TOWARD A THEORETICAL FRAMEWORK FOR PREDICTING OVERALL QUALITY EFFECTS OF INTERRUPTIONS ON CONTENT PRODUCTION TASKS

by

Nicole Elizabeth Werner
A Dissertation
Submitted to the Graduate Faculty of George Mason University in Partial Fulfillment of The Requirements for the Degree of Doctor of Philosophy Psychology

Committee:

_______________________________  Dissertation Director

_______________________________  Committee Member

_______________________________  Committee Member

_______________________________  Department Chairperson

_______________________________  Program Director

_______________________________  Dean, College of Humanities and Social Science

Date: __________________________ Fall Semester 2014
George Mason University Fairfax, VA
Toward a Theoretical Framework for Predicting Overall Quality Effects of Interruptions on Content Production Tasks

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Nicole Elizabeth Werner
Master of Arts
George Mason University, 2011

Director: Deborah A. Boehm-Davis, Professor
Department of Psychology and Dean, College of Humanities and Social Sciences

Fall Semester 2014
George Mason University
Fairfax, VA
This work is licensed under a creative commons attribution-noderivs 3.0 unported license.
DEDICATION

This dissertation is dedicated to my mother, Jacqueline Smollar, a continuous source of support, love, and inspiration for myself, and so many others.
ACKNOWLEDGEMENTS

First, I would like to thank my advisor and mentor, Deborah Boehm-Davis, without whom I would not be where I am today. Her excellent guidance, mentorship, and support have helped me grow in so many ways and have led me to incredible opportunities during my graduate studies, as well as moving forward in my professional career. I would also like to thank my committee members, Carryl Baldwin and Robert Youmans for their expert advice in shaping my dissertation. A huge thank you to all of my mentors: David Cades who helped me discover my love for research and who continues to help guide me through my career; Peter Doyle for introducing me to what would become my career, and for offering continued advice, expertise, and opportunities over the years; Richard Holden for constantly broadening my knowledge and for inspiring me to pursue a career in academia; Ayse Gurses for helping me learn more and more and for spending her limited time helping me; and Alicia Arbaje who has been instrumental in helping me shape my career path, who offers me so many incredible opportunities, and who continues to provide exceptional advice and expertise to help me learn and improve; and all of my mentors collectively for being so wonderfully influential in my life and for spending so much time to mentor me. I am eternally grateful to Cyrus Foroughi, my colleague and friend who was instrumental in the completion of this work, and who keeps me inspired about research and interrupted task performance. I would also like to thank Haneen Saqer, my colleague, friend, and voice of reason, who has been a gift to my life since day one of graduate school, and continues to provide an ear to listen and a shoulder to lean on. Thank you to my husband Edmund Jalinske for his love, patience, and support through this entire process, and for helping me through my most stressful moments. Thank you to all of my friends near and far for sticking with me even though I have disappeared for the past two years. And finally, I want to thank my family – especially my grandmother, my mother, my aunt Mary Ann Werner, my brother Michael Volpe and my sister Dawn Whitmore for their unconditional love and for believing in me even when I did not. It truly takes a village.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>viii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>vii</td>
</tr>
<tr>
<td>Abstract</td>
<td>ix</td>
</tr>
<tr>
<td>An introduction to Interrupted task performance</td>
<td>1</td>
</tr>
<tr>
<td>Theories of Interrupted Task Performance</td>
<td>4</td>
</tr>
<tr>
<td>Long-term Working Memory Theory</td>
<td>4</td>
</tr>
<tr>
<td>Memory for Goals Model</td>
<td>6</td>
</tr>
<tr>
<td>The Threaded Cognition Model</td>
<td>9</td>
</tr>
<tr>
<td>Theoretical Rationale</td>
<td>12</td>
</tr>
<tr>
<td>Experimental Overview</td>
<td>16</td>
</tr>
<tr>
<td>do interruptions affect the quality of work?</td>
<td>19</td>
</tr>
<tr>
<td>Theoretical Accounts of Interrupted Task Performance</td>
<td>19</td>
</tr>
<tr>
<td>Experimental Rationale</td>
<td>23</td>
</tr>
<tr>
<td>Experiment 1</td>
<td>24</td>
</tr>
<tr>
<td>Method</td>
<td>24</td>
</tr>
<tr>
<td>Participants</td>
<td>24</td>
</tr>
<tr>
<td>Task and materials</td>
<td>24</td>
</tr>
<tr>
<td>Design and Procedure</td>
<td>25</td>
</tr>
<tr>
<td>Measures</td>
<td>27</td>
</tr>
<tr>
<td>Results and Discussion</td>
<td>28</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>31</td>
</tr>
<tr>
<td>Method</td>
<td>32</td>
</tr>
<tr>
<td>Participants</td>
<td>32</td>
</tr>
<tr>
<td>Task and materials</td>
<td>32</td>
</tr>
<tr>
<td>Design and Procedure</td>
<td>32</td>
</tr>
<tr>
<td>Measures</td>
<td>32</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1. Sample Essay Prompt and Assignment</td>
<td>25</td>
</tr>
<tr>
<td>Table 2. Mean Scores</td>
<td>29</td>
</tr>
<tr>
<td>Table 3. Mean Scores</td>
<td>34</td>
</tr>
<tr>
<td>Table 4. Sample Essay Prompt and Assignment</td>
<td>46</td>
</tr>
<tr>
<td>Table 5. Mean Scores</td>
<td>55</td>
</tr>
<tr>
<td>Table 6. Concurrent verbal protocol instructions</td>
<td>61</td>
</tr>
<tr>
<td>Table 7. Description of coding scheme</td>
<td>64</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1. The Anatomy of an Interruption</td>
<td>3</td>
</tr>
<tr>
<td>Figure 2. The Anatomy of an Interruption</td>
<td>6</td>
</tr>
<tr>
<td>Figure 3. Activation model of interrupted task performance</td>
<td>8</td>
</tr>
<tr>
<td>Figure 4. The threaded cognition model</td>
<td>11</td>
</tr>
<tr>
<td>Figure 5. Time course of each phase and condition</td>
<td>27</td>
</tr>
<tr>
<td>Figure 6. Sample of the hierarchical outline structure</td>
<td>47</td>
</tr>
<tr>
<td>Figure 7. Time course of each phase and condition</td>
<td>48</td>
</tr>
<tr>
<td>Figure 8. Mean outline number of words</td>
<td>51</td>
</tr>
<tr>
<td>Figure 9. Mean outline number of primary points created</td>
<td>52</td>
</tr>
<tr>
<td>Figure 10. Mean outline number of secondary points created</td>
<td>53</td>
</tr>
<tr>
<td>Figure 11. Mean essay score</td>
<td>54</td>
</tr>
<tr>
<td>Figure 12. The Anatomy of an Interruption</td>
<td>62</td>
</tr>
<tr>
<td>Figure 13. Mean outline number of words</td>
<td>73</td>
</tr>
<tr>
<td>Figure 14. Mean outline number of primary points created</td>
<td>74</td>
</tr>
<tr>
<td>Figure 15. Mean Outline number of secondary points created</td>
<td>75</td>
</tr>
</tbody>
</table>
ABSTRACT

TOWARD A THEORETICAL FRAMEWORK FOR PREDICTING OVERALL QUALITY EFFECTS OF INTERRUPTIONS ON CONTENT PRODUCTION TASKS

Nicole Elizabeth Werner, Ph.D.
George Mason University, 2014
Dissertation Director: Dr. Deborah A. Boehm-Davis

This dissertation explores the effect of interruptions on the overall quality of work and proposes a new framework for predicting interrupted task performance. Five experiments are described that set about to determine if there is an effect of interruptions on content generation, and to explore a potential theoretical explanation for the underlying cognitive mechanisms that are driving this behavior.

The following chapters present two papers that sought to 1) determine whether interruptions affect the overall quality of work, and 2) explore the mechanisms that underlie interrupted task performance related to overall quality. We used a new interruption experimental paradigm in order to create a task where overall quality could be defined beyond the number of errors made, or the length of time to complete the task. Essay writing and essay outlining provided the necessary paradigm as a complex, creative thought task.
The first set of studies focused on whether interruptions affected the overall quality of work using an essay writing and outlining paradigm. The results suggested that interruptions during a writing task negatively impacted the overall quality of the writing. Moreover, when interruptions were experienced only during the planning of the essay, the overall quality of the essay suffered. Contrary to findings from interrupted task performance during procedural tasks, the second study determined this effect persisted even if participants were given as much time as needed to complete each outline and essay. It was not clear whether current theoretical models of interrupted task performance could account for the resulting performance.

The second paper sought to expand on this work by investigating underlying cognitive mechanisms that may be driving the detrimental effect on overall quality found in the first paper. In a series of three studies, we first quantified and compared outlines to determine how interruptions affect the outlining process and to explore the relationship between an outline created when interrupted and the resulting essay. In the first experiment, we found that interruptions led to less developed outlines and that this carried over to the essay. Next, we wanted to further explore the cognitive mechanisms underlying this performance through a Verbal Protocol Analysis of the outlining task during interruption. We were specifically interested in 1) exploring how people disengaged from the outlining task to start the interrupting task, and 2) determining the process used to resume and reengage the primary task when the interrupting task ended. We found support for a combination of underlying processes including a memory for the task location and the activation associated with the goal memory. We also found that
unlike resumption to a procedural task, participants had difficulty reconstructing pre-
interruption context at the point of resumption. This was likely due to the task context
being largely internal, rather than driven by the task itself. These results suggested a
spreading activation theoretical account for interrupted task performance based on the
activation and associative memory models. Finally, we provided further support for
spreading activation theory as a model for interrupted task performance in a content
creation task by making predictions for potential mitigation based on this theory. We
compared different interruption lengths with short or long interruptions. We found that
frequent interruptions lead to more disruption in a content production task compared to
one long interruption or no interruptions. It is possible that a longer interruption led to
complete decay of activation associated with the primary task such that participants were
forced to start from the beginning, recreating the initial context through spreading
activation from the original node. This is contrary to the frequent interruption conditions
in which participants tend to resume to the point at which they were interrupted, leading
to a change in internal context when activation is spread from the resumption point as the
new source node.

The following chapters describe the rationale, methodology, and findings from
these experiments.
AN INTRODUCTION TO INTERRUPTED TASK PERFORMANCE

Interruptions are an inescapable feature of daily life, invading work and social activities at all levels. Research has suggested that interruptions cost the American Economy 556 billion dollars in lost productivity (Basex, 2006), with workers interrupted up to 11 times per hour (Czerwinski, Horvitz, & Wilhite, 2004; Pitney Bowes, 1998; Werner et al, 2010), and shifting attention to different tasks as often as every three minutes (Gonzalez & Mark, 2004). Interruptions have become a permanent fixture of the daily landscape. Thus, it has become imperative to understand how performance is affected and what can be done to mitigate potential detrimental effects.

There has been a great deal of speculation in recent years about how interruptions affect the overall quality of work - that is, what are the detrimental effects of interruptions? When Barack Obama became president and indicated that he planned to continue using his Blackberry, there were concerns about an interruption in the Oval Office leading to a catastrophic event (Begley, 2009). More recently, radio talking heads and Internet bloggers alike have postulated that interruptions might have a detrimental impact on creativity (Dunbar, 1997; Smith, 2013). However, it is not yet clear whether these claims have merit.

This chapter presents the existing theoretical accounts of interrupted task performance. The interruption research paradigm incorporates an anatomy of an
interruption unfolding in a particular order of discrete events (Figures 1 and 2). The sequence begins with a primary task and is followed by a secondary (i.e., interrupting) task, which is typically preceded by an alert (Altmann & Trafton, 2002; Boehm-Davis and Remington, 2009). An alert can be anything from the ringing of a phone to someone knocking on the door prior to asking a question. The time between the onset of an alert and the first step on the interrupting task is termed the \textit{interruption lag}; this is the period of time in which a person disengages from the primary task. Once disengaged from the primary task, attention is shifted to the interrupting task and work begins on that task. At some point while working on the interrupting task, the goal associated with the interrupting task will either be completed, or a decision will be made to switch back to the primary task. Once this occurs, a person must disengage from the interrupting task and reorient attention to the primary task. To successfully resume the primary task, a person must recall what the next step is on that task - that is the step that was interrupted. The time period from the moment of disengagement from the interrupting task to the first step back on the primary task is termed the \textit{resumption lag}. Resumption lag is frequently used to measure the disruptiveness of an interruption. To determine whether and to what degree the interruption is disruptive, the resumption lag is compared to the inter-action interval – the time between two uninterrupted actions on the primary task. If the resumption lag is reliably longer than the inter-action interval, the interruption is considered disruptive to completion of the primary task.
To date, interruptions have been shown to be detrimental to task performance. They have been found to result in a longer total time spent on the primary task (Enrollee & Cellier, 2000; Gillie & Broadbent, 1989; Hodgetts & Jones, 2006a; Monk, Boehm-Davis, & Trafton, 2004; Trafton, Altmann, Brock, & Mintz, 2003; Zijlstra, Roe, Leonora, & Krediet, 1999;), time costs to recover from the interruption (Altmann & Trafton, 2002; Hodgetts & Jones, 2006; Iqbal & Bailey, 2006; Latorella, 1998; Miyata & Norman, 1986; Monk, Boehm-Davis, & Trafton, 2002; Monk et al., 2004), and an increased likelihood of error upon resuming the primary task (Bailey & Konston, 2006; Eyrolle & Cellier, 2000; Gillie & Broadbent, 1989; Zijlstra, Roe, Leonora, & Krediet, 1999). The findings of these studies provide strong evidence that interrupted task performance affects the time to complete that task and increases the likelihood of errors. The findings also suggest that

Figure 1. The Anatomy of an Interruption (from Boehm-Davis & Remington, 2009)
the solution to these detrimental effects is to provide additional time to correct errors and complete the task or to provide cues or context to signal where in the task to resume. However, this research has focused on procedural tasks and has not addressed the issue of whether interruptions have a negative effect on overall performance that cannot be mitigated simply by providing additional time.

