves to be more or less distractible than their peers. Fifty-one participants (~43%) self-reported being more distractible, fifty-one participants (43%) self-reported being less distractible and eighteen participants (15%) were unsure.

6.2.2. **Stimuli and Apparatus**

Stimuli. This section begins with a presentation of the driving scenario stimuli, how probes were determined and presented, the questionnaires used and how these questionnaires were generated.

A customized driving scenario was purchased from Beta Research, Inc. to fit the research requirements. The driving scenario requires segments that would vary driver workload by manipulating the number of steering inputs, speed changes and events, such as unexpected obstruction to the lane of travel, per unit of time. These segments were identified with the task analysis in Part II. For example, one scenario segment includes a long straightaway with no steering inputs, no traffic signals, no visual clutter, no events and one speed change which was intended to have minimal cognitive workload. Another segment designed to increase workload involves dense traffic scenes with visual clutter (e.g., parked cars, oncoming traffic, surrounding buildings, etc.) that requires lane changes, attention to traffic signals, and unexpected events. Figure 3 (above) is a diagram of the driving scenario and the driving route used in all three trials of the study.

Probe points were determined based on the cognitive load fluctuations determined in the task analysis reported in Part II. To validate workload fluctuations in the driving route, subjective mental workload was assessed on a sample of 30 undergraduate students who served as pilot participants with the same demographic as the intended study participants. At each of the five probes depicted as black stars in Figure 3, participants were asked to make a subjective mental workload rating using a modification of the NASA TLX (Hart & Staveland, 1988), called the Raw TLX (RTLX). The RTLX is the NASA TLX without pairwise comparisons;

it is the mean of all ratings for each of the six dimensions. The benefit to using the RTLX over the NASA TLX is that is more time efficient, simpler for participants to use and is sensitive to detecting workload fluctuations (Hart, 2006). A one-way, repeated measures ANOVA reveals that there were statistically significant differences between the mean RTLX ratings across all five probes points [$F(4,135)=3.42$, $p=.011$, $\eta^2=.303$]. Given missing cases of RTLX scores across probes, a Games-Howell post hoc test indicated a marginal significant difference between probes 1 and 2 ($p=.056$) and a significant difference between probes 1 and 3 ($p=.031$). Thus, there are no significant differences amongst Probes 2-5 but probe 1 has a lower significantly mean ratings than Probe3 and Probe 2. The NASA RTLX mean ratings are listed in Table 9. During each probe, participants were asked to verbalize their thought content, to verbalize whether they were or were not conscious of their thoughts (i.e., what is their level of meta-awareness), and to provide verbal ratings on a 0 to 100 point scale (0=low and 100=high) for temporal, functional and conceptual ratings of their thought content as it relates to the task.

During each test session, the participant had to drive a route in the simulation, and also had to complete a series of questionnaires. The ordering of questionnaire presentation was counterbalanced to minimize latent effects which are detailed in the procedure section below. Each participant completed a short form Beck Depression Inventory (BDI) questionnaire (measuring negative affect), an Aospan test (measuring working memory capacity), the Salient Life Events questionnaire (soliciting each person's six most current salient life events), the Mind Wandering Definition questionnaire (measuring each person's definition of mind wandering), and the Post-Driving questionnaire (solicited opinions, preferences and experiences with the use of each mind wandering metric). Copies of all questionnaires with the exception of the Aospan test and the BDI questionnaire are in Appendix H. The BDI questionnaire is an adapted paper-based survey (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) loaded into the Survey Gizmo™ tool. All questions and response formats were maintained in the electronic version. The Aospan, an application-based assessment produced by ePrime, was a ~20 minute letter span test where the maximum number of letters held in working memory was assessed. The Salient Life Events paper-based questionnaire asked participants to list six of their most important life events experienced on the day of the test. Participants were told that these events could have temporal quality (e.g., "My dishwasher broke this morning.") or not (e.g., "I joined the Christian faith.") as long as it was salient to their life on the day of the experiment. The participants wrote sentences describing each event on the numbered lines provided and then were asked to rate the importance of each numbered item on a dipole continuous response scale between 0 (not at all important) and 100 (extremely important). Participants were instructed to place the questionnaire responses out of the investigator's eye sight to maintain privacy. If a thought listed on this questionnaire came to consciousness, the participant was instructed to state the number associated with the event rather than the content.

Most of the questions in the Mind Wandering Definition Questionnaire and the Post-Driving Questionnaire were generated from discussions with the sample of 30 participants who provided workload ratings (NASA RTLX) for driving probes. The NASA RTLX participants were informally allowed to 'try out' various mind wandering metrics and discussions about their experiences led to questions that were included in the Post-Driving questionnaire (see

Appendix H). These questions were intended to understand how participants defined mind wandering and generally how they experienced mind wandering, if at all, across all three driving trials.

Apparatus. The driving simulation test and browser-based questionnaires were hosted on a Dell M4500 laptop while the Aoapan test was hosted on a Dell e600 laptop to reduce the bandwidth requirements on the M4500. The driving simulation software (BRDrivingSimEXM) executes driving scenarios using a Dell M4500 computer (see hardware specifications in Appendix I) with a Windows 7 Pro 32-bit operating system with 1024 x 768 screen resolution to save processing power. This LogiTech G27 simulator hardware includes both automatic and manual transmission, the steering wheel and foot pedal set and the software includes 30-fixed practice scenarios. The simulation software measures braking response time, number of collisions, intra-automobile distance, speed, and failure to signal. None of these 'canned' measures were used in this study because they were aggregate measures that could not be divided and associated with each probe. Also, the turn signal capability was disabled in the study because in pilot testing, it proved difficult to turn on and off.

6.2.3. **Procedure**

Counterbalancing. Two portions of this study were counterbalanced: the ordering of the study items and the ordering of the mind wandering metrics. Nine different orders existed which are listed in Appendix M. The ordering of the study questionnaires and driving tests was Latin square counterbalanced across subjects to minimize latent nuisance variables (e.g., test fatigue). The driving test consisted of one driving scenario that was driven three times by each participant; one time for each of the three mind wandering metrics (all probe, clicker, and probe/self-caught). The order of the mind wandering metric presentation was also Latin square counterbalanced (see Appendix N) such that each participant was exposed to all three metrics (all probe, clicker, and probe/self-caught) in a different order.

Procedure. This procedure described below represents one of nine possible counterbalancing orders explained in Appendix N. After informed consent was received from the participant, the ordering of the study items was verbally presented and any participant concerns were addressed before proceeding. For each questionnaire and test administered, oral and written instructions were provided. To reduce reckless driving behaviors common to driving video games, participants were told to approach the driving test with the same conscientiousness as the driving portion of a driver's license test. While participants completed various tests and questionnaires, the experimenter resided in an adjoining room to provide participant privacy. The Aospan test was first administered on the Dell e600 laptop and once complete, the participant moved to the larger Dell M4500 laptop for the remainder of the study session. The BDI questionnaire was administered online in the Mozilla Firefox browser followed by a paper-based Salient Life Events questionnaire and then browser-based Mind Wandering Definition questionnaire. Once the driving portion of the study was complete, the browser-based Post-Driving questionnaire was administered followed by a study debriefing.

The driving portion of the study comprised several steps as well. Prior to physically driving, participants placed their completed paper-based Salient Life Events questionnaire out of eye-sight of the investigator to maintain privacy. Participants could refer to that questionnaire at each probe and if any of salient events comprised their probed thought contents, they provided the number assigned to that life event rather than the content of the thought. The driving session began with a 15-minute driving tutorial to introduce simulator controls and vehicle handling in a high traffic density practice scenario. In addition, the mind wandering metrics (all probes, clicker, and probe/self-caught) were introduced and practiced before the test scenarios began. Immediately following the driving tutorial, the participant began the series of three test trials. During each probe, participants were asked to verbalize their thought content, including salient life events, to verbalize whether they were or were not conscious of their thoughts (i.e., meta-awareness), and to provide verbal ratings on a 0 to 100 point scale (0=low and 100=high) of the temporal, functional and conceptual relatedness of their thought content to the task. Immediately following the three trials, the participant was provided the online Post-Driving questionnaire to complete and they were debriefed.

6.3. Analysis and Results

The structure of the results begins with operational definitions of each variable studied in Part III (see section 6.3.1). Then, each research question or hypothesis investigated (e.g., section 6.3.2 to section 6.3.6). All subheadings for each section will include: list of the variables used to answer that question or hypothesis (e.g., section 6.3.2.1), the diagnostics conducted (e.g., 6.3.2.2), the analysis and results (e.g., 6.3.2.3), and discussion (e.g., 6.3.2.4).

6.3.1. Variables

The variables consist of five predictor variables (Absolute WMC Scores, RTLX scores, BDI scores, Probe number, and Participant number), five outcome variables (wandering, off-task, seconds of speeding, number of lane deviations and number of collisions) and three additional variables from the Post-Driving questionnaire. Each variable is separately discussed, below.

The Aospan software from ePrime® yields five measures per participant: absolute score, total correct, math errors, speed errors, and accuracy errors. The working memory capacity score for each participant was the Aospan's absolute score, which we renamed to the *absolute WMC score*. To the best of our knowledge, the absolute WMC score offered by ePrime's Aospan tool has not been normed. We divided the range of scores into low, medium and high working memory capacity. Low working memory capacity individuals have a range of 1-30, medium working memory capacity ranges from 31-49 and high working memory capacity individuals have scores above 50. The sample mean was 42.47 (*SD*=15.18).

The *RTLX score* is the mean rating across all six dimensions (mental, physical, temporal, performance, effort, and frustration) in one probe. Each rating was converted to a percentage of the total scale and the mean, called the RTLX score, was calculated for each probe that the participant was exposed to. A separate sample of thirty undergraduate student pilot participants drove the same route depicted in Figure 3 and were probed at each of the five locations depicted by the black stars in Figure 3. NASA TLX ratings on all six dimensions (mental, physical, temporal, performance, effort, and frustration) as well as an overall workload rating were collected. A Pearson Product Moment correlation was conducted (n~29) between the RTLX scores and the overall workload rating to assess whether the RTLX score was sensitive to changes in workload fluctuations collected on a different metric. Because the correlation coefficient $r = 0.756$ was high, only the RTLX scores were used in this analysis. The RTLX value ranges between 0.16 and 0.27 and the overall mean was 0.22 (*SD*=.15).

Table 9. Means and Standard Deviations (SD) of RTLX Scores For Each Probe

Probe	Mean RTLX	*SD*	*N*
1	0.16	0.11	29
2	0.26	0.16	29
3	0.27	0.16	29
4	0.24	0.17	26
5	0.17	0.14	27

The *BDI score* was the raw aggregate score from each participant's responses to the short form of the Beck Depression Inventory as a measure of negative mood, also called dysphoria. The test scoring guidance further decomposes scores into four levels of negative mood: minimum depression (score=0-9), mild depression (score=10-18), moderate depression (score=19-29) and severe depression (score=30-63). The raw scores range from 0 (no depression) to 63 (extreme depression). Very little normative information exists for the short version of the BDI. In one study conducted in New Zealand, for 16-19 year olds (n=28) the mean was 3.11, the *SD* was 2.99 and 90th, 95th and 99th percentile was 6.94, 8.01 and 10.08, respectively. For the 20-29 year olds (n=121), the mean was 1.81, the *SD* was 2.72, the 90th, 95th, and 99th percentile was 5.29, 6.27 and 8.15, respectively (Knight, 1984). In another study on college students using the long version of the BDI, the mean was 6.81 and *SD* was 5.52 (Hill, Kemp-Wheeler & Jones, 1986).

A *probe* occurred when the driving simulation was manually paused to collect participant's thought contents that occurred a few seconds before the probe. Each probe was numbered ranging from 1 to 10. Each participant drove the test-driving route three times, once per trial, using all three mind wandering metrics. One of the metrics (the clicker method) did not involve probing so each participant experienced ten probes, instead of fifteen probes. Regardless of the counterbalancing order of the metrics, probes #1-5 involved the self-caught and probe metric and probes #6-10 involved the all probe metric. Probes #1 and #6 occurred in the same physical location because the same driving route and probe location were repeated. Similarly, the following probe pairs all occurred in the same physical location: probes #2 and #7; probes #3 and #8; probes #4 and #9; and probes #5 and #10.
Participant number is the number assigned to each participant in the study.

Two similar dependent variables used to measure when a participant was not thinking about the task at hand were *off-task* (i.e., "yes/no I am off-task") and *wandering* (i.e., "yes/no I am mind wandering"). During pilot testing, participants often admitted they were off-task in their thinking but they did not consider themselves to be mind wandering. Thus, both variables were collected in the driving study and only two response options were available for each variable: "yes" (coded as 1) or "no" (coded as 0). Eight cases of missing data were discovered because some participants did not have enough time to finish all study probes; consequently, the missing data occurs during the last two probes of a study trial.

The seconds of speeding, number of lane deviations and number of collisions were manually tallied by reviewing each participant's video. This process began by calculating each participant's 14-second window for each probe experienced. The 14 -second window was exactly 14 seconds prior to the moment each probe was initiated because 14 seconds was identified as the mean interval between shifts of thought topics (Klinger, 1978). Smallwood, Beach, et al., (2008) and He, Becic, Lee and McCarley (2010) both used a 10-second window but a liberal 14-second window was chosen herein to capture performance decrements that may take a few extra seconds to develop. During this 14-second window, we tabulated seconds of speeding, number of lane deviations and number of collisions. Seconds of speeding were manually counted using the video player's timeline display. Even though 1 mph over the speed limit is by law 'speeding,' the subjective interpretation of speeding is different. Thus, participants were asked in the study debrief how many miles per hour over the speed limit is considered "speeding" and 6 mph or more had the highest frequency. Therefore, 6 + mph over the speed limit is "speeding" and the total number of seconds out of 14 the participant was "speeding" comprised the *seconds of speeding* variable.

The *number of lane deviations* incurred during that 14-second window was also manually tallied by reviewing each video performance of each participant. Two boundaries were taped to the monitor during this assessment to demark the right and left lane deviation boundaries and anytime the left or right lane demarcation crossed the boundary taped to the monitor, it was counted as a single lane deviation. The boundaries are depicted by the solid red lines in figures 9 and 10. In figure 9, the participant is not experiencing a lane deviation but in figure 10, the participant is.

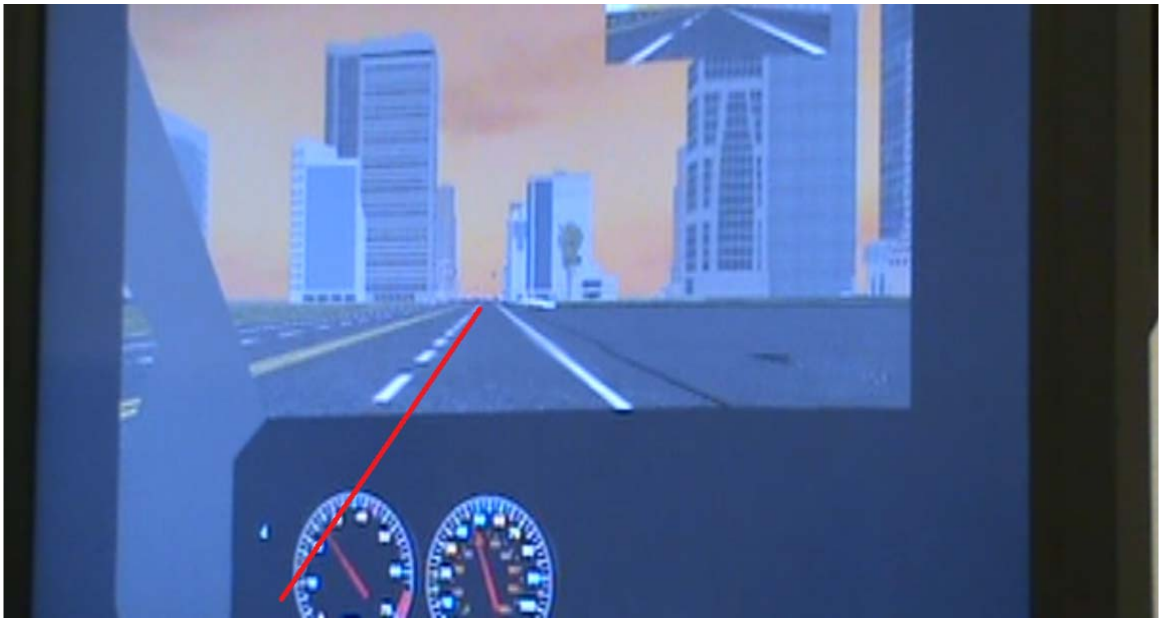


Figure 9. Participant is driving within their lane: the red line depicts Left Lane Deviation Boundary

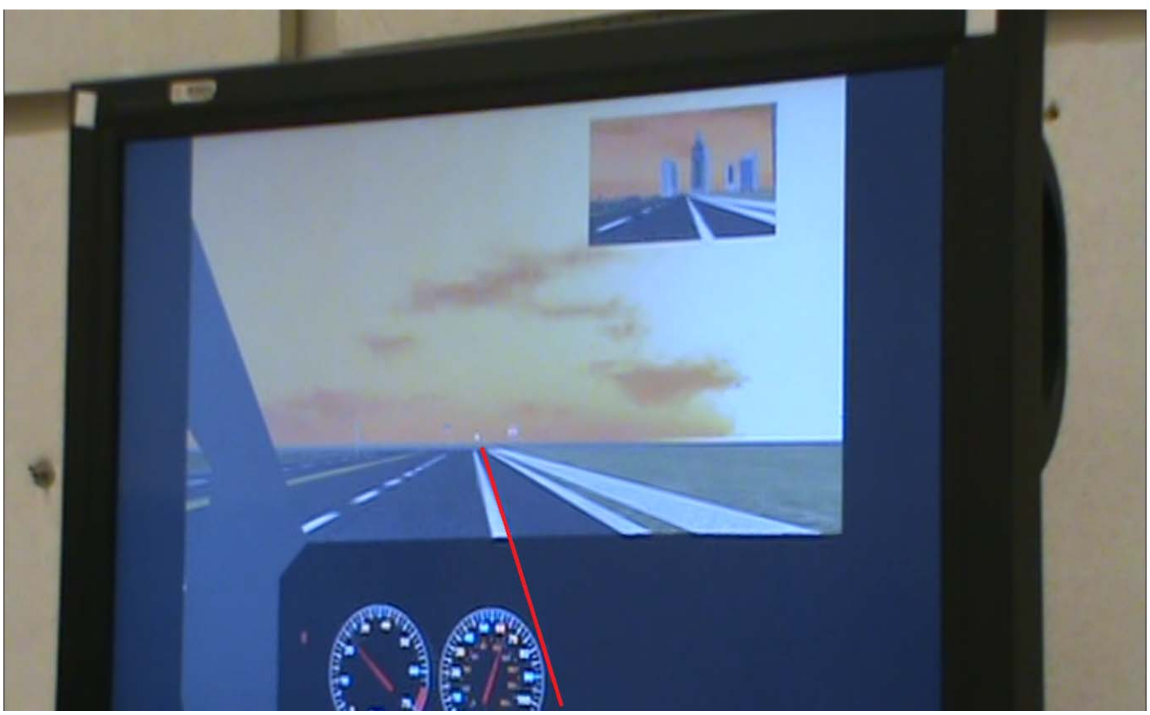


Figure 10. The participant is performing a lane deviation boundary: the red line depicts the Right Lane Deviation Boundary, which has been crossed by the right car wheels

The total *number of collisions* with trees, vehicles, road curbs, grass or pedestrians was tallied during each 14-second probe period but since only two collisions (participant 57, probe 5 and participant 29, probe 2) occurred across all participants, this dependent variable was dropped from the analysis.

In the Post-Driving questionnaire, participants were asked about their experiences and preferences with each of the three mind wandering metrics. *Meta-awareness* is the response to the question, "Which metric did you babysit your thoughts the most with?" Participants could choose one of the following four options: clicker, self-caught/probe, all probe and none- they all were the same. This question aids in identifying which metric may be correlated with higher frequencies of meta-awareness. The *Re-engagement* variable captures responses to, "Overall, which metric allowed you to re-engage the quickest with your original stream of conscious thought experienced before the probe?" Again four response options were available: clicker, self-caught/probe, all probe and none- they all were the same. The *Frequency Accuracy* variable captured responses to, "Which mind wandering metric was the best at accurately capturing how frequently you mind wandered?" And again, four response options were available: clicker, self-caught/probe, all probe and none- they all were the same.

6.3.2. **Hypothesis 1.**

Operators with low working memory capacity will mind wander more frequently during high task loads compared to operators with normal or high working memory capacity.

6.3.2.1. Variables

The independent variables are Absolute WMC Scores, RTLX scores, probe number and the dependent variables are off-task and wandering. Absolute WMC Score is an assessment of working memory capacity and RTLX score was an assessment of task load. Probe Number was an indexing

6.3.2.2. Diagnostics and Data Description

This section begins with a discussion of the diagnostics followed by a description of the data collected that is relevant to this hypothesis.

Diagnostics. Two logistic regression analyses were used to address hypothesis 1; one analyses on the off-task variable and other on the wandering variable. Diagnostics were conducted on observations to determine which observations had the most influence on the model fit. SAS offers a toolkit of diagnostics for logistic regression (listed in Appendix J) which are presented as a series of charts to visually identify influential observations. The observations that have been flagged as influential are also outliers in working memory capacity scores. The lowest working memory capacity score was more than two standard deviations below from the mean (i.e., Absolute WMC Score=3) while the highest score was approximately two standard deviations above the mean (i.e., Absolute WMC Score=75). The Absolute WMC Score was not normally distributed (Kolmogorov-Smirnov test $p<.001$) and the distribution was skewed left (Skew= - 0.145). However, when outliers were replaced with values that were +-1 SD from the mean, the model fit did not improve so the original data was used for the analyses. Another contributing factor to difficulties with model fit was the RTLX scores had a more flat, negatively skewed non-normal distribution; RTLX kurtosis was -.078 and skew= 0.838. No effort to transform the variables was made, because SAS modeling addresses non-normal data.

Describing the Data. Three variables are implicated in this analysis: working memory capacity, task load (RTLX scores) and mind wandering frequencies. We now review working memory capacity and mind wandering frequencies; RTLX scores are discussed in section 6.3.1, Table 9.

The distribution of working memory capacity scores had a mean of 42.47 (*SD*=15.24), was not normal (skewness = -.15 and kurtosis = -.30) with a negative tail and had a flat rather than multimodal curve. The mode was 44 and the two most frequent scores were 44 and 56. See Figure 11 for the histogram of working memory capacity scores for all 120 participants.

Figure 11. *Histogram of working memory capacity scores*

For each of the 120 participants, a percentage of all ten probes that the participant self reported off-task thinking was calculated. The mean percentage of off-task thoughts was 44.67% (*SD*=17.63%) indicating that the average number of probes that the person was off-task thinking was approximately 4 out of 10. A histogram of the total number of participants who experienced a certain percentage of off-task thinking is listed in Figure 12. Every participant had at least one off-task thought at one probe point, but none of the participants experienced off-task thinking across all probes.

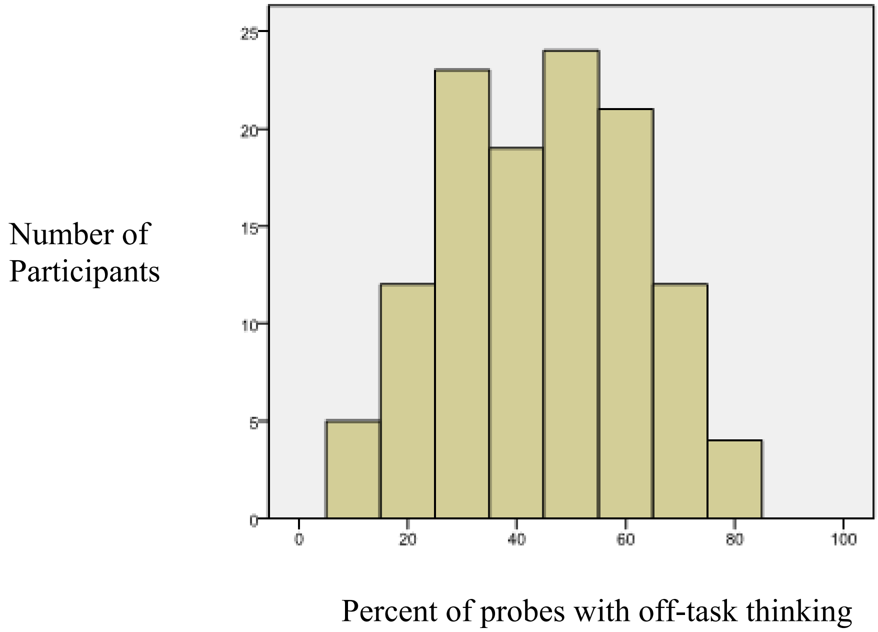
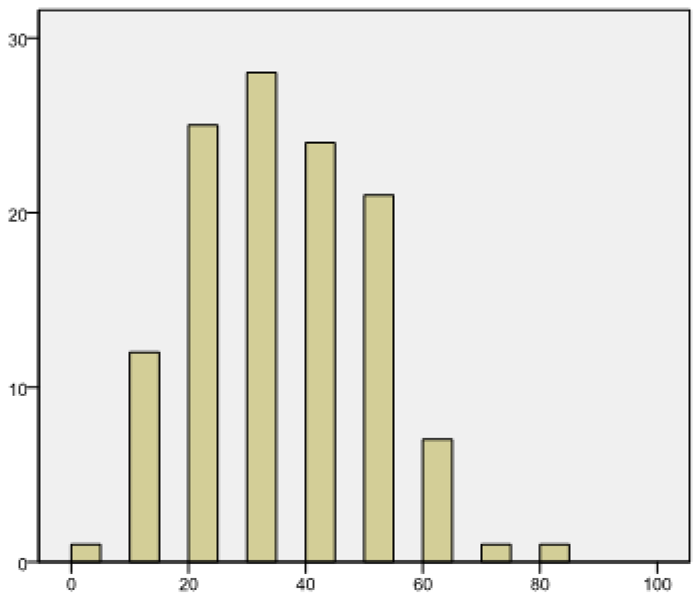


Figure 12. *Histogram of the number of participants who experiences a particular ratio of off-task thoughts out of 10 probes*

Remember that not every off-task thought was classified as a mind wandering episode. Thus, the histogram of percent mind wanderings across all ten probes has a slight positive skew (see Figure 13) and a mean of 33.68% (*SD*=15.25%) indicating that on average, participants experienced mind wandering at a little over 3 out of 10 probes.

Number of
Participants



Percent of probes with mind wandering

Figure 13. *Histogram of the number of participants who experienced a certain ratio of mind wandering incidents out of 10 possible probes*

How many times on average did participants proclaim that they were off-task but not mind wandering? The difference between the number of off-task and mind wandering episodes across all ten probes was calculated for each participant while most of the participants classified all off-task thoughts as wanderings (n=52), on average, participants experienced approximately one out of ten instances (M=1.10, SD=1.27, n=120) for probes where the off-task thought that was not subjectively interpreted as a case of mind wandering. Thus, most of the participants had at least one or more probes that was not considered mind wandering, but was considered off-task (see Table 10).

Table 10. *Frequency participants who experienced some number of probes out of ten that were subjectively interpreted as off-task but not mind wandering (n=120)*

Number of probes that were classified as off-task but not mind wandering	Percentage of Participants	Number of Participants
0	43.3	52
1	22.5	27
2	22.5	27
3	8.3	10
4	.8	1
5	2.6	3
Total	100	120

In addition to just frequencies of occurrences per participant, we computed the frequencies of off-task thinking and mind wandering across all 10 probes. Figure14 indicates the total number of participants (*n*=120) who at each probe experienced on-task (score=0) or off-task (score=1) thinking. Figure 15 indicates the total number of participants (*n*=120) whom at each probe experienced a mind wandering event (score=1) or did not experience a mind wandering event (score=0).

