Designing Instruction For Recovering Alcoholics: The Role Of Executive Function And Levels Of Guidance In Learning From Visually Complex Simulations

Jeffrey Gutkin
Graduate Center, City University of New York

Recommended Citation
Gutkin, Jeffrey, "Designing Instruction For Recovering Alcoholics: The Role Of Executive Function And Levels Of Guidance In Learning From Visually Complex Simulations" (2015). CUNY Academic Works.
http://academicworks.cuny.edu/gc_etds/572

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Graduate Works by Year: Dissertations, Theses, and Capstone Projects by an authorized administrator of CUNY Academic Works. For more information, please contact deposit@gc.cuny.edu.
DESIGNING INSTRUCTION FOR RECOVERING ALCOHOLICS: THE ROLE OF EXECUTIVE FUNCTION AND LEVELS OF GUIDANCE IN LEARNING FROM VISUALLY COMPLEX SIMULATIONS

by

JEFFREY S. GUTKIN

A dissertation submitted to the Graduate Faculty in Educational Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2015
2015
JEFFREY S. GUTKIN
Some rights reserved.
This work is licensed under a Creative Commons Attribution 4.0 United States License.
http://creativecommons.org/licenses/by/4.0/
This manuscript has been read and accepted for the Graduate Faculty in Educational Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Bruce Homer, Ph.D.

Date

Chair of Dissertation Committee

Alpana Bhattacharya, Ph.D.

Date

Executive Officer

Jan Plass, Ph.D.

Peggy Chen, Ph.D.

George Otte Ph.D.

Jeanette DeCastro, Ph.D.

Members of the Dissertation Committee

THE CITY UNIVERSITY OF NEW YORK
Abstract

DESIGNING INSTRUCTION FOR RECOVERING ALCOHOLICS: THE ROLE OF EXECUTIVE FUNCTION AND LEVELS OF GUIDANCE IN LEARNING FROM VISUALLY COMPLEX SIMULATIONS

by

JEFFREY S. GUTKIN

Advisor: Bruce Homer, Ph.D.

The present study examines the design of visually complex science simulations. Building upon an earlier study by Homer and Plass (2014), the current research determines under which circumstances adult learners, and alcoholics in recovery, would perform better from while learning with different levels of guidance. It was predicted that alcoholic adults in recovery would have impaired Executive Function (EF) as compared to controls selected from the general population and that EF would affect learning. An experiment investigated whether levels of EF predict learning from simulations that offered higher or lower levels of instructional guidance. Participants were 76 adults, half of which were alcoholics in recovery. They were randomly assigned to a treatment condition that taught about the Ideal Gas Laws from either a simulation that allowed them to freely explore the controls or one that used guided animation.

Analyses of variance revealed that the control group scored significantly better than the experimental group in EF on tests of processing speed (Stroop S). The experimental group performed slightly better than controls on tests of interference (Stroop I) and scored better on the Stroop (I) as their length of sobriety increased, but there was no significant difference on either. Age had a significant effect on the results of the Stroop. Both groups scored worse with age on the speed tests, but better with age on the interference test. Using a stepwise linear regression
analysis it was shown that the best predictor of performance on both tests of comprehension and transfer was the card rotation test (ETS S-1). There was no significant difference between groups on this measure. Results suggest that after a significant time away from a drink there is no difference in learning capabilities between recovering alcoholics and controls when level of education is controlled.
Acknowledgements

This paper is dedicated to my late mother, Dr. Joan Gutkin.
# Table of Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
</tr>
<tr>
<td>2</td>
<td>Review of the Literature</td>
</tr>
<tr>
<td></td>
<td>Theoretical Frameworks</td>
</tr>
<tr>
<td></td>
<td>Empirical Research</td>
</tr>
<tr>
<td></td>
<td>Research Questions &amp; Hypotheses</td>
</tr>
<tr>
<td>3</td>
<td>Method</td>
</tr>
<tr>
<td></td>
<td>Sample</td>
</tr>
<tr>
<td></td>
<td>Instrumentation</td>
</tr>
<tr>
<td></td>
<td>Measures</td>
</tr>
<tr>
<td></td>
<td>Data Collection</td>
</tr>
<tr>
<td>4</td>
<td>Results</td>
</tr>
<tr>
<td>5</td>
<td>Discussion</td>
</tr>
<tr>
<td></td>
<td>Relevance</td>
</tr>
<tr>
<td></td>
<td>Limitations Within the Study</td>
</tr>
<tr>
<td></td>
<td>Limitations and Future Direction</td>
</tr>
<tr>
<td></td>
<td>Appendixes</td>
</tr>
<tr>
<td></td>
<td>References</td>
</tr>
</tbody>
</table>
### List of Appendixes

<table>
<thead>
<tr>
<th>Appendixes</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Stroop Color-Word test</td>
<td>95</td>
</tr>
<tr>
<td>B Card Rotation Task (S-1)</td>
<td>100</td>
</tr>
<tr>
<td>C Working Memory, Inhibition and Shifting Task (WIST)</td>
<td>101</td>
</tr>
<tr>
<td>D Identical Pictures Test (P-1)</td>
<td>102</td>
</tr>
<tr>
<td>E Visualizer – Verbalizer (VVQ)</td>
<td>103</td>
</tr>
<tr>
<td>F Paper Folding Test (VZ-2)</td>
<td>104</td>
</tr>
<tr>
<td>G Approved Solicitation Posters</td>
<td>105</td>
</tr>
<tr>
<td>H Severity of alcohol dependence questionnaire (SADQ-C)</td>
<td>107</td>
</tr>
<tr>
<td>I Knowledge of Chemistry Posttest for Comprehension</td>
<td>109</td>
</tr>
<tr>
<td>J Knowledge of Chemistry Posttest for Transfer</td>
<td>113</td>
</tr>
<tr>
<td>K Scoring Rubric for Transfer Posttest</td>
<td>114</td>
</tr>
<tr>
<td>L Prior Knowledge of Chemistry Pretest</td>
<td>120</td>
</tr>
<tr>
<td>M Cognitive Load Survey</td>
<td>124</td>
</tr>
<tr>
<td>N Interest in Chemistry Survey</td>
<td>125</td>
</tr>
<tr>
<td>O Demographics Survey</td>
<td>126</td>
</tr>
<tr>
<td>P Mini International Neuropsychiatric Interview, Parts I and J (MINI 6.0.0).</td>
<td>127</td>
</tr>
<tr>
<td>Q Consent Form/ Information Sheet (WIST)</td>
<td>131</td>
</tr>
<tr>
<td>R Graduate School &amp; University Center (CUNY) IRB Approval</td>
<td>132</td>
</tr>
<tr>
<td>S Rutgers IRB Approval</td>
<td>133</td>
</tr>
<tr>
<td>Table</td>
<td>Title</td>
</tr>
<tr>
<td>-------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>1</td>
<td>Table of Measures</td>
</tr>
<tr>
<td>2</td>
<td>Descriptive Statistics of Participant Demographic Conditions</td>
</tr>
<tr>
<td>3</td>
<td>Comparison of Groups on Demographic and Pre-test variables</td>
</tr>
<tr>
<td>4</td>
<td>Correlations of All Variables</td>
</tr>
<tr>
<td>5</td>
<td>Analysis of Covariance with Stroop (I) as the Dependent Variable</td>
</tr>
<tr>
<td>6</td>
<td>Group by Treatment Means of Pretest and Posttests (Comprehension &amp; Transfer)</td>
</tr>
<tr>
<td>7</td>
<td>Multivariate Analysis of Covariance for All Main Effects and Interactions</td>
</tr>
<tr>
<td>8</td>
<td>Multiple Analysis of Variance With Stroop Interference as a Categorical Dependent Variable</td>
</tr>
<tr>
<td>9</td>
<td>Stepwise Regression Analysis For Variables Predicting Posttest Comprehension Scores</td>
</tr>
<tr>
<td>10</td>
<td>Stepwise Regression Analysis For Variables Predicting Posttest transfer Scores</td>
</tr>
<tr>
<td>11</td>
<td>Multiple Analysis of Variance With Education as a Categorical Dependent Variable</td>
</tr>
<tr>
<td>12</td>
<td>Means of Posttest Scores By Level of Education</td>
</tr>
</tbody>
</table>
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Cognitive Theory of Multimedia Learning</td>
<td>2</td>
</tr>
<tr>
<td>2.1</td>
<td>More Guidance Model</td>
<td>25</td>
</tr>
<tr>
<td>2.2</td>
<td>Less Guidance Model</td>
<td>25</td>
</tr>
<tr>
<td>3.1</td>
<td>Histogram of the Stroop Interference Distribution with Normal Curve</td>
<td>62</td>
</tr>
<tr>
<td>3.2</td>
<td>Histogram of the Stroop Speed Distribution with Normal Curve</td>
<td>63</td>
</tr>
<tr>
<td>3.3</td>
<td>Boxplot of Alcoholics’ Performance on Stroop Interference by Length of Sobriety</td>
<td>67</td>
</tr>
<tr>
<td>3.4</td>
<td>Boxplot of Alcoholics Performance on Stroop Speed by Length of Sobriety</td>
<td>68</td>
</tr>
<tr>
<td>3.5</td>
<td>Scatterplot of Alcoholics Performance on Stroop Interference by Length of Sobriety</td>
<td>69</td>
</tr>
<tr>
<td>3.6</td>
<td>Scatterplot of Alcoholics Performance on Stroop Speed by Length of Sobriety</td>
<td>70</td>
</tr>
</tbody>
</table>
Chapter 1