The purpose of the research presented in this chapter was to investigate whether interruptions have an effect on the overall quality of task performance and if so, to explore potential underlying cognitive mechanisms that may be driving this behavior.

**Theories of Interrupted Task Performance**

Previous research on interrupted task performance has been based on three theoretical frameworks: 1) long-term working memory theory, 2) the memory for goals model, and 3) the threaded cognition model. The use of the long-term working memory theory to account for interrupted task performance has been limited (Ericsson & Kintsch, 1995; Oulasvirta & Saariluoma, 2004; 2006). The most commonly leveraged theory is the memory for goals model, which has been adapted in recent years by Savilvucci and Taatgen (2008; 2010) to create an account of interrupted task performance known as the threaded cognition model. The following sections provide a discussion of the central tenets of each theory, how each accounts for interrupted task performance at the time and error level, and the extent to which each theory can predict the impact of interruptions on the overall quality of task performance.

**Long-term Working Memory Theory**

Long-term working memory theory, as it has been applied to interrupted task
performance, suggests that the quality of task performance after an interruption is contingent on whether information about the primary task has been encoded in long-term working memory (Ericsson & Kintsch, 1995; Oulasvirta & Saariluoma, 2004). In most models of memory, information being used to complete a task is kept in short-term memory or working memory unless it is encoded into long-term memory. Ericsson and Kintsch (1995) postulate another generic memory and retrieval structure, which they term long-term working memory. According to Ericsson and Kintsch, after information is in short-term memory for some period of time, the information is then encoded into a long-term working memory store (Ericsson & Kintsch, 1995). The length of time needed to move information into this store depends on the task being performed and the level of experience the user has with the task. The theory suggests that experts are able to encode task relevant information into long-term working memory stores more quickly than non-experts.

Oulasvirta and Saariluoma (2004) applied long-term working memory theory to interrupted task performance predicting that experts would be less likely to suffered negative consequences from interruptions. According to their interpretation of long-term working memory theory, a person must have encoded information about the primary task into long-term working memory prior to the interruption to reorient to the task quickly and without error. Thus, the central tenet of this theory is that once a task is encoded into long-term working memory, it is protected from the disruptive effects of interruptions. Moreover, because experts are faster at encoding, they will be able to more easily retrieve the information from the primary task following the interruption.
Oulasvirta and Saariluoma (2006) did not make overt predictions regarding global interrupted performance effects. However, drawing from its tenets and assuming that overall quality depends on being able to draw information from long-term memory, experts should be less affected than novices in completing a primary task since they would be able to encode information into long-term working memory more quickly than novices. With sufficient expertise, global performance effects would not be found.

Figure 2. The Anatomy of an Interruption (Altmann & Trafton, 2002)

**Memory for Goals Model**
The memory for goals model (Altmann & Trafton, 2002) is an activation-based model instantiated in the ACT-R (Anderson, et.al, 2004) architecture; it has been used to successfully predict interrupted task performance (Altmann, Trafton, Brock, and Mintz, 2003; Cades, et al, 2011; Hodgetts & Jones, 2006a; 2006b). The model describes tasks as being represented as goals in memory, with each goal having a certain level of activation associated with it. Working on a specific task (the *active* task) generates activation of that
task goal. In this framework, the most “active” goal drives behavior and the goal associated with the current task will always be the most highly activated goal. When an interruption occurs, the current goal (the primary task a person is working on) must be suspended and the goal associated with the interrupting task becomes the most highly activated. When this occurs, the activation associated with the primary task goal decays as a function of time; the longer you are away from the task, the more decay occurs and the lower the activation level becomes for that task (see Figure 3). Therefore, the longer you are away from a goal, the more difficult it will be to retrieve the goal. Due to the time factor, there are instances in which a task (goal) can decay beyond the ability to be retrieved. Retrieval of this task (goal) would require an outside mechanism to boost activation.

The memory for goals model provides possible mechanisms for boosting the activation level of goals. These mechanisms are: 1) strengthening, which suggests that mechanisms that increase the frequency and/or recency of goal activation (e.g., through rehearsal) can boost activation strength; and 2) priming, which suggests that environmental and mental cues or contexts can boost activation.
For example, an alert before an interruption such as the phone ringing may allow you to rehearse what you are doing before you answer the phone. The time between the alert and the beginning of the interrupting task can allow for a boost in the activation of the primary task. Alternately, a simple interrupting task may allow you time to think back to the primary task while still performing the interrupting task, boosting the activation of the primary task goal each time you think about it. Priming allows for boosts in activation through environmental context or cues. For example, if a lengthy interrupting task has led to activation of the primary task that is below the retrieval threshold, a cue in the environment such as a highlighted last sentence, or a cursor at the location where you left off can help boost that activation level back above threshold. The boost in activation can support successful retrieval (Cades, Ratwani, Boehm-Davis, & Trafton, 2008, Hodgetts & Jones, 2006a; Iqbal & Horwitz, 2007; Ratwani & Trafton, 2008).
The memory for goals model has typically been used to explain resumption performance directly following an interruption as a person attempts to resume the primary task. Broadly speaking, the memory for goals model makes robust predictions for resumption performance in terms of time to resume, with longer interruptions generally leading to longer resumption periods (subject to the constraints discussed above). Furthermore, the model makes the prediction that longer interruptions are more likely to result in errors upon resumption of the primary task as activation of the last action may be too low due to decay to recall. However, given the correct environmental context and enough time, the level of decay can be overcome in most instances (Cades, Ratwani, Boehm-Davis, & Trafton, 2008, Hodgetts & Jones, 2006b; Ratwani & Trafton, 2008).

There is an abundance of research on interrupted task performance and the memory for goals model is robust in accounting for observed time and errors. However, it is unclear whether the memory for goals model can be used to explain the effects of interruptions on the overall quality of task performance. Based on the predictions of the memory for goals model, some might postulate that given enough time, decrements to performance caused by interruptions could be overcome.

The Threaded Cognition Model

The threaded cognition model is a more recent theory developed to explain the detrimental effects of interruptions on task performance that is also instantiated in the ACT-R cognitive architecture (Salvucci & Taatgen, 2008; 2010). Although initially the threaded cognition model was applied to concurrent multitasking, Salvucci and Taatgen
(2008) have expanded the memory for goals model. They added both the task thread (i.e., thread as opposed to goal) and the problem state (a single mental resource that represents the current state of a task) components to attempt to more fully account for interrupted task performance (Salvucci & Taatgen, 2008). The threaded cognition model adopts many of the same assumptions as the memory for goals model. However, each task or thought is represented as a thread. In contrast to the memory for goals model, the threaded cognition model does not talk about activation, but rather describes the ways in which tasks are *threaded*. As depicted in Figure 3, rather than modeling interruptions as involving suspension of a task, the threaded cognition model suggests that tasks occur on a continuum. ACT-R posits that only one goal can be active at a time. Similarly, the memory for goals model posits a one goal at a time activation structure in which an interrupting task causes the suspension of a primary task. Only the current task serves as the active goal, while the suspended goal of the initial task decays. Conversely, the threaded cognition model proposes an architecture with more than one active goal at a time. As a function of the multiple goal-state, the model does not propose the time lags described in the memory for goals model. Instead, the threaded cognition model suggests that some processes, such as rehearsal of the primary task and interrupting task actions, occur concurrently. Concurrent actions are possible because the theory assumes that threads will continue based on the availability of resources in what is referred to as a greedy/polite manner. Greedy indicates that threads will use whatever resource they need; polite indicates that once a thread does not need the resource, it is returned to the general resource pool.
Another assumption of the threaded cognition model is that there is a separate resource that is dedicated to remembering the primary task when you are working on a separate task (Anderson et al., 2004; Salvucci & Taatgen, 2010). This is known as the problem state (Salvucci & Taatgen, 2010). The problem state resource is where interference issues can occur. Because there is only one resource dedicated to the problem state of a task, if someone is performing a primary task and an interrupting task with each requiring a problem state, there will be interference and the model predicts that an error is likely to occur (Salvucci & Taatgen, 2010). However, the prediction is that these vulnerable points of interference are going to occur as you try to resume to the primary task. As with the memory for goals model, the threaded cognition model can make predictions for errors following interruption resumption on a procedural task, but cannot be extrapolated to make predictions regarding overall quality.
Theoretical Rationale
The above-described theoretical accounts of interrupted task performance make disparate predictions about what to expect when trying to recover from an interruption. However, they do share one commonality – none makes predictions for, or even discusses the potential for, global performance effects due to interruptions. Thus, it is unclear whether these theoretical accounts can be used to explain interruption performance on the overall quality of a task. Both of these theories are based on modeling procedural tasks and therefore it is not clear if they can be used to make predictions about paradigms based on non-procedural tasks such as writing.

According to the memory for goals model, when people work on a task, goals (and sub-goals) receive activation during task execution. As people work on a specific sub-goal, that sub-goal receives a certain amount of activation. In addition, the next highest goal (or sub-goal) in the hierarchy receives some activation, but not as much as the current goal. This is the ACT-R mechanism known as back-propagation (Anderson & Lebiere, 1998). Each successive step of a hierarchical process feeds activation to each prior step, with diminishing returns the further back up the chain you go. Therefore, if the time spent working on the primary task prior to an interruption is long enough, the highest level goal may have received a sufficient amount of activation through back-propagation to be active enough to have a high probability of directing behavior upon resumption. That is, a person may be more likely to return to the top level of the task based on retrieval from long-term memory than to return to the specific sub-step on which they were working prior to the interruption, as it would have decayed from working memory.
For example, people may recall which overall goal they were working on (e.g., I was working on a word document), but not exactly where in the document they were working (e.g., I was editing paragraph two). Therefore, they may be able to return to the overall task, but would need more specific cues to activate the sub-goal associated with the specific sub-task on which they had been working when interrupted.

Knowledge of the overall task could lead back to the location of the task. Once in the location, they would have to search the environment for cues. For example, suppose you are called away from your office for a task that interrupted you from your work on a report. If you recall that you were working on the report, you can return to the task location and then search the area for cues such as a document on the desk or the cursor on the computer screen. The cues can boost the activation for the goal so that resumption is possible.

In resuming a task, a person must remember what the primary task or the overall goal was. If the goal memory for that task has decayed, then some kind of search must be performed in order to recall the primary task. Resumption to a procedural task typically includes some combination of a search of memory and a search of the environment. In procedural tasks, if this search is successful, primary task resumption will occur without error. If the search is not successful, it may take longer to reorient to the primary task or an error may occur. What remains unclear is an understanding of the type of information that needs to be remembered to successfully resume a content production task. Is it the memory for the goal, or is the memory for the location of the task sufficient for successful resumption?
The memory for goals model has been proposed as one theory to account for the pattern of performance when people resume a task following an interruption. Memory for goals argues that the active task (i.e., primary goal being worked on) is represented as a chunk in memory with a certain level of activation and that the goals decay as a function of time. Therefore, the memory for goals model posits that longer interruptions would be associated with longer resumption times because the primary task goal activation will have decayed more during a longer interruption period. However, activation level decay happens rapidly, and the memory for the goal will presumably fall below the threshold needed to re-activate the goal after a lengthy interruption. When decay occurs, it may be necessary to use some other mechanism to facilitate resumption. For example, people may use the environmental context and/or primary task cues in order to boost the activation of the primary task goal above threshold, making resumption possible. Although the memory for goals model discusses ways in which activation may be increased, the theory does not describe the specific mechanisms used, or steps that are followed, in attempting to boost activation.

An extension of the memory for goals model suggests that environmental context can act as a cue to increase the probability of a successful resumption (Ratwani & Trafton, 2008). Specifically, Ratwani and Trafton (2008) argue that the memory for spatial information is more resistant to decay and distraction than goal memory, allowing for successful resumption to occur beyond the duration of goal memory. This hypothesis is based on eye-tracking data that showed as participants attempted to resume the primary task following a longer interruption; they tended to go back to roughly the last place they
looked prior to the interruption (regardless of whether or not this was the correct resumption location).