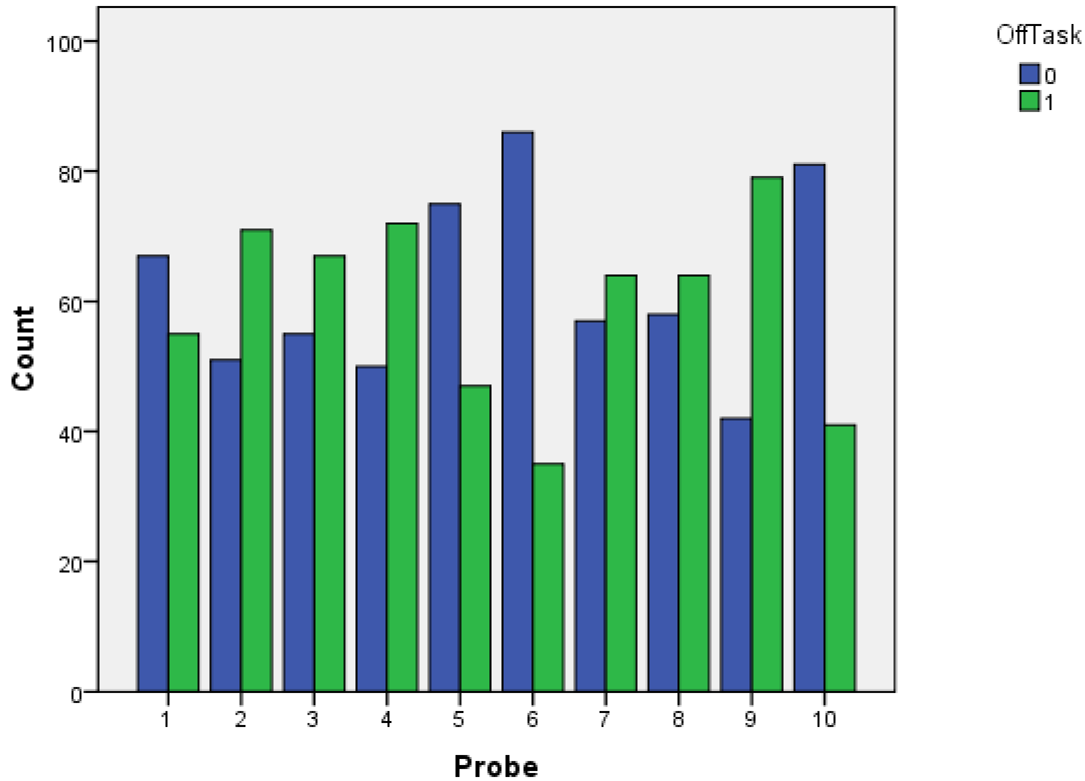


Figure 14. *Frequencies of participants who experienced off-task vs. on-task thinking at each of the 10 probes.*

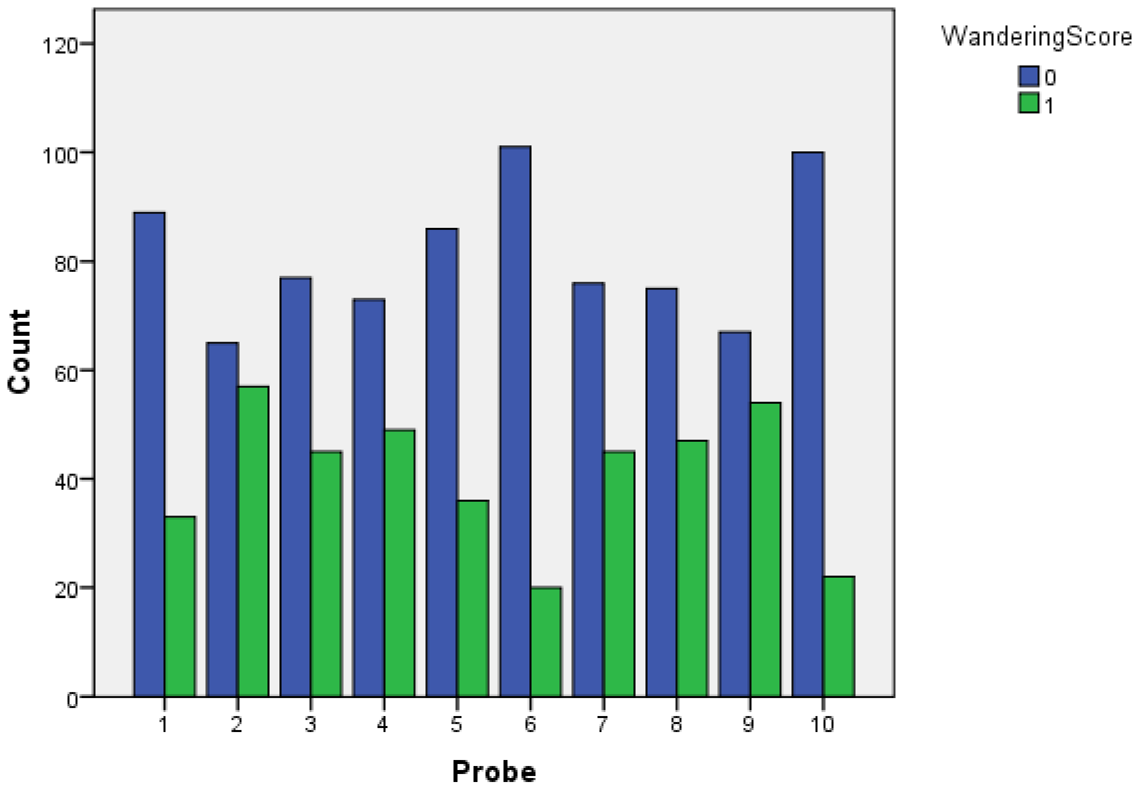


Figure 15. *Frequencies of participants who experienced mind wandering vs. non-mind wandering at each of the 10 probes*

6.3.2.3. Analyses and Results

To address this hypothesis, the RTLX scores from the pilot sample of ~29 participants were analyzed independently from the two logistic regression analyses (*n*~120) evaluating the relationships between working memory capacity scores and off-task thinking frequencies as well as working memory capacity scores and mind wandering frequencies. The RTLX analysis discussed in section 6.2.2 above is reiterated below. Each analysis is introduced by the dependent measure below: *Dependent Variable = RTLX scores*, *Dependent Variable = Off-task* and *Dependent Variable= Wandering.*

Dependent Variable = RTLX scores. A one-way (Probe #) repeated measures ANOVA (n~29) reveals that there were statistically significant differences between the mean RTLX ratings across all five probes points [$F(4,135)=3.42$, $p=.011$, $\eta^2=.303$]. Given unequal sample sizes across probes, a Games-Howell post hoc test indicated a marginal significant difference between probes 1 and 2 ($p=.056$) and a significant difference between probes 1 and 3 ($p=.031$). Thus, there are no significant differences amongst Probes 2-5 but probe 1 has a lower significantly mean ratings than Probe3 and Probe 2. The RTLX mean ratings are listed in Table 9 and Table 11.

The table of means and standard deviations for each probe's RTLX scores by probe are listed in Table 11 as well as the percentages of off-task and mind wandering thoughts. From the Games-Howell post hoc test, probe 1 (and probe 6) has a significantly lower mean RTLX rating than probe 3 (and probe 8).

Dependent Variable = Off-task. Because the dependent measure is a dichotomous response variable, a logistic regression was conducted with one independent variable, absolute WMC score, regressed on off-task and the overall omnibus model was non-significant [$\chi^2(1)=1.15$, $p=.283$]. The overall model fit statistic was 823.91 for -2 log Likelihood.

Dependent Variable =Wandering. A second logistic regression was conducted with absolute WMC score regressed on the wandering variable. The overall model fit statistic was 738.56 for -2 log Likelihood. The overall omnibus model was non-significant [$\chi^2(1)=0.536$, $p=.464$]. Scatter plots relating off-task thinking and absolute WMC score as well as mind wandering and absolute WMC score are shown in Figures 16 and 17, respectively. A score of 0 indicates a non-occurrence of off-task thinking or mind wandering and a score of 1 indicates an occurrence.

Table 11. *Percentage of Off-Task Thoughts and Mind Wanderings as well as Mean RTLX Scores for Each Probe*

Probe	% of Participants with Off-task thoughts(*n*=120)	% of Participants with Mind Wanderings (*n*=120)	Mean RTLX (*n~29*)	RTLX *SD* (*n~29*)
1	45.1	27.0	0.16	0.11
2	58.2	46.7	0.26	0.16
3	54.9	36.9	0.27	0.16
4	59.0	40.2	0.24	0.17
5	38.5	29.5	0.17	0.14
6	28.9	16.5	0.16	0.11
7	52.9	37.2	0.26	0.16
8	52.5	38.5	0.27	0.16
9	65.3	44.6	0.24	0.17
10	33.6	18.0	0.17	0.14

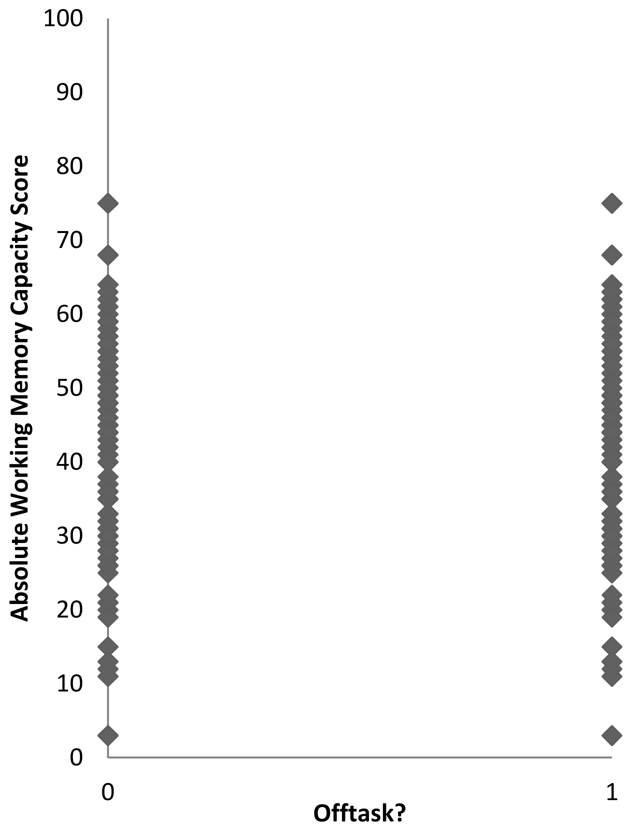


Figure 16. *Scatterplot of off-task thinking (1=yes, 0=no) and absolute WMC score (n=1200)*

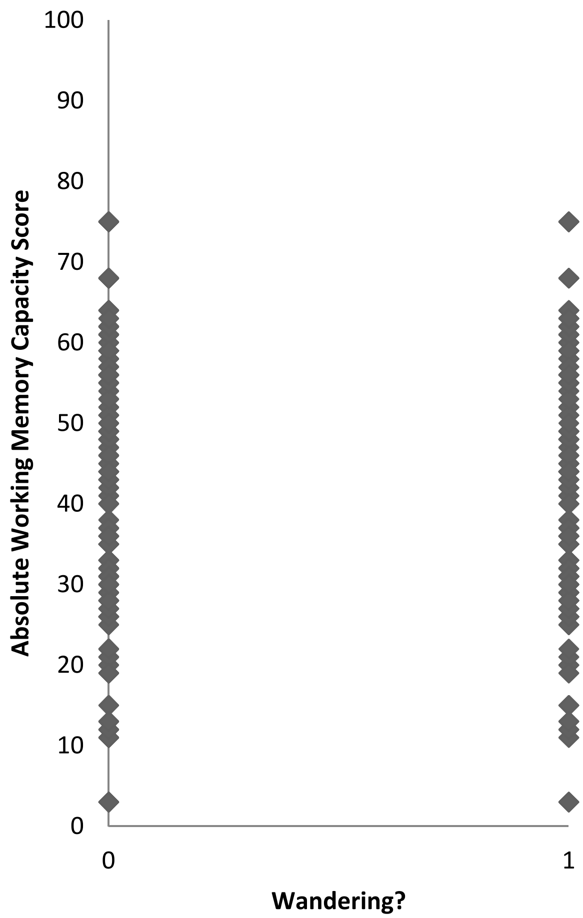


Figure 17. *Scatterplot of mind wandering (1=yes, 0= no) and absolute WMC score (n=1200)*

6.3.2.4. Discussion

In an attempt to assess data trending, a fixed effect logistic regression with RTLX scores and working memory capacity scores were loaded into the same statistical model, even though the inclusion of the RTLX scores into the model was inappropriate. No significant model could be found. Thus, RTLX and Absolute WMC Scores were separately evaluated with respect to incidents of off-task thinking and mind wandering. First, it was hypothesized that an inverse relationship between RTLX, an indicator of task load, and occurrences of mind wandering existed which was addressed in Table 11. When comparing the min/max mean RTLX ratings with the min/max percentages of off task and mind wanderings for probes 1-5 (which used the self-caught and probe metric), it appears that a direct relationship rather than an inverse relationship may exist. The same trend is apparent in probes 6-10 so the discussion is limited to probes 1 through 5. When the workload was the lowest at probe 1, one would expect to see the maximum amount of participants mind wandering if an inverse relationship occurred but the opposite was true. At probe 1, the percentage of participants experiencing mind wandering were also the lowest. When workload was highest, the highest frequencies of mind wandering also occurred. Since two different samples were used to collect the data depicted in Table 11, there is no certainty that this effect is valid. An obvious limitation of our work is that subjective mental workload ratings were not collected at each probe for fear that this information, in addition to all the other information collected, would become too disruptive. Thus, future work needs to assess the relationship between subjective mental workload and mind wandering frequencies.

The relationship between absolute working memory capacity and mind wandering was not statistically significant even though it was hypothesized to impact the relationship between task load and mind wandering frequencies. Raw data of the top and bottom most extreme scores of working memory capacity are listed in Table 12. The data fails to show that a relationship exists between mind wandering frequencies and working memory capacity scores. People at the min and max Absolute WMC capacity scores mind wandered at different frequencies.

Table 12. *The top and bottom three scores of working memory capacity and their percent of off-task and wandering frequencies out of ten probes (n=6)*

Percent of probes with off-task thinking	Percent of probes with mind wandering	Absolute WMC score
10	10	3
40	20	11
50	50	15
40	40	75
40	40	75
30	10	75

When reviewing the occurrences of mind off-task and mind wandering frequencies alone, it appears that another variable is influencing these frequencies at probes1 and 6 which were 45.1% and 28.9% respectively. Given that probes 1 and 6 were in the same exact physical location in the driving simulation, similar frequencies of off-task and mind wandering were expected. The only difference between probes 1 and 6 is the metric used; probe 1 involved the self-caught and probe metric while probe 6 involved the all probe metric. Yet, metrics used first were counterbalanced across participants. If metrics impacted the frequencies across probes, all of the probes, not just probes 1 and 6, would be affected. However, if metrics did impact frequencies of wandering, the all probe metric used in probes #6-10 would have higher incidents of off-task and mind wandering compared to self caught and probe (probes #1-5). The self caught and probe metric theoretically may involve more meta-awareness allowing participants to be more conscious of their off-task thinking which in turn might bias them to wander at different frequencies than natural. It might be higher than normal, because participants may spend more time thinking about whether or not a thought was off-task, or it may be lower than normal, if more resources available for wandering are consumed with monitoring thoughts. We anticipated that the all probe metric, compared to the self caught and probe metric, allowed people more resources available to think off-task, because participants were not consuming these resources as much during the monitoring of conscious thought. However, the data does not support this idea. The total number of off-task occurrence and mind wandering occurrences during the self caught and probe metric (probes #1-5) was 312 and 220, respectively, compared to the total number of off-task (281) and mind wandering (187) occurrences during the all probe metric (probes #6-10). This was the opposite from what we expected. If meta-awareness consumes attentional resources required for off-task thinking and mind wandering, we'd expect fewer occurrences during these probes, not more.

To understand if participants actually believed to be expending more resources in the self caught and probe metric compared to the all probe, a question was posed in the Post-Driving questionnaire. Participants were asked during which metric did he/she believe to monitor their own thinking most often (e.g., high meta-awareness requirements) and 40.0% said all probe, 37.5% said self caught and probe and 35.0% said the clicker metric (7.5% non-response). Even though these percentages are relatively similar, participants believed they were monitoring their thinking more in the all probe yet we expected the opposite. In review of their comments on the metrics used, no explanation for why participants felt this way was provided. To further support this unexpected finding, we asked participants in the Post-

Driving questionnaire, "Which metric they felt was the most accurate at capturing how frequently he/she mind wandered?" and 54.2% selected the clicker metric, 32.5% selected the self caught and probe metric and 10.0% selected the all probe (n=120). Thus, the all probe metric was subjectively the most arduous to experience and least accurate and may have contributed to subjective mental workload more than the other metrics.

6.3.3. **Hypothesis 2**

Dysphoric operators who experience negative emotional affect will mind wander more frequently than those who have normal or positive affect.

6.3.3.1. Variables

Two independent variables (BDI score and Participant number) and two dependent variables (wandering and off-task) were used in this analysis. Participant number was modeled as a random effect variable nested within BDI score.

6.3.3.2. Diagnostics and Data Description

Diagnostics. Prior to the analysis, diagnostics were conducted on observations to determine which had the most influence on the model fit. The SAS statistical software offers a diagnostics for logistic regression toolkit (listed in Appendix J) and the most influential observation had extreme values for the BDI score (participant #87 with a BDI score of 22). The distribution of BDI scores is skewed right (skew=+1.3) which supported the finding that the distribution was not normal (Kolmogorov-Smirnov test p<.005). No data transformations were executed, because SAS modeling addresses non-normal data.

Data Description. The distribution of BDI raw scores is close to normal distribution with a slightly positive skew (skewness=1.3, kurtosis=3.0), a mean of 4.57 (SD=0.35) and a range of scores from 0 to 22. According to the guidelines provided for classifying negative affect, 89.8% is minimally depressed, 9.8% is mildly depressed and 0.8% is moderately depressed. No cases of severe depression were discovered. These results were similar to the means found with college students (*M*=6.81, *SD*=5.52) (Hill, Kemp-Wheeler & Jones, 1986) and another study involving 16-19 year olds (*M*=3.11, *SD*=2.99) (Knight, 1984). Our standard deviation was much smaller comparatively.

A one-tailed Pearson correlation between BDI score and off-task yielded a non-significant correlation (*p*=.11) while a significant one-tailed correlation between BDI score and wandering was significant (*p*=.05). See Figures 18 and 19 below respectively.

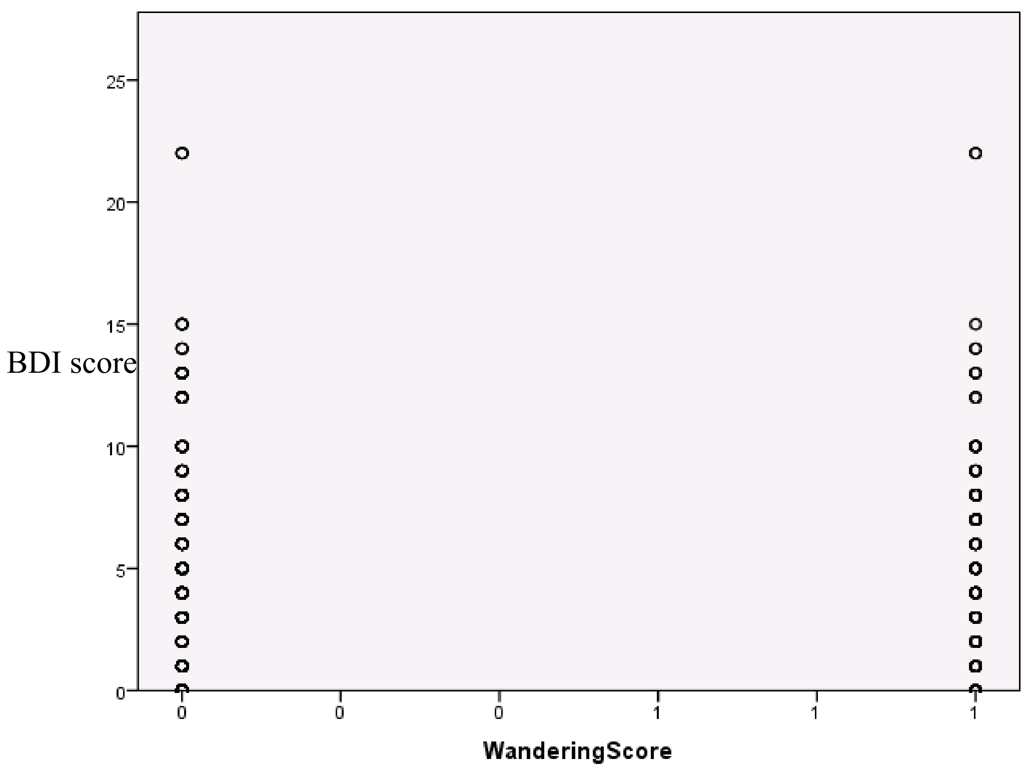


Figure 18. *Scatterplot of occurrences of mind wandering by BDI scores*

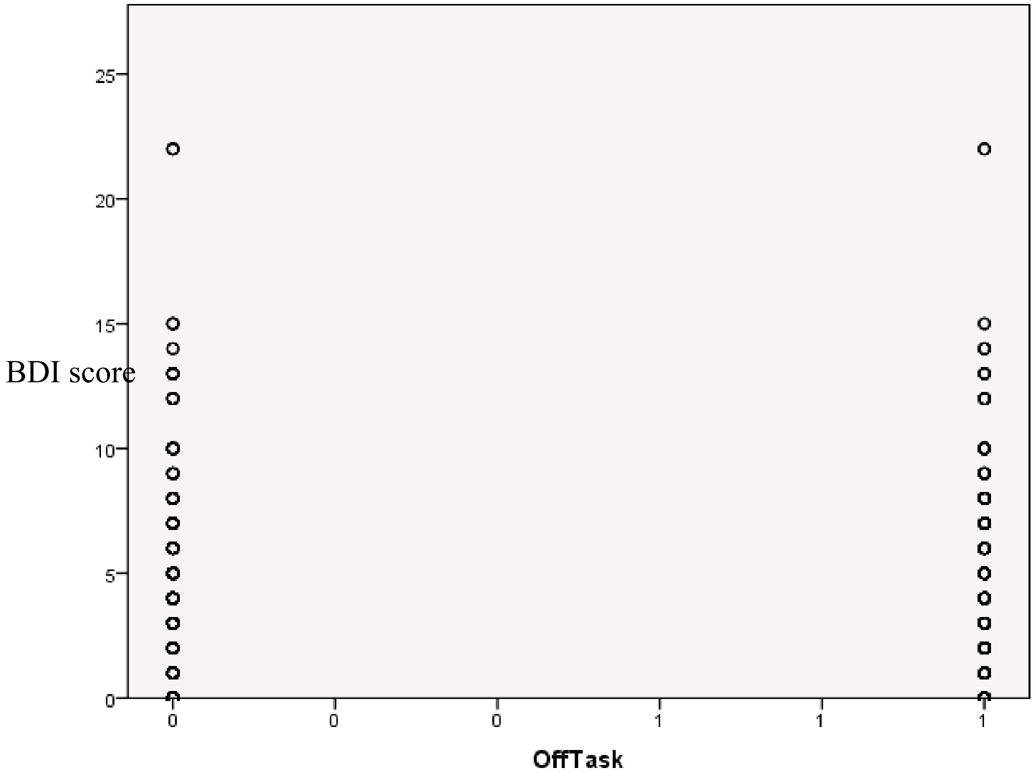


Figure 19. *Scatterplot of occurrences of off-task by BDI scores*

6.3.3.3. Analyses and Results

Two fixed effects logistic regression analyses were conducted; one for each dependent variable. Each analysis is introduced by the dependent variable involved.

Off-task. A hierarchical log linear model with BDI score nested within participant number (fixed effects) and participant number (random effect) was initially embarked in order to assess the variance of the dependent measures attributed to the BDI scores controlling for the effects of the participants. However, model convergence failed and all random effects were removed and a main effects model resulted. After model convergence was achieved, a stepwise method was used with all main effects (participant number and BDI score) and the interation initially loaded into the model. The interaction term was not statistically significant and thus, was removed. The resultant model included a significant Type III test of fixed effects for participant number [$F(1,1194)=30.23$, $p<.001$] and BDI score [$F(1,1194)=2.11$, $p=.0347$]. Figure 20 shows the probability of a BDI score correlated with occurrences of off-task thinking. In summary, the BDI score significantly predicted off-task thinking such that those individuals with smaller BDI scores (minimal levels of depression) had significantly smaller probabilities of being off-task than those with higher affect. In addition the confidence interval, represented by the shaded area around the regression line, is small for the lower BDI scores but as the BDI scores increase, the confidence interval increases also increase. Thus, at lower BDI scores (less depression), the variability of responses is less than the variability at higher scores. The resultant equation with the parameter estimates are:

Y(off-task) =.03872 – 0.00936(participant number) + 0.03291(BDI Score)

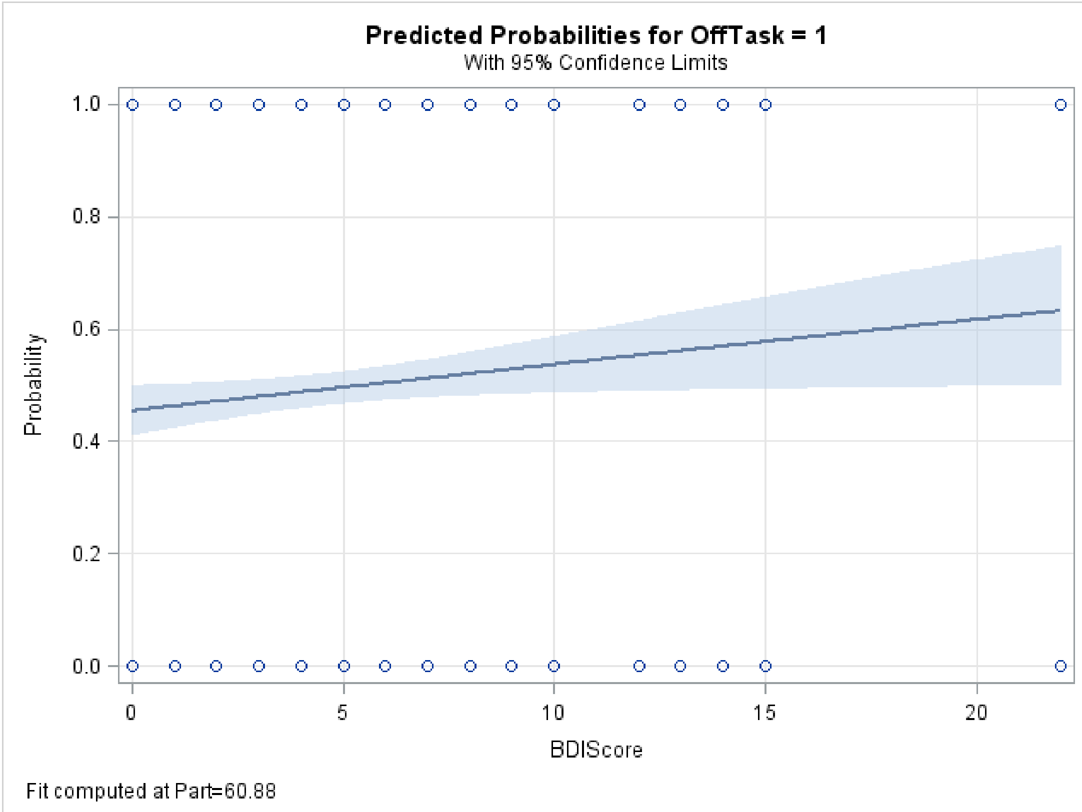


Figure 20. *Regression Line and Confidence Intervals of the Probability of Each BDI Score Correlated with Off-task Thinking*

Mind Wandering. The same model used for the off-task dependent variable was also used with mind wandering. All main effects (participant and BDI score) and the interaction were initially loaded into the fixed effect model but the interaction term was not statistically significant and thus, was removed. The resultant model includes a non-significant Type III test of fixed effects for participant number [$F(1,1194)=1.45$, $p=.225$] and a non-significant effect for BDI score [$F(1,1194)=3.47$, $p=.0629$]. Even though this is a statistically non-significant model, the relationships between the dependent and independent variables follow the same trends as when the dependent variable was off-task. Figure 21 depicts the regression line and confidence intervals for the relationship between mind wandering and participants' BDI score.

Y(wandering) =.-0.6868 – 0.002126(part) + 0.02979(BDI Score)

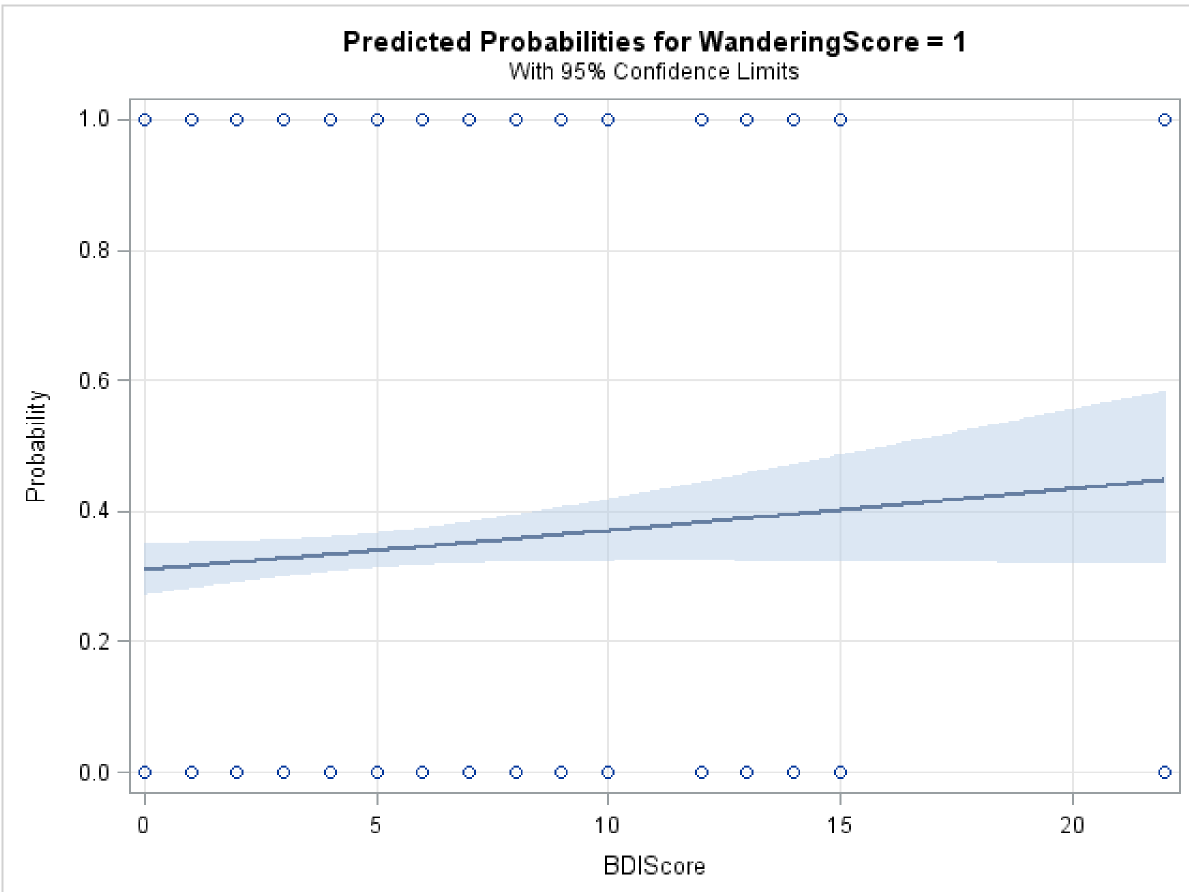


Figure 21. *Regression Line and Confidence Intervals of the Probability of Each BDI Score Correlated with Mind Wandering Occurrences*

6.3.3.4. Discussion

In summary, higher self-reports of depressive or negative mood (i.e., higher BDI scores) significantly correlates with higher incidents of off-task thinking but not with mind wandering. The relationship between BDI scores and mind wandering occurrences follows the same trend as off-task thinking, the lack of significance may be due to the fewer counts of mind wandering. Recall that off-task frequencies included all mind wandering frequencies because mind wandering was a form of off-task thinking. Thus, the frequencies of off-task thinking were always higher than mind wandering frequencies.