The potential for learning from multimedia has increased dramatically with the ubiquitous availability of online computer-based devices such as cell phones, tablets, and laptops. Multimedia, in its most simplistic form, is pictures and words (Mayer, 2001). Several theories have arisen over the past ten to fifteen years which seek to explain the effects of multimedia on learning from a learner centered approach, each of which considers the human cognitive architecture and how it processes information. Three leading theories; a Cognitive Theory of Multimedia Learning (CTML: Mayer, 2001, 2005), Cognitive Load Theory (CLT: Sweller, 1986) and an Integrative Theory of Pictures and Text Comprehension (ITPC: Schnotz, 2004) have operated on a minimum assumption that humans have a limited capacity working memory (Baddeley, 1986; Miller, 1956), and that auditory and visual information is encoded in two different channels (Paivio, 1971). These theories share a common thread in that understanding the limitations of working memory is crucial to designing instruction that helps learners integrate new knowledge with existing knowledge. In order to answer empirical questions in the area of working memory and learning, this paper investigates the design of multimedia learning materials within a framework of CLT (Sweller, 1999).

A Cognitive Theory of Multimedia Learning (Mayer, 2001, 2005) states that learners are active, in that they actively select, organize, and integrate relevant information (Wittrock, 1989). In this theory, working memory processes information as representations simultaneously across two channels (Paivio, 1971); one processes words (logogens) and the other processes images (imagens) (see Figure 1.1). representations are held in working memory and two mental models are created, one for verbal and one for pictorial. Those representations are then integrated with prior knowledge into a limitless long-term memory store. Non-verbal (e.g., pictures and icons)
are reproduced in visual memory, but it takes more effort to process verbal representations because they need to be decoded and are processed sequentially. For example, when a learner reads the word stop, they process it as an image, then create a sound representation. This extra step utilizes cognitive resources and can create a challenge for information processing.

Figure 1.1. Information processing in a Cognitive Theory of Multimedia Learning

Since each channel can hold about 7 bits of information only for a few seconds (Miller, 1956), by introducing spoken words and images at the same time instead of written words and images, the amount of information that can be processed at one time is increased. The improper overloading of the channels can detract from and inhibit learning. In addition, introducing pictures and words through visual sensory memory alone creates visual complexity, and that is one type of visual complexity explored in this paper.

One advanced type of multimedia, dynamic simulation, helps learners to understand complex ideas through interaction. Dynamic simulations provide an environment where learners can move slides, push buttons, enter variables or drag objects in order; for example, to visualize the effect of manipulating interactions between variables in chemistry. Despite the positive
effects that these simulations can have, they are visually complex and learning might not be achieved if the simulation is not properly designed to match a cognitive human architecture (Homer, Plass & Lee, 2006; Homer & Plass, 2014; Plass, Homer & Hayward, 2009; Plass et al., 2009; Plass, Homer, Milne, Jordan, & Kim et al., 2007). The design of visually complex simulations has been guided by CLT as a factor.

Cognitive load theory is an information processing and capacity theory which suggests that building schema is a main function of learning. Schema theory states that “people represent knowledge as networks of connected facts and concepts that provide a structure for making sense of new information” (Anderson and Bower, 1983). Schema building is facilitated by designing instruction that lowers the workload a task imposes on an individual, described here as cognitive load (Plass, Moreno & Brunken, 2010). When intrinsic load is high learners are asked to hold too many elements in working memory at a single time, high cognitive load occurs and schemata are not created. In order to better understand how learning occurs, one factor that has been investigated quite rigorously is how working memory performs when instructional materials offer more or less guidance (Sweller, 1999).

Dynamic simulations are usually accessed through one-on-one computer interaction and are described as necessitating self-direction and exploration during learning. There are ways; however, to incorporate simulations into the curriculum that add scaffolding and support to make them less exploratory and more guided (de Jong & Joolingen, 1998). According to CLT, one critical feature in designing multimedia to lower cognitive load is through direct instructional methods such as the use of worked examples (Kirschner, Sweller & Clark, 2006; Sweller, 1988). Central to CLT is the idea that using worked examples reduces cognitive load and that giving students the opportunity to discover solutions through exploration (less guidance) increases load
(Sweller & Chandler, 1994). Contrary evidence; however, shows *discovery models* of instruction have been successful in fostering critical thinking (Bruner, 1961; Hmelo-Silver, Duncan & Chinn, 2006).

Knowing which processes guide learning and having the ability to measure those cognitive processes is necessary in order to further investigate the design of materials. According to Brunken, Seufert and Paas (2010, pp. 199) very little has been accomplished in finding a direct measure of cognitive load, and this has placed the development of CLT at an impasse. Much of the research based on a model of working memory has focused on measuring the capacity of the phonological loop and visuospatial sketchpad (Baddeley, 1986). Recently the focus of working memory research has been on the *central executive* and its three separate cognitive control functions defined as: *inhibition of pre-potent responses, switching* and *updating* as primary factors in the function of the central executive (Miyake, Friedman, Emerson, Witzki, & Howarter, 2000). Based on this framework, cognitive load theorists have explored how EF can predict learning from dynamic visually complex simulations (Homer & Plass, 2014).

Homer and Plass (2014) reported that one EF, inhibition (as measured by the Stroop color-word test: Stroop, 1935), interacted with level of guidance, and that learners with lower scores on the Stroop test performed better in environments which provided higher levels of guidance, and learners with higher EF scores performed better with less guidance. According to CLT (Kirschner, Sweller & Clark, 2006; Sweller & Cooper, 1985) direct instructional methods induce lower cognitive load than means-end methods, though other research (Hmelo-Silver, Duncan & Chinn, 2007; Schmidt, Loyens, van Gog &Paas, 2007) showed that learners who attempt to solve problems learn better. This phenomenon has also been subject to expertise
reversal (Kalyuga, Ayres, Chandler, & Sweller, 2003) and EF (Homer & Plass, 2014; Plass, Milne & Jordan, 2009). What explanation could be offered for these interactions?

According to Mayer (2001) the prior knowledge structures that a learner brings to a learning situation can trigger a cognitive conflict between existing knowledge and novel information. It is possible that higher functioning of the inhibitory processes in EF can control and moderate this conflict, or that instructional design which places less of a burden on working memory also takes the burden off of inhibitory functions. It is thought that EF varies as a function of natural development (St Clair-Thompson, & Gathercole, 2006), though not enough is known about EF. EF in its broadest terms is thought to be higher level control mechanisms of more basic functions such as attention and memory, as well as the planning and control of mental activities and behaviors. Measuring EF is complex, because there is no single test that can measure all aspects of EF and several tests have been found to measure the same aspects (e.g., Miyake et al., 2000). In this circumstance the ability to control inhibitory function, and a limited working memory may play a direct role in explaining this phenomenon, so the Stroop task directed at measuring these specific attributes was chosen. Since the Stroop test has been shown to measure inhibition and attention it was well suited for this study.

Theory about the central executive can guide the development of learning materials for those who have cognitive impairments. One population that is highly functional, but suffers from impairments of EF is alcoholics. For alcoholics in treatment, or very early in their sobriety, learning potentially life-saving information is hampered by cognitive impairments of memory and EFs (Pitel et al., 2007; Zinn, Stein & Swartzwelder, 2004). Pitel et al., found significant deficits in working memory (verbal, visual and multimodal span tests) and EF (organization, inhibition, and updating) between control participants and patients from an inpatient 30-day
rehabilitation center. Zinn et al. (2004) also found that similar patients were deficient in the areas of abstract reasoning, memory discrimination, and effectiveness on timed tasks. According to Pitel et al., alcoholics taxed their working memory to compensate for impaired EF during learning causing them to have difficulties acquiring novel information. It is likely, therefore, that this group of learners would not perform as well as controls on the Stroop color-word test. These findings suggest that designing instruction for alcoholics should consider their impaired EF and that studying the cognitive functioning of this group could answer several questions about the central executive itself.