The question is whether these theories can be applied to non-procedural tasks. The underpinnings of both the threaded cognition model and the memory for goals model are focused on activation and associative memory. It is possible that these tenets can be applied to content production tasks to determine how interruptions affect the way in which activation is spread through an associative network of memory. We will look to spreading activation theory to provide the basis for these predictions (Anderson, 1983; Collins & Loftus, 1975; Quillian, 1962; Quillian, 1967). Spreading activation is a model of semantic processing in which nodes or concepts in the semantic network within memory are activated based on their levels of association with other concepts in the semantic network (Collins & Loftus, 1975). The semantic network is comprised of a complex set of associations between nodes connected by weighted associative pathways. When a concept is sampled, the corresponding node activates and that activation spreads across the associative pathways to nearby nodes. Similar to the memory for goals model, associations – that is, those nodes most closely related to each other – are driven by the level of activation of related nodes, and activation decays across associative pathways over time. Moreover, when a node is activated, activation for related nodes is propagated across the network from the source node.

For example, when the concept red is activated, it will forward propagate first to those concepts most highly associated with the word “red.” From there, the activation will continue to forward propagate based on the activation levels of associated concepts.
The semantic network is organized based on semantic association, thus nodes with closer activation levels will be grouped together if mapped, and are more likely to be sampled given the activation of one closely associated node. For example, the words “red” and “rose” could represent two highly-associated concepts that would be closely connected. When one node is activated, activation would quickly spread to the other and with less decay than the words “red” and “flower.”

We predicted that the spreading activation framework could provide an account for interrupted task performance in a content production task. We hypothesized that interruptions cut off the spread of activation through the associative semantic network. Unlike a procedural task where the only knowledge needed at resumption is the next correct action, a content production task likely requires a re-creation of the internal context from prior to the interruption. The necessity to recreate internal context means that traditional mitigation strategies such as cueing may not be as effective in supporting resumption within this type of paradigm. In other words, knowledge of the task itself may not be enough to support successful resumption.

**Experimental Overview**

The following chapters present two papers that sought to 1) determine whether interruptions affect the overall quality of work and 2) explore what mechanisms underlie interrupted task performance related to overall quality. We used a new interruption experimental paradigm in order to create a task where overall quality could be defined beyond the number of errors made, or the length of time to complete the task. Essay
writing and essay outlining provided the necessary paradigm as a complex, creative thought task.

The first paper focused on whether interruptions affected the overall quality of work using an essay writing and outlining paradigm. The results suggested that interruptions during a writing task negatively impacted the overall quality of the writing. Moreover, when interruptions were experienced only during the planning of the essay, the overall quality of the essay suffered. Counter to findings from interrupted task performance during procedural tasks, the second study determined the negative impact on quality persisted even if participants were given as much time and they wanted to complete each outline and essay. It was not clear whether current theoretical models of interrupted task performance could account for the resulting performance.

The second paper sought to expand on this work by investigating underlying cognitive mechanisms might be driving the detrimental effect on overall quality found in the first paper. In a series of three studies, we first quantified and compared outlines to determine how interruptions affect the outlining process and to explore the relationship between an outline created when interrupted and the resulting essay. In the first experiment, we found that interruptions led to less fully developed outlines and that the negative impact carried over to the essay. Next, we wanted to further explore the cognitive mechanisms underlying interrupted task performance of a content production task through a Verbal Protocol Analysis of the outlining task during interruption. We were specifically interested in how people disengaged from the outlining task to start the interrupting task, and then the process used to resume and reengage the primary task
when the interrupting task ended. We found support for a combination of underlying processes including a memory for the task location and the activation associated with the goal memory. We also found that unlike during resumption to procedural tasks, participants had difficulty reconstructing pre-interruption context at the point of resumption. The difficulty in resuming was likely due to the task context being largely internal, rather than driven by the task itself. These results suggested a spreading activation theoretical account for interrupted task performance based on the activation and associative memory models. Last, we provided further support for spreading activation theory as a model for interrupted task performance in a content creation task by making predictions for potential mitigation based on spreading activation. We compared different interruption lengths with short or long interruptions. We found that frequent interruptions lead to more disruption in a content production task compared to one long interruption or no interruptions. It is possible that a longer interruption led to complete decay of activation associated with the primary task such that participants were forced to start from the beginning, recreating the initial context through spreading activation from the original node. This is contrary to the frequent interruption conditions in which participants tend to resume to the point at which they were interrupted, leading to a change in internal context when activation is spread from the resumption point as the new source node.

The following chapters describe in detail the rationale, methodology, and findings from these experiments.
DO INTERRUPTIONS AFFECT THE QUALITY OF WORK?¹

Imagine that you need to finish writing a report for a presentation tomorrow morning. You sit down and begin typing. Your smartphone rings indicating you have received a new email. You read it and respond. You attempt to collect your thoughts and begin working on your report when your boss comes in and asks you a question. Once again, you have been interrupted and before you know it, thirty minutes have passed and you have gained no ground on completing your report. In office environments, employees are interrupted up to six times per hour (Pitney Bowes, 1998) and may shift tasks every three minutes (Gonzalez & Mark, 2004).

With such an impact on daily life, it is not surprising that there is an abundance of research focusing on interruptions. Interruptions have been well-documented in office environments (Czerwinski, Horvitz, & Wilhite, 2004; Gonzalez & Mark, 2004), in aviation (Disuses, Young, & Sumwalt, 1998; Loukopoulous, Dismukes, & Barshi, 2001), while driving (Strayer & Johnston, 2001), and in the health care industry (Chisholm, Collison, Nelson, & Cordell, 2000; McGillis, Pedersen, & Fairley, 2010; Tucker & Spear, 2006; Westbrook, Woods, Rob, Dunsmuir, & Day, 2010a; 2010b).

Theoretical Accounts of Interrupted Task Performance

A majority of traditional interruption research uses time and errors as measures of disruption as they are often the most appropriate metrics for determining the effects that
interruptions have on task performance. Research focusing on time has shown that interruptions increase completion time on tasks (Eyrolle & Cellier, 2000; Hodgetts & Jones, 2006a; Monk, Boehm-Davis, & Trafton, 2004; Trafton, Altmann, Brock, & Mintz, 2003; Ziljstra, Roe, Leonora, & Krediet, 1999; Gillie & Broadbent, 1989). In some cases, the delay or time lost as the result of an interruption can lead to negative consequences. For example, interruptions have been shown to increase the chance to fail to detect traffic signals and reduce reaction time while driving (Strayer and Johnston, 2001). Research focusing on errors has shown interruptions reduce overall accuracy in task performance (Eyrolle & Cellier, 2000; Gillie & Broadbent, 1989), and increase post-completion errors (Ratwani & Trafton, 2008). They have also been shown to disrupt overall surgical flow in the operating room (Wiegmann, El Bardissi, Dearani, Daly, & Sundt III, 2007), where a common error is repeating or skipping a step during a procedure.

However, time and errors may not be appropriate measures of disruption for all domains. For example, qualitative domains such as reading comprehension and writing may be better served by analyzing the quality of the final work product as the primary measure of disruption. If an interruption only causes a loss of time when writing an essay and time is not essential, then the interruption has not negatively affected the quality of the work. However, if an interruption has resulted in the lessening of the quality of the essay, then the interruption has had a real impact on the work product. We have been unable to identify any research that attempts to measure the negative effect of interruptions on the quality of the final product. Since research has shown that interruptions negatively affect performance on tasks as measured by time and errors, we
hypothesize that interruptions will also negatively affect the overall quality of the work product.

The Memory for Goals model (Altmann & Trafton, 2002) has been used in several experiments to successfully predict interrupted task performance (Cades et al., 2011; Hodgetts & Jones, 2006a). Memory for goals is an activation-based model instantiated in the ACT-R, a computational modeling framework (Anderson, 1993; Anderson et al., 2004; Anderson & Leibere, 1998). The model describes tasks as being represented as goals in memory, with each goal having a certain level of activation associated with it. Working on a specific task (the active task) generates activation of that task goal. In the ACT-R framework, the most “active” goal drives behavior and the goal associated with the current task will always be the most highly activated goal. When an interruption occurs, the current goal (the primary task a person is working on) must be suspended and the goal associated with the interrupting task becomes the most highly activated. When the interrupting task goal is activated, the activation associated with the primary task goal decays as a function of time; thus the longer you are away from the task, the more decay that occurs and the lower the activation level becomes for that task. Therefore, the longer you are away from a goal, the more difficult it will be to retrieve. Due to the time factor, there are instances in which a task goal can decay beyond the ability to be retrieved. Retrieval of the decayed primary task would require an outside mechanism to boost activation. The Memory for Goals model does provide such mechanisms for boosting the activation level of goals. The first mechanism is strengthening, which suggests that the frequency and/or recency of a goal can boost
activation. The second mechanism is priming, which suggests that environmental and mental cues or contexts can boost activation. Strengthening can be explained as the ability to rehearse during an interruption, which would increase the activation of that goal and thus lead to a higher likelihood that the goal would be retrieved. For example, an alert before an interruption such as a knock at the door may allow you to rehearse what you are doing before you answer the door, thus boosting the activation of the primary task. Similarly, a simple interrupting task may allow you to think back to the primary task while still performing the interrupting task, boosting the activation of the primary task goal upon each rehearsal. Moreover, the priming constraint allows for boosts in activation through environmental context or cues. For example, if a lengthy interrupting task has led to a decrease of activation of the primary task below a threshold in which it is able to be retrieved, a cue in the environment such as a highlighted last sentence, or a cursor on the computer screen can help boost that activation level back above threshold, allowing for retrieval (Cades, Ratwani, Boehm-Davis, & Trafton, 2008; Hodgetts & Jones, 2006b; Iqbal & Horwitz, 2007; Ratwani & Trafton, 2008).

This model has typically been used to explain resumption performance directly following an interruption as a person attempts to resume the primary task. Broadly speaking, the Memory for Goals Model makes robust predictions for resumption performance in terms of time to resume with longer interruptions leading to longer resumption periods while accounting for the constraints discussed above. Furthermore, the model makes the prediction that longer interruptions would be more likely to result in errors upon resumption of the primary task as activation of the last action may be too low
due to decay to recall. However, given the correct environmental context and enough time, this handicap can be overcome in most instances (Cades, Ratwani, Boehm-Davis, & Trafton, 2008; Hodgetts & Jones, 2006b; Ratwani & Trafton, 2008).

These times delays and errors are important to consider and to protect against. Much research has been devoted to understanding these predictions and how to mitigate the effects of decay. We currently have an abundance of research on interrupted task performance on the few steps following the interruption and the Memory for Goals model is robust in accounting for this performance. However, it is unclear whether this model can be used to explain interruption performance on the overall quality of a task. Based on the predictions of this model, some might postulate that given enough time, decrements caused by interruptions could be overcome. However, multiple time delays and sequential errors could potentially lead to goal confusion, interference, and error-laden performance.

**Experimental Rationale**

We were interested in whether the overall quality of work suffers when interrupted. To study this, we needed a task where quality could be defined beyond the number of errors made or time to complete the task. We selected a complex, creative thought task that mirrors a common, real-world task, outlining and writing an essay. To mirror interruptions and their effects in the real world, the interruptions needed to occur without warning. This reduces the time available for transitioning from the primary task to the interrupting task to a very small timeframe. We were also interested in whether the timing of the interruption mattered. Thus, interruptions were introduced at different
phases of the experiment, either when the participants were outlining their essay or when they were writing it.

**Experiment 1**

**Method**

**Participants**
Twenty-seven students from George Mason University participated for course credit. The data from one participant were excluded as both graders felt the participant did not take the assignment seriously. The participants (18 females and 9 males) had an average age of 23.6 years and were fluent in English.

**Task and materials**
The primary task required participants to outline an essay using pen and paper and then to write the essay on a computer using Microsoft Word. The three essay prompts and assignments (see Table 1) came from a stock bank of essay topics created by The College Board (2012B).
The interrupting task consisted of answering a series of unrelated questions using pen and paper. These questions included basic arithmetic, American history, problem solving, and word descrambling. The answer format was either fill in the blank, multiple choice, or open response. Participants completed as much of the interrupting task as they could in the given time. The interrupting task was self-paced, and it continued throughout the interruption interval. The researchers ensured participants actively completed the interrupting task and scored the accuracy of participant responses to the interrupting task.

**Design and Procedure**

The experiment consisted of a planning phase immediately followed by an execution phase. During the 12-minute planning phase, participants outlined their essay. During the 12-minute execution phase, participants typed their essay. The experiment used a within-subjects design where two factors were manipulated: presence or absence of an interruption and placement of the interruption, (planning or execution phase).
Participants served in all three conditions (planning phase interruption, execution phase interruption, and no interruption) and wrote all three essays. Participants were assigned to conditions using a Latin Squares design that counter-balanced the essay type (prompt 1, 2 and 3) and interruption location (planning, execution, and control). During an interruption condition, the participants were interrupted three separate times for 60 seconds each time. The interruptions occurred at the three, seven, and eleven minute marks. The three minutes lost to interruptions were added to the total time available for that phase so that participants would still have a total of 12 minutes to complete that phase (see Figure 5). The interruptions occurred without warning. A piece of paper with the interrupting task questions was placed over of the participants’ outlines. They were instructed to answer as many questions as possible. Once the interruption time was complete, the paper was removed and the participants continued working on the primary task. Participants were not informed that this study focused on understanding interruptions. They were informed that they were supposed to complete each task (outline or essay and the interruption) to the best of their ability.
Measures

Word counts were calculated for both the outlines and final essays. A graded score, as averaged by two independent graders, was used to evaluate the final essays. The interrupting task was scored for accuracy.