The relationship between BDI scores and occurrences of off-task thinking follow the literature such that those who are more depressed, tend to mind wander more frequently than those who are minimally depressed. The variance around the regression lines is the within-subject variance because the scores were collapsed across all ten trials for each participant. The variance is the highest at the higher levels of depressive states, which indicates that the participant did not consistently experience off-task thinking with every probe. At lower BDI scores, the smaller variance around the regression lines indicates that participants were more consistent in the responses: either they consistently were off-task thinking or not.

6.3.3.5. Research Question 3

What is the best metric for measuring mind wandering in the cockpit such that the disruptiveness is minimized and the ability to recall the thought content is high?

6.3.3.6. Variables

Three questions in the Post-Driving questionnaire provided data for three different variables. *Meta-awareness* was the variable used to represent responses to the question, "Which method did you babysit your thoughts the most with?" Three response options were provided: clicker, all probe, and self caught and probe. *Re-engagement* was the variable used to represent responses to "Overall, which measurement method allowed you to re-engage the quickest with your original stream of conscious thought you experienced before the probe?" Four response options were provided: clicker, all probe, self caught and probe, and none- they were all the same. *Frequency Accuracy* was the variable to represent responses to the question, "Which mind wandering measurement method was the best at accurately capturing how frequently you mind wandered?" Four response options were provided: clicker, all probe, self caught and probe, and none- they were all the same.

6.3.3.7. Diagnostics and Data Description

Diagnostics were not performed. No data descriptions provided.

6.3.3.8. Analyses and Results

Table 13 lists the frequencies of each metric across each of the three variables studied: Re-engagement, Frequency Accuracy, and Meta-awareness.

Table 13. *Frequency of Response Categories for Three Questions About The Efficacy of Mind Wandering Metrics (n=122)*

Response Options	Re-Engagement (%)	Frequency Accuracy (%)	Meta-Awareness (%)
Clicker	61 (50%)	65 (53%)	42 (34%)
Self Caught/Probe	38 (31%)	39 (32%)	45 (37%)
All Probe	13 (11%)	12 (10%)	48 (39%)
None, they were all the same	10 (8%)	6 (5%)	n/a

A Wald chi-square test was used to assess whether significant differences occurred between each of the three methods for task re-engagement lags. The model fit was significant with a likelihood ratio testing of the global null hypothesis [χ^2(1, n=122)=74.1, p<.001]. A significant difference occurred between the counts for each method [χ^2(1, n=122)=59.95, p<.001] and the odds ratio point estimate for method was .408 with a Wald 95% Confidence Limit between .33 and .51. Using contrast statements, the clicker metric was significantly different from self caught and probe [Wald χ^2 (1, n=122)=59.9, p<.0001]

A Wald chi-square test was used to assess whether significant differences occurred between each of the three methods for Frequency Accuracy. The model fit was significant with a likelihood ratio testing of the global null hypothesis [χ^2(1, n=122)=99.7, p<.001]. A significant difference occurred between the counts for each method [χ^2(1, n=122)=74.35, p<.0001] and the odds ratio point estimate for each method was .338 with a Wald 95% Confidence Limit between 0.26 and 0.43. Using contrast statements, the clicker metric was significantly different from self caught and probe [Wald χ^2 (1, n=122)=74.3, p<.0001]

For Meta-awareness, a Wald chi-square test was used to assess whether significant differences occurred between each of the three methods. The model fit was significant with a likelihood ratio testing of the global null hypothesis [χ^2(1, n=122)=99.7, p<.001]. A non-significant model arose such that there are no significant differences across methods for meta-awareness counts [χ^2(1, n=122)=.63, p=.43] .

6.3.3.9. Discussion

According to the study participants, the clicker metric most accurately captured mind wandering frequency, allowed for the quickest task re-engagement and was the least likely to facilitate conscious monitoring of one's own thoughts. Thus, it was considered the best performer. It was a surprise to find that there were no significant differences between the three metrics in terms of which metric facilitates the conscious monitoring of thoughts the most. Since the clicker metric requires participants to self declare when they are mind wandering, it would be reasonable to assume that this, in addition to the self caught and probe metric, would be the most highly chosen metric to be correlated with meta-awareness. Conversely, because the all probe metric does not require the participant to monitor conscious thought while driving, it was expected to be the least rated metric on the meta-awareness variable. We expected the self caught and probe metric to be ranked somewhere close to the clicker on meta-awareness, but not as highly rated as the clicker metric because of the probe component of that metric. It could be argued that the question, 'babysitting your thoughts,' was not directly measuring meta-awareness. However, during the first week of testing, many participants did not understand what "babysitting your thoughts" meant so it became standard protocol to verbally explain that it meant "monitoring what you are thinking" to reduce measurement error. In addition, "babysitting your thoughts" might only be targeting the mental effort required to monitor thinking. A person may answer this question using two pieces of information, which may in turn be subjected to an array of memory biases (e.g., availability heuristic, suffix effect, etc.). First, the total amount of perceived time spent monitoring relative to the total time spent could be one source of information used to answer this question. Presumably, some ratio of conscious to non—conscious thought is formulated in the participant's mind but how can the participant know how much time has passed in the semi-conscious or unconscious state? A participant may be conscious of their thinking (awareness) but not conscious of thinking about what they are thinking about (meta-awareness). Second, the amount of load experienced during bouts of conscious monitoring may help with memory formation; hence, people remember monitoring their thoughts when a certain threshold of effortful cognitive load is reached or exceeded. If the effortful episodes were salient experiences, it may be remembered more often than non-salient experiences (von Restorff, 1933).

Initially, the metric subjectively interpreted as the least intrusive to the task of driving and the most accurate at capturing mind wandering frequency, would be used in Part IV. However, if the clicker truly was a superior option, using it in the flight simulation in Part IV would imply that thought content could never be vocalized. Additionally, the clicker metric in theory measured thinking with meta-awareness because it would be difficult to recognize mind wandering in the absence of monitoring one's thoughts. We did not however, compare what participants preferred compared to their driving performance.

Even though the clicker method was subjectively interpreted as the best metric, we chose the all probe metric for Part IV so that we could capture a person's thought content. We did not choose the self caught and probe method, because during a self-caught episode in the driving simulation, several participants spent extra time explaining *why* they were having specific thoughts as opposed to *what* they were thinking. We feared that this expanded explanation could distract participants from their flight path more than in the driving simulation.

6.3.4. **Research Question 1**

Does a relationship exist between task type and mind wandering?

6.3.4.1. Variables

The independent variable Probe number and six dependent variables (off-task, wandering, mental processes, cognitive abilities, sensory abilities, psychomotor abilities) were used to answer research question 1 across two separate analyses.

The dependent variables with respect to abilities resulted from the efforts reported in Part II. According to O*Net^{TM}, an ability is an individual's enduring attribute that influences that person's performance as opposed to a *mental process*, which is the cognitive function that involves planning, problem-solving, decision-making, and innovating activities that are performed with job-relevant information. Abilities can be further decomposed into *cognitive abilities* (i.e., those that influence the acquisition and application of knowledge in problem solving), *sensory abilities* (i.e., those that influence visual, auditory and speech perception) and *psychomotor abilities* (i.e., those that influence the capacity to manipulate and control objects).

6.3.4.2. Diagnostics and Data Description

The same diagnostics used for hypothesis 1 were applied here. To describe the data, the means and standard deviations of off-task thoughts and mind wandering across all 10 probes are provided in Table 14 below. Because all off task and mind wandering responses were coded as either a 0 or 1, the means will fall between 0 and 1. Two one-way ANOVAs with probe number as the independent variable and off-task and wandering both as the dependent variables were conducted to understand the data trending. Both ANOVAs were significant: $F(9,1216)=7.21$, $p<.0001$ for off-task and $F(9,1216)=6.15$, $p<.0001$ for mind wandering ($n=1217$).

Figures 22 and 23 below depict the mean off-task thoughts and mind wanderings by probe and since all responses were coded as 0 and 1, all means will be below 1. Significant mean differences exist across all probes in Tukeys HSD post hoc analyses in both figures. In Figure 22, there are significantly mean differences between probe six and the group of probes #2-4 and #7-9. Probe #9 is significantly different from probes #1, #5, #6, and #10. There are significant differences between #1 and #2, but probes #6 and #10 are not significantly different.

Table 14. *Means and standard deviations for off-task and mind wandering variables across all probes.*

Probe	Mean Off-task Thoughts (*SD*)	Mean Mind Wandering *(SD)*
1	.45 (.50)	.27 (.45)
2	.58 (.50)	.47 (.50)
3	.55 (.50)	.37 (.48)
4	.59 (.49)	.40 (.49)
5	.39 (.49)	.30 (.46)
6	.29 (.46)	.17 (.37)
7	.53 (.50)	.37 (.49)
8	.52 (.50)	.39 (.49)
9	.65 (.48)	.45 (.50)
10	.34 (.47)	.18 (.39)
Total	.49 (.50)	.34 (.47)

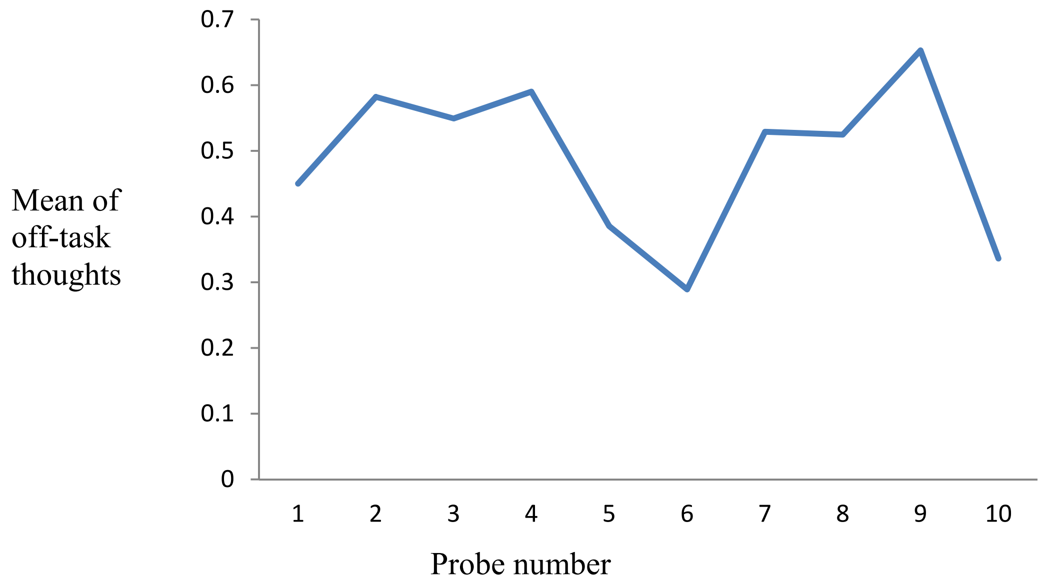


Figure 22. *Mean of off-task thoughts by probe*

Figure 23 has a similar pattern of significant mean differences. Probe #6 is significantly different from probes #2-4 and #7-9. Probe #9 is significantly different from probes #6 and #10 and probe #10 is significantly different from probes #2-4 and #7-9.

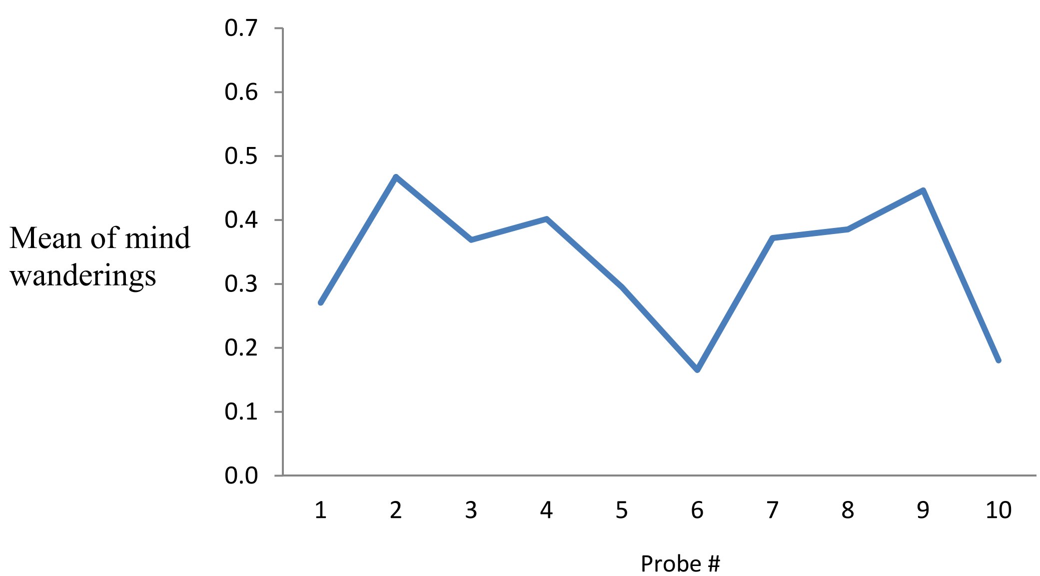


Figure 23. *Means of mind wandering thoughts by probe*

6.3.4.3. Analyses and Results

It is unclear from the literature how to properly classify tasks into task types. To circumvent this limitation, we use probe numbers to index a group of tasks in the driving sequence and then break down the probes into the respective profile of abilities and mental processes. The profile represents the task type. The first analysis described here assessed, "During what probes off-task and mind wandering occurrences were most likely to occur?" using logistic regression. Then, ability and mental process profiles were created for each probe based on the task analytic results obtained from Part II. Statistically significant differences in the frequencies of cognitive abilities and mental process requirements comprising each probe are determined using a contingency table logit analyses.

Logistic Regression. First, two identical fixed-effect logistic regression models were conducted with probe number as the fixed effect and off-task in one model, and probe number and mind wandering in the other model. The results yield a statistically significant main effect of probe on frequencies of off-task [$F(9,1207)=6.65$ $p<.001$] [model fit: -2 log likelihood fit statistics: 1623.25] and a significant main effect of probe number on frequencies of mind wandering [$F(9, 1207)=5.62$, $p<.0001$ model fit: -2 log likelihood fit statistic=1496.40]. All estimates, standard errors and *p*-values are listed in Table 15; results are grouped by each dependent variable. Recall that participants drove the same test route more than once; thus, probes #1-5 are identical in geographic location to probes #6-10. Table 16 lists the results for all ten probes, but identical probes are juxtaposed in the table. Notice that both dependent variables had similar data trending. Paired probe combinations #2 and #7, #3 and #8, #4 and #9 were all statistically significant with positive estimates on both dependent variables. The occurrence of significance for both probes in a probe combination (e.g., #2 and #7) gives some indication in the robustness of the result. Probes #6 and #10 had negative estimates.

Table 15. *Estimates, Standard Errors and p-Values for Two Logistic Regressions with Probe as the Independent Variable and Off-task and Wandering as the Dependent Variable*

	DV=Off-task			DV=Wandering		
Probe	Estimate	SE	p-value	Estimate	SE	p-value
1	0.48	0.26	.0676	0.52	0.31	.0940
6	-0.22	0.28	.4318	-0.11	0.34	0.7567
2	1.01	0.27	.0001 *	1.38	0.30	<.001 *
7	0.80	0.26	.0026 *	0.99	0.34	.0010 *
3	0.88	0.26	.0009 *	0.98	0.30	.0012 *
8	0.78	0.26	.0032 *	1.05	0.30	.0005 *
4	1.05	0.27	<.0001 *	1.12	0.30	.0002 *
9	1.31	0.27	<.0001 *	1.30	0.30	<.0001*
5	0.21	0.27	.4242	0.64	0.31	.0369 *
10	-0.68	0.19	.0004 *	-1.51	0.24	<.0001*

* $p<.05$

Contingency Table Logit Analysis. The purpose of this analysis is understand the task characteristics (i.e., cognitive abilities and mental processes) that may be contributing to the robust significant results related to probe pairs #2 and #7, #3 and #8, #4 and #9. From Part II, each task within the two-second period surrounding a probe, and their respective abilities (A) and mental processes (MP) requirements were identified and summed for each probe.

Abilities were further decomposed into cognitive (CA), sensory (SA), and psychomotor (PM). Appendix F lists the CAs, SAs, PMs and MPs of each task within each probe and Table 16 lists the aggregate frequency for each probe. A contingency table logit analysis, which is a multinomial logistic regression, was conducted to determine the relationship between probe number and frequencies of CAs, SAs, PMs, and MPs. The overall model was statistically non-significant [$\chi^2(3, n=208)=1.30, p=.73$]. Thus, there were no significant differences in counts of cognitive abilities and mental process requirements (within columns) across all five probes.

Table 16. *Contingency Table Abilities and Mental Processes Across 5 Probes*

Probe	Cognitive Abilities (CA)	Sensory Abilities (SA)	Psychomotor Abilities (PM)	Mental Processes (MP)	Total
1 & 6	10	13	8	24	55
2 & 7	7	13	0	25	45
3 & 8	4	8	0	15	27
4 & 9	9	4	6	19	38
5 & 10	7	9	5	22	43

* $p<.05$

6.3.4.4. Discussion

Probes #2-4 all had statistically significant relationships with occurrences of off-task thinking and mind wandering. If different probes had different frequencies, was the difference due to the type of task the participant was experiencing? We found no significant difference in the frequencies of CAs, SAs, PMs and MPs at each probe. However, the total number of abilities and mental processes did decline from probes #1 to #3 and then rise from probes #4 to #5, but even that result was not significant across all five probes. To explore the lack of significance further, let's investigate what the participant is doing that requires such abilities and processes.

Generally, probes #1 and #5 occurr at moments of relatively higher traffic density and steering maneuver requirements such that the participant was more likely to be actively engaged in the task of driving. In contrast, probes #2 and #3 occur after long straightaways of driving with little traffic density; thus enabling the participant to become more mentally disengaged from the task. Probe #4 occurs right after a steering input execution, but also after a long straightaway of driving. Probes #1, #4 and #5 require steering maneuvers which contribute to their psychomotor ability requirements compared to probes #2 and #3, which do not require these maneuvers. It is possible that psychomotor requirements, which require a level of attentiveness for executing behaviors, raise the awareness level just enough to inhibit off-task thinking in this environment. This possibility is supported at probe #4, when the participants experienced several seconds along an uneventful straightaway, before they complete a turn and finally experience the probe. Many participants self-reported that off-task thinking occurs in the straightaway, but at the time of the probe, participants were just re-engaging with task-based thinking as they were coming out of the turn. Since re-engagement is not always abrupt, it requires additional time to complete a 90-degree turn to bring individuals back to task engagement. Further research is needed to determine whether tasks with psychomotor ability requirements, as well as other ability requirements, inhibit off-task thinking.

6.3.5. **Research Question 2**

Is there a probabilistic relationship between certain types of mind wandering, certain behaviors and adverse consequences?

6.3.5.1. Variables

The independent variable is probe number. Three dependent variables are chosen to approximate the adverse performance consequences of distracted thinking: seconds of speeding, number of lane deviations and number of collision during the probe period. Because only two collisions occurred across all 1200 observations, the number of collisions was dropped from this analysis. All video files are manually reviewed to assess the seconds of speeding and number of lane deviations. Because this process was labor intensive, data is collected across only five probes when participants are using the all probe metric. The all probe metric is assessed because the self caught and probe condition often have participants verbalizing self caught thoughts for several seconds leading up to the probe, which potentially interferes with the data collection at each probe point.

6.3.5.2. Diagnostics and Data Description

The distributions of seconds of speeding and number of deviations are both assessed for model fit. Seconds of speeding is a zero inflated distribution with a range of values between 0 and 14, a mean of 5.02 (n=590), *SD* of 4.79, and a positive skew (skew= 0.514). The distribution is almost tri-modal with a heavy concentration of values at 0, 6 and 14 seconds. The number of lane deviations distribution is also zero inflated with a range of values between 0 and 11, a mean of 1.45 (*n*=590), *SD* of 1.46, and a positive skew (1.73). The analysis includes zero-inflated, negative binomial distributions.

6.3.5.3. Analyses and Results

Two sets of analysis address this research question. The first evaluates whether significant differences in mean numbers of adverse consequences (e.g., seconds of speeding, number of lane deviations, number of collisions) occur across probes. The second is a qualitative assessment of the types of off-task thoughts (e.g., count of off-task thoughts that lack meta-awareness, whether salient thoughts were recalled, etc.) that occur at each probe. Each statistical analysis is introduced by the dependent variable below.

Seconds of Speeding. A mixed model regression analysis (fixed effect=probe #, random effect=participant) was conducted using a zero-inflated negative binomial distribution to represent the seconds of speeding distribution. The -2 log likelihood fit statistic was 3128.46 and the -2 log likelihood for the conditional distribution (seconds of speeding/r. effects) was 2713.5. A significant main effect resulted for probe number [$F(4,466)=176.44$, $p<.0001$, $n=590$] and the covariance parameter estimate for participant number was 0.59, ($SE=.10$). The estimates, standard errors and adjusted *p*-values for each of the five probes are listed in Table 17. Probes #2- 5 had significant positive estimates meaning that these probes significantly explained the variance in the seconds of speeding. In order to assess the significant mean differences between each probe for seconds of speeding, a Studentized Maximum Modulus (SMM) post hoc analysis is conducted and results are listed in Table 18. SMM post hoc analyses are used in conjunction with zero-inflated negative binomial distributions to adjust for the correct degrees of freedom and the adjusted p-values are reported in Table 18. The result indicated that probes #3 and #4 had no significant differences in the number of seconds of speeding but all other mean differences were significant. Probe #1 is negatively related to all other probes meaning that it has significantly less seconds of speeding than any other probe. Probe #2 has the greatest number of seconds of speeding and each probe following had respectively less. When rank ordering in descending order the greatest number of seconds of speeding to the least, probe #2 would be the highest number of seconds of speeding followed by probe #3, probe #4, probe #5 and then probe #1. All ranks would be significantly different with the exception of probes #3 and #4.

Table 17. *Estimates and Standard Errors (SE) in the Seconds of Speeding For Each Probe (n=590)*

Probe	Mean Seconds of Speeding	Estimate	*SE*	p-value
1	1.17	-.06	.11	.58
2	9.47	2.02	.08	<.0001*
3	5.07	1.40	.08	<.0001*
4	5.53	1.49	.08	<.0001*
5	3.79	1.12	.09	<.0001*

**p*<.05

Number of Deviations. A mixed model regression analysis (fixed effect=probe number, random effect=participant) was conducted using a zero-inflated Poisson distribution. The -2 log likelihood fit statistic was 1728.07and the -2 log likelihood for the conditional distribution (seconds of speeding/r. effects) was 1560.37. A significant main effect for probe number [$F(4,466)=29.04$, $p<.0001$, $n=590$] was found and the covariance parameter estimate for participant was 0.148, ($SE=.0437$). The estimates, *SE* and adjusted *p*-values for each of the five probes are listed in Table 19.

Table 18. *Mean Differences of Seconds of Speeding Across Between Probes (n=590)*

Probe	Probe	Mean Seconds Of Speeding Differences	Estimate Differences	*SE*	Adj. p-value
1	2	-8.3036	-2.0857	0.09015	<.0001*
1	3	-3.9855	-1.4641	0.09427	<.0001*
1	4	-4.3582	-1.5515	0.09363	<.0001*
1	5	-2.61399	-1.1835	0.09791	<.0001*
2	3	4.4081	0.6216	0.05033	<.0001*
2	4	3.9454	0.5342	0.04912	<.0001*
2	5	5.6897	0.9022	0.05685	<.0001*
3	4	-0.4627	-0.0874	0.05629	0.7228
3	5	1.2816	0.2805	0.06314	0.0001*
4	5	1.7443	0.3679	0.06211	<.0001*

* $p<.05$

Probes #2-5 were statistically significant; probe #2 had significantly less deviation lanes and probes #3-5 had significantly more lane deviations. The results of a Studentized Maximum Modulus (SMM) post-hoc analysis listed in Table 20 indicated that probes #3 and #4 had no significant differences in the number of seconds of speeding but all other mean differences were significant. Probes #4 and #5 are not significantly different from each other, forming a group. Probes #1 and #2 are not significantly different from each other, forming another

Table 19. *Estimates and Standard Errors (SE) in the Number of Deviations For Each Probe (n=590)*

Probe	Mean Number of Deviations	Estimate	*SE*	*p*-value
1	1.00	-0.08	.10	.4046
2	0.74	-0.40	.12	.0005*
3	1.40	0.24	.09	.0077*
4	2.26	0.72	.08	<.0001*
5	1.85	0.52	.08	<.0001*

*$p<.05$

Table 20. *Mean Differences of Number of Lane Deviations Between Probes (n=590)*

Probe	Probe	Mean Differences in Number of Deviations	Mean Differences in the Estimates	*SE*	Adj. *p*-value
1	2	0.260	0.3174	0.1406	0.2189
1	3	-0.3915	-0.3222	0.1199	0.0719
1	4	-1.2520	-0.801	0.1102	<.0001
1	5	-0.8424	-0.6033	0.1149	<.0001
2	3	-0.6605	-0.6396	0.1316	<.0001
2	4	-1.5210	-1.1184	0.1229	<.0001
2	5	-1.1114	-0.9206	0.1271	<.0001
3	4	-0.8605	-0.4788	0.09838	<.0001
3	5	-0.4509	-0.281	0.1036	0.067
4	5	0.4096	0.1977	0.09217	0.2799

group. Generally, probes #1 and #2 have significantly less number of lane deviations than probes #4 and #5. Group 3 is not significantly different from group 1 and 5, but it is significantly different groups 2 and 4.

The analysis involves assessing the types of off-task thoughts occurring at each probe to begin to understand how these types of thoughts related to adverse consequences of off-task thinking. To answer the research question, we need to build on the above results by adding other information about what types of mind wandering do participants experience and at what probes were these wanderings experienced.

Table 21. *Characteristics of Thoughts at Each Probe in the All Probe Metric Condition*

Probe #	Off-task+ Wandering (%)	Total Off-task only (%)	Total Wandering (%)	Wandering Estimates	Total Salient thoughts (% of wanderings)	Total Lacked meta-awareness (% of wanderings)	Mean Temporal Rating (*SD*)	Mean Functional Rating (*SD*)	Mean Conceptual Rating (*SD*)	Mean Overall Rating (*SD*)
1	37 (31%)	15 (13%)	22 (18%)	-0.11	10 (45%)	8 (36%)	43.6 (36.4)	30.8 (31.7)	43.8 (35.8)	45.9 (35.8)
2	63 (53%)	17 (14%)	46 (38%)	0.99*	21 (46%)	15 (33%)	43.9 (38.5)	38.1 (33.0)	34.9 (34.9)	38.1 (28.3)
3	64 (53%)	17 (14%)	47 (39%)	1.05*	18 (38%)	12 (26%)	45.9 (40.2)	41.5 (34.8)	43.0 (37.3)	47.8 (35.9)
4	74 (62%)	21 (18%)	53 (44%)	1.30*	22 (42%)	17 (32%)	53.3 (38.8)	41.6 (32.9)	37.7 (33.4)	44.6 (31.3)
5	41 (34%)	19 (16%)	22 (18%)	-1.51*	6 (27%)	10 (45%)	54.2 (35.7)	43.7 (34.9)	48.1 (33.4)	50.6 (33.1)

* *p*<.05 from analysis in Table 15

The last four columns on the right of Table 21 display the mean percentages ratings and standard deviations of each type of wandering by probe number. The ratings were on a scale of 0-100 and high ratings indicate that the thought content the participant was experiencing had high relatedness to task related thinking. Low ratings mean that there is low relatedness. Ratings were provided verbally by each participant at each probe point as a mechanism to classify their own thought contents. For example, a low rating for temporal relatedness means that the thought content was about an event that occurred in the distant past or distant future from the present moment. A high rating for functional relatedness means that the probed thoughts involved a series of thoughts, each built upon the previous thought. A low rating would indicate that thoughts were not a function of the previous thoughts. High conceptual relatedness means that the thought content in the probe were driving related thoughts. A low overall rating means that overall, these thoughts discovered in the probe were minimally related to driving related thoughts.

In Table 21, two totals were included to the left of the four relatedness ratings columns: the total number of participants at each probe who thought about at least one of their salient thoughts on their Salient Life Events questionnaire, labeled total salient thoughts, and the total number of participants who experienced thoughts that lacked meta-awareness or conscious awareness of their own thoughts at each probe, labeled total lacked meta-awareness. In addition, a percentage of the total salient thoughts out of the total number of mind wanderings were included in the "Total salient thoughts" column in parenthesis. Also, the percentage of the total number of thoughts that lacked meta-awareness out of the total number of wanderings was included in the "Total lacked meta-awareness" column in parenthesis.

Recall that all mind wanderings are redundantly coded as off-task thoughts; however, some off-task thoughts are not mind wanderings. The left column titled "Off-task +wanderings" is the total number of participants at each probe who experience either an off-task thought or a wandering. The next column titled, "Total off-task only" are the number of participants whose thoughts are self-classified as off-task, but not as mind wandering. For example, a person who is 100 feet from a stop light and is probed might explain that he is thinking about a turn he made two intersections ago and the person classifies this thought as off-task, because he is not thinking about the approaching stop light. Given this situation, an example of a mind wandering thought would be thoughts about getting a hair cut in the coming week. The "total wandering" column is the total number of participants at each probe whose thoughts were self-classified as mind wanderings. The estimates from Table 15 called "Wandering estimates" (Probes #6-10) are also included to indicate the significant mean mind wandering frequencies at each probe.

No statistical analyses could be conducted to assess the relationship between types of wandering and adverse consequences. One would speculate that an analysis could be conducted on the relationship between probes numbers and mean ratings for temporal, function, conceptual and overall relatedness. However, this is risky because mean ratings at each probe had drastically different values of *n*, which would be calculated from the total number participants who experienced of off-task thoughts at that probe. Not every participant experienced an off-task or mind wandering occurrence at each probe. ANOVA models are fairly robust to different group sizes as long as each group has homoscedasticity of variance, which at a glance, could be possible. But if all probes have similar values for means and standard deviations, what do the results mean? Thus, no formal analyses were conducted beyond the results reported in Tables 17-21.