Based on these findings, this paper investigated the interaction between EF and instructional guidance by extending the research conducted by Homer and Plass (2014) to a new population. This paper used two similar simulations developed by the CREATE lab at NYU that modeled the idea gas laws (see Figures 1.2 & 1.3) because of their availability. One offers more guidance and the other offers less guidance, and both were similar in the subject matter and the types of representations used to label pertinent controls (Icons and Symbols). The populations chosen were recovering alcoholics because they have been tested to have impaired cognition and would benefit from research that would help them learn and a control group. Participants were randomly selected and randomly assigned in a two by two factorial design to one of the two simulations. Each participant completed a battery of cognitive and performance tests and each completed the Stroop color-word test on a computer. The results are listed in Table 1.
Table 1.

*Table of Measures*

<table>
<thead>
<tr>
<th>Demographic Data</th>
<th>Experimental (N = 40)</th>
<th>Control (N = 36)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment A</td>
<td>Treatment B</td>
</tr>
<tr>
<td><strong>M</strong></td>
<td><strong>SD</strong></td>
<td><strong>M</strong></td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td>45</td>
<td>49</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Female</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td><strong>Education (Years)</strong></td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td><strong>Length of Sobriety (days)</strong></td>
<td>4696</td>
<td>3818</td>
</tr>
<tr>
<td><strong>Processing Speed</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stroop (S)</td>
<td>1096.92</td>
<td>329.52</td>
</tr>
<tr>
<td><strong>Executive Function</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stroop (I)</td>
<td>49.21</td>
<td>11.19</td>
</tr>
<tr>
<td><strong>Spatial Abilities</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paper Folding Task</td>
<td>8.68</td>
<td>3.43</td>
</tr>
<tr>
<td>Card Rotation Task</td>
<td>96.68</td>
<td>38.30</td>
</tr>
<tr>
<td><strong>Knowledge of Chemistry</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>3.95</td>
<td>1.56</td>
</tr>
<tr>
<td>Posttest of Comprehension</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Posttest of Transfer</td>
<td>5</td>
<td>3</td>
</tr>
</tbody>
</table>
Chapter 2

Literature Review

The following review of the literature identifies the key theoretical constructs of CLT and research that has been conducted within this framework. It draws from the domain of cognitive psychology in an attempt to answer questions which will inform learner-centered instructional design literature; specifically, the idea that in order to design successful computer-based instruction the human cognitive architecture needs to be considered (Mayer, 2001, Sweller, 1999), most specifically, the idea that the effectiveness of the design of learning materials is dependent upon the learner (Plass, Kalyuga, & Leutner, 2010). This study explores adults that suffer from cognitive impairment as the result of alcoholism and compares them to a control group of normal adults to learn more about the effect of different levels of guidance and the role of EF on learning.

The following theories and models that are pertinent to this question are discussed; first, a model of working memory is described (Baddeley, 1986, 2000; Baddeley & Hitch, 1974) and then Dual-Coding Theory (Paivio, 1971, 1986), which explains how information is processed and represented within that model. Next, the paper will describe a Cognitive Theory of Multimedia Learning (Mayer, 2001, 2005), which outlines the basic design principles of multimedia learning environments based on a limited capacity working memory, and an Integrative Model of Text and Picture Comprehension that discusses how representation and integration of learning materials occurs (Schnotz, 2005). A CLT, which explains how to design instruction for a limited capacity working memory (Sweller, 1988, 1999) will be discussed, and how cognitive deficits in alcoholics can impede their capability of learning novel information. After these theories have been defined, the following section will report the pertinent research
findings within these theories to include; comparing levels of guidance in learning, measurement of cognitive load and EF in controls and alcoholics, design of visually complex simulations, and the design of multimedia learning for alcoholics.

**Working Memory**

The basics of how human memory works (Baddeley, 1986, 2000) is defined first for the purpose of this paper because WM is responsible for learning and this memory model is central to most cognitive research. The roots of working-memory research are based on a limited short-term memory as described by George Miller (1956) early in the cognitive revolution which measured the capacity to recall items from short-term memory at seven bits of information (plus or minus two) and the ability to hold them in memory for a very limited duration (Peterson & Peterson, 1959). This short-term memory concept was later described in a tripartic, information processing model of memory that included a sensory, short-term and long-term memory (Atkinson & Shiffrin, 1968). In this model, learning occurs when iconic and echoic memories are buffered in the sensory register and held for short-term memory where they are selected and rehearsed and finally integrated into a limitless long-term store. Baddeley’s (1986) model is non-unitary in that it there are separate modules (stores) that each have a different function, unlike a unitary model that approaches memory from the premise that working memory is a single activated area of long-term memory (e.g., Cowan in Miyake & Shah, 1999).

Short-term memory, later called working memory because of its ability to work rather than act as a passive store, consisted of two slave systems and a control system; a phonological loop which echoes incoming sounds in order to allow rehearsal of verbal information for long-term encoding, a visuo-spatial sketchpad (VSSP) for processing and manipulating incoming
visual information, and a central executive (CE) which controls attention and action (Baddeley, 1986). Later, Baddeley (2000) added an episodic buffer to the model.

The phonological loop. The phonological loop, or articulatory loop, is described as a two-part system (1) which temporarily stores incoming verbal traces and (2) an articulatory function which reverberates or echoes the sound of the information to refresh (rehearsal) the store while it is learned (Repovs & Baddeley, 2006). Incoming auditory information enters the store directly and can be rehearsed using subvocal repetition in order to keep it active long enough to be integrated into long-term memory. Visual information, such as words or representations enter the system through recoding whereas it is articulated and then stored. Processing is linear in nature and highly limited in capacity as shown through five distinct characteristics; (1) a limited span, (2) the phonological similarity effect, (3) irrelevant sound effect, (4) word length effect, and (5) articulatory suppression (Repovs & Baddeley).

Studies testing limited span usually have subjects listen to or read lists of numbers or words and then recall the information. One common reasons that explain this limitation is poor integration during the learning phase that causes an inability to recall the information during the retrieval phase. It has been found that the nature of the information affects the capacity of this store. For instance, phonological similarity studies have determined that similar sounding objects are harder to remember than dissimilar objects regardless of their meaning (Baddeley, 1966). This is because they are processed based on sound and not meaning. This has led to the conclusion that semantic encoding does not take place in the short-term memory store, and is processed further along in the integration process. This phenomenon also suggests that the temporary verbal information is stored in a code unique to verbal information. This is further explained by Dual-Coding Theory; (Paivio, 1986) which will be discussed later in this paper.
When subjects are presented with irrelevant sound during or after learning words their ability to recall those learned words are decreased greatly (Repovs & Baddeley, 2006). This would be consistent with the idea that by preventing rehearsal, memory traces in the phonological loop are not maintained. Word length was also tested and found that as the word gets longer the ability of recall declines. If this is the case, it confirms the idea that words have to be rehearsed in order to be remembered, with longer words taking more time. Articulatory suppression provides strong evidence for a phonological loop. Articulatory suppression is accomplished by having a subject repeat an unrelated word during learning. By interrupting the rehearsal (articulatory) process during learning with a secondary task incoming information cannot be processed using subvocalization.

**Visuospatial Sketchpad.** The VSSP is thought to process visual information that does not include textual representation, only non-verbal images. Using a dual-task interference test to fractionate the VSSP has demonstrated it has two clearly different functions; (1) as a store of images and (2) as a manipulation of incoming information. Tests using double dissociation have been performed to try to isolate the subsystems. For example, tests of spatial interference have been found to hinder spatial processing and not visual pattern learning, conversely, using a visual interference task has been shown to hinder performance on a visual patterns task, but does not interfere with spatial processing (Duff & Logie, 1999; Logie, 1995).

Mohr and Linden (2005) investigated a possible fractionation between spatial and visual memory by using a dual-task to demonstrate that the two stores are active (spatial) and passive (visual) in nature. Visual information consists of form, color, and texture and the spatial domain handles spatial relationship and movements. These findings suggest that active spatial processing is dependent upon the central executive because of its dynamic nature while the visual storage
component is not. Although the VSSP is analogous in makeup to the phonological loop on several levels, one distinct difference is that the sketchpad and the visual store have separate and independent rehearsal mechanisms for both spatial and visual information and another is that the VSSP stores information relationally and in parallel rather than in a linear fashion.

Visual memory capacity is limited to about four items, but when these items are combined the number of objects that can be retained increase (Luck & Vogel, 1997). Luck and Vogel showed that although only four colors or four orientations could be remembered in visual memory at one time, when combined, or integrated into one object the number of items recalled grew to eight and up to 16 total items if there are four objects and four features. These findings suggest that each item is stored independently and then integrated to form a larger object.