Two independent graders evaluated all essays using the College Board *Essay Scoring Guide* (The College Board, 2012a). Scores range from zero to six. A zero represents complete failure and indicates that the participant did not write at all or did not write on the appropriate topic. A six represents “clear and consistent mastery” of the essay that “effectively and insightfully develops a point of view on the issue.” (The College Board, 2012a). The measure of agreement between graders was tested using Cohen's Kappa. The results of the inter-rater reliability were Kappa = .705. The graders were blind to the experimental conditions. They only received the final essay text. No other identifying information was attached.

The two independent graders were trained on how to properly evaluate essays before scoring the essays from the participants. Both reviewed the College Board *Essay*
Scoring Guide (see the Appendix) and read pre-evaluated essays from all the score ranges (0-6). Each grader practiced scoring on pre-evaluated essays comparing their scores to those official scores (The College Board, 2012a). Lastly, the graders independently scored pre-evaluated essays and then discussed their own evaluations together in an effort to ensure consistency. The training period lasted approximately seven days. They were paid for their work and had no other affiliation with the experiment.

Results and Discussion
Mauchly’s Test of Sphericity was used to verify the assumption of homogeneity of variance for all the repeated measure analyses of variances (ANOVA) reported in this experiment. The results showed that the assumption was met for homogeneity of variance in all cases, $p > .10$.

A repeated measures ANOVA was conducted to examine whether the overall quality of work was affected by interruptions. Each final essay score was averaged from the two scores of the independent graders. The analysis revealed a significant difference among the scores across the three conditions, $F(2, 50) = 21.638, p < .001$. Post-hoc tests using a Bonferroni correction revealed that the scores in both interruption conditions were significantly lower than the scores in the non-interruption condition ($p < .01$ for both comparisons) with roughly a 1/2-point decrement on the six-point scale in the interrupted conditions. There was no difference in quality ($p > .10$) between the two interruption conditions (see Table 2).
Before conducting further analysis, we wanted to ensure that fatigue and type of essay (Essay prompts 1, 2, and 3) did not influence the lower scores. A repeated measures ANOVA revealed no significant reduction in quality over time, $F(2, 50) = 1.143$, $p > .10$ and no significant difference as a function of type of essay, $F(2, 50) = .267$, $p > .10$.

To understand the source of the quality differences, the number of words produced in each condition was examined. This analysis revealed a significant difference in the final word count of each essay across the three conditions, $F(2, 50) = 4.86$, $p = .012$. Post-hoc tests using a Bonferroni correction revealed that the word count was significantly lower when interruptions occurred during the execution phase ($M = 281.6$) when compared to the non-interruption ($M = 306.8$) and interruption during planning phases ($M = 315.3$). That is, interruptions during the execution phase led to a reduction in the amount of content produced.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Essay score ($M$)</th>
<th>Average outline word count</th>
<th>Average essay word count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planning phase interruptions</td>
<td>3.13**</td>
<td>96.12</td>
<td>306.81</td>
</tr>
<tr>
<td>Execution phase interruptions</td>
<td>3.06**</td>
<td>102.19</td>
<td>281.58*</td>
</tr>
<tr>
<td>No interruptions</td>
<td>3.71</td>
<td>103.38</td>
<td>315.31</td>
</tr>
</tbody>
</table>

*Note. *$p < .05$, **$p < .001$.}
The number of words produced in each outline was also examined. This analysis revealed no significant difference in the final word count of each outline across the three conditions ($F(2, 50) = .661, p > .10$). Post-hoc qualitative analysis of the outlines revealed that participants used many different formats to plan their essays including a traditional bullet format, a pros and cons list, and a web diagram.

A correlation coefficient was computed to assess the relationship between the word counts of the outlines, final essays, and the final score. There was a significant correlation between the word count of each essay and the final score of the essay, $r(26) = .723, p < .01$. Participants who were able to produce more content in their essays scored higher. This result is consistent with real-world results from actual essays scored for the SAT. That is, data analyzed by Dr. Perelman of the Massachusetts Institute of Technology from real SAT essays and scores associated with those essays revealed that word count significantly correlated with final score (Winerip, 2005). Moreover, the Princeton Review highlights essay length as a tip to scoring higher on that portion of the SAT (The Princeton Review, 2013). There was not a significant correlation between the word count of each outline and the final score of the essay, $r(26) = .304, p > .10$.

Accuracy for the interrupting task was also measured. The average score was 81% with scores ranging from 64% to 98%. A correlation coefficient was computed to assess the relationship between the accuracy of the interrupting task performance and the final score. There was no significant correlation between interrupting task performance and the final score of the essay, $r(26) = .07, p > .10$. 

30
These data suggest that interruptions lead to a reduction in quality for a complex, creative writing task. Quality scores for the essays were lower in both interrupted conditions. In addition, word count was reduced when the execution phase was interrupted. However, it is possible that the reduction in quality may have been due to lost time in resuming the primary task. Previous research suggests that interruptions lead to a resumption lag – the time it takes to resume the primary task following completion of the interrupting task (Altmann & Trafton, 2004). Thus, participants may have suffered delays while collecting their thoughts to resume after an interruption. Unfortunately, resumption lag times would have been difficult to measure because the resumption process occurred internally rather than with a physical process such as clicking or typing. One possible remedy would be to remove any time restrictions and allow participants as much time as they needed to complete the task.

Experiment 2
The results from Experiment 1 suggested that interruptions during a creative, complex task affect the overall quality of work. However, it is possible that participants did not have enough time to complete the assignment effectively due to resumption lags. It is also possible that participants anticipated the interruptions and this played a role in the outcome. In an effort to further investigate these findings, two changes were implemented for Experiment 2. First, the time at which the participants were interrupted was randomized instead of being presented at regular three-minute intervals. This was done to abate any potential expectancy effects. If the participants were able to anticipate incoming interruptions, they may have been able to use rehearsal or other means to help
retain information during the interruptions. The second change extended the overall time that participant's had to complete their essay in the execution phase. This change was implemented to ensure that participants had the necessary time to complete their essay.

**Method**

**Participants**
Twenty-seven students from George Mason University participated for course credit. The participants (22 females and 5 males) had an average age of 18.7 years and were fluent in English.

**Task and materials**
The primary task, interrupting task, and materials were identical to Experiment 1.

**Design and Procedure**
The overall design and procedure was identical to Experiment 1 with two exceptions. First, instead of interruptions occurring at regular, three minute intervals evenly spread across the total time (3:00-4:00, 7:00-8:00, and 11:00-12:00), interruptions occurred at random intervals. Therefore, the time in which the interruptions occurred varied within and between subjects. Second, the overall amount of time the participants had to complete each essay was extended. Instead of having 12 base minutes and three minutes of interruptions, participants had up to 20 minutes to complete their essays in the interrupted conditions. Participants also had up to 20 minutes in the no interruption condition.

**Measures**
The measures were identical to Experiment 1. The measure of agreement between graders was tested using Cohen's Kappa. The results of the inter-rater reliability were
Kappa = .768. Again, the graders were blind to the experimental condition associated with each essay. They only received the final essay text and no other identifying information was attached.

**Results and Discussion**

Mauchly’s Test of Sphericity was used to verify the assumption of homogeneity of variance for all the repeated measure analyses of variances (ANOVA) reported in this experiment. The results showed that the assumption was met for homogeneity of variance, $p > .10$.

A repeated measures ANOVA was conducted to examine whether the overall quality of work was affected by interruptions. Each final essay score was averaged from the two scores of the independent graders. The analysis revealed a significant difference among the scores across the three conditions, $F(2, 52) = 7.742, p < .001$. Post-hoc tests using a Bonferroni correction revealed that the scores in both interruption conditions were significantly lower than the scores in the non-interruption condition ($p < .05$ for both comparisons), with roughly a 1/2-point decrement on the six-point scale for the interrupted conditions. There was no difference in quality ($p > .10$) between the two interruption conditions (see Table 3).
Before conducting further analysis, we wanted to ensure that fatigue and type of essay (Essay prompts 1, 2, and 3) did not influence the lower scores. A repeated measures ANOVA revealed no significant reduction in quality over time, $F(2, 52) = 2.001, p > .10$ and no significant difference as a function of type of essay, $F(2, 52) = .908, p > .10$

To understand the source of the quality differences, the number of words produced in each condition was examined. This analysis revealed a significant difference in the final word count of each essay across the three conditions, $F(2, 52) = 9.266, p < .001$. Post-hoc tests using a Bonferroni correction revealed that the word count was significantly lower when interruptions occurred during the execution phase ($M = 269.4$) when compared to the non-interruption ($M = 317.3$) and interruption during planning phases ($M = 304.9$). That is, interruptions during the execution phase led to a reduction in the amount of content produced.

The number of words produced in each outline was also examined. This analysis revealed no significant difference in the final word count of each outline across the three

<table>
<thead>
<tr>
<th>Condition</th>
<th>Essay score ($M$)</th>
<th>Average outline word count</th>
<th>Average essay word count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planning phase interruptions</td>
<td>3.30**</td>
<td>85.56</td>
<td>304.89</td>
</tr>
<tr>
<td>Execution phase interruptions</td>
<td>3.17**</td>
<td>87.85</td>
<td>269.37*</td>
</tr>
<tr>
<td>No interruptions</td>
<td>3.74</td>
<td>89.11</td>
<td>317.33</td>
</tr>
</tbody>
</table>

*Note. *$p < .05$, **$p < .001$.  

Table 3. Mean (M) Essay Score, Outline Word Count, and Essay Word Count for Experiment 2
conditions. $F(2, 52) = 1.724, p > .10$. Post-hoc qualitative analysis of the outlines revealed that participants again used many different formats to plan their essays.

A correlation coefficient was computed to assess the relationship between the word counts of the outlines, final essays, and the final score. There was a significant correlation between the word count of each essay and the final score of the essay, $r(27) = .567, p < .01$. Again, this result is consistent with real-world results from actual essays scored for the SAT (Winerip, 2005). There was not a significant correlation between the word count of each outline and the final score of the essay, $r(27) = .134, p > .10$. Once again, participants who were able to produce more content in their essay scored higher overall.

Accuracy for the interrupting task was also measured. The average score was 78% with scores ranging from 61% to 95%. A correlation coefficient was computed to assess the relationship between the accuracy of the interrupting task performance and the final score. There was no significant correlation between interrupting task performance and the final score of the essay, $r(27) = -.02, p > .10$.

Experiment 2 allowed participants to take up to 20 minutes to complete their essays with interruption time factored out. The average completion time was not significantly different across condition, $F(2,52) = 1.232, p > .10$.

**General Discussion**

Previous work on interruptions has focused on the impact interruptions have on the ability to resume the primary task as measured by time and errors. The goal of this
study was to investigate the effect that interruptions have on the overall quality of the primary task work product.

Our data suggest that interruptions lead to a reduction in quality for a complex, creative writing task. Quality scores for the essays were lower in both interrupted conditions. In addition, word count was reduced when the execution phase was interrupted. This may have been due to lost time in resuming the primary task – that is, delays suffered by the participants while collecting their thoughts to resume after an interruption (Altmann & Trafton, 2004). Specific resumption lag times would have been difficult to measure because resumption occurred within the participants’ thought process rather than with a physical process such as clicking or typing. However, if the reduction of quality was caused solely by resumption lags, the increased time given to participants in the second experiment should have resulted in higher quality scores. Since quality did not increase in the second experiment, we believe the interruptions are causing more than just a delay in the thought process. It may be that the interruptions are causing a complete disruption in the participant's train of thought and that this disruption is causing a reduction in the overall amount of content produced, resulting in a lower-quality final product.

Attempting to apply the Memory for goals (Altmann & Trafton, 2002) model to these results is difficult, as the model does not speak directly to an impact on quality of the work product. However, goal activation is an important component of the model that may be relevant in understanding these results. The primary goal of this experiment is to complete the essay. In turn, each argument the participant develops represents a sub-goal
that needs to be completed in service of completing the larger goal of completing the essay. When participants are interrupted, both the higher-level goal (writing the essay) and the sub-goal (developing a specific argument) are interrupted.

Memory for goals posits that the activation of a goal must be greater than the activation of the interference level to drive the action upon resumption from an interruption. The level of interference will depend on the amount of “mental clutter” to which a participant has recently been exposed. In our experiment, interference arises from two primary sources, the interruptions themselves, and, less obviously, the number of previously sampled sub-goals. That is, all previously sampled sub-goals up to that point represent mental clutter that will increase the interference level, which in turn reduces the likelihood of resumption of the proper sub-goal being sampled pre-interruption. To successfully resume the sub-goal being sampled pre-interruption, the participant would need activation of that goal to be greater than the interference level. This becomes increasingly difficult in our experiment over time as the interference level is continuously growing due to repeated interruptions and the creation of new sub-goals across the experiment. Therefore, we believe that the Memory for goals model can account for the lack of proper activation of the most recently sampled goal, or sub-goal, post-interruption. However, we believe it is too far of a stretch to say that activation directly speaks to changes in quality of work.