6.4.Discussion

To answer this research question, we triangulated several analyses on human performance with the type of mind wanderings experienced at each probe. Generally, probes #2-5 had the most number of performance decrements. The seconds of speeding seemed positively correlated with the number of lane deviations with the exception of probe #2, in which a long straightaway occurs with no steering inputs and, thus, the participant has less lane deviations and several seconds of speeding. Coincidentally, probe #2 has the lowest overall relatedness ratings (38.1) indicating that on these long straightaways, the thoughts that were probed were less related to driving-related thoughts than other probes. At probe #2, a total of 63 participants experienced off-task thinking and 73% (46 participants) of those participants classified their thoughts as mind wandering. Compare that to the highest overall relatedness rating at probe #5 with a mean rating of 50.6, and out of the 41 participants who experienced off-task thinking, ~54% or 22 participants classified their thought content as mind wandering. To further understand the types of thoughts experienced at probes #2 and #5, we'll assess the temporal, functional and conceptual relatedness ratings at each of the probe points. If you rank order the relatedness ratings from highest to lowest at probe #2, temporal relatedness has the highest mean (43.9) followed by functional (38.1) and conceptual relatedness (34.9). Even though the standard deviations are high, they are relatively the same. The probe #5 rank order would be temporal (54.2), conceptual (48.1) and functional (43.7). Comparatively, thoughts at probe #2 were about events that occurred further in time as indicated by the lower temporal rating. Thoughts in probe #2 were less likely to be related to previous thoughts as indicated by the lower functional rating and thoughts experienced were less likely to be related to driving related thoughts. To give an example of a thought classified as such at probe #2, a participant may at one point be thinking about driving and then suddenly remember a wedding that occurred five years ago. At probe #5, the participant may be experiencing more driving-related thoughts that were built off of a previous thought and were more conceptually related to driving thoughts. For example, that participant may be remembering the wrong turn made during the practice driving session, which was cued off of the scenery in the driving simulation. That participant may have experienced a few thoughts that were all cued from the previous thought (i.e., higher functional relatedness). If probe #2 compared to probe #5 had higher frequencies of thoughts unrelated on three different dimensions to the driving task, perhaps a higher frequency of wanderings to salient life events would be possible. According to the data, 46% of the wanderings involve participants self-reported salient life events at probe #2 compared to probe #5 (27%).

Another relationship between human performance metrics (e.g., lane deviations and seconds of speeding), mind wandering incidents, the percentage of those wanderings that focus on life salient events is whether the participant was experiencing any awareness of their thoughts (e.g., Total lacked meta-awareness). At probe #2, the minimum number of mean lane deviations occurred (0.74) with the highest mean number of seconds of speeding (9.47) and at this probe, 46 participants experienced mind wandering occurrences and 33% of those participants (15) were not conscious of their thinking until the probe. Compare this to probe #5, in which the mean number of lane deviations (1.85) and mean number of seconds of speeding (3.79) are related to 22 participants experiencing mind wandering episodes and of those 22, 10 or 45% of those individuals were not conscious of their thoughts until probed. Thus, it is not clear whether long straightaways of driving (e.g., probe 2) facilitated occurrences of mind wanderings that lacked meta-awareness (e.g., 33% of the wanderings) compared to the series of driving simulation inputs (e.g., braking, turning) required at probe #5 (48% of the wanderings).

Two hypotheses based on the results of this work might be useful to study in the future. Hypothesis 1: Long straightaways with minimal system inputs compared to driving with more system inputs may be correlated with higher mean frequencies of mind wandering than with thought content that are more cognitively disengaged from task-related thinking on three different dimensions (e.g., temporal, functional and conceptual). Hypothesis 2: Greater performance decrements occur with greater frequencies of mind wandering.

7. Part IV. The Flight Simulation Study

We now describe Part IV, which is the study of off-task thinking in a flight simulation. This section consists of a discussion of the study purpose, method (i.e., participants, apparatus and stimuli), procedure, results and discussion. Because this is a qualitative data analysis, the analysis, results and discussion are combined into one section.

7.1.Purpose

The purpose of Part IV is to measure mind wandering occurrences that pilots may experience in the cockpit and to describe the types of thoughts these pilots experience.

7.2.Method

This section begins with a description of the sample, the apparatus and stimuli used in this study, and the procedure used.

7.2.1. Participants

Participants were required to possess a current fixed-wing certificate to participate in this study. Participants were recruited from two local flight schools at local airports in the Southeast portion of the United States. Recruitment occurred through recommendation and endorsement of a flight instructor, flight school listserv announcements and flyers posted at the local airports. The target sample size was 30 completed records, however non-response reduced the yield to 20 total participants: three participants served as beta testers and 17 participants served as the study sample. The data reported here is limited to the 17 participants.

All 17 participants were native English speaking males, with a mean age of 44.9 years (*SD*=15.8 years, min=18 years, max=68 years). All participants had a current certificate but two were not actively flying and one individual self-reported to be employed by a carrier. Eight participants were currently working on a certificate: 5 individuals on the Instrument Flight Rules (IFR) certificate, 1 individual on the Certified Flight Instructor (CFI) certificate, 2 individuals on their private pilot certificate, 7 participants were not working on a certificate and 2 did not respond. The mean number of accumulated flight hours was 813.2 (*SD*=1396.9 hours, min=60, max=5300). When asked how many days in a single month participants typically fly, the mean was 4.6 days (*SD*=4.7, min=0 and max=20). When asked what certificates the participant currently held, 11 individuals self-reported to have had a private pilot certificate, 4 had a commercial pilot certificates, 2 had a CFI, 2 had a CFI Instrument (CFII) certificate, 1 had a Multi-Engine Instructor (MEI) certificate, 0 had a complete flight engineer certificate, but 3 had the written flight engineer certificate, and 0 had an Airline Transport Pilot (ATP) certificate. When asked how many hours a month on average the participant flew under the various FARs, the mean for Part 91 was 8.2 hours (*SD*=6.8 hours), while nobody accrued hours under Part 121and part 135.

7.2.2. Stimuli and Apparatus

Stimuli. Microsoft Flight Simulator X was purchased to fit the research requirements. Since we expected that all pilots had trained on a Cessna 172, which is an aircraft commonly used in training, we chose this type of aircraft to reduce sampling constraints. A simulated 40-minute route was chosen between two local general aviation airports that was not only relatively easy to fly (e.g., no mountain ranges, no weather-related difficulties, etc.), but familiar to test participants. This route borders on restricted airspace surrounding military installations that requires participants to circumvent the military airspace, but without performing difficult flight route deviations. All settings related to the simulated flight route are described in the Apparatus section below.

All participants completed a consent form, a Post-Flight questionnaire and the adult version of the Life Stress Scale, which is also called the Social Readjustment Rating Scale, (Holmes & Rahe, 1967). The Post-Flight questionnaire collects basic demographic information as well as other experiences and opinions during the flight simulation. The Life Stress Scale assesses the level of risk for developing mental illness depending on the number of stressful life experiences that participants have experienced within the last six months. High scores suggest high levels of life stress, which may induce various forms of physical or mental illness. All scores attributed to each inventory item are summed and scores of 300 or more are "at risk of illness", scores of 150-299 are of moderate risk of illness and scores less than 150 are of slight risk for developing illness. Copies of the Life Stress Scale and the Post-

Flight questionnaire are located in Appendix L. In addition, a map was provided to participants as well as a checklist for the Cessna C172SP Skyhawk, but these artifacts are not included in Appendix L.

Apparatus. Microsoft Flight Simulator X allows the customer to not only select the type of aircraft to fly but also allows customers to adjust flight settings which potentially affect flight performance (e.g., time of day, weather, air traffic, etc.). In this study, all participants flew a non-glass cockpit Cessna C172SP Skyhawk, one of the most common aircraft flown by general aviators. Microsoft Flight Simulator X has many different software settings. The display settings control graphics quality, scenery quality, weather quality, etc. Each of these categories under the display settings is represented with adjustable sliders with a descriptive term (e.g., low, medium, high, etc.) but no numerical readouts. Therefore, a percentage of the total slider length was used to represent the slider value. The simulator was configured as follows: the graphics quality was set to *low* (45%), the aircraft was set to *very low* (28%), the scenery was set to *medium low* (57%), the weather was set to *medium low* (58%), and the traffic was set to *low* (40%). Other weather settings for this study included: 1) scattered cumulus clouds (setting 4/8) between 1500 and 3000 feet, 2) no precipitation, 3) visibility at 20 miles and 4) a light 8 knots wind speed at 29° from the west with light turbulence as participants travel in southeasterly direction. The temperature was set to 68° F, dew point 66° F and altimeter at 29.83 inches Hg. All participants flew at local time of 12:21 Easter Standard Time in December in a Southeastern state where weather is more stable. Visual Flight Rules were required and auto pilot was disabled. No failures were set for any portion of the flight. All participants began the flight with maximum allowable fuel load, the engines started, the parking brake set, a pre-set DME (Distance Measuring Equipment) and Air Traffic Control communications were turned off.

The hardware included a Dell Latitude E6500 that displayed the flight simulation on a 42” Vizio Class LED smart wall-mounted monitor. The flight controls consists of Saitek three lever pro throttles, Saitek pro yoke and Saitek pro flight rudder pedals. The yoke and throttle were affixed to the desk in each airport conference room where the study was conducted.

7.2.3. Procedure

All participants underwent the same procedure. After informed consent was provided, participants were received a verbal overview of the study activities. Then, participants were introduced to the probing method and the type of information that they could expect to be collected in a single probe. Due to time constraints, participants were not provided a practice trial for probing. During a single probe, participants were asked to classify whether off-task thinking and mind wandering was experienced at the time of the probe, and what the thought content was at the time of the probe, including the 14 seconds leading up to the probe. If off-task thinking or mind wandering was experienced, participants then provided ratings on temporal, functional, conceptual and overall relatedness to the task at hand. These types of relatedness were thoroughly explained with examples before the flight began. Participants were then introduced to all controls in the cockpit simulator by a confederate. All pilots received a paper map of the flight route as well as a flight checklist for the C172SP. Pilots who preferred to use their own personal mapping function on their iPad were permitted.

It was originally intended to allow pilots to fly the complete route from take off to touch down, including all pre-take-off checklist items and a few instrumentation failures to see if these were noticed during times of off-task thinking. However, in beta testing, pilots were too busy to experience off-task thinking. Beta testers tended to become anxious when difficulties arose, thus reduced the workload to a minimal amount to facilitate off-task thinking. For example, instrumentation failures were removed and the simpler flight plan enabled participants to take off more quickly and reach an altitude where they could fly straight and level for at least 30 minutes. To ensure pilots were flying the simulation with as much rigor as they would in live flight, pilots were asked to treat this study like a "check ride" which is a FAA flight test that must be passed in order to receive a certificate. They were also told to "pretend" that I was not present in the cockpit and at each probe, they were instructed to refrain from explaining rationale for their actions and refrain from defining terminology. Pilots were then instructed to fly the shortest route between two airports, to fly VFR (visual flight rules), to ascend to 3000 feet and maintain altitude until otherwise instructed. To maximize the amount of straight and level flight, pilots were instructed to not descend to the destination airport, but to maintain altitude and cruise over the airport until the experiment was terminated by the researcher. Each participant began the flight with the aircraft engine started and the parking brake set, the DME set, and the aircraft positioned at the point of the take off roll on the runway. Pilots were asked to talk aloud while they completed the checklist starting from the engine start portion of the checklist provided. Pilots were probed every 7 minutes starting from the takeoff roll. Probes were set at 7-minute waypoints, because every pilot did not fly the same route to the destination airport.

After completing 5 probes (35 minutes of flight), the scenario was halted, regardless of whether the participant in the vicinity of the destination airport.

Immediately following the flight, the participants were given a Life Stress Scale inventory and the Post-Flight questionnaire in a separate room to maintain privacy. Then pilots were compensated, debriefed and released.

7.3. Analyses and Results

Because the sample size was small and the occurrence of mind wandering was low, only descriptive statistics will be reported here, starting with a discussion of the thought content collected at each probe, a discussion of the demographics, and then associations between thought content and demographics. Example thoughts are provided, however, personally identifiable information was replaced with generic information in square brackets "[…]".

Thought content. During a probe, participants were asked if they believed they experienced off-task thinking at the time of the probe and whether that thinking was considered mind wandering. A total of 17 participants were probed five times for a total of 90 probes. Fifteen probes out of 90 probes or 16.7% involved off-task thinking. Out of those 15 probes, 7 probes were classified as mind wandering occurrences. Table 22 presents a break down of the total number of off-task and mind wandering events across all five probe points. However, after the simulated flight, participants were verbally asked how many times they might have mind wandered in between the probes, which yielded an additional 75 reported incidents across 17 participants (*M*=4.17 times per participant, *SD*=4.90 times).

Table 22. *Total number of off-task and mind wandering occurrences by probe according to self-classifications by 17 participants*

Probe	Off-task?	Mind wandering?
1	4	2
2	0	0
3	3	1
4	2	2
5	6	2

Given the relatively small number of mind wandering occurrences and wide variance in the ratings, no descriptive statistics of temporal, functional, and conceptual relatedness provided by the participant at the probe were assessed. However, in the Post-Flight questionnaire, participants were asked to characterize the off-task thoughts that they experienced during the simulated flight, regardless of whether they experienced the thought at the time of the probe or between probes. On average, participants self-reported that: 17.59% (*SD*=15.82%) of their off-task thoughts involved the re-visitation of past flying events, 17% (*SD*=24.29%) involved a sense of familiarity to a place visited in the past but no tangible thoughts came to mind, 15% (*SD*=26.81%) were ‘zone outs’, 4.76% (*SD*=7.77%) involved planning for future flying events, 4.41% (*SD*=9.98%) involved thinking that was driven by physiological cues such as hunger, thirst, fatigue, pain, etc., 2.65% of the thoughts (*SD*=7.93%) were of stressful events and circumstanced in their personal life, and 2.35% (*SD*=7.52%) involved humming songs in their minds. We also asked participants what percent of their time while in flight did they spend monitoring their thoughts and the mean percentage was 38.41% of their flight time was spent monitoring their own thoughts (*SD*=39.45%). The mean percentage of off-task thoughts that they experienced while their state of consciousness about the task transitioned from unconscious to conscious and vice versa was 26.35% of off-task thoughts (*SD*=27.33%).

The above descriptions of the thoughts experienced were from the vantage point of the participant. All on-task and off-task thought content was coded from all 90 probes. Codes were aggregated into categories; however, not all categories were mutually exclusive, so a single code could be counted among multiple categories. For rigor, the content of each probe was broken down into the number of thoughts comprising each probe. The following example from a single probe was counted as four thoughts; each thought numbered in parentheses.

> *I'm looking at the speed (1). Some problems here (2). The VOR is not sensing right (3). The DME is not sensing the right speed (4).*

Most of the thoughts were denoted by a single sentence, but if one sentence had more than one idea, each idea was counted as a thought. A total of 326 individual thoughts were collected across all 90 probes, averaging 3.62 thoughts per probe. Figure 24 provides the breakdowns in the total thoughts per probe.

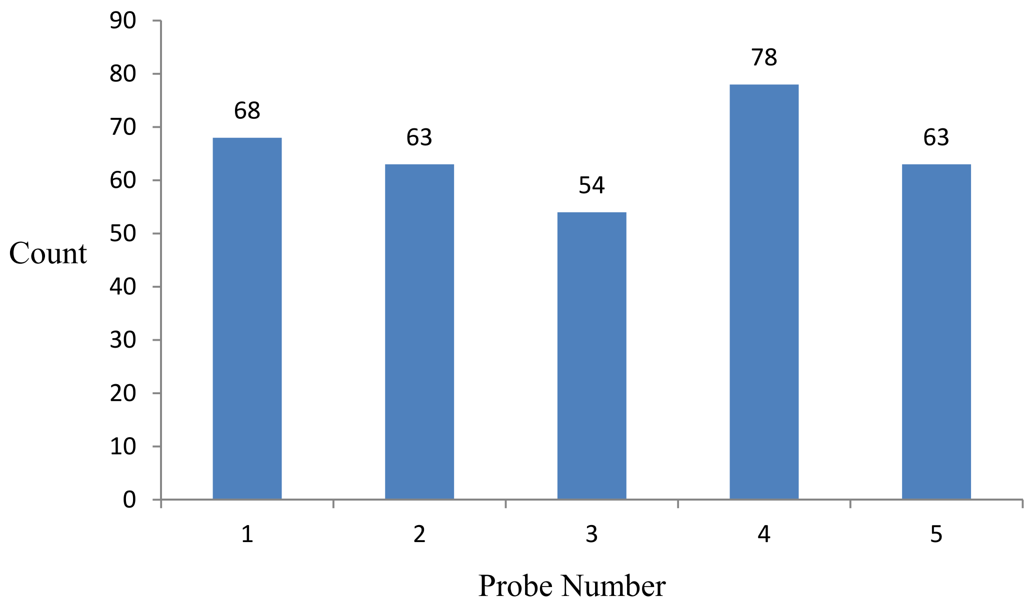


Figure 24. *The number of thoughts per probe (n=326)*

The first thought category is called ‘thought tense’ and the frequencies of prospective, retrospective and present thoughts across all five probes appear in Figure 25. A total 79 out of 326 thoughts (24.23%) were about musings in the present tense, 41 thoughts (12.58%) were prospective or about events in the future, 36 (11.04%) were retrospective or about events in the past, and 170 thoughts (52.15%) had no temporal quality.

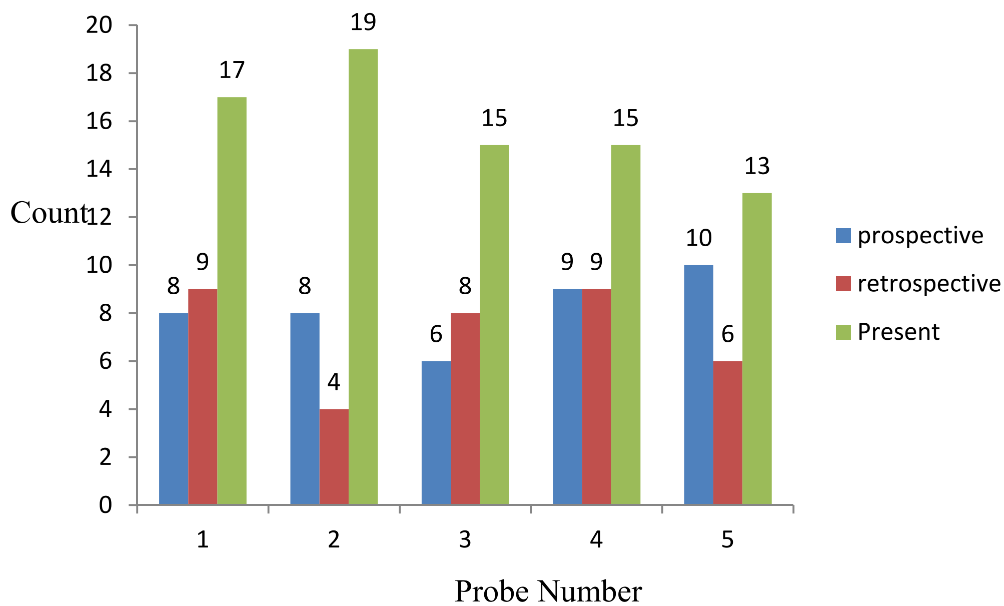


Figure 25. *Thought Tense; Frequencies of prospective, retrospective and thoughts about the present moment across all five probes (n=326)*

Second, thoughts were coded by level of awareness that participants self reported, labeled "Degrees of Awareness" (see Table 23). During a probe, participants would disclose their thought content and the researcher would inquire about their level of awareness, if the level was not mentioned in the content. Because participants often had different levels of consciousness for each thought, especially if the participant was drifting in and out of consciousness across a chain of thoughts, each thought was coded separately in this category. Mutual exclusive codes do not exist here. If a participant's self-report was "full awareness," he or she could be fully aware of their own thinking and still be "between thoughts". In fact, we find that "between thoughts" could be redundantly coded with all codes except "zoning out." However, it is not likely that a participant's self-report could be "noticing" and "zoning out" on the same thought, as these categories are separate by definition.

Table 23. *Degrees of Awareness; Thoughts grouped by degrees of awareness (n=326)*

Code	# of thoughts (percentage)	Most common Probe of occurrence > least common probe	Definition
Full awareness	311 (95.40%)	Probe 2 > Probe 4	The person self-reports to be fully aware or conscious of one's thoughts, feelings, and musings.
Noticing	14 (4.29%)	Probe 3, 4, & 5 (tied rank) > Probe 1	The person is attending to items in the environment and the person feels alert. The person is not experiencing any monologues in their minds about anything while noticing is occurring.
Between thoughts	2 (0.62%)	Probe 3 & 4 (tied rank) > Probe 1, 2, & 5 (tied rank)	At the time of the probe, the person often said they were thinking about "nothing" but they just finished a thought and were about to start thinking about another topic.
Zone Out	1 (0.31%)	Probe 4 > Probe 1, 2, 3 & 5 (tied rank)	This person completely loses awareness of thought processes even though thinking is occurring. At the time of the probe, the thoughts are brought forth to consciousness but seconds before the probe, the thoughts were not conscious.

Third, the thoughts were categorized with respect to the thought's semantic content. In this coding exercise, each thought per probe was assigned one or more codes; thus, *n* is 90 probes rather than 326 thoughts. In Table 24, the columns from left to right are: the code assigned, the percentage of probes with the assigned code, the probes that most commonly and least commonly contain the codes, the definition of the code and an example statement(s) from participants. None of these codes are mutually exclusive: for example, a thought about an action or behavior might also be about an event.

Generally, most of the probes contained thoughts about actions and behaviors executed inside and outside of the task environment. One would expect to see consistent patterns about the probes that involve contemplation (i.e., planning, staging, rehashing), but this was not the case. Probe #2 had the highest frequency of rehashing and staging, but the lowest frequency for planning. Probe #2 also had the least common occurrence of thoughts about events and actions. At probe #2, participants were the least likely to think about conceptually unrelated ideas as well as events and actions or behaviors, but were still more likely to be thinking about staging and rehashing. Thus, probe #2 must be a watershed between the kinds of thoughts experienced at probes #1 and #3-5. To frame this thinking in terms of events related to the flight, we'll discuss flight-related factors that may have contributed.

Not all participants were at the same physical location in the flight at each probe because probes were temporally defined at 7-minute intervals rather than geographically markers. In general, probe #1 was executed when pilots were either on the ground doing checklists, during the take off roll or slightly after the take off roll; thus, it makes sense that navigation was not crossing the pilot's mind at this probe point, but instead their thoughts would be about actions and behaviors related to take off. Probe #2 was during the ascent, which involves an approximately 160-degree clockwise turn after take off to navigate to the destination airport. Thus, navigation and staging was important at this probe point, but planning was not as important. Also, some participants experienced difficulties getting adjusted to the controls of the aircraft which might have led to higher frequencies of rehashing as they revisit their past performance. For example, many participants verbally commented that the yoke was more sensitive than they were accustomed to in a real cockpit. In addition, the task load may be comparatively higher at probe #2 with check-listing, trimming the aircraft and navigation, forcing participants to engage with the task of flying and potentially contributing to the lowest frequencies of conceptually unrelated thoughts to flight. Probe #3-5 were taken during straight and level flight but at probes #4 and #5, some participants were starting to visually look for the destination airport and to reference aviation maps. Thus, at probe #4, we see high levels of planning, thoughts about actions and behaviors and the second highest frequency of navigational thoughts. At probe #5, some of the participants already reached their destination and relaxed their minds to think off-task; thus the highest frequencies of conceptually unrelated thinking and events occurred, here.

Table 24. *Semantic content of thoughts collected at each probe (n=90)*

Codes	# of examples (percentage)	Most common Probe of occurrence > least common probe	Definition	Example
Action/Behavior	73 (81.11%)	Probes 1, 3, 4 (tied rank) > Probe 2	Thoughts about an action or behavior the pilot executed	"I was actually thinking about my scan and I was checking out the engine instruments making sure they were all in the green"

Table 24, Continued

Navigation	49 (54.44%)	Probe 2 > Probe 1	Thoughts about the aircraft navigation or way-finding	"I was looking at the nav 1 there and seeing that it's not in…"
Rehashing	28 (31.11%)	Probe 2 > Probe 5	The re-visitation of past flight performance in this simulation	"It seems like the aircraft is out of balance to the…it's been wing-heavy on the right and I'm constantly having to keep right aileron."
Planning	27 (30.00%)	Probe 4 > Probe 2	Thoughts about making plans or strategizing for some event in the future.	"I was thinking about making my flight a little more efficient. I'll adjust the mixture a little bit."
Event	26 (28.89%)	Probe 5 > Probe 2	Thoughts about an occurrence or happening	"I was thinking about how my friends make fun of my accent. I saw where it said that this (screen) was donated by the pilots association. And then I was thinking that I'll have to work the next time they have a meeting… and that's next

Table 24, Continued

				Saturday and that is where they make fun of the way i talk.”
Conceptually external	19 (21.11%)	Probe 5 > Probe 2 & 4 (tied rank)	Content was about events and items outside of the concept of flight or flying simulation	“I was thinking of my dog (Roxy) because it looks kinda like the cloud.”
Staging	15 (16.67%)	Probe 2 > Probe 4	Planning or doing something now that will enable something else to happen in the future	“I'm trying to set the stabilizer so I can focus on where I'm going and looking out the window.”

At probe #2, we mentioned that participants were commonly rehashing and we speculate that these participants were thinking about their performance prior to lift off. This idea is not fully supported by the data. Probe #1 has the highest frequencies of thoughts related to simulation difficulties that compromised performance and these thoughts may have continued to be rehashed at probe #2. That said, such thoughts were not verbalized in probe #2. Table 25 presents three codes comprising the category ‘self assessments’ in the first column, followed by the number of examples out of 90 probes, the most and least common probe these codes occurred on, and an example statement from the data. In addition, probes #4 and #5 appear to have an associated quality, such as reduced workload, by which participants can self reflect upon their performance, whether this was their aggregate performance or performance from a single moment.

Table 25. *Self Assessments. Frequencies and examples of self assessment thoughts (n=90)*

Codes	# of examples (percentages)	Most common Probe of occurrence > least common probe	Example
Performance was due to simulation difficulties experienced	19 (21.11%)	Probe 1 > Probes 3 & 4 (tied rank)	"The planes not working right. I know I'm safe."
Negative self judgments of own performance	5 (5.56%)	Probe 4 & 5 (tied rank) > probe 2 & 3 (tied rank)	"I don't have a great course laid out."
Positive self judgment of own performance	2 (2.22%)	Probe 4 & 5 (tied rank) > probe 1, 2 & 3 (tied rank)	"I was kinda laughing at myself because when I'm flying, i don't like to take a lot of steep turns and my instructor again a couple of weeks ago was belly aching at me because he wanted me to do a 30 degree bank and i don't do those kind of turns. i just like to fly. I did that kinda turn here and if [Instructor X] were here he'd be yelling at me."

Finally, thoughts were also coded with respect to their origin: for example, were thoughts cued by an aspect of the person's experience or not cued at all. Information about how thoughts arose was volunteered by the participants and not solicited by the researcher. Table 26 contains four codes under the category 'idea origin' listed in the left most column: cued by environmental visual stimuli, cued by the research investigator, mind pops, and physiologically driven. The five columns in Table 26 include the code, the number and percentage of examples out of 90 probes, the most and least common probe wherein these types of thoughts were experienced, a definition of the code and an example from the participants. These codes are mutually exclusive in Table 26: a thought that was 'cued by the environmental visual stimuli' was not redundantly coded with 'cued by the research investigator.' The code 'cued by the environmental visual stimuli' also could have had auditory cueing, but the simulator sounds provided by Microsoft Flight Simulation X were not clearly audible and all of the cueing in the probe appeared to be visual. The next code, 'cued by the research investigator' was created because on several occasions during the flight simulation, the participant would remark about something I said or did, or the participant would attempt to engage in conversations with the researcher. One participant explained that it was natural to engage in conversations with passengers at a certain altitude to help maintain awareness and to put the passenger at ease. Other times, the participants found themselves explaining to the researcher the definition of certain terms or the reasoning for certain actions taken and these were also counted. Thoughts cued by the environment were more prominent on probes #1 and #3, which are also moments where novel encounters appear in the simulation under reduced workload. At probe #1, the flight simulation was likely novel and noticing information in the simulation might cue other thoughts. After take off and close to cruising altitude, participants may have had the opportunity to look at the view outside the cockpit window. At probe #3, three participants made similar statements to this, "I'm at the point where I'm at the altitude where I want to be (3500 feet) and I'm trying to set the stabilizer so I can focus on where I'm going and looking out the window

(Participant 16)." Probe #5 had the highest frequency of "cued by the research investigator' without any reasons why. One example of this cue occurred at probe #5, because I asked participants to treat this flight like a "checkride" and this thought appeared by Participant 10 in response, "I was thinking, 'Yeah, I'm going to nail this [flight simulation] just like I did my checkride.'" At times, participants muse about when I was going to probe next. For example, Participant 12 said, "I wonder when she's going to stop me next. Then what am I going to say when she does stop me. I am mostly on-task so how can I make her experiment very interesting."

Since the research investigator did not explicitly ask participants where their thoughts originated from, it was unclear what the true frequencies of origination are. As an alternative approach to study true origination, self-reports were analyzed in the Post-Flight questionnaire. Table 27 compares the self-reports in the Post-Flight Questionnaire to the coding formulated by the research investigator.