Spatial processing encompasses the ability to hold objects and perform actions on these objects. Miyake, Friedman, Rettinger, Shah and Hegarty (2001) conducted a path analysis in order to best define the constructs of spatial abilities with standard field measures and describe their reliance on the central executive to function. More specific sub-processes are; spatial visualization, spatial relations, and visuospatial perceptual speed (Miyake et al., 2001). Of these, spatial visualization, e.g., the ability to rotate an object, is the most investigated. One psychometric test that aligns with spatial visualization is the paper folding task (see Appendix F). In order to perform the paper folding task, there are several intermediate steps which require internal visualizations to be constructed and held in memory, necessitating much input from the central executive. The card rotation task (S-1)( Appendix B) does not require intermediate steps to solve the problem and directly measures spatial relations. This test requires less input from CE to perform the task. The mechanism of CE, EF, is measured by the Stroop Color-Word test (Stroop, 1935; Appendix A). This test quantifies the ability to inhibit prepotent responses and
shifting tasks, and can possibly predict performance in simulations with high spatial working memory demands as found in multimedia learning environments (Homer & Plass, 2014). According to Wu and Shah (2004) students that can mentally manipulate information better, tend to perform better in chemistry tasks. The card rotation test accurately tests the spatial ability associated with this task. Perceptual speed tests of the identical pictures test measure the speed it takes to match two pictures with each other and probably places the least amount of demand on the central executive.

Perception in the visual divide is subject to both top-down processing where previously stored information cues the recognition and processing of novel information through transfer from long-term memory to the working store and bottom-up processes when grouped according to Gestalt principles. Using a change detection task, Rensink, O’regan and Clark (1997) showed that long-term memory stores certain selected aspects of a scene and later directs attention back to those aspects of interest that have been stored in visual working memory.

Neuropsychologists have used the dual-task using other modes of interference, such as motor skills, to isolate spatial functionality. Some of the common research paradigms utilize eye movement, tapping of fingers, and arm movement. Motoric movement seemed to interfere with spatial processing but not visual processing, suggesting that planning and executing voluntary movement shared the same resources as spatial memory. It was later found that visualizing these movements without actually performing them caused interference in that a shift in attention alone could produce these effects. These finding also suggests that this store is strongly linked to an attentional unit such as the central executive (Smyth, 1996).

**Episodic Buffer.** The episodic buffer (Baddeley, 2000) was added as a third sub-system in order to explain several functions which could not previously be explained by the tripartic
system. For example the ability to; relate meaning between words, chunk data, relate information to and from long term memory, combine multi-modal information between the two slave systems, and allow for the immediate recall of memory while delayed recall was impaired, could not been explained in the previous model. In general the buffer is responsible for the integration and temporary storage of multidimensional and multi-modal information. There is some controversy in the literature regarding the ability of the episodic buffer to handle the diverse set of instructions it would have to handle in order to function as described.

**Central Executive.** This unit is responsible for the control of the sub-systems of working memory. The central executive has been described as a *homunculus or ragbag* (Baddeley, 1986, 2000) because whenever there was a function that could not be explained in WM it was said that it was a function of the central executive. Over the last ten years, or so, many studies have been successful in fractionating and further identifying the functionality of the central executive in working memory. Baddeley (1986) incorporated the Norman and Shallice (1986) supervisory activating system (SAS) as a possible model for the central executive. In the Norman and Shallice model routine tasks can be controlled through schemata in what is called contention scheduling; however, novel tasks need to be thought out and processed by combining information from long-term memory (existing knowledge) and external stimuli, and could be better explained as a function of the SAS. From the SAS model Baddeley (1996) defined the EFs as the ability to focus, divide and switch attention, and to transfer the content of working memory to long-term memory.

Baddeley (1996) argued for the possible fractionation of three specific functions; (1) *dual-task*; the ability to coordinate function on two separate tasks (2) *switching*; the capacity to switch retrieval strategies such as in random generation, and (3) *inhibition*; the capacity to attend
selectively to one stimulus and inhibit the disrupting effect of others. Miyake et al. (2000) did not find a significant path distinguishing the dual-task as a separable function; however, did find a third distinguishable function for *Working Memory*, but subsequently renamed the function “working memory” to *Updating* because it was more representative of its function.

Measures for *Updating* (refreshing information in working memory), *shifting* (the ability to switch between mental models), and *inhibition* were found to have very low inter-correlations within-subjects suggesting each measure taps a different construct in what Miyake et al. (2000) terms *frontal lobe*, or executive tasks. There was a high inter-correlation between-subjects with tests which are purported to measure the same construct. The role of updating is to keep track of items in the working memory stores, update the stores as new information is needed and manipulate relevant items. Tasks used to measure working memory include the *keep track task*, the *letter memory task*, and the *tone monitoring task*. These three tests involve the constant monitoring and updating of information in working memory, but the type of information and goals are different. Miyake et al. (2000) define shifting as switching back and forth between two tasks and found three measurements; the *plus-minus*, *number-letter* and *local-global* tasks correlate as measures of shifting. Inhibition is defined as the deliberate inhibition of dominant, automatic, or prepotent responses when necessary. It is measured by the *Stroop test*, the *anti-saccade task*, and the *stop-signal task*.

According to recent evidence (Miyake et al., 2000) there is enough variance between constructs to substantiate the functional fractionation of the central executive into a diverse set of roles, but it is also suggested that there might be other EFs that have not yet been explored, and they might or might not be conducted by the central executive.
Theoretical frameworks

**Dual-coding theory.** Utilizing the phonological loop and the VSSP, *Dual-Coding Theory* (DCT; Clark & Pavio, 1991; Paivio, 1986) posits that working memory codes pictures and text differently on two separate channels and using both channels at once increases learning capacity. Dual-coding theory explains how information is processed on a surface level, which is the mechanical conversion of external information to an internal representation, through *referential links* and *associative structures*, and then processed on a semantic level (higher-order).

According to Mayer (2001), two mental models, one for images and one for words are constructed from those representations, though according to Schnotz (2005) there is one mental model created in the VSSP. In both models integration takes place between working memory and long-term memory. Higher order processes are considered to include *intention* and integration as learners actively seek information (Mayer, 2001, 2010). In these models symbolic information (words) are first processed through the visual cortex, then converted to a verbal model in the phonological loop. The mental labor invested in this conversion process is thought to be one basic cause of cognitive load.

**A Cognitive Theory of Multimedia Learning.** A Cognitive Theory of Multimedia Learning (CTML: Mayer, 2001, 2005), used interchangeably with the term Multimedia Learning Theory, also considers three types of cognitive load (DeLeeuw, & Mayer, 2008), and incorporates Dual-coding theory (Paivio, 1986), a model of working memory (Baddeley, 1986) and a Theory of Generative Processing (Wittrock, 1989) into its framework. In this theory learning takes place on two levels, perceptual and cognitive. Mayer (2001) proposes a theoretical framework for studying the effect of instructional content on the human cognitive architecture and posits twelve principles of design based on the central tenet that “people learn better from
words and pictures than from words alone” (pp.1). This tenet, the multimedia principle, is based on a cognitive framework that processes verbal and non-verbal information in parallel. By utilizing both channels at the same time, the amount of information that can be processed is increased, however, the improper overloading of the channels can detract from and inhibit learning. In CTML, a key characteristic of working memory is that it creates two mental models, one for verbal and one for non-verbal representations, with which, learners actively select, organize and integrate information from a sensory memory into long-term memory.

A few principles of multimedia design are present in this paper that are focused on reducing cognitive load through multimedia design (Mayer, 2005). The Split Attention principle states that when designing instruction multiple sources of information should be integrated to eliminate the need for integrating that information. Material that has high element interactivity is more likely to cause cognitive load when taught in isolation for expert learners because of the mental effort needed to integrate the elements. For low prior knowledge learners it is likely to be beneficial and for materials low in intrinsic difficulty this principle might not be a factor. Spatial Contiguity can be controlled for by moving corresponding information in a better proximity to each other on a page, and Temporal Contiguity, presenting auditory information at the same time as corresponding visual information. This reduces the need for learners to scan their eyes back and forth across a page, or remember what was said, both of which processes can use up cognitive resources.

The Guided Discovery principle suggests design factors when using discovery learning for learning science and designing science simulations (de Jong, 2005). This principle suggests that the learner’s prior knowledge should be taken into consideration when deciding the level of guidance to be offered. Furthermore the prior knowledge should be in scientific discovery. Low
prior knowledge learners might not benefit from discovery learning simulations and therefore guidance should be adaptable to the level of the learner.