In line with the argument above, participants may have difficulty re-activating the specific sub-goal. Instead, participants may re-activate another sub-goal or create a new sub-goal. That is, when an interruption occurs in the planning phase, the participant may
be interrupted before fully developing the idea they were mentally planning or currently writing. When the interruption comes to an end and, participants resume the task, they may not necessarily return to the last argument (sub-goal) on which they were working. Anecdotally, the proctors noticed that some participants appeared to return to the beginning of their outline or essay after an interruption occurred. In either case, it is likely that the sub-goal being sampled pre-interruption was not fully developed post-interruption. As such, the quality of that argument (sub-goal) would be lower. This return to the highest level is consistent with modeling that was done examining the impact of interruptions on completion of checklists (Diez, Boehm-Davis, & Holt, 2002). This is also consistent with anecdotal evidence in the real world. For instance, when people are interrupted when telling a story, they often do not resume where they left off, perhaps replaying some of the story in their head or repeating the last few lines. That is, "Where was I?" followed by "Oh, that is right, this happened and then this happened" is often seen in real-world settings.

**Conclusion**

This research represents a crucial first step in understanding how interruptions affect the quality of work. Adults today are fond of multi-tasking, and they feel that they can do so with no impact on the individual tasks in which they are engaged (Aratani, 2007). This work suggests that this is not the case.

Future research in this area is needed to further explore these results. The types of tasks that are negatively impacted by interruptions need to be further examined. Our results have shown that interruptions negatively impact the overall quality of work in a
complex, creative writing task. However, it is unclear if the same decrements would occur in other types of tasks. The relationship between expertise on a task and quality of work could also be explored. In procedural tasks, overall disruption decreases as skill increases (Altmann & Trafton, 2007). It is unclear if this would translate to quality of work.

Once a better understanding of the types of tasks is understood, the mechanisms behind the decrement should be explored more fully. This might be done through an evaluation of physiological measures or performance tests, which can be used to assess changes in the body and brain. For example, research has shown a strong correlation between interrupted task performance as measured by resumption times and working memory capacity, as measured by the operation span (Foroughi, Werner, Cades, Boehm-Davis, 2014; Werner et al., 2008). Therefore, assessing working memory capacity may provide insight into individual performance while interrupted. Other research has shown that anodal stimulation of the dorsolateral pre-frontal cortex using transcranial direct current stimulation (tDCS) can increase neural firing in that area, leading to increased performance on a host of tasks (Nitsche et al., 2008). Therefore, using anodal stimulation over the dorsolateral pre-frontal cortex to increase neural firing may improve task performance while interrupted.

Once the mechanisms by which interruptions affect performance are better understood, researchers can begin to focus on creating and testing procedures and techniques that can overcome reductions in quality of work. For example, can an alert of an incoming interruption be effective at overcoming the reductions in quality of work? Or
would encouraging an approach that has been effective at reducing errors, such as rehearsal, be better? As technology continues to increase the number of interruptions we experience daily, answers to these questions will become increasingly important.
HOW DO INTERRUPTIONS AFFECT CONTENT PRODUCTION? AN ASSOCIATIVE ACTIVATION ACCOUNT

Activation-based models provide robust accounts for interrupted task performance on procedural tasks. Specifically, computational models instantiated in the ACT-R (Adaptive Control of Thought - Rational) (Anderson, 1993; Anderson et al., 2004; Anderson & Leibere, 1998) architecture of performance have successfully predicted that interrupting a procedural task will lead to time lags associated with the resumption of the original task (Altmann & Trafton, 2002; Hodgetts & Jones, 2006; Monk, Boehm-Davis, Trafton, 2004; Trafton, Altmann, Brock, & Mintz, 2003; Trafton, Altmann, Brock, 2005), an increase in overall time to complete the original task (Gillie & Broadbent, 1989; Mansi & Levi, 2013; Marulanda-Carter & Jackson, 2012), and a likelihood of error on the task at the point of resumption (Bailey & Konston, 2006; Cellier & Eyrolle, 1992; O’Conaill & Frohlich, 1995; Zijlstra et al., 1999).

However, it is not clear whether these models can predict performance on non-procedural tasks. Thus, the objectives of this research were to 1) determine how non-procedural tasks, such as a creative content production task, are affected by interruptions, and 2) examine whether an associative activation model can be used to account for interrupted task performance in this new domain. We start with an examination of two activation-based accounts of interrupted task performance – memory for goals and threaded cognition.
**Memory for Goals Model**

The memory for goals model (Altmann & Trafton, 2002) is an activation-based computational model. The model describes tasks as being represented as goals in memory, with each goal having a certain level of activation associated with it, and with the most active goal driving behavior on the task. The model also indicates that activation levels for any given goal decay as a function of time. When an interruption occurs, the active goal (i.e., the primary task a person is working on) must be suspended and the goal associated with the interrupting task becomes the most highly activated. When the primary task is suspended, the activation associated with the primary task goal decays as a function of time. The longer you are away from the task, the more the activation for that task decays. Decay leads to a reduction in activation. As the activation level decays, the goal will become more difficult to retrieve. The memory for goals model provides mechanisms for boosting the activation level of goals. These mechanisms are: 1) strengthening, which suggests that mechanisms that increase the frequency and/or recency of goal activation (e.g., through rehearsal) can boost activation strength; and 2) priming, which suggests that environmental and mental cues or contexts can boost activation.

**The Threaded Cognition Model**

The threaded cognition model is a more recent theory developed to explain the detrimental effects of interruptions on task performance that is also instantiated in the ACT-R cognitive architecture (Salvucci & Taatgen, 2008; 2010). Expanding upon memory for goals, Salvucci and Taatgen (2008; 2010) added two components – a task thread (i.e., thread as opposed to goal) and the problem state (a single mental resource
that represents the current state of a task) – to attempt to more fully account for interrupted task performance (Salvucci & Taatgen, 2008). Rather than goals, each task or thought is represented as a thread. In contrast to the memory for goals model, the threaded cognition model does not talk about activation, but rather describes the ways in which tasks are threaded. The threaded cognition model suggests that tasks occur on a continuum, where more than one task thread can be active at one time at a time. As a function of the multiple task-state, the model does not predict the time lags described in the memory for goals model. Instead, the threaded cognition model suggests that some processes, such as rehearsal of the primary task and interrupting task actions, occur concurrently. Concurrent actions are possible because the theory assumes that threads will continue based on the availability of resources in what is referred to as a greedy/polite manner. Greedy indicates that threads will use whatever resource they need; polite indicates that once a thread does not need the resource, it is returned to the general resource pool.

Another assumption of the threaded cognition model is that there is a separate resource that is dedicated to remembering the primary task when you are working on a separate task (Anderson et al., 2004; Salvucci & Taatgen, 2010). This is known as the problem state (Salvucci & Taatgen, 2010). The problem state resource is where interference issues can occur. Because there is only one resource dedicated to the problem state of a task, if someone is performing a primary task and an interrupting task with each requiring a problem state, there will be interference and the model predicts that an error is likely to occur (Salvucci & Taatgen, 2010). However, the prediction is that
these vulnerable points of interference are going to occur as you try to resume the primary task. As with the memory for goals model, the threaded cognition model can make predictions for errors following interruption resumption on a procedural task, but cannot be extrapolated to make predictions regarding overall quality.

**Experimental Rationale**

The question is whether procedurally-based models can be used to account for interrupted task performance of a content production task. The computational framework of the memory for goals model is well-suited to procedural tasks, but may not be useful for making predictions for content production tasks in which the next step is unknown. In addition, although the problem state rehearsal feature of the threaded cognition model may be helpful in modeling tasks where the next action on the primary task is predetermined, it is not clear that the features that boost resumption in procedural tasks will be useful in a content production task.

Recent findings suggest that interruptions negatively affect the overall quality of work on a content production task (Foroughi, Werner, Nelson, & Boehm-Davis, 2014). The experiment used a within-subjects design where two factors were manipulated: presence or absence of an interruption, and placement of the interruption in either the planning (outlining) or execution (essay writing) phase. The following findings resulted from the experiment: 1) a decrease in final essay quality as a function of producing fewer total words, which occurs when interrupted during the execution phase, and 2) a decrease in final essay quality when interrupted during the planning phase. It is particularly intriguing that when interrupted during planning, participants were able to produce the
same number of words for the essay as in the non-interrupted condition, but the quality was lower (Foroughi et al., 2014).

Unfortunately, the Foroughi et al (2014) work did not allow for an examination of the processes underlying performance. Once we can examine the processes, we may have a better understanding of the underlying cognitive mechanisms leading to a reduction in the ability to produce content. We designed a set of experiments to explore whether the traditional goal-activation account of procedurally-based interrupted task performance, specifically, the memory for goals model, can be applied to a content production task.

**Experiment 1**

To understand how interrupting tasks were affecting content production, we wanted to examine the planning process. Foroughi et al (2014) found a decrease in final essay quality when interrupted during the planning phase, but participants were able to produce the same number of words for the essay as in the non-interrupted condition. Unfortunately, the outlines produced in their experiments were such that they could not be compared. Thus, we designed an experiment to test the effect of an interrupting task on the planning process.

We hypothesized that interruptions during planning would decrease the amount of content produced. Further, we predicted that interruptions in either phase (planning or execution) would reduce the final quality of the essays. We also predicted that a comparison of the outlines produced in an interruption condition versus a non-interruption condition would provide insight into the mechanisms underlying interrupted task performance during content production, and that traditional theoretical accounts of
interrupted task performance would not be able to account for this performance.

**Method**

**Participants**

Twenty-seven students from George Mason University participated for course credit. The participants (18 females and 9 males) had an average age of 22.1 years, and were fluent in English.

**Task and Materials**

The primary task required participants to outline an essay following a structured outline format using pen and paper and then to write the essay on a computer using Microsoft Office Word\textsuperscript{TM}. Three essay prompts and assignments came from a stock bank of topics created by The College Board (2012b; see Table 4).

<table>
<thead>
<tr>
<th>Table 4. Sample Essay Prompt and Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Essay Prompt</td>
</tr>
<tr>
<td>Essay Assignment</td>
</tr>
</tbody>
</table>
The structured outline format required for outlining was based on Hacker’s (2006) *Sample Outline in an MLA paper* (see Figure 6).

The interrupting task consisted of answering a series of arithmetic questions using pen and paper. Participants completed as much of the interrupting task as they could in the given time. The interrupting task was a continuous self-paced task. The researchers ensured that participants actively completed the interrupting task.

![Sample Outline in an MLA Paper (Levi)](image)

*Figure 6. Sample of the hierarchical outline structure (from Levi, 2009)*
Design and Procedure

The experiment consisted of an outlining phase immediately followed by a writing phase (see Figure 7). All participants were first instructed on how to construct a formal, structured outline. They were shown three examples and asked if they understood how to outline in the format. Participants had 12 minutes to outline and an additional 12 minutes to write the essay. The experiment used a within-subjects design in which two factors were manipulated: presence or absence of an interruption and placement of the interruption (outline or writing phase). Participants served in all three conditions (outlining interruption, writing interruption, and no interruption) and wrote all three essays. A Latin Squares design was used to counterbalance the essay type (prompt 1, 2, and 3) and interruption location (outlining, writing, and none).

![Figure 7. Time course of each phase and condition. Twelve minutes were allotted for each phase.](image)

In an interruption condition, the participants received three separate 60-second interruptions at the three, seven, and eleven-minute marks. The three minutes lost to
interruptions were added to the total time available to complete the outline for a total of 15 minutes. The interruptions occurred without warning. A piece of paper with the interrupting task was placed over the participants outline during the interruption period.

When the interruption time ended, the paper was removed and participants resumed the primary task. Participants were informed that they were to complete each task to the best of their ability. Participants were not informed about the purpose of the study.

**Measures**

Word counts and the numbers of supporting arguments produced (primary and secondary) were used to evaluate the total content produced in the outlines and essays. A graded score, as averaged by two independent graders, was used to evaluate the final essays. The interrupting task was scored for accuracy.

Two independent graders were trained on how to evaluate essays before scoring the essays from the participants. Both reviewed the College Board *Essay Scoring Guide* (Appendix A) and read pre-evaluated essays from all the score ranges (0-6). Each grader also practiced scoring on pre-evaluated essays, and compared their scores to those official scores (The College Board, 2012a). Lastly, the graders independently scored pre-evaluated essays and then discussed their own evaluations together in an effort to ensure consistency. The training period lasted approximately seven days. The graders were blind to the experimental conditions, were paid for their work, and had no other affiliation with the experiment. They only received the final essay text. No other identifying information was attached. Inter-rater reliability was sufficient (Kappa = .741).
Results and Discussion

Mauchly’s Test of Sphericity was used to verify the assumption of homogeneity of variance for all repeated measure analyses of variances (ANOVA) reported. The assumption was met for homogeneity of variance in all cases, $p > .10$.

We first wanted to ensure that certain individual essay prompts (1, 2, and 3) and fatigue did not influence the results of the outline word counts or the results of the final essay scores. A repeated measures analysis of variance (ANOVA) revealed no reliable difference as a function of the individual prompts, $F(2, 52) = 1.32, p > .10$, and no reduction in quality over time, $F(2, 52) = 1.09, p > .10$.

We also wanted to confirm that participants were actively completing the interrupting task. The average score was 86%, with scores ranging from 75% to 96%. A correlation coefficient was computed to assess the relationship between the accuracy of the interrupting task and the overall word count of the outlines. There was no significant correlation, $r(27) = .11, p > .10$.