Table 26. *Thought origin. Frequencies and examples of thought origins (n=90)*

Codes	# of examples (percentage)	Most common Probe of occurrence > least common probe	Definition	Example
Cued by environ mental visual stimuli	18 (20.00%)	Probe 1 & 3 (tied rank) > Probe 2	These are thoughts in which the participant stated that some artifact in the task environment cued their thinking described in the probe.	“That looks like [X Air Force Base]. I was hoping that was [X Air Force Base]. I wanted to look at the runway to make sure I was facing the same way as it did in the map. I was also thinking about the time that when I first got my pilots license I landed at [YCity], that's where I used to be from, to pick up my dad for a flight and there was a time where I was flying not quite on this path but on a different path from [X] airport which is where I used to fly out of and I was sitting in the cockpit and I heard this rumbling and I was a new pilot. So I was thinking, 'what the heck is going on with the airplane' and a C130 just flew right over me on a landing at [X Air Force Base]. When I saw [X Air Force Base], I was thinking about the C130. ”
Cued by the research investig ator	8 (8.89%)	Probe 5 > Probe 1, 2, & 3 (tied rank)	These are thoughts that were cued by interactions with the researcher.	“I’m wondering what you’re thinking because I'm so good now. I know everything is right. I'm at 3500 feet. I'm on course. I'm in good shape. I've got nothing to do and now you can talk. "

Table 26, Continued

Mind pop	7 (7.78%)	Probe 5 > Probe 1, 2 & 4 (tied ranks)	These are thoughts that seem to lack any origin.	"My favorite saying is "half as much, twice as slow" and I was trying to remember where I heard that"
Physiologically driven	1 (1.11%)	Probe 2 > Probes 1, 3, 4, & 5 (tied ranks)	Thoughts that are driven by the person's physiology; e.g., hunger, thirst, pain, etc.	"I was also thinking about how warm it is in here."

Table 27. *Thought origin comparison between participant self-reports and the investigator's classification*

Self-Reports (*n*=17 participants)			Investigator's Classification (*n*=90 probes)		
Mean % of thoughts cued by environment (*SD*)	Mean % of mind pops experienced (*SD*)	Mean % caused by measurement of thoughts (*SD*)	# of examples cued by the environment (%)	# of mind pop examples (%)	# of examples cued by researcher (%)
39.76% (41.33%)	17.94% (29.72%)	10.71% (18.80%)	18 (20.00%)	7 (7.78%)	8 (8.89%)

Possible correlates with these findings.

Several questions were posed in the Post-Flight questionnaire to explain some of the findings in the thought coding. One explanation might be the level of expertise, individual differences with respect to habits and traits, and so on, which we now discuss.

A pilot's expertise gleaned from the Post-Flight questionnaire was evaluated. The mean number of self-reported hours spent flying in a non-simulated Cessna 172 was 157.39 hours (*SD*=203.56 hours) with self-reports that ranged from 1 to 800 hours. All but one participant was strictly a GA pilot. The mean number of self-reported hours in a simulated flight environment was 922.26 hours (*SD*=2631.56 hours), ranging from 0 to 10,000 hours. The mean number of accrued flight hours in any aircraft type was self-reported to be 813.24 hours (*SD*=1396.85) ranging from 60 to 5300 hours of accrued flight time. Participants also self-reported that they flew on average 4.65 days in a single month (*SD*=4.70 days) and had flown more than one aircraft throughout their flight careers (*M*= 6.47 different types of aircraft, *SD*=5.23 aircraft, ranging from 1-20 different aircraft flown). In fact, 16 out of 17 participants flew more than one type of aircraft in their flight history, including models by Cessnas, Remos, Pipers, Robinsons, and Beech.

Individual differences were also identified with respect to habits, traits and beliefs. First, we asked participants about their level of distractibility compared to their peers, when they were most likely distracted and possible explanations as to the cause of distraction. Nine of the participants felt they were more distractible and four participants felt they were less distractible than their peers. Participants were asked about the stability of their distractibility throughout the day. Fourteen out of 16 participants self reported that their distractibility was variable throughout the course of the day and when they believed to feel more distractible (see Table 28). The most common response was that the afternoon was the most distractible time, followed by experiences of boredom and disinterest. In Part III, it was discovered that some participants seek out distractions from task related thinking to boost alertness. Thus, pilots in Part IV were asked how frequently they intentionally mind wandered or sought out distractions to boost alertness during flight. Four response options were provided: never, occasionally, frequently and very often. Seven participants self-reported occasionally, six reported never, two did not respond because they believed that they do not mind wander, one reported frequently and one reported very often. Thus, some of the pilots occasionally had this habit while others never did. Because fatigue was identified as a possible reason for elevated mind wandering frequencies in Part III, we asked how many hours of sleep participants had the night before their study session. The average was 7.18 hours of sleep with a *SD* of 1.35 hours of sleep.

Also, it was believed that stressful life events can be correlated with distractibility and given time constraints, self-reports of distractions to salient life events during probes could not be discussed; however, at the end of the study, we asked participants to list the salient life events in an attempt to understand how they might relate to performance. The mean score on this assessment was 109.41, SD=76.38, and six of the participants were ranked as moderate risk of developing illness while the rest were mild risk. Correlations were conducted between the Stressful Life Events scores and the number of self-reported occurrences of mind wandering reported in the Post-Flight questionnaire and this result was non-significant ($r = -0.22$, $p > .05$). However, correlations between the Stressful Life Events scores and the percentage of off-task thoughts that were mind pops were significant (r=0.41, $p < .05$). Despite the low sample size, this suggests that people with more stressful life events might have a greater frequency of mind pops (e.g., un-cued, spontaneous thoughts) than thoughts cued by the flight environment.

Table 28. Number of pilots who self report having more distractibility during the following tasks or times of day (n=17)

Category	**Sub Category**	**# of participants**
Time of Day	Early morning	1
	Afternoon	5
	Evening	2
	During work	1
	After work	1
Characteristics of tasks	Boring tasks	2
	Low mental needs	1
	Don’t require a lot of thought	1
Types of tasks	Attending uninteresting classes	2
	Reading	
	During monotonous tasks	1
	After spending a significant amount of time in front of a computer	1
Other	During unstructured or idle time	2
	When in a state of boredom or fatigue	2

7.4.Discussion

In summary, most participants were moderately experienced and so it is difficult to determine if our results were due to a lack of automatize flight processes and procedures or to the novel environment of the flight simulation. We hypothesize that the reduction of cognitive demands may allow participants sufficient cognitive and attentional resources to experience off-task thinking and mind wandering. While there were no noticeable shifts in the incidents of self-reports of off-task incidents or mind wandering occurrences at each probe, there was an observed shift in the thought patterns as flight time progressed. The thoughts in the beginning of the flight appear to be concentrated on behaviors and actions taken during task execution and thought sequences were usually triggered by the visual environment when the cockpit was still fairly novel. At probe #2, we speculate that pilots experience the highest workload even though subjective mental workload was never assessed. Pilots at probe #2 are in the ascent phase of flight and they tend to think about their performance difficulties with the simulation environment while making a sharp ~160 degree right turn to navigate to the destination airport. Some participants had unexpected difficulties interpreting why the aircraft behavior was not aligning with their expectations and spent time repeatedly cross checking gauges for safety. Also, the yoke was reported by participants as being more sensitive than yokes in a real aircraft. Therefore at probe #2, it is possible that participants were still trying to alter their yoke inputs to obtain the outcome behavior desired from the aircraft. Thus, higher frequencies of thoughts pertaining to navigation, staging and rehashing past performance appear at probe #2. Probes #3-5 were at cruise altitude and we expect to see that participants would begin to gradually, mentally disengage from the task and to identify thought semantics that shift to planning thoughts, navigational thoughts and eventually conceptually unrelated ideas. Perhaps with more flight time, those same participants may have started to self-report off-task and mind wandering occurrences.

In beta testing, we instructed all pilots to fly a prescribed flight route, but we quickly realized that the route was difficult to follow and pilots became obsessive about flight path deviations. Thus, we decided to make the flight task easier and to allow participants to fly their own route. Performance would be judged by crashes and inability to get to the destination airport. None of the off-task thinking and mind wandering incidents led to a crash. In one incident, participant 8 began "chatting" with the research investigator in between probes and he interrupted his chatting with a visual identification of the destination airport. At that moment, the participant accidentally pitched the aircraft into a dive for a few seconds, but then recovered. Seven out of 17 pilots visually sited the destination airport accurately.

We also sought to understand which habits, traits and beliefs were correlated with higher and lower frequencies of mind wandering. Participants indicate that they are likely to be distracted during the afternoon and during boring or un-stimulating activities. Since the frequencies of self reports for mind wandering and off-task thinking were low at each probe, a relationship between occurrences of mind wandering and other explanatory variables, such as "distractibility compared to peers," could not be assessed. Even though we did not observe differences in off-task and mind wandering episodes across flight, flight can become boring, monotonous and un-stimulating at times and so this may need further investigation to discover incidents of mind wandering. Some participants admit that they occasionally induce mind wandering to boost alertness, a compensatory strategy for monotonous flight. Therefore, the effects of fatigue and lack of sleep on mind wandering and performance is an opportunity for further study, as we know that a relationship exists between high levels of fatigue and performance decrements in flight simulations (Caldwell, Caldwell, et al., 2003).

8. Conclusions

This research sought to investigate the relationship between mind wandering, task execution and operator performance in driving and flight simulations. In addition, we attempted to understand what factors (e.g., negative mood, stressful life events, etc.) correlate with frequent subjective experiences of mind wandering.

We discovered several unexpected, research findings that impacted subsequent phases of our studies. From the sample of 30 pilot participants in Part III, we learned that participants believe that off task thinking and mind wandering are not conceptually the same idea. Some participants believed that he/she was off task in their thinking, but that they were not mind wandering. Thus, both off task thinking and mind wandering frequencies were incorporated into later stages of the research in order to determine when a thought was considered off task versus mind wandering. In later phases of the study, we found that the prototypical example of mind wandering was muddied by fringe cases. For example, in beta testing in Part III, participants described thoughts that were difficult to classify as mind wandering. These thoughts include humming, thinking about past simulation performance unrelated to the task at hand, strategizing for better future performance, the state of being between thoughts, etc. To better classify these thoughts, we tested these fringe cases in the full study plan in Parts III and IV.

8.1.Addressing the hypotheses

Responses to all hypotheses and research questions are briefly reviewed below.

Hypothesis 1 questions the relationship between task load, mind wandering and working memory capacity. This hypothesis was not clearly supported by the data. There was no inverse relationship found and working memory appears to have no effect on the relationship between task load and mind wandering frequency (see Figure 26). However, a direct relationship between subjective mental workload and frequencies of mind wandering may exist. The lack of overall significant findings may be due to the way task load was represented as an independent variable, to the task chosen or to the cognitive demands this task does or does not impose.

Hypothesis 2 was partially supported by the data. Participants with higher self-reports of negative mood were significantly more likely to report off task thinking, but not more likely to report mind wandering (see Figure 26). Prior literature does not distinguish between off task and mind wandering thoughts and, if the two constructs were combined together to represent 'mind wandering,' these findings would be significant.

Research question 1 was not clearly supported by the data. This question has two parts: first, we assessed the relationship between probe number, which was a proxy or index for task type, and mind wandering; and second, we evaluated whether significant differences exist between cognitive ability and mental process requirements across all five probes. The cognitive ability and mental process requirements were used to approximate "task type" at each of the five probes. A significant relationship between probe number and frequencies of off task thinking and mind wandering was observed, but no significant differences between cognitive ability and mental process requirements across all five probes were found (see

Figure 26). Thus, the representation of task type as cognitive ability and mental process requirements may be inaccurate, or some other phenomenon is driving the significant relationship between probes and off task thinking and mind wandering.

Research question 2 queries the relationship between types of mind wandering, task performance and adverse consequences, in which performance and consequences were aggregated. This relationship was studied in two parts: first, the assessment of mean differences in seconds of speeding and lane deviations across probes; and second, the qualitative assessment of the ways that participants categorize their off task thinking at each probe (e.g., temporal, functional and conceptual). Mean differences exist across probes for both seconds of speeding and lane deviations in the first part, such that probe #2 had the largest mean seconds of speeding and the lowest lane deviations. At probe #2, participants explained that their thoughts had the least amount of conceptual overlap with the task of driving which is also a significant main effect for classification of thoughts as mind wandering in Part I. Probe #4 had the highest number of lane deviations and the second highest mean number of seconds of speeding and the second lowest ratings for conceptual relatedness. Thus, the two probes into thoughts that were the most conceptually unrelated to the task of driving had the greatest performance decrements with respect to speeding and lane deviations. It is not clear whether there is a causal relationship or a corollary relationship so bidirectional arrows are used in Figure 26.

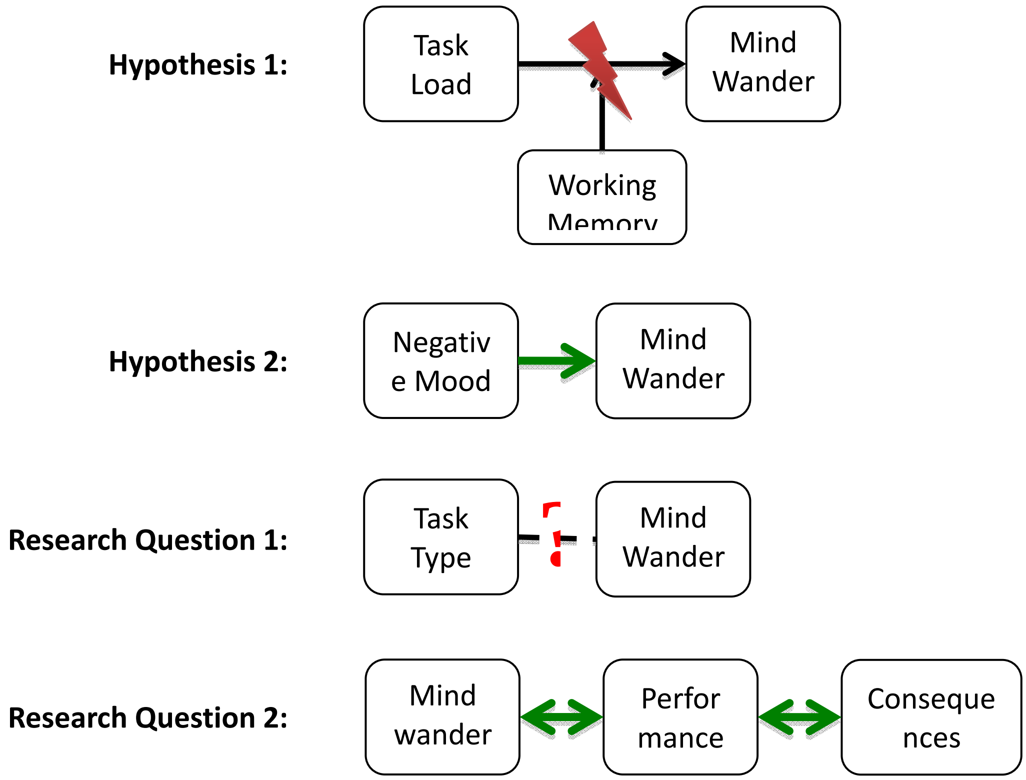


Figure 26. *Responses to hypotheses and research questions.*

Research question 3 asks which metric was the most appropriate measure of mind wandering in the cockpit that minimizes disruption to the task and enhances the ability to recall probe content. According to participants' judgments, the best performing metric was the clicker with respect to its ability to foster the fastest reengagement with the task, to accurately capture mind wandering incidents and to reduce the least amount of metric-induced monitoring of conscious thoughts.

8.2.Building Theory

Part of the qualitative aspects of this research is to generate theory. Based on all of the accrued observations, three hypotheses are proposed for future research.

Our results suggest that there may be a shift from tight coupling, which is thinking about actions and events immediately experienced, at higher workload and novel situations toward a gradual decoupling from task during lower workload scenarios (e.g., due to planning and staging types of thoughts). This decoupling might occur while participants still believe they are on task. From probe #1 to probe #5, the decoupling appears to increase, but it is unclear if this observation has to do with workload, which also decreases with later probes, or with time on task, or both. It is also unclear whether decoupling is on a continuum in which a decoupling threshold is reached that determines when the participant realizes that he/she is off task or mind wandering. Based on our findings, we can hypothesize that a certain threshold on a continuum of decoupling needs to be crossed before participants begin classifying their thoughts as off task and then as mind wanderings. The decoupling might be related to conceptual relatedness between thoughts and the task, or the time spent unengaged with the task.

It is also hypothesized that mind wandering may not be just a passive experience, rather it could be deliberate for certain reasons. In this research, mind wandering was not a random occurrence, rather some cognitive system seem to be a gatekeeper for when off task thinking and mind wandering are allowed. In addition, mind wandering may serve different functions. For example, some participants in Part I indicated that they mind wander during stressful events while other participants in Part IV indicated that they mind wander during boring or monotonous events or to boost alertness or arousal levels. If the mind has a homeostasis for arousal levels and the arousal level is too high (e.g., during stressful events), does mind wandering serve to reduce arousal levels back to a homeostatic level? If the person's arousal levels are too low during boring or monotonous tasks, does wandering reset the arousal level to homeostasis? On the lowest end of arousal is sleepiness and fatigue and a few participants in both the driving and flight simulations explained that they were more likely to mind wander during times of fatigue or sleepiness. Thus, future work might investigate whether this homeostatic arousal level exists and whether mind wandering is one mechanism to maintain homeostasis.

The last point with respect to building theory is that it is still unclear how mind wandering fits in with other concepts like vigilance, automatic and controlled processing, multitasking and distracted driving. Table 29 lists the definitions of each concept from the literature for reference. It is unclear whether mind wandering is a vigilance decrement or a result of multiple vigilance decrements. Can mind wandering occur with automatic processing, or is it equally likely to occur with both automatic and controlled processing? Does mind wandering change with expertise such that novices, who have high cognitive burdens from learning a new task, have lower frequencies of mind wandering than experts? In Part IV, moderately experienced pilots comprised our sample and we observed relatively low frequencies of mind wandering during the Part III driving simulation. But novice drivers in Part III experienced high frequencies of mind wandering compared to pilots so the relationship between expertise and mind wandering is unclear. Thus, is there a relationship between expertise, automatic processing of tasks that result from expertise, and mind wandering, or is there a stronger relationship between mind wandering and the task environment? Also, if mind wandering is deliberate, is the person multitasking or distracted and not mind wandering? Alternatively, does mind wandering include deliberate distractions and non-deliberate distractions? These unanswered questions demonstrate the additional depth and challenge remaining in this research area that can be investigated in the future.

Table 29. *Definitions of related concepts in research domain*

Concept	Definition	Citation
Vigilance	"the ability to sustain attention to a task for a period of time".	Parasuraman, 1998
Vigilance	"the state of readiness to detect and respond to certain specified small changes occurring at random time intervals in the environment,"	Mackworth, 1957
Automatic process	"An automatic process can be defined within such a system as the activation of a sequence of nodes with the following properties: (a) The sequence of nodes (nearly) always becomes active in response to a particular input configuration, where the inputs may be externally or internally generated and include the general situational context. (b) The sequence is activated automatically without the necessity of active control or attention by the subject (p. 2)" "Since an automatic process operates through a relatively permanent set of associative connections in long-term store, any new automatic process requires an appreciable amount of consistent training to develop fully. Furthermore, once learned, an automatic process is difficult to suppress, to modify, or to ignore (p. 2)" "A *controlled process* is a temporary sequence of nodes	Schneider & Shiffrin, 1977

Table 29, Continued

	activated under control of, and through attention by, the subject. Because active attention by the subject is required, only one such sequence at a time may be controlled without interference, unless two sequences each require such a slow sequence of activations that they can be serially interwoven (p. 2)”	
Multitasking	“In this paper, we define multitasking as the simultaneous conduct of two or more activities during a given time period.”	Kenyon & Lyons, 2007
Multitasking	“When humans multitask, they work on two or more tasks and switch between those tasks, either as individuals or within groups.”	Waller, 1997
Distracted Driving	In this sense, driver distraction results when drivers’ normal cognitive processes (i.e., attention-sharing) and adaptive strategies fail and drivers are no longer able to adequately divide their attention between the driving and secondary tasks and maintain driving performance at a satisfactory level. Distraction can occur either because the secondary task is so complex or compelling that drivers fail to allocate (or prioritize) sufficient attention to driving, or because the demands of the driving task are so high that they do not allow the performance of a secondary task at any level.”	Young & Reagan, 2007

REFERENCES

Anderson, B. F. (1975). *Cognitive psychology.* New York: Academic Press.

Anderson, B. F. (1985). *Cognitive psychology and its implications.* New York: Freeman.

Antrobus, J. S. (1968). Information theory and stimulus-independent thought. *British Journal of Psychology, 59,* 423–430.

Ammerman, H.L., Becker, E.S., Jones, G.W., & Tobey, W.K. (1987). *FAA. Air traffic control operations concepts: Volume I: ATC background and analysis methodology* (Report No. DOTIFAAJAP-87-Gl (VOL#I)). Washington, D.C.: U.S. Department of Transportation, Federal Aviation Administration.

Baddeley, A. & Hitch, G. J. (1974) Working memory. In G. Bower (Eds.) *Recent advances in learning and motivation, vol. 8,*Academic Press.

Bastik, T. (1982). *Intuition: How we think and act.* Chichester, England: Wiley.

Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Arch. Gen. Psychiatry, 4,* 561–71.

Binder, J. R., Frost, J. A., Hammeke, P. S., Bellgowan, S. F., Rao, S. M., & Cox, R. W. (1999). Conceptual processing during the conscious resting state: A functional MRI study. *Journal of Cognitive Neurosciences, 11,* 80–93.

Brannick, B. T. & Levine, E. L. (2002). *Job Analysis; Methods, research, and applications for human resource management in the new millennium.* Sage Publications, Inc. Thousand Oaks, CA.

Brannick, B. T., Levine, E. L., & Morgeson, F. P. (2007). *Job Analysis: Methods, Research, and Applications for Human Resource Management (2nd Ed.).* Thousand Oaks, CA: Sage.

Broadbent, D. E., Cooper, P. F., et al. (1982). The Cognitive Failures Questionnaire (CFQ) and its correlates. *The British Journal of Clinical Psychology, 21, 1,* 1–16.

Caldwell, J., Caldwell, J. L., Brown, D., Smythe, N., Smith, J., Mylar, J., Mandichak, M., & Schroeder, C. (2003). The effects of 37 hours of continuous wakefulness on the physiological arousal, cognitive performance, self reported mood, and simulator flight performance of F-117A pilots. (AFRL-HE-BR-TR-2003-0086).

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston, MA: Houghton Mifflin

Cheyne, J. A., Carriere, J. S. A., & Smilek, D. (2006). Absent-mindedness: Lapses in conscious awareness and everyday cognitive failures. *Consciousness and Cognition, 15*, 578–592.

Cicogna, P., & Nigro, G. (2005). Time-based prospective remembering: Interference and facilitation in a dual task. *European Journal of Cognitive Psychology, 17*, 2, 221-240.

Cohen, J. D., Botvinick, M. M., & Carter, C. S. (2000).Cheyne, J. A., Solman, G. J. F., Carriere, J. S. A., & Smilek, D. (2009). Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors. *Cognition, 111,* 998-113.

Converse, J. M., & Presser, S. (1986). *Survey Questions. Handcrafting the Standardized Questionnaire.* Thousand Oaks, CA: Sage Publications, Inc.

Cowan N. (1995). *Attention and memory: An integrated framework.* Oxford Psychology Series, No. 26. New York: Oxford University Press.

Davies, D. R. & Parasuraman, R. (1982). *The Psychology of Vigilance.* London: Academic Press.

Edmondson, A. C. (2004).Learning from mistakes is easier said than done: Group and organizational influences on the detection and correction of human error. *Journal of Applied Behavioral Science, 40*, 66-90.

Endsley, M. R. (1995a). Measurement of situation awareness in dynamic systems. *Human Factors, 37(1),* 65-84.

Endsley, M. R. (1995b). Toward a theory of situation awareness in dynamic systems. *Human Factors 37(1), 32*-64.

Eysenck, M,W., & Byrne, A. (1992). Anxiety and susceptibility to distraction. *Personality and Individual Differences, 13,* 793-798,

Eysenck, M,W., & Graydon, J, (1989). Susceptibility to distraction as a function of personality. *Personality and Individual Differences, 10,* 681-687,

Forster, S. & Lavie, N. (2009). Harnessing the wandering mind: The role of perceptual load. *Cognition, 111,* 345-355.

Gertman, D. I., Hallbert, B. P., Parrish, M. W., Sattision, M. B., Brownson, D.,Tortorelli, J. P., Trager, E. E. A., & Persensky, J. J. (2001). Review of findings for human error contribution to risk in operating events. Prepared for the Office of Nuclear Regulatory Research (INEEL/EXT-01-01166).

Giambra, L. M. (1989). Task-unrelated-thought frequency as a function of age: A laboratory study. *Psychology and Aging, 4(2),* 136-143.

Giambra, L. M. (1995). A laboratory method for investigating influences on switching attention to task-unrelated imagery and thought. *Consciousness and Cognition, 4*, 1-21.

Giambra, L.M, Grodsky, A., Belongie, C., & Rosenberg, E.H. (1994– 1995). Depression and thought intrusions, relating thought frequency to activation and arousal. *Imagination, Cognition and Personality, 14*, 19–29.

Glaze, A. L., & Ellis, J. M. (2003). *Pilot study of distracted drivers*. Richmond, VA: Survey and Evaluation Research Laboratory, Center for Public Policy, Virginia Commonwealth University.

Greenwald, D. F., & Harder, D. W. (1995). Sustaining fantasies, daydreams and psychopathology. *Journal of Clinical Psychology, 51,* 719– 726.

Greenwald, D. F., & Harder, D. W. (1997). Fantasies, coping behavior and psychopathology. *Journal of Clinical Psychology, 53,* 1–7.

Grodsky, A., & Giambra, L. (1991). Task unrelated images and thoughts whilst reading. In J. Shorr, P. Robin, J. A. Connek, & M. Wolpin (Eds.), Imagery: Current perspectives. New York: Plenum Press.

Hackos J. T., & Redish J. C. (1998). *User and Task Analysis for Interface Design.* John Wiley & Sons, New York.

Hanowski, R. J., Perez, M. A., & Dingus, T. A. (2005). Driver distraction in long-haul truck drivers. *Transportation Research Part F: Traffic Psychology and Behavior*, *8*, 441–458.

Hart, S. G. (2006). NASA-task load index (NASA-TLX); 20 years later. *Proceedings of the 50th Annual Meeting of the Human Factors and Ergonomics Society (HFES),* San Francisco, CA.

Hart, S. G., & Staveland, L. E. (1988). Development of a NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload,* edited by P. S. Hancock and N. Meshkati, (pp. 139–83). Amsterdam, North-Holland.

Helmholtz, H. von. (1896). *Vortrage und Reden.* Braunschweig, Germany: Vieweg und Sohn.
He, J., Becic, E., Lee, Y.C., McCarley, J.S. (2011). Mind wandering behind the wheel: Performance and oculomotor correlates. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 53(13),* 13-21.

Heck, K., & Carlos, R. (2008). Passenger distractions among adolescent drivers. *Journal of Safety Research*, *39*, 437–443.

Helton,W. S., Kern, R. P., & Walker, D. R. (2009). Conscious thought and the sustained attention to response task. *Consciousness and Cognition, 18*, 600-607.

Hill, A.B., Kemp-Wheeler, S. M., & Jones, S.A. (1986). What does the Beck depression inventory measure in students? *Personality and Individual Differences, 7(1),* 39-47.

Hitchcock, E. M., Warm, J. S., Matthews, G., Dember, W. N., Shear, P. K., Tripp, L. D., et al. (2003).Automation cueing modulates cerebral blood flow and vigilance in a simulated air traffic control task. *Theoretical Issues in Ergonomics Science, 4*, 89–112.

Holmes, T.H., & Rahe, R. H. (1967). The Social Readjustment Rating Scale. *J Psychosom Res 11*(2), 213–8.

Horrey, W. J., & Wickens, C. D. (2006). Examining the impact of cell phone conversations on driving using meta-analytic techniques. *Human Factors*, *48*, 196–205.

Ishigami, Y., & Klein, R. (2009). Is a hands-free phone safer than a handheld phone? *Journal of Safety Research*, *40*, 157–164.

James, W. (1890). The Stream of Consciousness. In *Psychology,* Chap. XI. New York, NY: Henry Holt and Co.

Jacobsen, P.D. & Gostin, L.O. (2010). Regulation and education to avert traffic injuries and fatalities. *JAMA, 303(14),* 1419-1420.

Jennings, D. E. (1987). Judging inference adequacy in logistic regression. *Journal of the American Statistical Association, 81,* 471-476.

Kahneman, D. (1973). *Attention and Effort.* Engiewood Cliffs, NJ: Prentice Hall.

Klinger, E. (1978). Modes of normal conscious thought. In K. S. Pope & J. L. Singer (Eds.), *The stream of consciousness: Scientific investigations into the flow of human experience* (pp. 225–258). New York, NY: Plenum.

Knight, R.G. (1984). Some general population norms for the short form Beck Depression Inventory. *Journal of Clinical Psychology, 40*(3), 751-753.

Kvavilashvili, L., & Mandler, G. (2004). Out of one's mind: A study of involuntary semantic memories. *Cognitive Psychology, 48,* 47–94.

Lee, J. D., Caven, B., Haake, S., & Brown, T. L. (2001). Speechbased interaction with in-vehicle computers: The effect of speech-based e-mail on drivers' attention to the roadway. *Human Factors*, *43*, 631–640.

Loukopoulos, L. D., Dismukes, R. K., & Barshi, I. (2001). Cockpit interruptions and distractions: A line observation study. *Proceedings of the 11th International Symposium on Aviation Psychology, Columbus, OH.*

Loukopoulos, L. D., Dismukes, R. K., & Barshi, I. (2003). Concurrent task demands in the cockpit: Challenges and vulnerabilities in routine flight operations. *Proceedings of the 12th Internationl Symposium on Aviation Psychology, Dayton, OH.*

Loukopoulos, L. D., Dismukes, R. K., & Barshi, I. (2009). *The Multitasking Myth: Handling Complexity in Real-World Operations*. Burlington, VT: Ashgate.

Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology, 1,* 6-12.

Mackworth, N.H., 1957. Vigilance. *The Advancement of Science 53,* 389–393

Macwan, A., & Mosleh, A. (1994). A methodology for modeling operator errors of commission in probabilistic risk assessment. *Reliability Engineering and System Safety, 45*, 139-157.

Manly, T., Robertson, I. H., Galloway, M., & Hawkins, K. (1999). The absent mind: Further investigations of sustained attention to response. *Neuropsychologia, 37*, 661–670.

Mason, M. F., Norton, M. L., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macreae, C. N. (2007). Wandering minds: The default network and stimulus-independent thought. *Science, 315*, 393-395.