**An Integrated Model of Text and Picture Comprehension.** An Integrated Model of Text and Picture Comprehension (ITPC: Schnottz & Bannert, 2003; Schnottz, 2005) seeks to explain the use of multiple sensory modalities during learning (multimedia learning). Learning, in this theory, occurs when external information sources are used to construct internal representations in working memory and then stored in long-term memory. ITPC assumes a limited capacity working memory (Baddeley, 1986, 200), Dual-Coding Theory (Paivio, 1974), and a single mental model (Johnson-Laird, 1983; Friedman & Miyake, 2000). In this model there are only two types of representations, *descriptive* and *depictive*. Based on a semiotic approach, (Pierce, 1955) descriptive representations are symbols that “have no similarity with their referent” (Schnottz, pp. 52). To understand a symbol a learner needs to understand the convention in which it is meant to be used, for instance the word b-i-r-d does not look like an actual bird. A learner upon seeing this symbol for a bird (b-i-r-d) would decipher the meaning of the individual letters, create a propositional representation of the word in working memory and draw a visual mental model of a bird in based on one found in their own long-term memory. Depictive representations are iconic and in this case a picture of a bird found in a textbook or on a computer screen looks like an actual bird.

Similar to CTML learning takes place on two levels, perceptual and cognitive. The perceptual level consists of multiple sensory channels consistent with a sensory register that processes surface level information. The cognitive level consists of verbal and pictorial channels similar to that described by Paivio (1974), but in this case each channel can convey multiple representations of texts and pictures. The verbal channel can process spoken text and *sound*
images and the pictorial channel processes written text and visual images. ITPC distinguishes itself from other multimedia theories in that it posits that memory integrates verbal and visual information to create one mental model, rather than create a verbal and visual mental model and then integrate it. Sound images have not previously been described by CTML or CLT and refer to direct representations of sound. For instance, in music a “high C” has a distinct pitch that a trained musician can match externally and internally. When a musician hears a note, auditive working memory can match the exact pitch to a mental model of that note that the musician already has stored in long-term memory. Based on this description it is possible that auditory information can be iconic in nature in that it is not subject to semiotics.

ITPC does not seek to measure cognitive load in terms of three distinct types of load, it does suggest however, that a single mental model is influenced more by the capacity of the visuo-spatial sketchpad in working memory than verbal working memory (Friedman & Miyake, 2000). In a further analysis it is suggested that visual and spatial processing are done separately, and that mental model construction is primarily a function of spatial memory (Johnson-Laird, 1983). Measuring spatial memory ability should have some predictive ability of the ability to create a mental model. In order to compare this theory with CLT the use of the terms mental-model and schema need to be determined. For the purpose of this paper mental models and schema are similar in the sense that they can reside in both working memory and long-term memory, are dynamic and can change with learning, help to organize information, and are subject to a limited capacity memory (Johnson-Laird, 1989 in Jones, Ross, Lynam, Perez & Leitch, 2011; Sweller, 1988). During learning existing schema or mental models are brought into working memory from long-term memory and integrated with new information creating a broader knowledge in LTM. Taking this approach allows this paper to hypothesize that
measuring spatial working memory can predict cognitive load in both schema creation and mental model construction and that properly designing instruction will aid both mental model and schema creation.

**Cognitive Load Theory.** The focus of CLT is on two constructs (1) cognitive load and (2) learning. Cognitive load is the amount of capacity the learning materials place on limited memory resources; and, learning, the permanent storage of information in long-term memory. Cognitive load theory is an *information processing and capacity theory* that offers a comprehensive framework to explain the load placed on a limited working memory by instructional design during learning (Sweller, 1988; 1999). Central to CLT is the premise that instructions which introduce cognitive activities not directly related to learning are *extraneous* in nature and should be limited. Though CLT has produced a large body of literature about instructional design, it has not contributed to the general knowledge of working memory. The primary scope of CLT has been to inform instructional design in the consideration of (1) intrinsic, extraneous, and germane type load (Sweller, 1994); (2) expertise-novice learners (Kalyuga, Ayres, Chandler, & Sweller, 2003; Touvenin and Sweller, 1999) (3) use of modality (Mousavi, Low and Sweller, 1995); and, (4) instructional format *e.g.*, the worked-example effect (Sweller & Cooper, 1985).

**Information Processing Theory.** CLT bases its assumptions in part on a two-process theory of information processing that functions within two different modes: controlled and automatic processes (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). According to this theory, when a process has become automatic, it can be retrieved into working memory (brought into attention) from long-term memory effortlessly, places little or no load on working memory, is developed with extensive practice and is difficult to alter once learned. Conversely, controlled
processes are not yet automated so they are cognitively demanding of attention, dependent upon load, easily established, yet easily altered by a learner. CLT describes these automatic processes as schema and assumes the goal of learning is to automate schema in long-term memory so that attention is not dependent upon working memory during learning. By creating domain specific schemata in long-term memory functions are automated so that future calculations or problem solving do not tax working memory (Sweller, 1988).

For example, expert chess players have created extensive schema memorizing large sequences of moves and can recognize patterns of play. This allows them to react quicker than players who are faced with novel situations during play (DeGroot, 1966). When existing schema are retrieved from long-term memory they are treated as a single element in working memory. Since there is no limit to the possible complexity of a single element, working memory can process exponentially larger chunks of information if elements are encoded in long-term memory (Chi, Glaser & Rees, 1982). In this model long-term memory is virtually limitless in capacity and information can be brought from LTM to WM instantly and that EF is controlled by existing schema in LTM and that schema determine what information is needed in learning situations.

**Types of Cognitive Load.** Cognitive load theory explains three separate types of load as, extraneous, intrinsic and germane. In total the additive effect of these types of load, in addition to free cognitive resources are the total cognitive capacity, but that load can be increased exponentially because it is not linear in nature. In other words, cognitive load cannot be measured by counting the number of elements in the learning material; rather it is non-linear in the sense that it can increase exponentially dependent upon individual learners and instructional design (Whelan, 2006). Extraneous cognitive load consists of elements that do not contribute to schema creation and automation and this type load can be reduced through instructional design.
Reduction in extraneous load frees up resources for germane and intrinsic processing. The intrinsic load of learning materials is inherent of the subject matters complexity and cannot be manipulated according to Sweller (1999), though there is some question as to whether that is the case (e.g., Lee, Plass, Homer, 2006). Intrinsic difficulty stems from a combination or number of elements to be held in working memory at a single time. Germane load is considered the effort used toward learning. Recently theorists have determined that germane load might not be necessary because it can be redundant and indistinguishable from intrinsic load. This assumption can be made if it is also considered that intrinsic load can be adjusted through instructional design (Kalyuga, 2011). If this is the case then instructional design models should focus on optimizing instruction by balancing intrinsic and extraneous load.

*Worked example effect.* Studying worked examples results in better performance on subsequent tests of problem solving than solving the equivalent problems (Renkl, 2005). In order to reduce cognitive load Sweller and Cooper (1985) examined two methods for teaching math problems; worked example and means-end. The worked examples effect posits that worked example problems aid learning because fewer productions are held in working memory at a single time leaving resources available for schema production. In a means-end problem the learner becomes overloaded because of the need to keep the goal and the process in working memory while solving a problem. Sweller and Cooper found that problem solving skills are domain specific and not necessarily generalized to solving problems in other domains. Therefore, a novice learner in one domain can also be an expert learner in another.
Research

The following research is broken down into the following sections; levels of guidance in learning, the expertise reversal effect, measurement of cognitive load, the Stroop Effect and EF in alcoholics, and the design of multimedia learning in alcoholics.

These images (see Figures 1 & 2) describe two simulations that differ in the amount of guidance given to the learner during learning. They were developed as part of an ongoing line of Institute of Education Sciences (IES) research. The IES provides federal funds for research and education in most fields of science and engineering. This IES research was conducted by the “Minds and Molecules: Optimizing Simulations for Chemistry Education” project at the CREATE lab, at the Steinhardt School of Education, New York University (create.nyu.edu) (Homer & Plass, 2010; Homer et al., 2009; Lee, Plass & Homer, 2006; Plass, Homer and Hayward, 2009; Plass et al., 2009; Plass et al., 2007). This NSF research targets STEM sciences and was used as part of this dissertation to extend the research to a new population.

Chemistry is a highly visual science and learning chemistry is dependent upon spatial ability in working memory (Baker & Talley, 1972; Wu, Krajcik & Soloway, 2001; Wu & Shah, 2002). In order to study the interaction of chemical components and visualize their relationships chemists have developed a “language” of visual representations. Three levels of chemistry that have been identified are; macroscopic, microscopic and symbolic. Macroscopic level processes are exemplified as burning candles and are observable, but microscopic level phenomena such as molecules and atoms, and symbolic level items such as numbers and formulas are invisible and abstract and harder to learn (Wu et al. 2001). One difficulty in learning chemistry comes when novice learners are required to interpret the meaning of these three types of representations during learning. While expert chemists can give concrete examples, novices lack the basic
understanding of the concepts. Furthermore, learners need to hold multiple items in working memory at a single time and transform single dimensional knowledge into two and three dimensional knowledge to create chemical models. Designing instruction to aid chemistry comprehension should consider that holding and manipulating multiple representations in a limited working memory taxes learner’s resources and can cause cognitive load. Scaffolding learners while they internalize the meanings of those representations can help novices to process chemistry problems more efficiently by reducing the load in working memory during learning. In order to enhance and aid chemistry learning a computer-based tool such as a dynamic visual simulation or animation could be highly beneficial to help learners visualize and manipulate microscopic and symbolic representations.