To establish whether interruptions reduced content production while outlining, we conducted a repeated measures ANOVA, which revealed a main effect for final word count of each outline across the three conditions, $F(2, 52) = 24.66, p < .01$. Post-hoc tests using a Bonferroni correction revealed that word count was reliably lower when interruptions occurred during the outlining phases ($M = 117.12$, SD = 46.3) when compared to the non-interruption ($M = 146.71$, SD = 48.1) and interruption during writing phases ($M = 152.59$, SD = 54.9). That is, interruptions during outlining reduced the number of words produced (see Figure 8).
We conducted a repeated measures ANOVA to determine if interruptions reduced the number of primary points created in the outline. The analysis revealed a main effect across the three conditions, $F(2, 52) = 5.101, p < .01$. Post-hoc tests using a Bonferroni correction revealed that the number of primary points created was reliably lower when interruptions occurred during the outlining phases ($M = 3.0$, SD = .6) when compared to the non-interruption ($M = 3.4$, SD = .9) and interruption during writing phases ($M = 3.4$, SD = .9). That is, interruptions during outlining reduced the number of primary points created (see Figure 9).
To discover whether interruptions reduced the number of secondary points created in the outline, we conducted a repeated measures ANOVA, which revealed a main effect across the three conditions, $F(2, 52) = 25.15, p < .01$. Post-hoc tests using a Bonferroni correction revealed that the number of secondary points created was significantly lower when interruptions occurred during the outlining phase ($M = 6.1, SD = 2.9$) when compared to the non-interruption ($M = 7.9, SD = 3.6$) and interruption during writing phases ($M = 7.7, SD = 3.6$). That is, interruptions during outlining reduced the number of secondary points created (see Figure 10).

![Figure 9. Mean primary points produced with standard error.](image-url)
We also wanted to determine if quality of the final essays suffered as a result of interruptions. A repeated measures ANOVA revealed a reliable difference in the final scores across the three conditions, $F(2, 52) = 22.04, p < .01$. Post-hoc tests using a Bonferroni correction revealed that scores in both interruption conditions were significantly lower than the scores in the non-interruption condition ($p < .01$ for both) with roughly a ½ point reduction in final essay score based on a six-point scale. There was no difference in quality ($p > .10$) between the two interruption conditions (see Figure 11).
A correlation coefficient was computed to assess the relationship between the word counts of the outlines and the final essay score. There was a significant correlation between the word count of each outline and final essay score, $r(27) = .501$, $p < .01$. This revealed that the amount of content produced (word count) in the outlines can be used to measure the outcome of the final quality of the essay.

Table 5 provides an overall summary of the results.
The data supported all of our hypotheses. Interruptions during the planning phase reliably reduced the number of words created overall, the number of primary points created, and the number of secondary points created. Moreover, interruptions in both the outline and writing phases reliably reduced the final quality of the essays. Interruptions led to lower quality outlines and a reduction in quality of the plan carried over to lower quality essays.

The timing and error measures from both the memory for goals model and threaded cognition were not useful measures of performance in the case of content production. Contrary to what both models would predict, and as evidenced by the lower content production numbers in the outlines in the interruption conditions, the availability of the environmental context and cues did not aid resumption in the way that has been
shown previously in procedural tasks. One reason why these models were not useful predictors in these experiments is that there is not a specific well-learned connection between steps in content production, unlike what is seen in a procedural task. In a procedural task, we know the next step on a task and there is a correct next step and an incorrect next step. Conversely, in content production, plans are produced internally; thus, if the creation of internal context is interrupted, or if an interruption occurs prior to writing down the plan, it may be difficult to recreate that internal context after the interruption is completed. In other words, for content production, there is no correct or incorrect step at the point of resumption; you are just trying to re-establish the internal context from where you left off on the primary task.

Environmental context, cues, and the ability to rehearse are central tenets of procedurally-based models of interrupted task performance. However, those mechanisms did not support interrupted task performance in a content production task. Although the models as they stand cannot accurately account for these findings, the underpinnings of the models, which are based on associative memory, can potentially be used to understand interrupted content production performance.

**Experiment 2**

The findings from Experiment 1 suggest that both the memory for goals model and threaded cognition models are not able to fully account for interrupted task performance during a content production task. Although the theories designed to explain interrupted task performance on a procedural task cannot provide specific predictions for
the performance deficit, the tenets of the models can provide a potential guidepost for theoretical investigation.

**Spreading Activation Theory of Semantic Processing**

We looked to the concept of activation and associative memory within the ACT-R framework to provide guidance for making predictions. Within the notion of associative memory, the concept of spreading activation seemed to us to provide a potential framework for making predictions regarding interrupted task performance on non-procedural tasks (Anderson, 1983; Collins & Loftus, 1975; Quillian, 1962; Quillian, 1967).

Spreading activation is a model of semantic processing based on associative processes in which nodes or concepts in the semantic network within memory are activated based on their levels of association with other concepts in the semantic network (Collins & Loftus, 1975). The semantic network is comprised of a complex set of associations between nodes connected by weighted associative pathways.

Pertinent to interrupted task performance, when a concept is sampled, the corresponding node activates and that activation spreads, or forward propagates, through the associative pathways to nearby nodes (i.e., most highly associated). Associations – that is, those nodes most closely related to each other conceptually – are driven by the level of activation of related nodes. Activation decays across associative pathways over time. Moreover, when a node is activated, activation for related nodes is propagated across the network from the source node. As the activation of specific nodes is based on probabilities associated with the level of activation, activation of a source node does not
guarantee an identical associative pathway through the semantic network at any one activation. Therefore, once a primary task is interrupted (i.e., forward propagation is interrupted), the same pathway from the source node may not be taken when the task is resumed.

We hypothesized that an interruption to the spread of activation would cause activation to spread from a faulty cue upon resumption, leading to forward propagation through a semantic network that is different from the path established prior to the interruption. We explored the application of the spreading activation theory using Verbal Protocol Analysis (VPA) to gain insight into whether these specific cognitive mechanisms (i.e., forward propagation of spreading activation) were driving interrupted task performance of a content production task.

**Method**

**Verbal Protocol Analysis**

VPA, otherwise known as “think-aloud” protocols, have been used in various forms of behavioral research since the 1950s to determine how people are thinking and how they are solving problems, and to explore many other features of cognition and task performance (Austin & Delaney, 1998; Newell & Simon, 1972, Trickett & Trafton, 2009). VPA uses observable behaviors to provide insight into the underlying mechanisms that may be in play during performance of a task (Ericsson & Simon, 1987; 1993; Fox, Ericsson, & Best, 2011; Martin & Klimoski, 1990; Sanderson, Verhage, & Fuld, 1989; Schweiger, 1983).

Concurrent VPA requires that a person performing a task talks aloud throughout
the entire task so that each thought and mental process is verbalized as a task is
completed (Austin & Delaney, 1998; Trickett & Trafton, 2009; Youmans, Gonzalez,
Figueroa, & Bellows 2013; see Austin and Delaney, 1998 for a primer on VPA).

Concurrent VPA has been deemed by some to be the gold standard and has been
used to understand cognitive processes in more than 1500 published journal articles. The
reason for the popularity of VPA for understanding underlying cognitive processes
involved in a task is that the person is not subject to the short-term memory loss that
occurs when talking through a task retrospectively (Austin & Delaney, 1998; Ericsson &
Simon, 1993).

Some debate exists as to the disruptiveness of Concurrent VPA to performance on
a task (Austin & Delaney, 1998; Trickett & Trafton, 2009; Youmans, et al., 2013). Some
studies have found that concurrent protocols can promote performance on a task (Barry &
Braodbent, 1984; Birns, Joffre, Leclerc, and Paulsen, 2002; Ericsson & Simon, 1987;
1993), whereas others have found that concurrent protocols can be detrimental to task
performance (Rhenius & Deffner, 1990). However, a recent meta-analysis of 94
Concurrent VPA studies found that the Concurrent VPA method had no effect on
performance (Fox, Ericsson, & Best, 2011).

Participants
Five participants (3 male and 2 female) with an average age of 25 years and
English fluency volunteered to participate in this study. Due to nature of the in-depth data
collection, VPA data provide a large volume of rich information; because of this,
information can reach a saturation point using fewer participants than in experimental
trials (Austin & Delaney, 1998; Trickett & Trafton, 2009), and sample size is typically less than 10 people. VPA studies with as few as one single participant have been fruitful in generating hypotheses to guide further research (Trickett & Trafton, 2009). As this was an exploratory experiment aimed at gaining a deeper understanding of the underlying cognitive processes related to interruption of a content production task, the smaller sample size was sufficient to meet our research objective.

Task and materials
The primary task, interrupting task, and materials were identical to Experiment 1.

Design and Procedure
The outlining and interruption procedures were the same as those used in Experiment 1.

Additionally, participants were trained on the Concurrent VPA method (see Table 6) using simple scenarios (e.g., how many windows are in your house?), and were able to practice as many times as needed until they felt comfortable with the procedure.

Once participants understood the outlining and Concurrent VPA methods, the experimental trial began. Using a counterbalanced design, each participant participated in one trial of outlining with interruption and a no talk-aloud non-interrupted control trial. Participants were given as much time as they needed to complete the outline to the best of their ability.
Data Analysis

Video screen capture and audio data were recorded for analysis using Screenflick™ software for Mac. Verbal protocols were transcribed and coded. The initial segmentation of the data was conducted using a top-down approach based on the current theoretical models of interrupted task performance. Points of interest were defined by the currently accepted interruption anatomy as well as previous research on interruptions. These suggest that the points on the anatomy that have the greatest impact on interrupted task performance are the point at which a person disengages from the primary task and the period between disengaging from the interrupting task and reengaging with the primary task (Boehm-Davis & Remington, 2009; see Figure 12).
We then coded the data by modifying an existing VPA coding scheme (Goor, 1974), which we chose based on its ability to capture the creation of new information versus the development of existing information related to the task; an important feature in understanding behavior in a content production task (Ericsson & Simon, 1993; Goor, 1974; Goor & Sommerfield, 1975). The modified coding scheme included five codes: 1) surveying given or generated information, 2) primary task reengagement utterances 3) producing new ideas (related to the thesis, to primary arguments, or to secondary arguments), 4) changing directions from previous course, and 5) developing an existing idea. Each sentence from the transcripts was coded using these five codes (see Table 7).
Results and Discussion

Surveying Behavior and Reengagement Utterances

Surveying behavior seemed to play a critical role in the resumption process.

Following an interruption, participants often re-read the idea generated just prior to the interruption. For example, “I can’t remember what I was writing [before the interruption]. I’m reading this over again;” “I’m confused about what my next argument is supposed to be. I’m reading....;” “Okay I’m thinking about “why people are more happy if they focus on their goals...” [rereads prior point];” and, “I’m thinking about why people are more happy if they focus on their goals because they will not be preoccupied.”

Similarly, utterances reflected frustration in the difficulty of recreating the internal context from prior to the interruption: “I don’t know if I can do this,” “I’m not thinking well,” and “I was kind of confused before I started doing the other task and now I’m even more out of it.”
Previous research on procedural tasks has illustrated the importance of cues to the success of resuming from an interruption (Altmann & Trafton, 2002; 2004; Hodgetts & Jones, 2006b; Cades et al, 2008). Therefore, it is not surprising that participants surveyed the information they wrote immediately prior to the interruption, as this was the only

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Related Quote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surveying</td>
<td>Words indicating the active use of provided or created information</td>
<td>“I can’t remember what I was writing [before the interruption]. I’m reading this over again.”</td>
</tr>
<tr>
<td>Re-engagement utterances</td>
<td>Words uttered during the reengagement period from the end of the interrupting task to the first actions back on the primary task</td>
<td>“I was kind of confused before I started doing the other task and now I’m even more out of it.”</td>
</tr>
<tr>
<td>Producing new ideas</td>
<td>Words related to the creation of or thinking about creating new ideas</td>
<td>“When I think of narcissistic tendencies, I think of...uh...social media and selfies.”</td>
</tr>
<tr>
<td>Changing directions</td>
<td>Any words related to a change in course from a previously determined plan or idea</td>
<td>“I’m trying to think of the way to word the thesis. Actually, I’m going to do that last.”</td>
</tr>
<tr>
<td>Developing existing ideas</td>
<td>Words related to expanding upon a previously generated idea</td>
<td>“I’m thinking how you benefit. Umm...teaching, exchanging knowledge, and umm...food. You will have food.”</td>
</tr>
</tbody>
</table>
primary task cue available. However, procedural tasks provide a salient cue to the proper next step on the primary task. That is not necessarily the case in a non-procedural task. The results of the VPA suggest that primary task cues did not promote resumption. It is possible that the cues available were not enough to recreate the internal context of the primary task from immediately prior to the interrupting task.

**Changing Directions, Idea Production and Development of Ideas**

The primary task plan often changed direction following an interruption. For example, one participant had created a thesis argument and then was starting to think about the development of the first supporting point, “Okay point one...umm...okay...people...” At that point, the interruption occurred. Instead of resuming to the creation of the supporting point, the participant changed directions, “I think I should go back [to the thesis] and talk about or present kind of like my points in my thesis because that’s how I remember doing it.” The correct cue was not available to guide the participant that working on the first supporting point was the next step. The next step was determined internally prior to the initiation of the interrupting task. Instead, all that was available upon resumption was the thesis argument written on the outline, so that was used as the cue.