Matthews, G., Davies, D. R., Westerman, S. J., & Stammers, R. B.(2000). *Human performance: Cognition, stress and individual differences.* East Sussex, UK: Psychology Press.

McVay, J. C., & Kane, M. J. (2010). Does mind wandering reflect executive function or executive failure: Comment on Smallwood and Schooler (2006) and Watkins (2008). *Psychological Bulletin, 136,* 188–197.

McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *J. Exp. Psychol. Learn. Mem. Cogn., 35(1),* 196-204.

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior.* New York: Holt, Rinehart & Winston.

Morgeson, F. P., & Dierdorff, E. C. (2011). Work analysis: from technique to theory. In S. Zedeck (Eds.) *APA Handbook of Industrial and Organizational Psychology, (*Vol. 2, pp. 3-41). Washington, DC: Am. Psychol. Assoc.

Nachreiner, F., & Hanecke, K. (1992). Vigilance. In A. P. Smith&D.M.Jones (Eds.), *Handbook of human performance* (pp.262–288). San Diego: Academic Press.

Norman, D. A. (1983). Design rules based on analyses of human error. *Communications of the ACM, 26*(4), 254-259.

Parasuraman, R. (1987). Human-computer monitoring. *Human Factors, 29,* 695-706.

Parasuraman, R., 1998. In: Parasuraman, R. (Ed.), *The Attentive Brain.* The MIT Press, Cambridge, MA.

Payne S. J., Duggan G. B., & Neth H. (2007). Discretionary task interleaving: Heuristics for time allocation in cognitive foraging, *Journal of Experimental Psychology: General, 136*(3), 370-388.

Pounds, J., & Isaac, A. (2002). Development of an FAA-EUROCONTROL Technique for the analysis of human error in ATM. DOT/FAA/AM-02/12.

Reason, J. (1995). Understanding adverse events: Human factors. *Qual. Saf. Health Care, 4,* 80-899.

Reason, J. (2000). Human error: Models and management, BMJ, 320, 768-770.

Reason, J. T., & Mycielska, K. (1982). *Absent-minded?: The psychology of mental lapses and everyday errors.* Englewood Cliffs, NJ: Prentice- Hall.

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). "Oops!" Performance correlates of everyday attentionalfailures in traumatic brain injured and normal subjects. *Neuropsychologia, 35,* 747–758.

Rouse,W. B. (1977). Human–computer interaction in multi-task situations, *IEEE Transactions on Systems, Man and Cybernetics, SMC-7,* 384–392.

Sanchez, J. I. & Levine, E. L. (1999). Is job analysis dead, misunderstood, or both? New forms of work analysis and design. In A. Kraut & A. Korman (Eds.) *Evolving Practices in Human Resource Management*, (pp. 43-68). San Francisco, CA: Jossey-Bass.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing. I. Detection, search, and attention. *Psychological Review*, *84*(1), 1–66.

Schooler, J. W. (2002). Re-representing consciousness: Dissociations between experience and meta-consciousness. *Trends in Cognitive Science, 6,* 339–344.

Schooler, J. W., Falshore, M., & Fiore, S. M. (1995). Putting insight into perspective. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 589–597). Cambridge, MA: MIT Press.

Schooler, J. W., Reichle, E. D., & Halpern, D. V. (2005). Zoning-out during reading: Evidence for dissociations between experience and meta-consciousness. In D. T. Levin (Ed.), *Thinking and seeing: Visual metacognition in adults and children* (pp. 204–226). Cambridge, MA: MIT Press.

Schooler, J. W., & Schreiber, C. A. (2004). Consciousness, metaconsciousness, and the paradox of introspection. *Journal of Consciousness Studies, 11,* 17–39.

Seibert, P. S., & Ellis, H. C. (1991). Irrelevant thoughts, emotional mood states and cognitive performance. *Memory & Cognition, 5,* 507–513.

Shappell, S., Detwiler, C., Holcomb, K., Hackworth, C., Boquet, A., &Wiegmann, D. (July, 2006). Human error and commercial aviation accidents: A comprehensive, fine-grained analysis using HFACS. (DOT/FAA/AM-06/18)

Shiffrin, R. M., & Dumais, S. T. (1981). The development of automatism. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 111-140). Hillsdale, NJ: Erlbaum.

Shiffrin, R. M., & Schneider, W. (1984). Theoretical note: Automatic and controlled processing revisited. *Psychological Review, 1*(2), 269-276.

Smallwood, J., Beach, E., Schooler, J. W., & Handy, T. C. (2008). Going AWOL in the brain: Mind wandering reduces cortical analysis of external events. *Journal of Cognitive Neuroscience*, *20*, 458–469.

Smallwood, J., McSpadden, M., & Schooler, J. W. (2007). The lights are on but no one's home: Meta-awareness and the decoupling of attention when the mind wanders. *Psychonomic Bulletin & Review, 14*(3), 527-533.

Smallwood, J. M., McSpadden, M., & Schooler, J. W. (2008). When attention matters: The curious incident of the wandering mind. *Memory & Cognition, 36*(6), 1144-1150.

Smallwood, J. M., Baracaia, S. F., Lowe, M., & Obonsawin, M. (2003). Task unrelated thought whilst encoding information. *Consciousness and Cognition, 12*, 452-484.

Smallwood, J., Davies, J. B., Heim, D., Finnigan, F., Sudberry, M., O'Connor, R., & Obonsawin, M. (2004). Task engagement and disengagement during sustained attention. *Consciousness and Cognition, 13*, 657-690.

Smallwood, J., Fishman, D. J., & Schooler, J. W. (2007). Counting the cost of an absent mind: Mind wandering as an underrecognized influence on educational performance. *Psychonomic Bulletin & Review, 14(2),* 230-236.

Smallwood, J., Nind, L., & O'Connor, R. (2009). When is your head at? An exploration of the factors associated with temporal focus of the wandering mind. *Consciousness and Cognition, 18,* 118-125.

Smallwood, J., Obonsawin, M., & Heim, D. (2003). Task unrelated thought: The role of distributed processing. *Consciousness and Cognition, 12,* 169-189.

Smallwood, J., Obonsawin, M., Baracaia, S. F., Reid, H., O'Connor, R. C., & Heim, S. D. (2004b). The relationship between rumination, dysphoria and self-referent thinking: Some preliminary findings. *Imagination, Cognition and Personality.*

Smallwood, J., Obonsawin, M., & Reid, H. (2002-2003). The effects of block duration and task demands on the experience of task unrelated thought. *Imagination, Cognition and Personality, 22*(1), 13-31.

Smallwood, J., & Schooler, J. W. (2006). The Restless Mind. *Psychological Bulletin, 132(6)*, 946-958.

Smallwood, J., O'Connor, R. C., Sudberry, M. V., Haskell, C., & Ballantyne, C. (2004). The consequences of encoding information on the maintenance of internally generated images and thoughts: The role of meaning complexes. *Consciousness and Cognition, 13*, 789-820.

Smith, B. H (1995). Quality cannot always be quantified. *BMJ, 311,* 258a-258.

Strayer, D. L., & Drews, F. A. (2007). Cell-phone-induced driver distraction. *Current Directions in Psychological Science*, *16*, 128–131.

Strayer, D. L., Drews, F. A., & Crouch, D. J. (2006). A comparison of the cell phone driver and the drunk driver. *Human Factors*, *48*, 381–391.

Strayer, D. L., Drews, F. A., & Johnston, W. A. (2003). Cell phone induced failures of visual attention during simulated driving. *Journal of Experimental Psychology: Applied*, *9*, 23–32.

Strayer, D. L. & Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on cellular telephone. *Psychological Science, 12*(6), 462-466.

Stutts, J. C., Reinfurt, D. W., Staplin, L., & Rodgman, E. A. (2001). *The role of driver distraction in traffic*. AAA Foundation for Traffic Safety Crashes. Retrieved from http://www.aaafoundation.org/pdf/distraction.pdf

Teasdale, J. D., Dritschel, B. H., Taylor, M. J., Proctor, L., Lloyd, C. A., Nimmo-Smith, I., et al. (1995). Stimulus-independent thought depends on central executive resources. *Memory and Cognition, 23,* 551–559.

Teasdale, J. D., Proctor, L., Lloyd, C. A., & Baddeley, A. D. (1993). Working memory and stimulus-independent thought – Effects of memory load and presentation rate. *The European Journal of Cognitive Psychology, 5,* 417–433.

Thurstone, L.L. (1931). Multiple factor analysis. *Psychological Review*, *38*, 406–427.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28*, 127–154.

Von Restorff, H. (1933). The effects of field formation in the trace field. *Psychological Research 18* (1), 299–342.

Waller, M. J. (1997). Keeping the pins in the air: How work groups juggle multiple tasks. In M. M. Beyerlein & D. A. Johnson (Eds.), *Advances in interdisciplinary studies of work teams* (Vol. 4, pp. 217–247). Stamford, CT: JAI Press.

Wegner, D. M. (1989). *White bears and other unwanted thoughts: suppression, obsession and the psychology of mental control*. New York, NY: Viking.

Wharton, C., Bradford, J., Jeffries, J., & Franzke, M. (1992). Applying Cognitive Walkthroughs to more Complex User Interfaces: Experiences, Issues and Recommendations. *CHI,* 381–388.

Wiegmann, D. A. Shappell, S. A. (2001). A human error analysis of commercial aviation accidents using the Human Factors Analysis and Classification System (HFACS). (DOT/FAA/AM-01/3)

Wiener, E. L., & Curry, R. E. (1980). Flight-deck automation: promises and problems, *Ergonomics, 23,* 95–1011.

Wickens, C. D. (1992). *Engineering Psychology and Human Performance*, (2nd Ed.). New York, NY: HarperCollins.

Woods, D. D., Johannesen, L. J., Cook, R. I., & Sarter, N. B. (1994). Behind human error: Cognitive systems, computers, and hindsight. (CSERIAC # SOAR 94-01). Wright-Patterson AFB, OH: Crew System Ergonomics Information Analysis Center.

Young, L. R. A. (1969). On adaptive manual control*, Ergonomics, 12*, 635–657.

Young, K. & Regan, M. (2007). Driver distraction: A review of the literature. In I.J. Faulks, M. Regan, M. Stevenson, J. Brown, A. Porter & J.D. Irwin (Eds*.). Distracted driving* (pp. 379-405). Sydney, NSW: Australasian College of Road Safety.

APPENDICES

APPENDIX A. *Aviation Safety Reporting System (ASRS) Reports of Mind Wandering By Pilots in the Cockpit*

Time of Day	Altitude	Crew Size	Mission	Flight Phase	Narrative
Night	3200	2	Personal	Climb	PIC FAILED TO PAY ATTN AS PAX FLEW INTO BOTTOM OF TCA ON INITIAL CLB OUT OF RENTON ARPT. PIC DAYDREAMING AND LOOKING AT VIEW. NO LIGHT IN ALTIMETER ON PANEL. PIC SUDDENLY REALIZED PLANE WAS STILL CLBING AND SUSPECTED TCA INCURSION; TUNED VOR TO SEA 116.8 AND DETERMINED ACFT WAS ON 007 RADIAL AT 3200 MSL. BY THE TIME INCURSION WAS CONFIRMED THE ACFT HAD EXITED THE TCA. WE THEN TURNED E TO AVOID FURTHER INCURSION. PIC GOT VERY NERVOUS ABOUT TCA INCURSION AND BEING WORRIED ABOUT POSSIBLE CERTIFICATE ACTION, TURNED OFF MODE C HOPING THAT ATC WOULD NOT NOTICE THE INCURSION. THIS WAS SIMPLY BAD JUDGEMENT. THE PIC ALSO DECIDED TO FLY AMONGST THE HILLS HOPING TO LOST ATC IF ATC WAS AWARE OF THE INCURSION; MORE BAD JUDGEMENT. THE ALTIMETER IN THE ACFT WAS UNLIGHTED AND UNCERTIFIED AS THE ACFT'S REGULAR ALTIMETER WAS BEING SERVICED. THE ALTIMETER INSTALLED IN THE ACFT WAS A LONER AND BEING UNLIGHTED WAS NOT EASY TO SEE IN THE DARK COCKPIT. THE PIC SHOULD HAVE CHKED ALL OF THIS BEFORE THE FLT RATHER THAN AFTER THE FLT. TECHNICALLY, THE ACFT WAS PROBABLY NOT AIRWORTHY DUE TO NONCERTIFICATION OF THE ALTIMETER. THE FLT SHOULD HAVE BEEN MADE IN AN AIRWORTHY CRAFT OR POSTPONED. THE PIC IS A CFI WORKING TOWARDS A CAREER IN COMMERCIAL AVIATION. THE SPECTER OF POSSIBLE SUSPENSION OR REVOCATION OF LICENSE DUE TO THE TCA INCURSION WAS FEARSOME. PIC

Time of Day	Altitude	Crew Size	Mission	Flight Phase	Narrative
					WAS VERY NERVOUS ABOUT THE REPUTED 'MANDATORY' 60 DAY SUSPENSION OF LICENSE. HAD THE SUSPENSION NOT BEEN 'MANDATORY' THE PIC WOULD HAVE CALLED ATC, RPTED POSITION AND EXITED TCA ACCORDING TO ATC INSTRUCTIONS (HAD HE STILL BEEN IT IT).
Daylight		2	Passenger	Parked	OUR COMPANY HAS A POLICY AND APPROVED TRAINING ALLOWING FO PLTS TO PERFORM ENG STARTS. MKE HAS NO GND PWR UNIT AND HAVE NOT HAD FOR THE YR WE'VE BEEN OPERATING HERE. ALL ENG STARTS ARE PERFORMED USING BATTERY PWR. FO STATES THAT DURING A BATTERY START OF ENG #2 THAT ITT LIMIT OF 950 DEGS C WAS EXCEEDED (HE ESTIMATES APPROX 970 DEGS FOR 2 SECONDS). THOUGH NOT AT THE IDEAL VANTAGE POINT (I WAS OUTSIDE AT THE FRONT OF THE ACFT WITH A RAMPER) I COULD SEE MY FO VISUALLY MONITORING THE START (AS OPPOSED TO DAYDREAMING OR GIRL WATCHING). I HEARD THE STARTER ENGAGE, THE IGNITORS CLICK THE LIGHT OFF OF THE ENG AND THE IMMEDIATE SHUTDOWN OF THE ENG FOLLOWED BY THE FO HAND-SIGNALING THAT HE HAD ABORTED THE START. I ADD THIS TO POINT OUT THAT I FEEL THE FO WAS CONSCIENTIOUS DURING THE START. WHILE WRITING THIS, MAINT HAS ARRIVED FROM OUR HOME BASE. HE HAS TESTED THE FLT DATA INFO UNIT (FDAU) WHICH STATES FDR AND OTHER INFO. HIS DIAGNOSTIC EQUIP SHOWS THAT NO ENG LIMITS WERE EXCEEDED. MY LABOR UNION COUNSEL HAS TOLD ME THAT EVEN THOUGH THE FO IS APPROVED FOR ENG STARTS, EVEN IF I AM NOT PRESENT, THE FAA MIGHT STILL ATTEMPT TO FIND FAULT WITH ME OR TAKE ACTION AGAINST ME. THUS, THIS RPT.

Time of Day	Altitude	Crew Size	Mission	Flight Phase	Narrative
Daylight	2700	2	Cargo / Freight	Descent; Descent	I WAS BEING VECTORED FOR AN ILS APCH RWY 13, AT MDT IN IMC CONDITIONS. I WAS CLRED FOR DSCNT FROM 4000 FT TO 3000 FT. I DSNDED TO 2700-2650 FT BEFORE CLBING BACK TO 3000 FT. AT 2700 FT THE ALT CHIME SOUNDED. AT 4000 FT, THE CAPT DIALED IN 3000 FT IN THE ALT ALERTER AND BECAUSE OF MY DSCNT BELOW 4000 FT THE ALT CHIME DID NOT SOUND A 1000 FT WARNING. THIS IS SOMETHING I'M USED TO HEARING. I HAD BEEN AWAKE SINCE XA00 AND HAD STARTED MY TRIP IN SALT LAKE CITY. I WAS WELL WITHIN DUTY AND FLT TIME LIMITS, BUT I WAS TIRED. UPON NEARING MY FINAL LNDG OF THE DAY, I GOT COMPLACENT, MAYBE EVEN DAYDREAMING. IN ANY CASE, I LET MY ALT SLIP, AND CORRECTED IT WITH NO ADVERSE AFFECTS. AS I'VE TOLD MYSELF AND OTHERS MANY TIMES BEFORE, WHEN YOU'RE TIRED YOU MUST BE MORE ALERT AND DON'T LET COMPLACENCY ENTER THE COCKPIT.
Night	27000	2	Passenger	Cruise; Cruise	ENRTE ON JET AIRWAY, AT FL290, ATC ISSUED A CLRNC TO CROSS FAK VOR AT FL270. WE WERE APPROX 70 NM FROM FAK AND ONLY 2000 FT TO DSND. BOTH THE FO AND MYSELF FOLLOWED COMPANY PROCS BY SETTING 27000 FT IN THE ALT ALERTER AND ACKNOWLEDGING XING RESTR. BECAUSE IT WAS AN ABSOLUTELY CLR, BEAUTIFUL NIGHT WITH UNLIMITED VISIBILITY, AND DSCNT NEED NOT START IMMEDIATELY, I LOOKED OUT THE WINDOW TO SIGHT-SEE AND ENJOY THE SUNSET. NEXT THING I HEAR IS ATC ADVISING US TO DSND TO FL270. I THEN NOTICED WE WERE 3 NM FROM FAK VOR. ONCE WE LEVELED AT FL270 WE WERE 10 NM THE OTHER SIDE OF THE VOR. WE HAD MISSED OUR XING RESTR BY 10 NM! TO MY KNOWLEDGE, NO TFC CONFLICT EXISTED, NO

Time of Day	Altitude	Crew Size	Mission	Flight Phase	Narrative
					EVASIVE ACTION WAS TAKEN AND ATC NEVER QUESTIONED OUR DSCNT. ALTHOUGH BOTH THE FO AND MYSELF FOLLOWED ALL CRM AND ALT AWARENESS PROCS, I FAILED IN LETTING MY GUARD DOWN TO ENJOY A BEAUTIFUL EVENING OF FLYING. IN THE FUTURE, I SHOULD LEAVE THE DAYDREAMING (NIGHT DREAMING) TO OUTSIDE THE COCKPIT!
Daylight	3500	2	Agriculture	Cruise	I FLY A 2 PLT CREW ENVIRONMENT AND I AM THE FO AND FOR THE OCCURRENCE, ALSO THE PNF. IT WAS AN XA00 TO XB00 LCL AND WE WERE ON THE SECOND FLT OF THE DAY. WE WERE WORKING N/S PASSES OVER PDZ VOR (PARADISE) BTWN 2 NM N OF ONT AND PLEASANTS PEAK. MY CAPT IS AN EXCELLENT CAPT AND I FEEL THIS DAY HIS MIND WAS NOT AT ALL ON FLYING. ON A NBOUND PASS AT 2 NM S OF PDZ SOCAL CALLED TFC 11 O'CLOCK POS MOVING TO 12 O'CLOCK POS SE AT OUR ALT -- IT WAS A TFC ALERT. I SAW THE ACFT (A C172) AS SOCAL WAS GIVING THE MESSAGE AND RESPONDED, 'IN SIGHT, WE'LL DSND AND PASS BEHIND.' MY CAPT MAINTAINED ALT AND HEADING. I ASKED 'DO YOU SEE HIM?' HE REPLIED 'YES.' 'WE NEED TO DSND' I REPLIED. HE SAID 'YES.' I THEN SAW IT WAS GETTING TOO CLOSE -- I CALLED OUT 'CAPT DSND!' AND PUSHED FORWARD ON THE STICK. IT WAS NOT SEVERELY LIFE THREATENING, JUST WAY TOO CLOSE FOR COMFORT AND SAFETY. AFTER WE PASSED THE TFC MY CAPT SEEMED TO BE JARRED OUT OF A TRANCE. HE SAID 'I'M SORRY.' I WAS TOO IRRITATED TO RESPOND. I THINK THE PROB CAME FROM THE CAPT HAVING TOO MUCH ON HIS MIND. I DON'T BELIEVE IT WAS INTENTIONAL, JUST LACK OF ATTN. EVERYONE HAS OTHER PARTS OF LIFE THAT INFLUENCE THEIR FEELINGS -- FLYING IS NOT THE TIME TO

Time of Day	Altitude	Crew Size	Mission	Flight Phase	Narrative
					BE DAYDREAMING.
Daylight		1	Personal	Taxi	I HAD JUST COMPLETED LNDG AND WAS APPLYING BRAKES TO TURN OFF THE RWY. AS I REACHED TO RETRACT FLAPS, MY HAND WENT TO THE GEAR RETRACTION LEVER INSTEAD. BEFORE I COULD REACT, THERE WAS A SINKING FEELING AS THE PLANE, AND MY STOMACH SETTLED ONTO THE RWY. THERE WERE NO INJURIES, I WAS ALONE IN THE PLANE, AND NO PROPERTY DAMAGE, OTHER THAN MY PROP. I CANNOT FAULT THE PLANE. IT WAS MY OWN LAPSE. I GUESS I WAS DAYDREAMING ABOUT HOW NICE A DAY IT WAS TO FLY.
Daylight	11000	2	Passeng er	Initial Climb	DEPARTED DAL WITH AN IFR CLRNC FOR RADAR VECTORS FOR THE JOE POOL.3 DEP. ON CLBOUT, RECEIVED INSTRUCTION FOR IMMEDIATE EVASIVE ACTION, 'TURN TO HDG 270 DEGS.' SECONDS LATER, RECEIVED AURAL AND VISUAL WARNING OF 'TFC' FROM ONBOARD TCAS. CTLR ADVISED THAT I HAD ONLY BEEN CLRED TO 10000 FT MSL AND HAD CREATED A 'NEAR MISS' AS I CLBED THROUGH 110000 FT MSL. I THOUGHT I HAD BEEN CLRED TO 12000 FT MSL, AND UPON HIS WARNING VERIFIED THAT I HAD WRITTEN DOWN AND ENTERED 12000 FT IN THE ALT PRE-SELECTOR. CTLR MAINTAINED THAT ANOTHER ACFT HAD BEEN CLRED TO 12000 FT AND THAT THEY HAD ACKNOWLEDGED THE CLRNC. I WOULD LIKE TO HEAR THE AUDIO RECORDING OF THE CLRNC AND ACKNOWLEDGEMENT XMISSIONS AS I REALLY HAD NO DOUBT ABOUT BEING CLRED TO 12000 FT. IF I WAS IN FACT ONLY CLRED TO 10000 FT AND ACKNOWLEDGED SUCH, THEN MY BRAIN WAS COMPLETELY DISCONNECTED. I HAVE HAD TIMES WHEN I WAS DAYDREAMING OR UNSURE OF A CLRNC

Time of Day	Altitude	Crew Size	Mission	Flight Phase	Narrative
					FOR WHATEVER REASON, BUT THIS WAS NOT ONE OF THEM. NEVERTHELESS, IN HINDSIGHT THE ONE THING I REALIZE I SHOULD HAVE DONE BETTER IS MONITORING THE TCAS DISPLAY. RECEIVING A TFC WARNING SHOULDN'T BE A SURPRISE IF YOU ARE APPROPRIATELY MONITORING THE MOVEMENT OF TARGETS, WHICH I WAS NOT. EVEN IF I WAS CLRED TO 12000 FT, SEEING THE APCHING TFC WOULD HAVE CAUSED ME TO QUESTION THE CTLR THUS POINTING OUT OUR DIFFERENCE OF UNDERSTANDING AND ALLOWING A MORE TIMELY CORRECTION OF THE SITUATION. WE ARE THE BENEFICIARIES OF A NUMBER OF EXTRAORDINARY PIECES OF TECHNOLOGY, BUT THEY ARE ONLY BENEFICIAL TO THE EXTENT WE UTILIZE THEM.

APPENDIX B. *Demographics of Factor Scenario Questionnaire Sample (n=114)*

Variable	Level	Value
Native Language	English	37
	German	2
	Persian	1
Student Pilot?	Yes	5
	No	35
Certificates held	ATP	6
	Private	27
	CFI	11
	CFII	8
	Commercial	19
	Flight Engineer	0
	Flight Engineer (Written)	1
	Instrument	24
	MEI	7
Hours/month under:	FAR Part 91	*M*=1180.80 hours
		SD=6887.83 hours
	FAR Part 121	*M*=6.82 hours
		SD=22.61 hours
	FAR Part 135	none
Currently employed by air carrier?	Yes	0
	No	39
# of participants who	Air Taxis	0

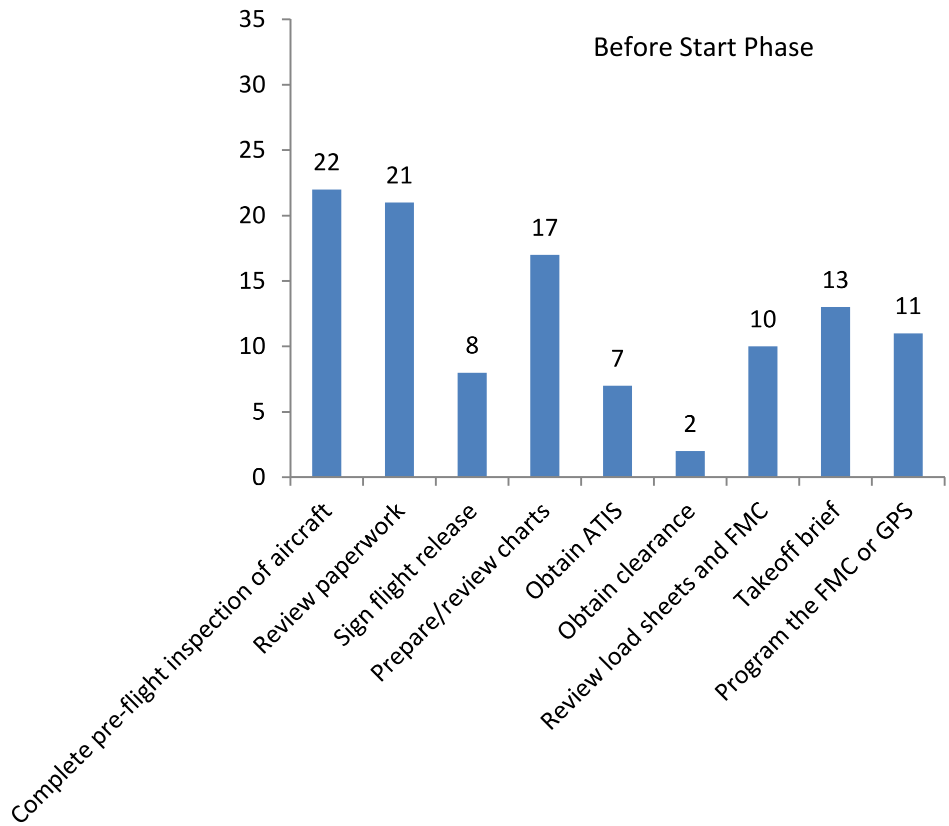
Variable	Level	Value
predominantly fly for the following:	Commuter/Regional	1
	Major Commercial	1
	Cargo	0
	Military	2
	UAV	0
	GA	17
# aircraft types flown	Mean	5.97 types
	SD	5.85 types
Total # flight hours	Mean	1477.98 hours
	SD	3038.41 hours
Ave # days/month of flight	Mean	9.38 days
	SD	31.56 days

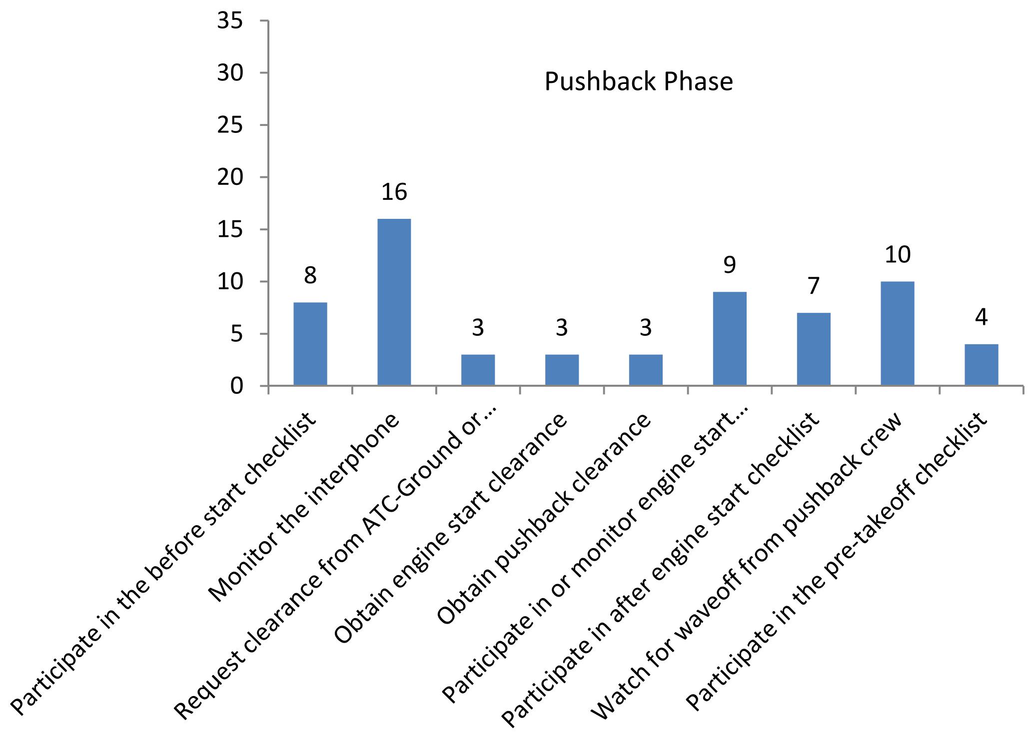
APPENDIX C. *Factor Scenario Questionnaire Results: Participants' Mean Ratings for Each Dimension (Temporal, Functional, Conceptual, Tense) of Mind Wandering Construct*