In this paper, and prior papers, different design factors are explored that can help learners build a better understanding of the subject matter using interactive simulations that reduce cognitive load during learning. In both of the models in this paper, important controls are represented by icons as well as symbols. In other studies key information has been represented by icons only or symbols only, or symbols and icons together. This model should benefit novice and expert learners because in high-complexity environments adding icons to symbols help both type learners. The simulations are high in intrinsic complexity because of the integration of the three variables (temperature, pressure and volume) of the ideal gas laws on one page, and also visually complex because they present a chart alongside the simulation (see Figures 2.1 and 2.2.)
**Figure 2.1. More Guidance Model**

In this section, you will study the relationship between two different variables at a time, with the remaining third variable fixed. Suppose you wish to see how the gas pressure changes when you change the volume of the gas. You would require to keep the temperature constant. Work through the following example. For each step in the simulation, click the numbered box to indel the number. Enter the result you obtain for pressure and volume. You can run the video again by clicking the button again.

- Select a value for the temperature and freeze it.
- Select a value for the volume.
- Click the value to be one-half of its value.
- What happens to the pressure?
- Decrease the volume button.
- What happens to the pressure?
- Double the volume.
- What happens to the pressure?

**Figure 2.2. Less Guidance Model**

Welcome to the gas law simulation!

Before you explore the gas law simulation, please do some exercises using the sliders and blocks in the simulation.

Change the value of temperature, pressure, and volume by using the sliders. You can also type the values directly into the corresponding boxes. Observe that the quantity of the other variable changes as a result of your input. The change is also shown in the graph.

You can lock the value of any variable by using the button next to the slider. Click on it to see what happens. Try using the sliders and the locks yourself and observe any changes to the quantities and the graph.

When you feel confident with manipulating the slides and locks, click on the "Continue" button.
In the more guided version learners are led through nine steps of learning, each step asking them to either watch the simulation, manipulate a variable, or check the results of a manipulation on the accompanying chart. In the less guided simulation learners have the ability to explore the variables by themselves. Though one has more guidance than the other, both allow for user control of the pace of the simulation. Both simulations use iconic and symbolic representation together.

Levels of Guidance. Although there is evidence that guided instruction (worked-examples) reduces cognitive load, and thereby aids schema construction (Kirschner, Sweller & Clark, 2006; Mayer, 2004), there is also evidence that Problem-based, Inquiry and Discovery learning are just as beneficial if they scaffold the learner (Bruner, 1961; Hmelo-Silver, Duncan & Chinn, 2007). The following sections describe two approaches that are at the center of this research. The goal of this paper is not to distinguish which method is better, or to describe how to design simulations that favor either method over the other, but rather, to understand whether cognitive differences in EF will affect a learner’s ability to learn better from either one.

More guidance. Sweller & Cooper (1985) argue that schema creation should be the focus when teaching learners algebra. In one experiment they tested students in algebra for their knowledge of Einstellung as an indicator of expertise. Einstellung is defined by Sweller and Cooper as a phenomenon that occurs when a “previously acquired schema is inappropriately used because a problem is incorrectly perceived as belonging to a familiar category that requires the use of that particular schema” (78). In tests of memory where participants were asked to recall random characters there was no difference in their memory’s capacity to do so, but there was a difference between year 9 participant’s ability and the remaining students in recalling strings of algebraic equations. This would suggest that the differences in existing schematic
knowledge of algebra affected the amount of strings of that could be recalled. This further suggests that more experienced learners can hold more elements in working memory within that domain than novice learners. They also found that the experimental group that learned from worked examples took less time and made less mathematical errors than the group that learned through worked examples. They suggest, therefore that initial learning should focus on the creation of schema so that that knowledge can be used later to solve problems, and that guidance needs to be given to novice learners to ensure that the proper solution is used for the proper category of problem.

Klahr and Nigam (2004) tested whether learners in a direct–instruction or discovery condition learned better by asking the learners to design four experiments that determine how far a golf ball or rubber ball will roll after varying the steepness and length of an incline. Learners in the direct group were shown how to conduct an experiment while the discovery group designed their own experiment. Although in both learning conditions the learners were actively participating, the difference was in the amount of instructor led guidance provided. In this study the direct instruction group was able to outperform discovery based learners in demonstrating expertise in control-of-variables strategy (CVS). Although the direct instruction group performed better, there were still significant gains made by the discovery group which led the researchers to suggest that individual differences, or specific characteristics of some learners make discovery learning less or more effective than other types.

In a recent article Kirschner, Sweller and Clark (2006) argued that minimally guided instruction is not effective and does not consider the human cognitive architecture. The argument raised in this article is that expertise in problem solving comes from a vast store of information already stored in long-term memory and that explanations (such as problem-based methods) that
Less guidance. Constructivism, based in part on Piaget’s research on accommodation and assimilation, has taken on many forms, including problem-based, (PBL; Barrows and Tamblyn, 1980) inquiry-based (Rutherford, 1964) and discovery learning (Bruner, 1961). Based on research by Piaget, Bruner suggested that children attempt to understand their environment through interaction with the environment, and therefore that the environment should provide opportunity for interaction with and construction of knowledge. Bruner discussed the need to transform what is learned by reconstructing knowledge in new ways. This method, known as discovery learning, suggested that learners should discover rules for problem solving on their own instead of through memorization. Although Bruner advocated discovery learning, he also understood the importance of laying a groundwork of knowledge first that considered the limited cognitive framework described by Miller (1956) and suggested that 7 bits of information should first be “gold and not dross” (Bruner, 1964 pp. 123). Through further investigation researchers identified pure discovery (no teacher guidance), guided discovery (some teacher guidance) and expository (problem and answer are provided) (Hmelo-Silver, Duncan & Chinn, 2007).

According to Shulman and Keisler, (1966) guided discovery is most effective because active learners gain appropriate knowledge, rather than potentially inappropriate methods.

Mayer (2004) suggests that learners are active in that they select, organize and integrate novel information and rather than learning by doing they learn by thinking. In this sense,
guidance involves helping students accomplish the above goals and pure discovery learning allows for misconceptions and confusion. Although Mayer does not advocate for pure discovery learning, CTML is founded on constructivist principles in that learners actively select information. This is predicated on the works of Wittrock (1989) who proposes that learning is a generative process where learners make connections between content and then between themselves and the content.

According to Schmidt, Loyens van Gog and Paas (2007) problem-based learning (PBL) does consider the human cognitive architecture. PBL allows levels of guidance to be manipulated by altering just-in-time information and monitoring the group closely to guide in the right direction as needed. In discussing this line of research it is important to understand that not all constructivist based teaching approaches are unguided and that PBL and Inquiry models should not be lumped in with unguided “pure discovery” models because by using scaffolding and guidance PBL can be highly effective (Hmelo-Silver, Duncan, & Clark, 2007).

Based on the above description, the approach used in this research for the “less guided” simulation would be considered less guided because it does not scaffold the learner or offer just in time information by the designers and allows them to freely explore the simulation. This design was used in order to properly represent two different sides of the spectrum; therefore, it is likely that novice learners will have significantly more difficulty on tests of comprehension and transfer than those receiving more guidance during learning. One factor that determines the amount of guidance that is beneficial to a learner is prior expertise in a subject domain.

**Expertise Reversal and the Prior Knowledge Principle.** Consistent with information processing theory, CLT studies show that prior knowledge moderates the effectiveness of most instructional designs. According to Sweller (1988) there are three distinctions between experts
and novices; “memory of problem state configurations; problem solving strategies; and features used in categorizing problems” (pp. 258). One example of a “memory of problem state configurations” expert can be demonstrated by chess players who have already committed hundreds of sequences of moves to memory for easy recall (DeGroot, 1966). “Problem solving strategy” domain experts have demonstrated the ability to work forward while novices use a less efficient means-end approach. “Features used in categorizing problems” are schema based. Experts will categorize problems based on their solutions, while novices might use a surface-structure to categorize problems. These distinctions make up the difference between the way novice and expert learners approach problem solving.

The effects of individual differences and instructional format on learning was explored by Touvenin and Sweller (1999) who conducted a 2x2 factorial experiment with prior experience on two levels (higher or lower levels) and instructional format on two levels (worked-example or exploration group). An interaction effect was reported where learners with less experience scored better from worked examples, but those with previous experience learned better from exploration. The interaction was more pronounced for novice learners suggesting that until schema are developed, more scaffolds such as more guidance is needed. Here, expert learners were less affected by instructional format.