The inability of the cues to recreate the internal context from prior to the interruption often led to a change in direction from the initial plan. By the third interruption, the direction continued to move further away from the original thesis argument. As participants resumed from the interruption, they used the last item created prior to the interruption to guide idea production and development. However, the internal
context from prior to the interruption was not reproduced following the interruption. For example, in an outline about happiness, one participant stated, “I want to describe narcissism and what that is,” but they never wrote the idea down. They were interrupted before they had the opportunity. Following the interruption, perhaps cued by an idea written about narcissism and social media one person commented, “maybe I will talk about how narcissism would be detrimental to happiness.” The change in direction following an interruption points to the potential impact of interruptions on the creation of a plan. That is, if a thought process is interrupted, it may be difficult to recreate that thought process at the point of resumption.

Discussion
These results have two important implications for understanding interrupted task performance on a content production task: 1) spatial cues from the primary task seem to drive resumption behavior, and 2) at resumption, participants do not seem to be able to recreate the internal context that existed prior to the interruption. Instead, participants tended to go back to the environmental task cue, which was the portion of the outline they were working on prior to the interruption, and continue their thought process from that point. However, using this cue in resuming the primary task tended to result in them moving away from their original plan (i.e., the plan formulated prior to the interruption).

Theories designed to explain performance on procedural interrupting tasks, such as the memory for goals model, cannot account for these findings. Step-to-step known actions on the task are characteristic of procedural tasks. Thus, successful resumption requires matching the internal context with real-world information, such as a cue to the
next known action as they would in a procedural task. These results suggest that in content production, cues are used for resumption, but are not helpful. This may be because the cues did not point to the next known action as in a procedural task. In content production, there is no real-world preexisting structure. The structure of the task is based on internal context; therefore, the only thing to rely on for resumption is reconstituting connections by starting from where you left off in the task.

**Cues and Environmental Context Drive Resumption**

Based on previous research in the area of interruptions in general, it is not surprising that cues are an important mechanism for resumption. Ratwani and Trafton (2008) posit that memory for spatial information related to the primary task is more resistant to decay and distraction than goal memory, allowing for successful resumption beyond the duration of goal memory. This hypothesis is based on eye-tracking data showing that when participants attempted to resume the primary task following a long interruption; they tended to go back to roughly the last place they looked prior to the interruption, regardless of whether or not this was the correct resumption location.

Further support for the importance of spatial information in resumption can be found in work related to the role of environmental context and cues in resuming an interrupted task. This includes availability of environmental context of the primary task during the interruption, and/or the availability of environmental context or cues during the resumption lag. Ratwani, McCurry, Andrews, and Trafton (2008) showed that auditory interruptions in which the primary task was still visible led to faster resumption than visual interruptions in which the primary task was completely occluded. Hodgetts
and Jones (2006b) lend additional support to the role of environmental context on interrupted task performance. Using a computerized Tower of London task, participants were to move five disks from one peg to another, following specific constraints to complete the task. In some of the conditions, participants were returned to the same task at the same step but the colors of the disks were changed. The change in environmental context led to longer resumption times, providing support for the importance of environmental context at the resumption step.

Resumption cues have also been suggested as supporting resumption performance. Ratwani, McCurry, and Trafton (2008) used a complex computer interface task in which participants follow a hierarchy of steps. Participants’ eye movements were tracked during the experiment and, in one experimental condition, the primary task was interrupted just prior to the last step on the primary task. When participants were returned to the primary task following the interruptions, researchers used eye movement data to see if participants were looking at the correct location on the screen (i.e., the last step on the primary task). Errors (i.e., forgetting the last step on the primary task at the point of resuming from the interruption) were predicted if the participant did not look at the correct step or if they looked at it for too long without taking action. Receiving an explicit cue at those potential error points reduced the number of errors made (Ratwani et al., 2008).

A Spreading Activation Account

Associative activation through spreading activation of the semantic network provides a robust account for the findings of interrupted task performance on a content
production task. Participants relied on primary task cues to aid resumption after they were interrupted. However, the cues were not related to the internal context established immediately preceding the onset of the interrupting task. Upon resuming from the interrupting task, the available primary task cue became the new source node for spreading activation. Differences in the plan may have resulted from spreading activation through a different pathway based on the new source node rather than the original source node of the task.

**Experiment 3**

This experiment was designed to provide further support for a spreading activation account of interrupted task performance during a content production task. Spreading activation theory would predict that frequent interruptions would make forward propagation through the semantic network difficult, and, as we have shown in previous experiments, the overall quality of the work will suffer.

Shorter frequent interruptions may lead to perseveration on ideas formed just prior to the interruption. The perseveration can be explained through a mechanism known as back propagation. According to ACT-R, each sub-goal that is being worked on receives a certain level of activation and this also gives a certain level of activation to the overall goal and to the sub-goal that was being worked on prior to this step. This is known as back propagation. Simply put, each successive step of a hierarchical process feeds activation to each step prior with diminishing returns the further back up the chain you look. Therefore, if a task is interrupted frequently, the highest-level goal (i.e., the source node for the associative pathway of the primary task) may have received enough
activation through back propagation to be active enough to have a high probability of directing behavior. In other words, a person would be return again and again to the source node of the task based on retrieval from memory.

Conversely, research in the area of design fixation has shown that interrupting a design task can support creative performance on that task (Jannson & Smith, 1999; Kohn & Smith, 2009). Longer interruptions have been shown to release the fixation (Jannson & Smith, 1999; Kohn & Smith, 2009). Thus, as opposed to the negative impact of longer interruptions found in previous interruptions research based on procedural tasks, it is possible that a longer interruption could support the creation of ideas in a content production task.

Spreading activation theory may provide the best account for both short and long interruptions. A longer interruption could lead to a complete decay in activation of the associative pathway related to the essay prompt. Unlike when experiencing frequent interruptions, where the original node is consistently receiving activation, during a long interruption, activation will continue to decay as a function of time. Instead of reconstructing context from the point of interruption, participants may be forced to start again from the beginning (i.e., recreating the plan), allowing for natural forward propagation through the semantic network from the source node.

Based on spreading activation theory, we predicted that one long interruption would lead to the production of more ideas than frequent/short interruptions, and would produce no reliable difference in outline quality from not receiving an interruption.

**Method**
Participants

Twenty-seven students (5 males, 22 females) with an average age of 21 years and fluent in English from George Mason University participated for course credit.

Task and materials

The primary task, interrupting task, and materials were identical to those used in Experiment 1.

Design and Procedure

The overall design and procedure were identical to Experiment 1 with two exceptions. First, instead of interruptions occurring three times, there were two interruption conditions, one in which the participant experienced three shorter 1-minute interruptions, and one condition in which the participant experienced one longer 3-minute interruption. There also was a condition in which no interruptions occurred. A Latin Squares design was used to counterbalance the essay type and condition. Second, the participants only completed an outline for the essay; they did not write an essay. Participants were not informed about the purpose of the study.

Results and Discussion

Mauchly’s Test of Sphericity was used to verify the assumption of homogeneity of variance for all the repeated measure analyses of variances (ANOVA) reported in this experiment. The results showed that the assumption was met for homogeneity of variance in all cases.

We also wanted to ensure that participants were actively completing the interrupting task. The average score was 87%, with scores ranging from 76% to 100%. A
correlation coefficient was computed to assess the relationship between the accuracy of the interrupting task and the overall word count of the outlines. There was no significant correlation, \( r(27) = -.08, p > .10. \)

The number of words produced in each outline was examined to see if the total amount of content produced was affected. The analysis revealed a main effect for final word count of each outline across the three conditions. \( F(2, 46) = 60.45, p < .05 \) (see Figure 9). Post-hoc qualitative analysis of the outlines revealed that participants followed the hierarchical format required to plan the essays. Post-hoc tests using a Bonferroni correction revealed that the number of words created was significantly lower for the 3 interruptions condition (\( M = 136.28, \ SD = 50.20 \)) when compared to both the non-interruption condition (\( M = 179.72, \ SD = 51.64 \)) and one long interruption condition (\( M = 158.67, \ SD = 52.37 \); see Figure 13).
Next, we wanted to determine whether there was a difference in the number of primary points created across conditions. The analysis revealed a main effect for the number of primary points created across the three conditions. \( F(2, 46) = 6.60, p < .05 \). As expected, post-hoc tests using a Bonferroni correction revealed that the number of primary points created was reliably lower for the 3 interruptions condition \( (M = 2.63, SD = .65) \) when compared to the non-interruption condition \( (M = 3.04, SD = .55) \). In support of our predictions, we also found a significant difference between the three interruptions condition \( (M = 2.63, SD = .65) \) and one long interruption condition \( (M = 3.00, SD = .72) \), with fewer ideas created in the three interruptions condition, and no difference between the number of ideas created in the one long interruption condition versus no interruption (see Figure 14).
We also wanted to establish whether there was a difference in the number of secondary points created in the outline across conditions. This analysis revealed a main effect for the number of secondary points created across the three conditions. $F(2, 46) = 16.46\ p < .05$. Post-hoc tests using a Bonferroni correction revealed that the number of secondary points created was significantly lower for the 3 interruptions condition ($M = 5.54, SD = 2.04$) when compared to the non-interruption condition ($M = 7.92, SD = 2.62$) and one long interruption condition ($M = 6.83, SD = 2.48$; see Figure 15).

Figure 14. Mean outline number of primary points created
We found that frequent interruptions lead to greater disruption in a content production task compared to one long interruption or no interruption. When a procedural task is interrupted, it has been shown that as the length of an interruption increases, the time to resume and the likelihood of making an error upon resumption increases (Altmann & Trafton, 2002; Cades et al., 2008; Monk et al., 2008). However, in the content production task used for this experiment, a longer interruption had no impact on performance.

**Discussion**

These data provide further support for a spreading activation account for interrupted task performance of a content production task. Spreading activation would posit that more frequent shorter interruptions cause only minimal decay to the associative pathways formed. When returning to the primary task from a shorter interruption, the last
idea written on the page may serve as a cue and become the source node. Spreading activation may occur from this new source node (rather than the original from prior to the interruption), leading to the creation of a different plan, or difficulty in developing ideas due to a continuous starting over of forward propagation.

Contrary to this, it is possible that a longer interruption may have resulted in a complete decay of activation for the primary task associative pathway. The decay of activation may have forced participants to use the original source node as their resumption point, which would allow for activation to forward propagate in accordance with the original plan.

**General Discussion**

The goal of these studies was to determine 1) how interrupting the process of creating an outline for an essay affected the subsequent formulation of the plan at resumption, and 2) what cognitive mechanism underlies this performance.

First, we wanted to determine the affect of interruptions on content production. We found in our first two experiments that interruptions have a generally detrimental impact on the overall quality of a content production task. However, our last experiment showed that this is conditional and some interruptions have no affect on content production performance.

Next, we wanted to examine what cognitive mechanisms could account for our findings. We found that the available theories of interrupted task performance could not account for our findings. For example, once it was determined that there was a decrement to performance even for experts, long-term working memory theory was dismissed.
because the basic tenet of this theory is that experts at a task will not suffer negative effects when resuming from an interruption. We also found that the memory for goals and threaded cognition models, although useful for explaining performance on procedural tasks, cannot be used to account for our findings with regard to behavior on a content production task. One reason why these models were not useful predictors in these experiments is that, unlike in a procedural task, in content production, there is no specific well-learned connection between steps. In a procedural task, we know what the next step is and there is a correct next step and an incorrect next step. Conversely, in content production, plans are produced internally and if the creation of internal context is interrupted, or if an interruption occurs prior to writing down the plan, it would be difficult to recreate that internal context after the interruption is completed. In other words, for content production, there is no correct or incorrect step at the point of resumption; you are just trying to re-establish the internal context from where you left off on the primary task.

These theoretical models did provide insight into the process of activation, which is an important factor related to interrupted task performance. We know that activation is fundamental to the ability to resume in procedural tasks. Although the overall task activation may not have played a role in the performance decrement we found (certainly participants were not forgetting they were working on an essay), activation of the separate ideas (or sub-goals) associated with the task seemed to play a role in resumption of the primary task. Both the memory for goals and threaded cognition models are instantiated within the ACT-R architecture, a computational model of memory based on
associative memory and the spreading of activation through association. It follows that overall quality effects may have been due to some sort of inhibition of the spread of activation to new ideas.

We also considered the role of cues and environmental context in interrupted task performance. An extension to the memory for goals model suggests that environmental context can act as a cue to increase the probability of a successful resumption (Ratwani & Trafton, 2008). We found that environmental cues were used to resume an interrupted content production task, but that these cues did not necessarily promote resumption performance as they have been shown to in procedural tasks. This is because cues in content production tasks do not necessarily point to the next action on the task as they do in procedural tasks.