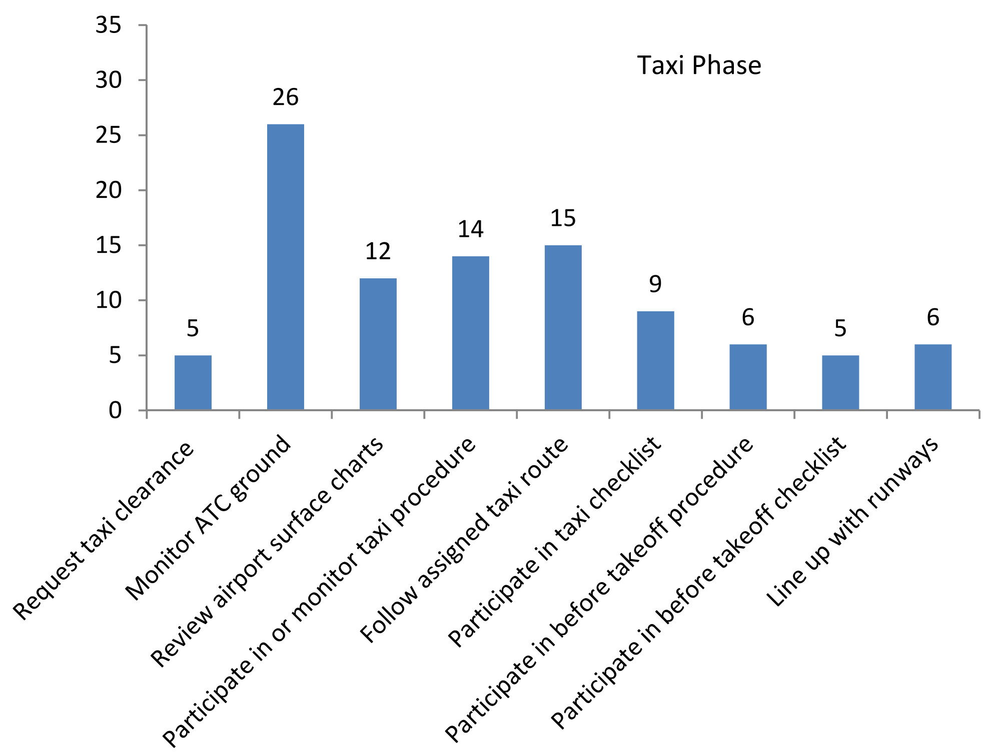
Scenario	N	Tense	Temporal	Temporal *M*	Functional	Functional *M*	Conceptual	Conceptual *M*	Mind Wandering =yes
1	13	Retrospective	High	4.4	High	5.8	High	5.1	8
2	15	Retrospective	High	3.8	Low	5.8	High	5.8	10
3	15	Retrospective	High	6.0	High	5.3	Low	5.9	5
4	15	Retrospective	High	4.9	Low	3.4	Low	7.99	11
5	13	Retrospective	Low	3.7	High	4.5	High	2.7	10
6	4	Retrospective	Low	6.7	Low	4.0	High	2.0	3
7	13	Retrospective	Low	4.3	High	5.5	Low	5.9	11
8	15	Retrospective	Low	4.1	Low	4.0	Low	4.0	13
9	15	Prospective	High	6.3	High	7.8	High	8.4	1
10	13	Prospective	High	4.6	Low	5.6	High	5.7	9
11	4	Prospective	High	5.3	High	4.3	Low	5.8	1
12	13	Prospective	High	3.4	Low	3.3	Low	3.8	13
13	15	Prospective	Low	4.2	High	3.8	High	4.9	10
14	13	Prospective	Low	6.2	Low	4.0	High	3.6	8
15	4	Prospective	Low	4.3	High	3.8	Low	6.3	2

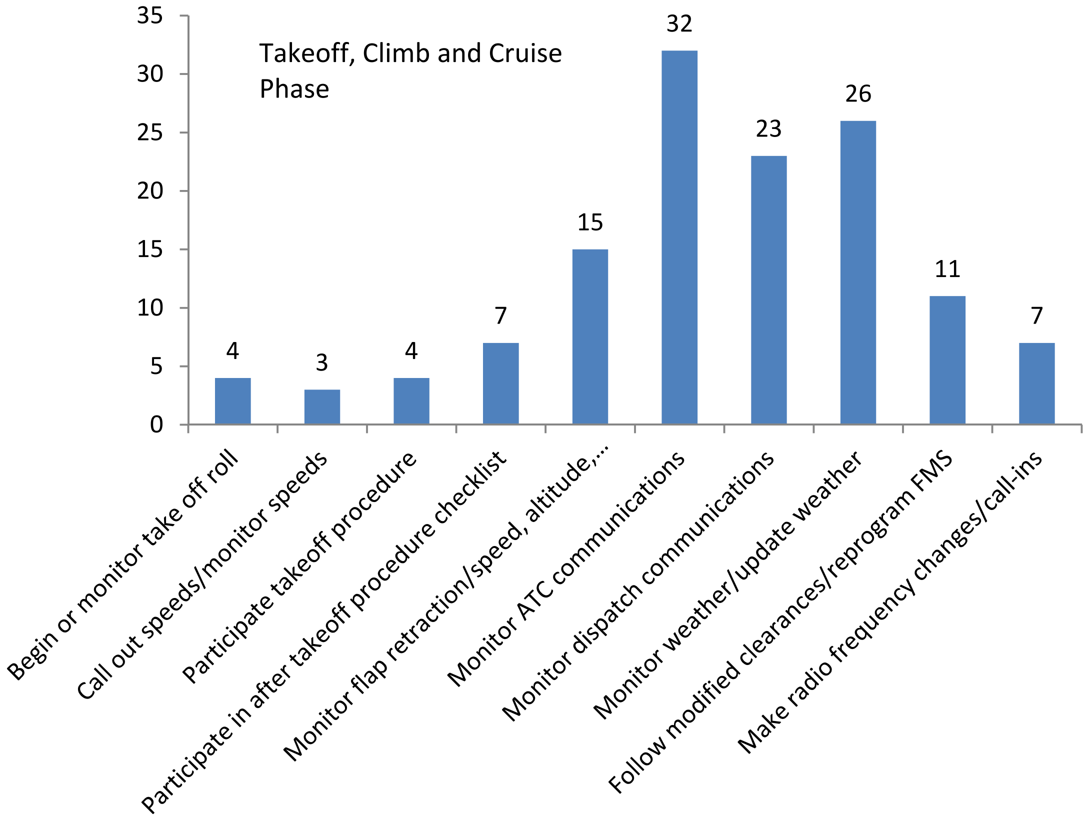
Scenario	N	Tense	Temporal	Temporal *M*	Functional	Functional *M*	Conceptual	Conceptual *M*	Mind Wandering =yes
16	12	Prospective	Low	4.3	Low	2.6	Low	5.2	12

APPENDIX D. *Series of Bar Charts Depicting the Frequencies of Mind Wandering for Each Flight Task within Each Phase of Flight*

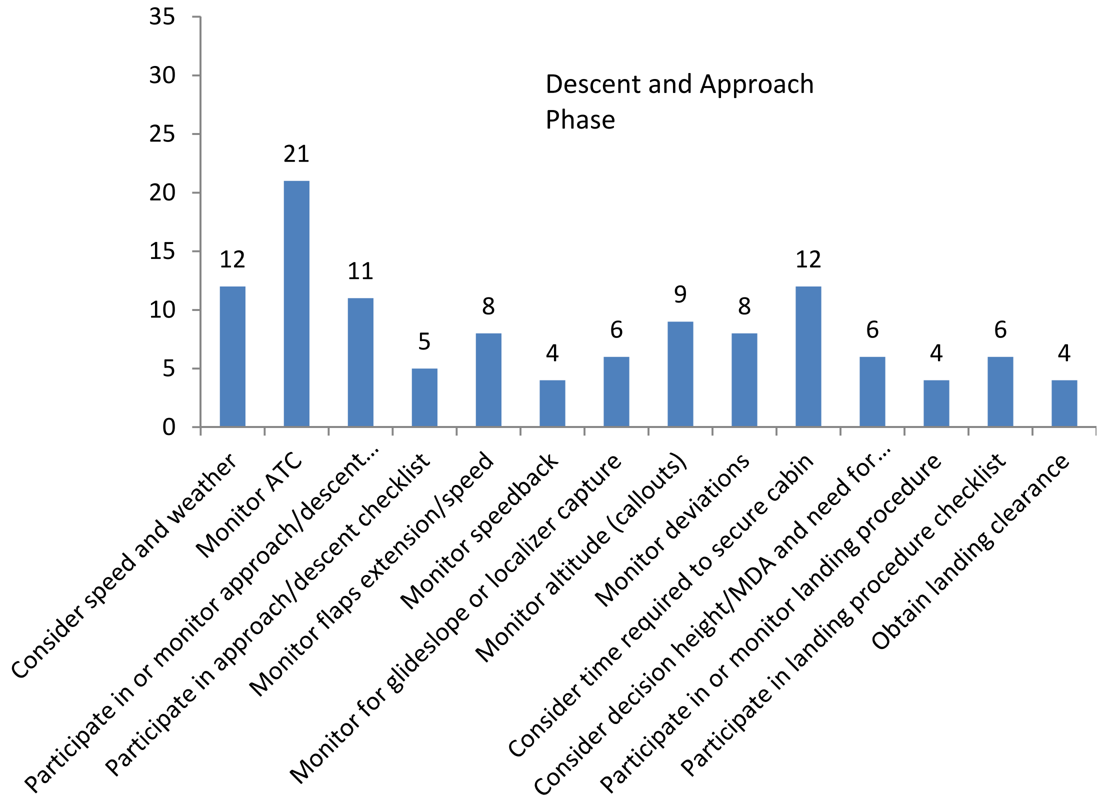


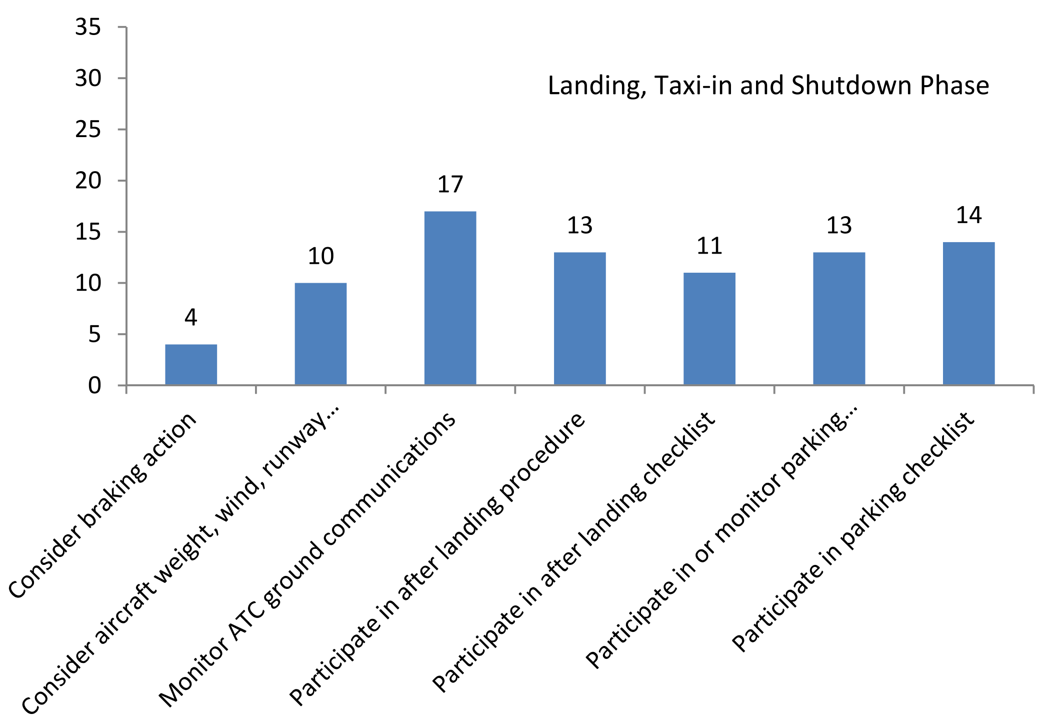

Pushback Phase
35
30
25
20
15
10
5
0
8
16
3
3
3
9
7
10
4
Participate in the before start checklist
Monitor the interphone
Request clearance from ATC-Ground or ...
Obtain engine start clearance
Obtain pushback clearance
Participate in or monitor engine start...
Participate in after engine start checklist
Watch for waveoff from pushback crew
Participate in the pre-takeoff checklist


Taxi Phase
35
30
25
20
15
10
5
0
5
26
12
14
15
9
6
5
6
Request taxi clearance
Monitor ATC ground
Review airport surface charts
Participate in or monitor taxi procedure
Follow assigned taxi route
Participate in taxi checklist
Participate in before takeoff procedure
Participate in before takeoff checklist
Line up with runways



Takeoff, Climb and Cruise Phase
35
30
25
20
15
10
5
0
4
3
4
7
15
32
23
26
11
7
Begin or monitor take off roll
Call out speeds/monitor speeds
Participate takeoff procedure
Participate in after takeoff procedure checklist
Monitor flap retraction/speed, altitude,...
Monitor ATC communications
Monitor dispatch communications
Monitor weather/update weather
Follow modified clearances/reprogram FMS
Make radio frequency changes/call-ins


Descent and Approach Phase
35
30
25
20
15
10
5
0
12
21
11
5
8
4
6
9
8
12
6
4
6
4
Consider speed and weather
Monitor ATC
Participate in or monitor approach/descent…
Participate in approach/descent checklist
Monitor flaps extension/speed
Monitor speedback
Monitor for glideslope or localizer capture
Monitor altitude (callouts)
Monitor deviations
Consider time required to secure cabin
Consider decision height/MDA and need for…
Participate in or monitor landing procedure
Participate in landing procedure checklist
Obtain landing clearance


Landing, Taxi-in and Shutdown Phase
35
30
25
20
15
10
5
0
4
10
17
13
11
13
14
Consider braking action
Consider aircraft weight, wind, runway…
Monitor ATC ground communications
Participate in after landing procedure
Participate in after landing checklist
Participate in or monitor parking…
Participate in parking checklist

APPENDIX E. *Definitions of Each Ability and Mental Process from the US Department of Labor's O*Net Site*

Element Name	Description
Abilities	Enduring attributes of the individual that influence performance
Cognitive Abilities	Abilities that influence the acquisition and application of knowledge in problem solving
Knowledge	Organized sets of principles and facts applying in general domains
Mental Processes	What processing, planning, problem-solving, decision-making, and innovating activities are performed with job-relevant information?
Psychomotor Abilities	Abilities that influence the capacity to manipulate and control objects
Physical Abilities	Abilities that influence strength, endurance, flexibility, balance and coordination
Skills	Developed capacities that facilitate learning or the more rapid acquisition of knowledge
Spatial Abilities	Abilities related to the manipulation and organization of spatial information

APPENDIX F. *Table of O*Net's Abilities and Mental Processes Required by Probe*

Probe	Tasks	Abilities Required	Total Abilities	Mental Processes	Total Mental Processes
1	Interpret car feedback, (Is the steering and acceleration going as planned?)	Time Sharing (CA), Near vision (SA), perceptual speed (CA)	3	Analyzing Data or Information (MP)Processing Information (MP)	2
	Use feedback to make steering adjustments	Information ordering (CA), Control precision (PM), Multi-limb coordination (PM), Inductive reasoning (CA)	4	Processing Information (MP)Updating and Using Relevant Information (MP)	2
	decide when to terminate acceleration	Inductive Reasoning (CA), Far vision (SA), Inductive reasoning (CA)	3	Analyzing Data or Information (MP) Scheduling work and activities (MP) Organizing, planning and prioritizing work (MP) Making Decisions and Problem solving (MP) Updating and Using Relevant Information (MP)	5
	Notice the vehicles on	Near vision (SA), Depth	2	Processing Information	3

Probe	Tasks	Abilities Required	Total Abilities	Mental Processes	Total Mental Processes
	the right and monitor for safety concerns	Perception (SA)		(MP) Updating and using relevant information (MP) Develop Objectives and Strategies (MP)	
	Terminate acceleration	Multi-limb coordination (PM) Control precision (PM),	2	Updating and Using Relevant Information (MP)	1
	check speed	Control precision (PM) Near vision (SA), problem sensitivity (CA)	3	Processing Information (MP)	1
	Visually scan environment for stop sign	Depth Perception (SA), Far vision (SA), Visual Color Discrimination (SA)	3	Processing Information (MP) Updating and using relevant information (MP)	2
	Notice stop light	Far vision (SA), Color Discrimination (SA)	2	Processing Information (MP) Updating and Using Relevant Information (MP)	2
	Decide when to begin braking (if at all in preparation for the light changing to	Rate Control (PM), Reaction time(PM), Depth Perception (SA), Far Vision (SA), Visual Color	8	Analyzing Data or Information (MP) Developing objectives and strategies (MP) Organizing, planning and	6

Probe	Tasks	Abilities Required	Total Abilities	Mental Processes	Total Mental Processes
	red),	Discrimination (SA), Spatial Orientation (CA), Inductive reasoning (CA), perceptual speed (CA)		prioritizing work (MP) Making decisions and solving problems (MP) Processing information (MP) Updating and Using Relevant Information (MP)	
	place foot on brake	control precision (PM)	1		0
2	Check to make sure I did not coast past speed limit	Near vision (SA), Deductive Reasoning (CA)	2	Judging the qualities of things, services or people (MP) Processing Information (MP) Updating and Using Relevant Information (MP)	3
	Visually scan for road sign ahead	Far vision (SA), Depth Perception (SA), Color discrimination (SA),	3	Processing Information (MP) Updating and using relevant information (MP)	2
	Notice intersection approaching	Far vision (SA), Depth Perception (SA)	2	Processing Information (MP) Updating and using relevant information	3

Probe	Tasks	Abilities Required	Total Abilities	Mental Processes	Total Mental Processes
				(MP) Processing Information (MP)	
	Recall name of street to turn onto	Memorization (CA), perceptual speed (CA)	2	Processing Information (MP)	1
	Identify white road sign ahead	Far vision (SA), Depth Perception (SA), Color discrimination (SA)	3	Processing Information (MP) Updating and using relevant information (MP)	2
	Determine that this is not a road sign but speed limit sign	Color discrimination (SA), Inductive reasoning (CA)	2	Analyzing Data or Information (MP) Making Decisions and Problem solving (MP) Organizing, planning and prioritizing work (MP) Updating and Using Relevant Information (MP)	4
	Check blind spots as I approach intersection	Near vision (SA), problem sensitivity (CA)	2	Judging the qualities of things, services or people (MP) Organizing, planning and prioritizing work (MP) Processing Information	4

Probe	Tasks	Abilities Required	Total Abilities	Mental Processes	Total Mental Processes
				(MP) Updating and Using Relevant Information (MP)	
	Notice strange item in the roadway	Near vision (SA)	1	Processing Information (MP) Updating and Using Relevant Information (MP)	2
	Determine what the strange item is	Near vision (SA), perceptual speed (CA), problem sensitivity (CA)	3	Analyzing Data or Information (MP) Judging the qualities of things, services or people (MP) Making Decisions and Problem solving (MP) Updating and Using Relevant Information (MP)	4
3	Scan visual environment to locate all items in environment	Far vision (SA), Depth perception (SA)	2	Processing Information (MP) Updating and using relevant information (MP)	2
	Scan rear view for traffic behind	Near vision (SA)	1	Processing Information (MP) Updating and Using Relevant Information	2

Probe	Tasks	Abilities Required	Total Abilities	Mental Processes	Total Mental Processes
				(MP)	
	Notice automobiles on the right	Far vision (SA), depth perception (SA)	2	Judging the qualities of things, services or people (MP) Processing Information (MP) Updating and Using Relevant Information (MP)	3
	Check speed	Near vision (SA)	1	Processing Information (MP) Updating and Using Relevant Information (MP)	2
	Determine if these automobiles are parked or waiting to merge into lane of travel	Far vision (SA), Depth perception (SA), Inductive reasoning (CA), visualization (CA,)perceptual speed (CA), problem sensitivity (CA)	6	Analyzing Data or Information (MP) Developing objectives and strategies (MP) Organizing, planning and prioritizing work (MP) Making decisions and solving problems (MP) Processing information (MP) Updating and Using Relevant Information	6

Probe	Tasks	Abilities Required	Total Abilities	Mental Processes	Total Mental Processes
				(MP)	
4	begin countersteer input	Multi-limb coordination (PM), control precision (PM),	2		0
	Use feedback to make steering adjustments	Information ordering (CA), Multi-limb coordination (PM), control precision (PM), Inductive reasoning (CA)	4	Processing Information (MP) Updating and Using Relevant Information (MP)	2
	Determine if I have a lane deviation concern	Inductive reasoning (CA), Near vision (SA), perceptual speed (CA), problem sensitivity (CA)	4	Judging the qualities of things, services or people(MP) Making Decisions and Problem solving(MP) Organizing, planning and prioritizing work (MP) Updating and Using Relevant Information (MP)	4
	Monitor acceleration rate	Near vision (SA)	1	Judging the qualities of things, services or people(MP) Processing Information (MP) Updating and Using Relevant Information	3

Probe	Tasks	Abilities Required	Total Abilities	Mental Processes	Total Mental Processes
				(MP)	
	Determine when to terminate acceleration	Inductive reasoning (CA)	1	Analyzing Data or Information (MP) Scheduling work and activities(MP) Organizing, planning and prioritizing work (MP)Making Decisions and Problem solving (MP) Updating and Using Relevant Information (MP)	5
	Terminate acceleration	Multi-limb coordination (PM), control precision (PM), Deductive reasoning (CA)	3		0
	Scan visual environment to locate all items in environment	Far vision (SA), Depth perception (SA)	2	Processing Information (MP) Updating and using relevant information (MP)	2
	decide to stay in the left lane initially	Inductive reasoning (CA), Time Sharing (CA),	2	Analyzing Data or Information (MP) Making Decisions and Problem solving (MP) Organizing,	3

Probe	Tasks	Abilities Required	Total Abilities	Mental Processes	Total Mental Processes
				planning and prioritizing work (MP)	
5	Initiate acceleration	Multi-limb coordination (PM), control precision (PM)	2		0
	Begin countersteer input	Multi-limb coordination (PM), Control precision (PM)	2		0
	Determine if I have a lane deviation concern	Near vision (SA), Inductive reasoning (CA), problem sensitivity (CA)	3	Analyzing Data or Information (MP) Judging the qualities of things, services or people(MP) Making Decisions and Problem solving (MP) Organizing, planning and prioritizing work (MP) Updating and Using Relevant Information (MP)	5
	Scan visual environment to locate all items in environment	Far vision (SA), depth perception (SA)	2	Processing Information (MP) Updating and using relevant information (MP)	2
	Identify automobiles	Far vision (SA), depth	2	Processing Information	2

Probe	Tasks	Abilities Required	Total Abilities	Mental Processes	Total Mental Processes
	on right side of street	perception (SA)		(MP)Updating and using relevant information(MP)	
	Determine if these automobiles are parked or waiting to merge into lane of travel	Far vision (SA), depth perception (SA), Inductive reasoning (CA), perceptual speed (CA) problem sensitivity (CA)	5	Analyzing Data or Information (MP) Judging the qualities of things, services or people (MP) Making Decisions and Problem solving(MP) Organizing, planning and prioritizing work(MP) Updating and Using Relevant Information (MP)	5
	Create a safety plan to circumvent collision if cars pull into lane	Response orientation (PM), Deductive reasoning (CA), problem sensitivity (CA)	3	Developing objectives and strategies (MP) Organizing, planning and prioritizing work (MP) Making decisions and solving problems (MP)Processing information(MP) Judging the qualities of things, services or people (MP)	5

Probe	Tasks	Abilities Required	Total Abilities	Mental Processes	Total Mental Processes
	Notice woman waiting in between cars.	Far vision (SA), depth perception (SA), Reaction time (PM), Speed of limb movement (PM), perceptual speed (CA), problem sensitivity (CA), visualization (CA)	7	Processing Information (MP) Updating and using relevant information (MP) Judging the qualities of things, services or people (MP)	3
	Identify stop sign	Far vision (SA), depth perception (SA), color discrimination (SA)	3	Processing Information (MP) Updating and using relevant information (MP)	2

APPENDIX G. *Table of Demographics for Part III Driving Simulation Study Sample (n=120)*

Questionnaire Section	Variable	Value	
Basic Demographics	Age	*M*=19.05 years	
		SD=1.25 years	
	Gender	Male	77 (64%)
		Female	43 (36%)
	Fist Language	English	109 (91%)
		Spanish	4 (3%)
		Gujarati	2 (2%)
		Other	5 (4%)
Driving Demographics	Valid license?	Yes	118 (98%)
		No	2 (2%)
	Number of years with license	*M*=2.38 years	
		SD=1.90 y ears	
		"don't know"	12 (10%)
	While in college, on average how often do you drive an automobile?	I do not drive while I'm attending college	34 (28%)
		0- 1 round-trip drives per week	23 (19%)
		2-4 round-trip drives per week	25 (21%)
		5-6 round-trip drives per week	10 (8%)
		7+ round-trip drives per week	28 (23%)

Questionnaire Section	Variable	Value	
	What percentage of your driving, while in college, is in high-density (e.g., urban) traffic?	*M*=47.38% *SD*=37.18%	
	What percentage of your driving, while in college, is in low-density (e.g., rural) traffic?	*M*=24.28% *SD*=26.42%	
	Access to automobile	I own a vehicle	80 (67%)
		I don't own but I have access	33 (28%)
		No access	7 (6%)
	Type of transmission you predominantly drive	Manual	14 (12%)
		Automatic	106 (88%)
Distractibility	Would you consider yourself to be generally more distractible than your peers?	Yes	51 (43%)
		No	51 (43%)
		Not sure	18 (15%)

APPENDIX H. *Copies of Consent Form and All Questionnaires Used in Part III Driving Simulation Study*

Consent Form

Thinking While Driving
RESEARCH INFORMATION AND CONSENT FORM
Primary Investigator: Jennifer Cowley
Faculty Advisor: Doug Gillan

Introduction:
You are invited to participate in a research study investigating what people think about when they are driving an automobile. You were selected as a possible participant in this research because you have some experience with driving an automobile. Please read this form and ask questions before you decide whether to participate in the study.

Background Information:
The purpose of this study is to understand what you may be thinking about while you are driving. Approximately 150 people are expected to participate in this research.

Procedures:
If you decide to participate, you will be asked to first complete a few short evaluations to help us understand how you think and to assess what you believe task-unrelated thinking (e.g., mind wandering) is. Then you will be introduced to the simulator to allow you to practice driving it in. Then, you will be given a set of directions and expected to drive the designated route. However, right before you begin driving, you will be asked to list a certain number of prominent and important life events on a piece of paper such that you can tell us when, if at all, you think about them during driving. After the driving is complete, we will provide a short follow-up questionnaire and discuss informally with you some aspects of the driving task. This study will take a little more than 1.5 hours of your time.

Risks and Benefits:
The study has four risks. First, during the working memory test, participants may feel as if they've failed because the test does not provide performance information. This may affect people who have test anxiety and it is unclear what percentage of our participants will have test anxiety. Second, participants are taking the Beck Depression Inventory (BDI) to measure mood. It is likely that some individuals may be reminded of negative personal events while taking this test and so the participant is advised to tell the researcher if he/she wishes to forgo the test. Should the test become overwhelming, resources to professional services will be provided (e.g., http://www.suicide.org/hotlines/north-carolina-suicide-hotlines.html, the counseling center at NCSU). The participant will be informed that the BDI results are confidential and that the answers will not be analyzed but rather the total score will be used. The BDI will never be used as a clinical diagnosis for mood disorders and the answers provided in this test will not be reviewed by the researcher; only the total score. Third, participants will list several salient life events which are a personal. In addition, when the participant admits that they are mind wandering, he/she may willingly elaborate on each salient life event yet the participant may feel uncomfortable disclosing personal information with the researcher. Fourth, the participant will be videotaped but the camera will be pointing to the back and side of their head. The focus of the camera will be on the keyboard and screen and never will the camera be pointing at the participant's face unless they turn and look directly at the camera; which identifies the participant.

There are no direct benefits to participating in this study.

Compensation:
If you have participated in this study through the Experimetrix program, you will receive 1 credit for every 30 minutes of your time in addition to $10 for the entire session. Each participant will be paid at the end of the study but if the participant terminates the study, he/she will still be paid for the time they spent in the study and will be compensated full credit on the Experimetrix site. The monetary compensation will not be pro-rated; he/she will receive five dollars per hour, even if he/she completes the study in less time. If you are recruited through class announcements or flyers in the hall, you will be paid $10 for your participation and again, this amount will not be pro-rated.

Confidentiality:
Any identifying information obtained in connection with this research study will be kept confidential. In any written reports or publications, no study participant will be identified or identifiable and only group data will be presented. All data collected, including audio and video files, will be analyzed on a password protected computer or secured in a locked storage cabinet in 740 Poe Hall when not being analyzed. Only the principal investigator and faculty advisor will be allowed access to all data but some of the faculty collaborators will be allowed to see subsets of the data that are relevant to their interests. All data will be analyzed by Fall of 2011.

Voluntary nature of the study:
Participation in this research study is voluntary and should you refuse at any time to participate, you will not be financially or educationally penalized.

New Information:
If during course of this research study I learn about new findings that might influence your willingness to continue participating in the study, I will inform you of these findings.

You may keep a copy of this form for your records.

Statement of Consent:
You are making a decision whether or not to participate. Please place an "X" in the check boxes next to the statements you consent to below. Your signature indicates that you are at least 18 years of age, that you have read this information, and your questions have been answered. Even after signing this form, please know that you may withdraw from the study at any time and no further data will be collected.

*

- [ ] I have read and I understand the above information.
- [ ] I agree to participate in this study and confirm that I am at least 18 years of age.
- [ ] I do not wish to participate in this study.