Although expert learners are characterized by their existing schema, it does not always guarantee they will perform better. The expertise reversal effect (Kalyuga, Ayers, Chandler & Sweller, 2003) has demonstrated that as learners gain a higher level of expertise in a subject, they will need less instructional support. It has been shown that offering too much support to an “expert” that has already developed schema of that domain content, can detract from their learning. How can this be explained? Learners with no prior knowledge have to search their
memory extensively to solve a problem when they lack domain specific schema in long-term memory. This can cause overload. Learners with high prior knowledge who are given external support have to reconcile it with the guidance given by their own EF and this too can cause overload. A similar expertise reversal effect has been found to occur when expert learners are given too many visual scaffolds. In order to understand this effect and design materials that can adjust to compensate for learner differences, Kalyuga (2005 pp. 325) describes the need for future research to develop measurements that will give real-time feedback of cognitive processes to evaluate learner progress. In order to do so, cognitive structures need to be identified so that different measures are tapping the same construct across individual learners.

In an attempt to understand the role of multimedia in visual complexity, Lee, Plass & Homer (2006) added icons to symbols (words & numbers) in a science simulation. Adding icons to symbols could cause cognitive load because they convey redundant information. For example, a picture of a lock, and the word l-o-c-k appear to be the same information, but in essence do not cause load because, according to a semiotic approach, icons and symbols are processed differently (Pierce, 1955). Lee et al. (2006) predicted that the use of varied representations in a visual simulation would interact with learner’s prior knowledge because novices in a domain first understand iconic modes of representation and then over time develop an understanding of symbolic representation. The researchers found that the design elements intended to support the learning of low-prior knowledge students (icons) ended up increasing the cognitive load placed on more knowledgeable learners (icons added to symbols) and hindered their learning outcomes. This outcome was attributed to an expertise reversal effect for representation (Kalyuga et al., 2003). In another recent experiment using visually complex simulations (Homer & Plass, 2014) the expertise reversal effect did not occur for junior high school learners, but did for high school
learners when learning from visually complex simulations. In that study icons were effective for middle school children with both high and low prior knowledge, and effective for high school learners with low prior knowledge. These findings indicate that the expertise reversal effect can be mediated by cognitive development and other factors, not just domain specific prior knowledge. In a study by Lee (2007) spatial abilities were measured using the identical pictures test and the card rotation test. In that study learners with high spatial ability performed as well on either treatment (visual or no visual cues) and learners with low spatial ability only scored well with visual cues. These findings are consistent with both the individual principles effect and the spatial contiguity effect (Mayer, 2001). According to Homer and Nelson (2005) there is a developmental trajectory where children first learn the meaning of iconic representations and then build symbolic meaning. In domain specific areas, such as science, novice learners first build iconic representations and then develop symbolic representations regardless of age (Deacon, 1997).

Research has shown that performance is highly dependent on individual differences such as prior knowledge and EF. Although CLT has explained this difference in terms of prior knowledge, Homer and Plass (2014) found that EF interacted with instructional format as well as representation, specifically, inhibition as tested by a Stroop color-word test. It is argued that EF controls the intentional allocation of mental resources and therefore can compensate for cognitive load placed on learners by varied instructional formats. Furthermore, Homer and Plass found that level of guidance interacted with the intrinsic complexity of a simulation. When learning about kinetic theory, which is of relatively low complexity because it only models two variables, the exploratory group performed significantly better on tests of transfer. When learning about the Ideal Gas Laws, however, which has a higher intrinsic complexity, not all learners
performed better on tests of transfer given the ability to explore. Only those learners with higher levels of EF benefitted from exploration.

Although there is preliminary evidence of this interaction, there is very little research on the subject. Further testing with diverse populations could uncover more about the subject and is a primary focus of this paper.

**Measurement of Cognitive Load.** Initial work in cognitive load measurement used a production system approach, whereas “the number of statements in working memory, productions, number of cycles to solution and the total number of conditions were calculated” (Sweller, 1988, 263). This type measure focuses on the intrinsic difficulty of the learning materials, *mental load*, and did not consider individual differences in learners. Since the introduction of CLT a reliable measure of the construct has been sought. There is still no agreed upon standard for measurement, though it is proposed that a direct-objective measurement is superior to indirect-subjective measure (Brunken, Plass & Leutner, 2003). Indirect measures are subject to user performance and can vary within and between users, while direct (usually physiological) measures are not subject to the same interference.

One subjective study conducted by Paas and van Merrienboer (1993) used a self-report survey taken by the learner after performing a learning task. This likert scale type report measured *mental effort* in a posthoc rating that considered mental effort. Mental effort is the actual cognitive capacity a learner uses during task performance. Later, an *objective* performance calculation was added to this measure which was described as *cognitive efficiency*. The resulting score controlled for the effort a learner reported using during learning with the actual amount of learning achieved, whereas high efficiency correlates with low effort and high performance, and low-efficiency correlates with high effort and low-performance (Paas, Touvinen, Tabbers, & van
Gerven, 2003). Though self-report measures have been very popular they are not always compatible with an instructional design that wishes to take a measure of load during learning. It is also difficult to compare between learners, causing a lowered internal consistency of the measure. While we do gain a multitude of information from these measures it is desirable to have a direct objective measure as a comparison.

There are several objective measurements which have been tested that include, learning outcomes (Mayer, 2005; Mayer & Moreno, 1998), time-on-task (Tabbers, Martens, & van Merriënboer, 2004), task complexity (Seufert, Janen & Brunken, 2007), physiological (Van Gerven, Paas, van Merriënboer, & Schmidt, 2004), dual task or secondary task (Brunken, Steinbacher, Plass & Leutner, 2002), and eye-tracking (Folker, Ritter, & Sichelschmidt, 2005) (Plass, Moreno & Brunken, 2010 pp.193). Learning outcomes are gathered using pre and posttest scenarios whereas, high performance on a posttest after engaging in a learning intervention supposes a lower level of cognitive load imposed by the design of the materials. Since learning is the main objective of instruction, this measure is used in this research design. Though outcomes are objective, they are not direct. For instance, it is thought that while learning novel information a physiological change in palm sweat, or pupil dilation occurs which cannot be controlled by the learner. Measuring this change can shed light on the amount of mental effort being used in working memory and can guide instruction. Though both of these physiological measures are promising the cost and availability of having this type equipment on hand makes in prohibitive outside of the laboratory.

The dual-task has been beneficial in the study of working memory by allowing researchers to explore the fractionation of memory subsystems. The premise is that if a subsystem is being accessed in an attempt to perform a task and a different task is introduced that
uses the same cognitive subsystem, the successful completion of the task will be diminished. If, however, a second task is introduced and performance is not diminished then there is evidence that it is utilizing a different subsystem, or that the subsystem has greater capacity. Brunken, Steinbacher, Plass and Leutner (2002) used a dual-task to assess the cognitive load placed on a learner while interacting with an audio-visual or visual simulation. In theory, if a learning scenario induces less cognitive load then learners’ response times to a secondary task should be faster. Brunken et al. (2002) found that response time to a secondary visual task was faster when presented in audio-visual mode because, based on the dual-coding principle, this scenario uses less cognitive resources. However, in one dissertation learning for the dual-task group was 5-15% lower than controls with no dual-task for groups in higher load scenarios (Whelan, 2006). Because participants watched the letter “A” in a box placed above the learning materials and pressed a space bar when it turned red, the lowered learning could have been caused by a split-attention effect (Chandler & Sweller, 1992), whereas learners in the dual-task group had to focus their attention on the “A” and not the learning scenario.

In another study designed to test the dual-task as a measurement of cognitive load, Brunken, Plass and Leutner (2004) added two secondary auditory tasks, background music and background music with narration. According to the coherence principle (Moreno & Mayer, 2000), when visual only or auditory and visual scenarios are presented, a decrease in response time to an auditory secondary task when seductive or unrelated music is added should occur. In the study by Brunken et. al. (2004), however, the addition of music when no other auditory information was being presented had little impact on performance of a secondary audio task, and when narrative information pertinent to learning was presented response time diminished. Brunken et al. suggest the lack of effect was due to time on task because in their experiment
learning was for approximately 15 minutes, and Moreno and Mayer only allowed their learners 3
minutes, so learners had more time to select relevant information and ignore non-relevant
information (Brunken et al.)