We found support for a combined resumption mechanism including spreading activation theory and memory for the spatial location. Interrupting tasks suspend forward propagation of an associative pathway through the semantic network leading to decay of activation for that pathway. The associative pathway represents the internal context for the primary task developed prior to the interruption. Our data suggest that cues related to the original task are used to try and resume the task after being interrupted, and these cues may promote forward propagation. However, because task context is internal and the next correct action on the task is not necessarily externally motivated by cues (as it is in procedural tasks), forward propagation may occur down a different associative pathway than the original pathway propagated prior to the interruption. Based on these findings, it is possible that the longer interruption can serve as a potential mitigation
strategy to protect against the negative impact on interruptions. A longer interruption may allow ideas to forward propagate through the semantic network without older goals causing interference and leading down alternative associative pathways. However, this strategy merits further exploration.

These findings provide support for an associative activation account of interrupted task performance during a content production task. However, further exploration of associative activation, forward propagation, and how they interact with specific features of interruptions is needed to further develop theory and mitigation strategies in this area.
CONCLUSIONS

The objectives of these experiments were to determine whether interruptions have an effect on overall quality of performance and, if so, to establish a theoretical basis for this effect. Although the available theoretical models of long-term working memory theory, memory for goals model, and threaded cognition model have been used to explain interrupted task performance, they have been used only to account for time and error performance on procedural tasks and not to explain overall quality of performance.

In the first set of experiments presented in Chapter 1, we found that interruptions have a generally negative effect on overall quality of performance on an essay writing task. We also found that the available theories could not account for this finding. Specifically, once it was determined that there was a decrement in performance even for experts, long-term working memory theory was dismissed because the basic tenet of this theory is that experts will not suffer negative effects when resuming from an interruption. We also found that the memory for goals and threaded cognition models, although useful for explaining performance on procedural tasks, cannot be used to account for our findings with regard to behavior on a content production task. This is because content production does not rely on performing the task through a set of sequential actions; rather, it relies on reproducing the context for the task internally. However, these theoretical models did provide insight into the process of activation, which we argue is an
important factor related to interrupted task performance. We found that activation is
fundamental to the ability to resume in procedural tasks. Although the overall task
activation may not have played a role in the performance decrement we found (certainly
participants were not forgetting they were working on an essay), activation of the
separate ideas (or sub-goals) associated with the task seemed to play a role in resumption
of the primary task. It follows that overall quality effects may have been due to some sort
of inhibition of the spread of activation to new ideas.

We also considered the role of cues and environmental context in interrupted task
performance. An extension to the memory for goals model suggests that environmental
context can act as a cue to increase the probability of successful resumption (Ratwani &
Trafton, 2008). We found that environmental cues were used to resume an interrupted
content production task, but that these cues did not necessarily promote “correct”
resumption as they have been shown to do in procedural tasks. This is because cues in
content production tasks do not necessarily point to the next action on the task as they do
in procedural tasks.

Based on the collective findings of these experiments, we have proposed and
found support for a theoretical account for interrupted task performance during content
production tasks – spreading activation theory. Content is created through the spread of
activation (i.e., forward propagation) through an associative pathway of the semantic
network until that process is interrupted at which point the activation of the associative
pathway decays as a function of time. We were able to determine that interruptions have
a detrimental effect on the quality of performance on a non-procedural task due to a
suspension of forward propagation through the semantic network. However, the impact of the interruption depended upon certain features of the interrupting task such as the length of time to complete the task and the frequency of the interruption. This is because the success of interrupted task performance on a content production task is based upon the ability to recreate the internal context of the primary task from prior to the interruption. We have shown, based on spreading activation theory, that different interruption features will affect the ability to resume successfully from an interruption.

**Implications and Future Research**

Our findings have many implications for future work in this field. First, they point to a need for research to deepen our understanding of the cognitive processes underlying resumption of a task when there are interruptions to content production tasks. These studies present the initial evidence that spreading activation can account for performance changes. However, this is only a first step, and we need to continue to develop our understanding of how spreading activation would explain behavior related to interruption of content production tasks and test further predictions related to resumption performance.

Future research should also focus on determining what other creative domains are impacted by interruptions. For example, does the work of an artist or visual designer lose originality if the artist is interrupted during the design phase? Existing research related to interruptions in the area of creativity has resulted in disparate findings. On one end of the spectrum, interruptions have been shown to support creative performance such as in
research in the area of design fixation (Jannson & Smith, 1999; Kohn & Smith, 2009; Youmans, 2011a).

Design fixation refers to the tendency of a designer to perseverate on previous designs, leading to the creation of similar designs (Youmans, 2011a). In design fixation studies, interruptions to the design process have allowed for a break in that perseveration (Jannson & Smith, 1999; Kohn & Smith, 2009; Youmans, 2011a; 2011b). In other words, studies have shown that if a person is experiencing design fixation, a break from the task (i.e., an interruption to the task) has promoted a break in the fixation (Jannson & Smith, 1999; Kohn & Smith, 2009; Youmans, 2011a; 2011b). Similarly, our results suggest that an interruption, or the ability of the interruption to clear the activation of a particular path in the semantic network can allow for propagation down a different associative path. This could promote the creation of a new idea.

However, we have shown that interruptions can reduce creativity by inhibiting the ability to produce content altogether. At first glance it seems we have shown conflicting results, but in fact a spreading activation approach has broader implications for the design fixation literature. It is possible that more frequent interruptions may lead to an over-activation of the original source node and this leads to a perseveration or “fixation” on the original idea, making it difficult to forward propagate through the semantic network. On the other hand, a long interruption may clear activation of the associative pathway that was activated during the primary task. In turn, there is no consistent back propagation of activation to the source node because there was only one interruption.
Thus, a person returns to the primary task, the source node is reactivated, and there is the potential to forward propagate through the semantic network through a new path.

Lastly, future work should also focus on systems that can be used to reduce the number of interruptions that might have a negative impact on performance and to determine the specific interruption features that make them more or less disruptive to quality performance on content production tasks. Considering once again the concept of design fixation, it has been shown that designers working in groups have been able to be more effective at creating new original designs than those working alone (Youmans, 2011b). Although these findings may seem counter-intuitive to our initial results from Experiment 1; Experiment 3 provided evidence that not all interrupting tasks are created equal. Some interrupting tasks may be neutral and some interrupting tasks may even be beneficial to performance. This is especially important to consider when designing work environments where collaboration would be crucial to success versus environments where interruptions may have a more consistently negative impact. However, more research in this area is necessary to determine these specific features.

These findings point to a set of potential strategies that can be further developed and tested for use in today’s interruption-driven environment. As technology increases, the potential for interruptions increases. Cell phone calls, text messages, emails, instant messaging, social media only serve to increase the interruptions that may exist in our daily routines at work, school, home, or anywhere in between. We showed that interruptions do more than increase time on task and increase errors; they also reduce the amount of content produced and the final quality of the product.
## College Board grading rubric for essay scoring

<table>
<thead>
<tr>
<th>Score</th>
<th>What The Score Represents</th>
</tr>
</thead>
</table>
| **6** | An essay in this category demonstrates clear and consistent mastery, although it may have a few minor errors. A typical essay:  
  - Effectively and insightfully develops a point of view on the issue and demonstrates outstanding critical thinking, using clearly appropriate examples, reasons and other evidence to support its position  
  - Is well organized and clearly focused, demonstrating clear coherence and smooth progression of ideas  
  - Exhibits skilful use of language, using a varied, accurate and apt vocabulary  
  - Demonstrates meaningful variety in sentence structure  
  - Is free of most errors in grammar, usage and mechanics |
| **5** | An essay in this category demonstrates reasonably consistent mastery, although it has occasional errors or lapses in quality. A typical essay:  
  - Effectively develops a point of view on the issue and demonstrates strong critical thinking, generally using appropriate examples, reasons and other evidence to support its position  
  - Is well organized and focused, demonstrating coherence and progression of ideas  
  - Exhibits facility in the use of language, using appropriate vocabulary  
  - Demonstrates variety in sentence structure  
  - Is generally free of most errors in grammar, usage and mechanics |
| **4** | An essay in this category demonstrates adequate mastery, although it has lapses in quality. A typical essay:  
  - Develops a point of view on the issue and demonstrates competent critical thinking, using adequate examples, reasons and other evidence to support its position  
  - Is generally organized and focused, demonstrating some coherence and progression of ideas  
  - Exhibits adequate but inconsistent facility in the use of language, using generally appropriate vocabulary  
  - Demonstrates some variety in sentence structure  
  - Has some errors in grammar, usage and mechanics |
<p>| <strong>3</strong> | An essay in this category demonstrates developing mastery, and is marked by ONE OR MORE of the following weaknesses: |</p>
<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
</table>
| 2     | An essay in this category demonstrates little mastery, and is flawed by ONE OR MORE of the following weaknesses:  
  - Develops a point of view on the issue that is vague or seriously limited, and demonstrates weak critical thinking, providing inappropriate or insufficient examples, reasons or other evidence to support its position  
  - Is poorly organized and/or focused, or demonstrates serious problems with coherence or progression of ideas  
  - Displays very little facility in the use of language, using very limited vocabulary  
  - Demonstrates frequent problems in sentence structure  
  - Contains errors in grammar, usage and mechanics so serious that meaning is somewhat obscured |
| 1     | An essay in this category demonstrates very little or no mastery, and is severely flawed by ONE OR MORE of the following weaknesses:  
  - Develops no viable point of view on the issue, or provides little or no evidence to support its position  
  - Is disorganized or unfocused, resulting in a disjointed or incoherent essay  
  - Displays fundamental errors in vocabulary  
  - Demonstrates severe flaws in sentence structure  
  - Contains pervasive errors in grammar, usage or mechanics that persistently interfere with meaning |
| 0     | Essays not written on the essay assignment will receive a score of zero |
APPENDIX B

SAT Prompts from the College Board

Essay Prompt 1. “People are happy only when they have their minds fixed on some
goal other than their own happiness. Happiness comes when people focus instead on the
happiness of others, on the improvement of humanity, on some course of action that is
followed not as a means to anything else but as an end in itself. Aiming at something
other than their own happiness, they find happiness along the way. The only way to be
happy is to pursue some goal external to your own happiness.”

Assignment:

Are people more likely to be happy if they focus on goals other than their own happiness?

Plan and write an essay in which you develop your point of view on this issue. Support
your position with reasoning and examples taken from your reading, studies, experience,
or observations.

Essay Prompt 2. “Heroes may seem old-fashioned today. Many people are cynical and
seem to enjoy discrediting role models more than creating new ones or cherishing those
they already have. Some people, moreover, object to the very idea of heroes, arguing that
we should not exalt individuals who, after all, are only flesh and blood, just like the rest
of us. But we desperately need heroes to teach us, to captivate us through their words and
deeds, to inspire us to greatness.”

Assignment:
Is there a value in celebrating certain individuals as heroes? Plan and write an essay in which you develop your point of view on this issue. Support your position with reasoning and examples taken from your reading, studies, experience, or observations.

**Essay Prompt 3.** “It is often the case that revealing the complete truth may bring trouble—discomfort, embarrassment, sadness, or even harm—to oneself or to another person. In these circumstances, it is better not to express our real thoughts and feelings. Whether or not we should tell the truth, therefore, depends on the circumstances.”

**Assignment:**

Do circumstances determine whether or not we should tell the truth? Plan and write an essay in which you develop your point of view on this issue. Support your position with reasoning and examples taken from your reading, studies, experience, or observations.
REFERENCES


Background Television as an Inhibitor of Performance on Easy and Difficult Homework Assignments. *Communication Research*, 27(3), 293-326.


Younan, R.J., Gonzalez, C.A., Figueroa, I.J., & Bellows, B. (2013). The Effects of
The entirety of this chapter has been published in the journal, *Human Factors*, under the title, “Do interruptions affect the quality of work?” The reference is listed within the reference section.
CIRRICULUM VITAE

Nicole Elizabeth Werner received her Bachelor of Science degree from George Mason University (GMU) in 2009 after which she began the doctoral program in Human Factors and Applied Cognition at GMU under the advisement of Dr. Deborah Boehm-Davis. She received her Master of Arts in Psychology with a concentration in Human Factors and Applied Cognition from GMU in 2011. Her research focuses on the development and application of Human Factors theory in applied environments, especially in healthcare systems. Her research interests include interrupted and multiple task management, communication and team coordination, error analysis and prevention, and macroergonomic evaluation and design. She has worked in several hospital systems including, the Johns Hopkins Health System, Inova Health System, Children’s National Medical Center, the Vanderbilt University School of Medicine Center for Research and Innovation in Systems Safety, and the Johns Hopkins Home Healthcare Group. Her work in these settings has focused on improving patient safety and system performance in a variety of areas including the hospital pharmacy, nursing medical/surgical units, pediatric intensive care, neonatal intensive care, the emergency department, medical device use and procurement safety, skilled home healthcare, the operating room, trauma resuscitation, and most recently in transitioning patients across healthcare settings. She is currently a senior Human Factors analyst at the Johns Hopkins University (JHU) School of Medicine (SOM) in the Division of Geriatric Medicine and Gerontology as well as the JHU School of Nursing Center for Innovative Care in Aging, and a member of the JHU SOM Armstrong Institute for Patient Safety and Quality. She is a 2014 Telluride Patient Safety Roundtable Scholar and recipient of the 2014 Human Factors and Ergonomics Society Student Member with Honors award.