Salient Life Events Questionnaire

Every day, people experience pleasant and unpleasant events that are meaningful to their lives. These include:

Serious illnesses	Marriages/weddings
Awards and distinctions	Accidents
Birth or death of a loved one	Profound events that changed your life
Academic achievements	Vacations in the past or in the future
Religious events	Relationships with friends or family

In the blanks provided below, we'd like you to list the top six most important or salient positive or negative life events that you are experiencing *today*. Please provide a one- or two-sentence description in the blank as well as number between 0 and 100 indicating your rating of importance.

|---|

0 — Not at all Important

100 — Extremely Important

1.____

____

Importance rating: ___________

2.____

____

Importance rating: ___________

3.____

____

Importance rating: ___________

4.____

____

Importance rating: ___________

5.____

____

Importance rating: ____________

6.____

____

Importance rating: ____________

Verbal Protocol for Data Collection for Each Probe

Any **salient thoughts on** your list? ___yes ___no

What number(s)? __________

Any other thoughts beyond your salient thoughts in your list? (please vocalize)

Were you **conscious** of your thoughts at the time of the probe? ___yes ___no

What **task** were you executing at the time your thoughts were stopped? (please vocalize)

Temporal relatedness

|--|

Not at all Temporally Related	Completely Temporally Related

Functional relatedness

|--|

Not at all Functionally Related	Completely Functionally Related

Conceptual relatedness

|--|

Not at all Conceptually Related	Completely Conceptually Related

Overall relatedness

|--|

Not at all Related	Completely Related

Is this an example of mind wandering? ___yes ___no

Mind Wandering Definition Questionnaire

Factor Scenario Survey-Abbreviated

Opinions About Mind Wandering

Overall instructions:

There are several sections to this survey. Please answer all questions. If you do not know the answer, please state that you do not know the answer.

Section I. What is Mind Wandering to You?

In the blank provided below, please define in your own words what "mind wandering" is.

Back Next

5%

Section I (cont.)

Instructions:

Four scenarios will be provided that describe a shift in a person's thinking while driving a car. For each scenario, rate how temporally, conceptually and functionally related (defined below) you think the two described thoughts are by selecting a number on the rating scale that represents your opinion. Then, for each scenario, please provide an overall rating of relatedness and state, "yes" or "no," whether you believe the person is mind wandering. Press the "next" button at the bottom of the screen to read the first scenario.

Definitions

Temporal Relatedness: Two thoughts are related because the content of these thoughts (e.g., events, memories, plans, activities, etc.) occur relatively close in time. For example, you may be concentrating on your drive home from work and suddenly remember that at the next traffic light, you need to turn into the shopping center to pick up an item at the grocery store. Stopping by the grocery store at the next intersection has some degree of temporal relatedness because it occurs in the next few moments.

Conceptual Relatedness: Two thoughts are related because the content of these thoughts (e.g., events, memories, plans, activities, etc.) share the same concept or main idea. For example, if someone is riding a horse and they think about a new saddle they tried out last week, the thoughts about riding and thoughts about a new saddle are from the same concept; horseback riding.

Functional Relatedness: Two thoughts are related functionally in two ways. First, when the thought contents (e.g., event, memory, activity, etc.) require the same action, activity or operation to accomplish a goal, these may be functionally related. For example, the computer keyboard shortcuts, Ctrl+C (copy) and Ctrl+V (paste) may be two functionally related tasks. Second, when these two thoughts are about two events, memories, activities, etc. that are dependent upon one another, these may also be considered functionally related. For example, you are concentrating on cooking dinner on the grill when you remember that your second bag of charcoal is in the car. The grill requires charcoal for cooking so there is some degree of functional relatedness.

Back Next

Scenario 1

You are driving home from work when you are detoured onto another route that you are semi-familiar with. At this very moment your thinking is focused on down shifting your manual transmission into third gear in order to make a detour-directed turn . Then, your thinking is interrupted by thoughts about when you put your car in reverse ten minutes ago and accidentally grinded your gears. The two thoughts are: 1) downshifting to make a turn and 2) grinding the gears 10 minutes ago.

How temporally, functionally and conceptually related are these two thoughts? Overall, how related are these two thoughts?
*

	Not At All	1	2	3	4	5	6	7	8	9	Extremely Related
Temporal Relatedness	○	○	○	○	○	○	○	○	○	○	○
Functional Relatedness	○	○	○	○	○	○	○	○	○	○	○
Conceptual Relatedness	○	○	○	○	○	○	○	○	○	○	○
Overall Relatedness	○	○	○	○	○	○	○	○	○	○	○

Is Scenario 1 an example of mind wandering? *

○ Yes ○ No

Back Next

Scenario 5

You are driving home from work when you are detoured to another route home that you are semi-familiar with. At this very moment your thinking is focused on shifting your manual transmission into third gear so you can make a left turn at the detour sign ahead. Then, you begin thinking about a drive you took last week where the "low oil" light came on unexpectedly. The two thoughts are: 1) downshifting to make a turn and 2) the drive you took last week when "low oil" light came on.

How temporally, functionally and conceptually related are these two thoughts? Overall, how related are these two thoughts? *

	Not At All	1	2	3	4	5	6	7	8	9	Extremely Related
Temporal Relatedness	○	○	○	○	○	○	○	○	○	○	○
Functional Relatedness	○	○	○	○	○	○	○	○	○	○	○
Conceptual Relatedness	○	○	○	○	○	○	○	○	○	○	○
Overall Relatedness	○	○	○	○	○	○	○	○	○	○	○

Is Scenario 5 an example of mind wandering? *

○ Yes ○ No

Back Next

35%

Scenario 12

You are driving home from work when you are detoured to another route home that you are semi-familiar with. At this very moment your thinking is focused on shifting your manual transmission into third gear so you can make a left turn at the detour sign ahead. Then, your thoughts shift to a conversation you will have on the cellular phone in ten minutes about how to operate the metal press gears at work. The two thoughts are: 1) downshifting to make a turn and 2) phone conversation in ten minutes about how to operate the metal press gears at work.

How temporally, functionally and conceptually related are these two thoughts? Overall, how related are these two thoughts? *

	Not At All	1	2	3	4	5	6	7	8	9	Extremely Related
Temporal Relatedness	○	○	○	○	○	○	○	○	○	○	○
Functional Relatedness	○	○	○	○	○	○	○	○	○	○	○
Conceptual Relatedness	○	○	○	○	○	○	○	○	○	○	○
Overall Relatedness	○	○	○	○	○	○	○	○	○	○	○

Is Scenario 12 an example of mind wandering? *

○ Yes ○ No

Scenario 16

You are driving home from work when you are detoured to another route home that you are semi-familiar with. At this very moment your thinking is focused on shifting your manual transmission into third gear so you can make a left turn at the detour sign ahead. Then, you remember that next week after the children start school, you will need to start planning for the family vacation because their school schedule will be available. The two thoughts are: 1) downshifting to make a turn and 2) planning for a family vacation next week.

How temporally, functionally and conceptually related are these two thoughts? Overall, how related are these two thoughts? *

	Not At All	1	2	3	4	5	6	7	8	9	Extremely Related
Temporal Relatedness	○	○	○	○	○	○	○	○	○	○	○
Functional Relatedness	○	○	○	○	○	○	○	○	○	○	○
Conceptual Relatedness	○	○	○	○	○	○	○	○	○	○	○
Overall Relatedness	○	○	○	○	○	○	○	○	○	○	○

Is Scenario 16 an example of mind wandering? *

○ Yes ○ No

Back Next

What is your participant ID#? *

Several fringe cases of mind wandering exist. We'd like to know whether you consider the following to be cases of mind wandering if they occur while you are driving. These cases are defined as follows:

Zone-outs: Thinking of nothing; sometimes your vision becomes blurry as your mind loses touch with its original stream of consciousness.

Physiologically driven thoughts: Thoughts that occur because of a feeling or sensation within your body (e.g., hunger pangs, pain, fatigue, etc.).

Humming: Singing a song in your head as you accomplish tasks.

A feeling of familiarity: Seeing something that reminds you of a period of time in your life or places you've been before (physically or mentally). You are not reminded of a specific event or memory but a generic memory. For example, you may see a sea shell that reminds you of the beach; not a specific beach but the beach.

Planning: Some driving-related event just occurred and now you would like to plan for similar future events to either prevent negative outcomes or enhance positive outcomes.

Rehearsing: Some driving-related event just occurred moments ago and you are rehashing the event in your mind while driving.

Are the following fringe cases examples of mind wandering? *

	Yes	No
Zone-outs	○	○
Physiologically driven thoughts	○	○
Humming a song in your head	○	○
A feeling of familiarity	○	○
Planning	○	○
Rehearsing	○	○

Back Submit

95%

Thank you for your help! If you ever have any questions regarding your participation in this study, please contact me personally; jacowley@earthlink.net.

Sincerely,

Jennifer Cowley

North Carolina State University

100%

Post-Driving Questionnaire

What is your participant ID number? Please write the number in the blank below. *

Mind wanderings can sometimes occur when the person sees something in the environment that reminds him/her of something else (i.e., cued by the environment). Other times, thoughts just pop into our minds at random times (i.e., mind pops). What percentage of your mind wanderings were cued by the environment and what percentage were mind pops? *

	% of wanderings you experienced in this study
Cued by environment	
Mind pops	

Measuring mind wandering is difficult and often times, the method we used to measure it accidentally causes people to mind wander. For example, one woman mentioned that because the study was about mind wandering, she sometimes looked for things to wander to rather than experiencing mind wandering naturally. What percentage of your wanderings, if any, do you believe were caused by the measurement of mind wandering? *

What percentage of your driving time did you spend babysitting your thoughts? *

Sometimes we begin mind wandering without any awareness that we are doing so. However, sometimes we eventually become conscious that we've been wandering to task-unrelated thoughts; i.e., mind wanderings were originally unconscious but then they surfaced into consciousness. What percentage of your mind wanderings, if any, involved these kinds of "unconscious-to-conscious" thoughts? *

Please estimate what percentage of your thoughts were the following types? (These percentages do not have to sum to 100% because you could experience more than one at the same time) *

	% of your thoughts
Zone-outs (e.g., not thinking of anything)	
Physiologically driven thoughts (e.g., hunger pangs, pain, etc.)	
Humming a song in your head	
Familiarity (e.g., sunset reminds you of good times at the beach)	
Planning for future driving events	
Rehashing driving events that just occurred	

In your daily life, are there certain tasks during which you are more likely to experience mind wandering? If so, please elaborate. *

In general, why do you think people engage in mind wandering or task-unrelated thinking? In your opinion, is there some purpose it serves? *

Next

0%

One problem with measuring mind wandering is that when the person is probed, their stream of conscious thought they experienced before the probe, was potentially interrupted. When the driving task subsequently resumes, it may take an individual some time to re-engage (if they can) with their original stream of conscious thought. Overall, which measurement method allowed you to re-engage the quickest with your original stream of conscious thought you experienced before the probe? *

- All probe
- Probe/Self caught/immediate recall
- Clicker method
- None; they were all the same

Which mind wandering measurement method was the best at accurately capturing how frequently you mind wandered? *

- All probe
- Probe/ self caught/ immediate recall
- Clicker method
- None; they were all the same

1. Which method did you babysit your thoughts the most with? *

- [ ] All probe
- [ ] Probe / self-caught / immediate recall
- [ ] Clicker method

Are there any other thoughts or opinions about these mind wandering methods that you'd like to share? Did you prefer one method over the another? *

Back Next

33%

What is your native language (e.g., English, Spanish, etc.)? *

Would you consider yourself to be generally more dis tractable than your peers (e.g., distracted by sights and sounds or distracted by your own thoughts)? *

- Yes
- No
- Not sure

Does your level of distractability stay constant throughout the day or not?

Are there certain times of the day or certain tasks you perform in which you are more or less distractable? *

Do you have a valid state (or country) driver's license to drive a vehicle? *

- Yes
- No

Approximately when did you get your driver's license? ***(Month/Year) If you do not remember, just write "don't know" in the blank provided.***
*

Before you came to college, how often did you drive an automobile? *

- I did not drive before I came to college
- 0- 1 round-trip drives per week
- 2- 4 round-trip drives per week
- 5- 6 round-trip drives per week
- 6 or more round-trip drives per week

While in college, on average how often do you drive an automobile? *

- I do not drive while I'm attending college
- 0- 1 round-trip drives per week
- 2- 4 round-trip drives per week
- 5- 6 round-trip drives per week
- 6 or more round-trip drives per week

What percentage of your driving, while in college, is in high-density (e.g., urban) traffic? *

What percentage of your driving, while in college, is in low-density (e.g., rural) traffic? *

What is your gender? *

- Male
- Female
- Transgendered

What is your current age? *

Do you currently own or lease an automobile? *

- Yes
- No

If you currently do not own a lease an automobile, do you have access to one? *

- Yes
- No
- I own an automobile

Do you predominantly drive a manual (or stick shift) or automatic transmission? *

- Manual (or stick shift)
- Automatic

Back Submit

67%

Thank you for your help! If you ever have any questions regarding your participation in this study, please contact me personally; jacowley@earthlink.net.

Sincerely,

Jennifer Cowley

North Carolina State University

100%

AOSPAN Example Questions

In this experiment you will try to memorize letters you see on the screen
while you also solve simple math problems.

In the next few minutes, you will have some practice to get you familiar with how the experiment works.

We will begin by practicing the letter part of the experiment.

Click the left mouse button to begin.

For this practice set, letters will appear on the screen one at a time.
Try to remember each letter in the order presented.

After 2-3 letters have been shown, you will see a screen listing 12 possible letters.
Your job is to select each letter in the order presented.
To do this, use the mouse to select each letter.

The letters you select will appear at the bottom of the screen.

Click the mouse button to continue.

When you have selected all the letters, and they are in the correct order,
hit the **EXIT** box at the bottom right of the screen.

If you make a mistake, hit the **CLEAR** box to start over.

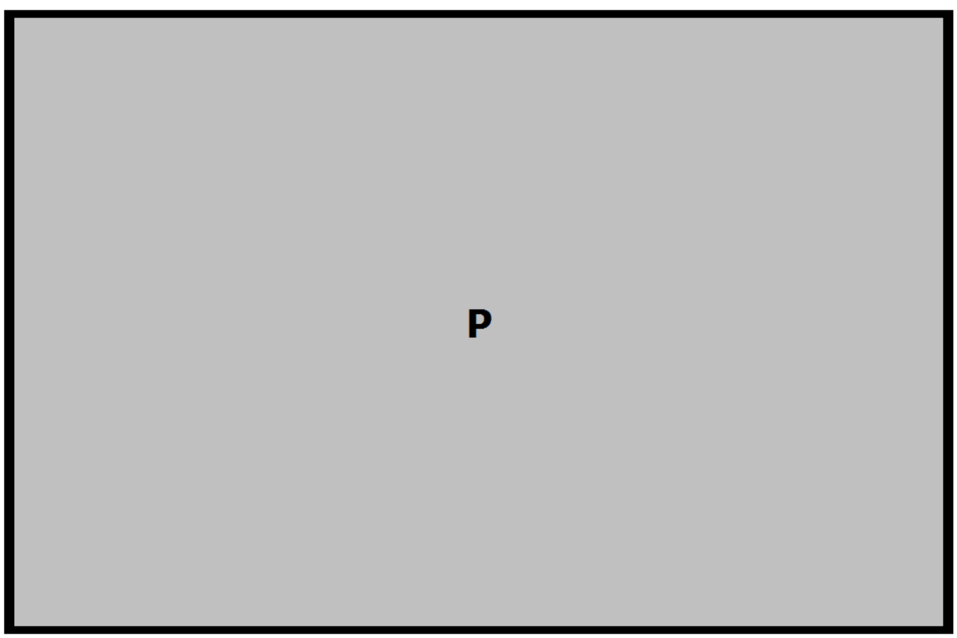
If you forget one of the letters, click the **BLANK** box to mark the spot
for the missing letter.

Remember, it is very important to get the letters in the same order as you see them.
If you forget one, use the **BLANK** box to mark the position.

Do you have any questions so far?

When you're ready, click the mouse button to start the letter practice.

Then letters are flashed on the screen like this:



Select the letters in the order presented. Use the blank button to fill in forgotten items.

F	H	J
K	L	N
P	Q	R
S	T	Y

BLANK

CLEAR

EXIT

When finished, click the "exit" button to move to the next block.

BDI Online Questionnaire (not a complete copy)

BDI

Instructions

Instructions:

This is a questionnaire assessing mood. Each item is a category (e.g., Mood, Pessimism, Sense of Failure, etc.) and you must select an answer underneath the category that best describes your mood at this moment in time. Then, select the radio button that best matches your answer. To maintain confidentiality of your answers, please do not share your answers with anyone in the test room.

What is your participant ID number? Please enter it in the blank below. *

Next

0%

A. Mood *

- 0 I do not feel sad
- 1 I feel blue or sad
- 2a I am blue or sad all the time and can't snap out of it
- 2b I am so sad or unhappy that it is very painful
- 3 I am so sad or unhappy that I can't stand it

B. Pessimism *

- 0 I am not particularly pessimistic or discouraged about the future.
- 1a I feel discouraged about the future
- 2a I feel I have nothing to look forward to
- 2b I feel that I won't ever get over my troubles
- 3 I feel that the future is hopeless and that things cannot improve

APPENDIX I. *Computer Hardware Specifications Required for Driving Simulator Software*

- Module Description
- Precision Mobile M4500Precision Mobile M4500
- Operating SystemGenuine WindowsR 7 Professional, 32-bit, no media
- ProcessorIntelR CoreT i5-520M Dual Core 2.40GHz 3MB
- Memory4.0GB, DDR3-1333MHz SDRAM, 2 DIMMS
- Internal KeyboardInternal English Keyboard
- GraphicsNVIDIAR Quadro FX 880M Graphics with 1GB dedicated memory
- Primary Storage250GB 7200rpm Hard Drive
- Fingerprint Reader OptionsNo Fingerprint Reader and No Contactless
- Smartcard Reader
- LCDs15.6" HD (1366 x 768) Anti-Glare LED Display
- ModemNo Modem
- AC Adapter130W 3P, A/C Adapter
- Primary Optical Device8X DVD with Cyberlink Power DVDT, no media
- Camera / MicrophoneNo webcam with microphone
- Wireless LAN (802.11)Dell WirelessT 1501 802.11b/g/n Half Mini Card
- Systems ManagementNo Intel vProT Technology's advanced management
- features
- Productivity SoftwareMicrosoftR Office Starter 2010
- System DocumentationNo Resource DVD or Quick Reference Guide
- Precision ONNo Precison ON
- 6-cell (60Wh) Lithium Ion Battery6-cell (60Wh) Lithium Ion Battery
- Hardware Support Services3 Year Basic Limited Warranty and 3 Year Next
- Business Day Onsite Service
- System RecoveryDell Back-up and Recovery Manager for Windows 7
- Processor BrandingIntel Core i5 Label

APPENDIX J. *List of Diagnostics Available to Model a Logistic Regression in 9.4 Base SAS*

- Model and Outlier Diagnostic (Pearson Chi-Square Residual and Deviance Residual),
- Leverage Diagnostics,
- Influence on Parameter Estimates (CI Displacements C, CI Displacements CBar),
- Influence on Model Fit (Pearson Chi-Square Residual and Deviance Residual),
- Influence on Estimates of Intercept (e.g., DfBetas),
- Influence on the estimate of independent variables (e.g., DfBetas),
- Pearson chi-square deletion differences as predicted probabilities for the dependent variable
- Deviance deletion differences by predicted probabilities for dependent variable
- Confidence interval displacement C by predicted probabilities for dependent variable
- Leverage by predicted probabilities for dependent variable
- Pearson chi-square deletion differences by leverage
- Deviance deletion differences by leverage
- Confidence interval displacements C by leverage
- Predicted probabilities for dependent variable by leverage
- Influence on model fit and parameter estimates for both Pearson chi-square deletion difference and deviance deletion difference

APPENDIX K. *Table of Demographics for Part IV Flight Simulation Study Sample (n=17)*

Questionnaire Section	Variable			
Basic Demographics	Age	*M*=44.88 years, *SD*=15.79 years		
	Gender	Male	17 (100%)	
	Native Language	English	17 (100%)	
Flying Demographics	Currently working on a certificate?	No	7 (41%)	
		Yes	10 (59%)	IFR =5 CFI=1 Private=2 Unknown=2
	Certificates held	ATP	0 (0%)	
		Private	11 (65%)	
		CFI	2 (12%)	
		CFII	2 (12%)	
		Commercial	4 (24%)	
		Flight Engineer	0 (0%)	
		Flight Engineer (Written)	0 (0%)	
		Instrument	3 (18%)	
		MEI	1 (6%)	
	Hours/month under:	FAR Part 91	M=8.21 hours, SD=6.78 hours	
		FAR Part 121	0	
		FAR Part 135	0	
	# of participants	Air Taxis	0	

	who predominantly fly for the following:	Commuter/Regional	0
		Major Commercial	0
		Cargo	0
		Military	0
		UAV	0
		GA	15 (88%)
	# aircraft types flown	*M*=6.47 types	
		SD=5.23 types	
	Total # flight hours	*M*=813.24 hours	
		SD=396.85 hours	
	Ave # days/month of flight	*M*=4.65 days	
		SD=4.70 days	
Distractibility	In real life, I ____ wander to boost alertness.	Never	8 (47%)
		Occasionally	7 (41%)
		Frequently	1 (6%)
		Very Often	1 (6%)
	More distractible than peers?	Yes	9 (53%)
		No	4 (24%)
		Not sure	2 (12%)
	Level of distractibility stay constant throughout day?	Yes	2 (12%)
		No	14 (82%)
		Not sure	1 (6%)

APPENDIX L. *Copies of Consent Form and All Questionnaires Used in Part IV Flight Simulation Study*

Consent Form

North Carolina State University
INFORMED CONSENT FORM
Thinking During Flying

Primary Investigator: Jennifer Cowley
Faculty Advisor: Doug Gillan

INFORMATION
You are invited to participate in a confidential research study if you are at least 18 years of age and hold a current pilot license/certificate for fixed-winged aircraft. The purpose of this study is to investigate a pilot's stream of conscious thought during flight.

RISKS
Since this survey asks for your personal experiences with off-tasking thinking, we encourage you to complete this survey in a private area and to prevent others from accessing your survey responses if you need to leave the computer before submitting your answers.

CONFIDENTIALITY
The information in the study records will be kept strictly confidential and you are asked not to provide your name or other identifying information. Data will be stored securely and will be made available only to persons conducting the study unless you specifically give permission in writing to do otherwise. No reference will be made in oral or written reports that could link your participation to this study

CONTACT
If you have questions at any time about the study or the procedures, you may contact the faculty advisor of this research, Doug Gillan, at doug_gillan@nscu.edu. If you feel you have not been treated according to the descriptions in this form, or your rights as a participant in research have been violated during the course of this project, you may contact Mr. Matthew Ronning, Assistant Vice Chancellor, Research Administration, Box 7514, NCSU Campus (919/513-2148).

PARTICIPATION
Your participation in this study is voluntary; you may decline to participate at any time.

CONSENT

- ☐ I have read and I understand the above information.
- ☐ I agree to participate in this study and confirm that I am at least 18 years of age.
- ☐ I do not wish to participate in this study.

Life Events Questionnaire

Instructions: Place a checkmark next to each life event that you've personally experienced in the last 12 months.

Life event	Life change units
Death of a spouse	
Divorce	
Marital separation	
Imprisonment	
Death of a close family member	
Personal injury or illness	
Marriage	
Dismissal from work	
Marital reconciliation	
Retirement	
Change in health of family member	
Pregnancy	
Sexual difficulties	
Gain a new family member	
Business readjustment	
Change in financial state	
Death of a close friend	
Change to different line of work	
Change in frequency of arguments	
Major mortgage	
Foreclosure of mortgage or loan	
Change in responsibilities at work	
Child leaving home	
Trouble with in-laws	
Outstanding personal achievement	

Life event	Life change units
Spouse starts or stops work	
Begin or end school	
Change in living conditions	
Revision of personal habits	
Trouble with boss	
Change in working hours or conditions	
Change in residence	
Change in schools	
Change in recreation	
Change in church activities	
Change in social activities	
Minor mortgage or loan	
Change in sleeping habits	
Change in number of family reunions	
Change in eating habits	
Vacation	
Christmas	
Minor violation of law	

Post Flight Questionnaire

1. What is your participant ID number? Please write the number in the blank below.

2. Please define “mind wandering” in your own words.

 ____

 ____

 ____

3. Mind wanderings can sometimes occur when the person sees something in the environment that reminds him/her of something else (i.e., cued by the environment). Other times, thoughts just pop into our minds at random times (i.e., mind pops). What percentage of your mind wanderings were cued by the environment during flight and what percentage were mind pops?

	% of wanderings you experienced in this study
Cued by environment	
Mind pops	

4. Measuring mind wandering is difficult and often times, the method we used to measure it accidentally causes people to mind wander. For example, one participant mentioned that because the study was about mind wandering, she sometimes looked for things to wander to rather than experiencing mind wandering naturally. What percentage of your wanderings, if any, do you believe were caused by the measurement of mind wandering?

 %

5. What percentage of your flying time did you spend babysitting (e.g., thinking about what you were thinking about) your thoughts?

 %

6. Sometimes we begin mind wandering without any awareness that we are doing so. However, sometimes we eventually become conscious that we've been wandering to task-unrelated thoughts; i.e., mind wanderings were originally unconscious but then they surfaced into consciousness. What percentage of your mind wanderings, if any, involved these kinds of "unconscious-to-conscious" thoughts?

 %

7. One problem with measuring mind wandering is that when the person is probed, their stream of conscious thought is interrupted. When the flying task resumes, it may take an individual some time to re-engage (if they can) with their original stream of conscious thought. On average, how many seconds did it take you to re-engage with flight tasks after a probe.

   ____________ seconds

8. How many times do you think you mind wandered during the entire flight?

   ____________ times

9. What percentage of your mind wanderings during flight were of stressful events and circumstances in your personal life?

   ____________ %

10. Please estimate what percentage of your thoughts during flight was comprised of the following types of thoughts? (These percentages do not have to sum to 100% because you could experience more than one at the same time)

	% of your thoughts
Zone-outs (e.g., not thinking of anything)	
Physiologically driven thoughts (e.g., hunger pangs, pain, etc.)	
Humming a song in your head	
Familiarity (e.g., sunset reminds you of good times at the beach)	
Planning for future flying events	
Rehashing flying events that just occurred	

11. Any additional comments about mind wandering during certain flight tasks?

   ____
   ____
   ____
   ____
   ____
   ____

Background Information

12. Are you currently working on a certification?

- Yes
- No

If "yes," what certificate?

13. What certificate/ratings do you currently hold?
Select all that apply.

What certificate/ratings do you currently hold?	Yes
ATP	
Private	
CFI	
CFII	
Commercial	
Flight Engineer	
Instrument	
Flight Engineer Written	
MEI	

14. How many hours in a typical month do you fly under the following FARs?

How many hours in a typical month do you fly under the following FARs?	# of Hours
Title 14 CFR Part 91	
Title 14 CFR Part 121	
Title 14 CFR Part 135	

15. Are you currently employed as a pilot for a particular air carrier?

☐ Yes

☐ No

16. Do you <u>predominantly</u> fly for:

	Yes
Air Taxis	☐
Commuter/Regional Airlines	☐
Major Commercial Airlines	☐
Cargo Operations	☐
Military	☐
UAV	☐
GA	☐

17. List all the aircraft (e.g., Boeing 737, Cessna 172) you have flown in the past:

____
____
____
____
____
____

18. Approximately how many total flight hours have you logged in your lifetime?
[ ] flight hours

19. On average, how many hours of sleep do you typically get the night before a flight?
[ ] hours of sleep

20. On average, how many days do you fly in a single month?
[ ] days

21. How often do you intentionally mind wander to boost your alertness level?

☐ Never ☐ Occasionally ☐ Frequently ☐ Very often

22. On average, what percentage of your flying time involves significant reliance on autopilot for flight?

    ______ %

23. Would you consider yourself to be generally more distractable than your peers (e.g., distracted by sights and sounds or distracted by your own thoughts)?

    - ☐ Yes
    - ☐ No
    - ☐ Not sure

24. Does your level of distractibility stay constant throughout the day?

    - ☐ Yes
    - ☐ No
    - ☐ Not sure

25. If you answered, “no” in the previous question, are there certain times of the day or certain tasks you perform in which you experience more distracted thinking?

    ____
    ____
    ____
    ____
    ____
    ____

26. What is your native language (e.g., English, Spanish, etc.)?

    ______

27. What is your gender?

    - ☐ Male
    - ☐ Female

28. What is your current age?

    ______ years old

29. Is there anything else you’d like to share with us about mind wandering in the cockpit?

    ____

Thank you for your help! If you ever have any questions regarding your participation in this study, please contact me personally; jacowley@earthlink.net.

Sincerely,
Jennifer Cowley
North Carolina State University

APPENDIX M. *Flight Simulation Settings and Flight Scenario Settings*

Microsoft Flight Simulator Settings

General Setting	**Level**
Aircraft Type	Cessna C172SP Skyhawk
Graphics	Medium low
Aircraft	Medium high
Scenery	Medium low
Weather	Medium low
Traffic	Low
Control Settings	Default
Realism Settings	Default
GA Traffic Density	5%
Airport Vehicle Density	Minimum

Flight Scenario Settings

Scenario Setting	Initial	Final

Scenario Setting	Initial	Final
Turbulence	None	Light
Time of Year	Fall, Daytime, 12:21pm	Fall, Daytime, 12:21pm
Icing Conditions	None	None
Precipitation	None	None
Cloud Type	Scattered Cumulus	Scattered Cumulus
Cloud Layer	1500-3000 feet	
Wind Speed and Direction	16 knots at 29 degrees	8 knots at 29 degrees
Visibility	20 miles	20 miles
Failures set	Aileron (_) and heading indicator (15-25 minutes)	None

APPENDIX N. *Counterbalancing of Part III Questionnaires*

Part #	**Ordering#**	**First**	**Second**	**Third**
73	1	BDI	Aospan	Condition C 1
74	2	Aospan	Condition C2	BDI
75	3	Condition C3	BDI	Aospan
76	4	BDI	Aospan	Condition C2
77	5	Aospan	Condition C3	BDI
78	6	Condition C1	BDI	Aospan
79	7	A (BDI	Aospan	Condition C3
80	8	Aospan	Condition C2	BDI
81	9	Condition C1	BDI	Aospan

Conditions

	Metric ordering		
Condition	First	Second	Third
C1	All probes	Clicker	Probe/Self caught
C2	Probe/Self caught	All probes	Clicker
C3	Clicker	Probe/Self caught	All probes