The above studies uncovered potential success for measuring cognitive load using the
dual-task, though it has been problematic. In an unpublished study Gutkin (2010) used the dual
task to measure cognitive load in a visually complex science simulation that presented pertinent
information in iconic only or iconic and symbolic form. In this study learners pressed a spacebar
on the computer keyboard when the background changed colors. Though learning was not
affected between groups despite the addition of the secondary visual task, participants did not
show significant knowledge gain in either condition. Despite some success in the area of dual-
task as a cognitive load measure, it has focused on measuring memory capacity rather than
further exploring how working memory functions. This study evaluates tests of executive
function and the interaction with levels of guidance in hopes to provide meaningful data that can
point to a better way to measure cognitive load.

**Stroop Color-word test:** It is likely that EF relates to learning. Blaire, Knipe and
Gamson (2008) related the ability to solve math problems to a potential *interference resolution*
problem such as that found on the Stroop test. Blaire et al., went on to say that EF is not fully
developed until mid-to late adolescence, suggesting a developmental difference between
learners, and also suggest that lack of prior knowledge rules could cause errors in calculations.
Homer and Plass (2014) tested EF and prior knowledge against representation and instructional
format on high school students and found an expertise reversal effect for representation, and an
interaction effect for EF and instructional format. Executive function in this article was measured
by the Stroop color-word test, which more specifically has been found to accurately predict
levels of “inhibition.” Inhibition is described as an intentional act of EF control (Miyake et al., 2000). Furthermore, since intention has been described by Mayer (2001) as a higher-order process in active learning, and linked to mental model construction, it is possible that on a semantic level the control of intrinsic load is dependent on more than working memory capacity for the integration of novel information and, possibly, dependent upon individual differences in executive functioning. If this is so, the measurement of EF becomes of central importance to multimedia learning and cognitive load theorists, more specifically, inhibition, because it might predict the use of the proper instructional format. The following section describes a population of learners that has been shown to suffer from impairments of inhibition, making them good candidates for this research and EF research in general.

Executive function in recovering alcoholics: According to Jellinek (1960) alcoholics have a psychological continual dependence on the effects of alcohol to relieve bodily or emotional pain and alcoholism has been classified as a disease in the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). Jellinek has shown that even after alcohol is removed from the body of alcoholics they still return to drinking days, months and years after. This action suggests some type of cognitive deficiency. According to Zinn, Stein and Swartwelder (2004) cognitive deficiency can affect the success rate for patients who might have difficulty learning information that can keep them sober. Zinn et al. go on to say that EF deficits are the cognitive impairments most likely to affect alcoholic’s success in achieving long-term sobriety.

Executive function deficits have been identified within several areas of addiction such as; Binge Eating Disorder (Duchesne et al, 2010), Cigarette Smoking (Hefferenan, Carling & Hamilton, 2014), Cocaine Dependence (Albein-Urrios, Martinez-Gonzalez, Lozano-Rojas & Verdejo-Garcia, 2014), Heroin (Chuang & Jiaxiu, 2004), Opiates and Pedophilia (Cohen, Nesci,
Steinfeld, Haeiri & Galynker, 2010), Internet Addiction Disorder (IAD) (Young, 1996),
Massively Multiplayer Online Role Playing Games (MMORPG) (Metcalf & Pammer, 2011),
Marijuana (Gruber, Sagar, Dahlgren, Racine, Lukas, 2012), and Methadone Maintenance
(Abidizadegan, Moradi & Famam, 2008). These studies utilize the Stroop color-word test to
uncover deficiencies in inhibitory responses. Interestingly, there are similar cognitive
deficiencies found between addicts who have abused substances and those that have abused non-
narcotic behaviors. Drawing from this comparison it is possible to say that the abuse of narcotics
or alcohol may not be the cause of the cognitive deficiency, rather, it is possible that the
deficiency is present before, during and after use. For example, there were significant cognitive
differences between MMORPG that reported to be heavy users and those who reported to be
addicted with addicted players scoring significantly lower on cognitive tests. The functional
difference between the two is that those who are addicted play many more hours until their lives
are affected by symptoms such as not showing up for work, or not eating and they can’t stop on
their own.

Pitel et al. (2007) found significant deficits in working memory (verbal, visual and
multimodal span tests) and EF (organization, inhibition, and updating) between control
participants and alcoholic patients in an inpatient 30-day rehabilitation center. The researchers
worked within a framework of working memory (Baddeley, 1986) and an Adaptive Control of
Thoughts model (ACT: Anderson, 1992). Pitel et al. cited the need for alcoholics to inhibit
customary behavior and adopt new habits in order to stop drinking; therefore, the Stroop color-
word test was used to assess inhibition. Controls scored significantly higher than alcoholics on
that test ($p = .0005$). Consistent with the findings of Pitel et al., Zinn et al. (2004) found that
similar alcoholic patients were deficient in the areas of abstract reasoning, memory
discrimination, and effectiveness on timed tasks.

Despite their being a plethora of research on the EF of alcoholics in their first 30 days of
abstinence, there is little research on the long term cognitive effects of alcoholism after they have
stopped drinking (Fein, Torres, Price and Sclafani, 2006; Fein, Shimotsu, Chu and Barakos,
2009). According to Fein et al. (2006) many of the cognitive deficits in long-term abstinent
alcoholics (6.7 years) were similar to controls, except for spatial processing abilities. Fein et al.,
found a significant difference between controls and alcoholics with controls scoring better on
tests of spatial abilities. Despite these findings the authors suggest more testing is needed on the
population of alcoholics with longer lengths of abstinence. In a second study, Fein et al. (2009)
found that spatial processing deficits were related to gray matter deficits in the parietal lobe,
suggesting that learning from visually complex simulations could be more difficult for alcoholics
because they heavily utilize the VSSP during learning.

Zinn, Bosworth, Edwards, Logue, and Swartzwelder (2003) found individual cognitive
processing across demographic factors found that impairments were fairly homogenous, though
identified two possible confounding characteristics to be gender and age. The greatest effects of
aging on cognition were from the 31–50 to 51–65 age groups, middle-aged (40–66) alcoholic
patients showed the most pronounced deficits and the 20–39 group showed clinically relevant
deficits as well. For gender, Huizinga, Dolan and van der Molen, (2006) found that females
scored better than males on the Stroop color-word test of attention. An effect size of .62 (p < .05)
was found. Fein et al. (2006) however, found that adding gender as a covariate in a
developmental study on EF did not alter the main effects or interactions.
This paper focused on designs that improve learning by reducing cognitive load. In order to do so the design needs to consider individual differences in EF. The literature shows alcoholics suffer from both limited working memory and EF skills; therefore alcoholics were chosen as an experimental group for this study (Pitel, et al., 2007; Zinn, et al., 2004). Based on the characteristics of this population we would expect to find exaggerated variance in EF measurements in alcoholics with shorter lengths of abstinence, and if EF interacts with levels of guidance there should be a pronounced effect between instructional formats across learners.

**Multimedia learning design for alcoholics.** There has been little work done in the area of specializing multimedia instruction for alcoholics (Alternman & Baughman, 1991). Though most alcoholics show cognitive deficits during and shortly after detoxification there is little research that explores the extent in which these patients can learn and retain information presented to them in treatment centers, or even years after they have stopped drinking, from multimedia based learning materials. There are no articles in the multimedia learning design literature that are focused directly on this population. Though there is some information about how computers can be used for alcohol screening (e.g., Murphy, Bijur, Rosenbloom, Bernstein & Gallagher, 2013), cognitive-behavioral therapy (e.g., Carroll et al., 2008), or cognitive recovery (Peterson, Patterson, Pillman & Battista, 2002), they are not focused on design factors that could improve leaning materials in general for this specific group of learners. This is unfortunate, because it is important that alcoholics in recovery learn about their disease and how to avoid relapse in their sobriety. Computer-based learning can provide individualized attention to patients, but there has been very little positive effect shown on learning from in-patient rehabilitation studies (Stalonas, Keane, & Foy, 1979; Meier, 1988; Alternman and Baughman, 1991) Each of these three tested alternative methods of instruction including computer-based,
videotape and written (traditional). Though all methods have been found minimally effective, there has been very little difference between them in terms of learning outcomes between alcoholic groups.

A study by Alternman and Baughman (1991) tested videotaped versus computer interactive learning modules with alcoholics in an outpatient alcoholism rehabilitation program and controls of similar demographics. In this study 91 male alcoholics were tested against 35 male non-alcoholics. It was hypothesized that the interaction in an active learning environment provided by a computer program could improve learning. The study utilized a 2 (group) x 2 (intervention) RM-ANOVA and found no significant difference between alcoholics and controls, or between videotaped versus computer-based treatments in knowledge acquisition and there was no significant learning found within the pre and posttests as well for either groups. There are a few explanations for these findings; first, the population tested was from a lower socio-economic environment with the average education being close to the high-school level suggesting that they may not have been experienced learners. Second, the learning materials were primarily verbal in nature and did not tax the visuo-spatial areas of cognition. Since the VSSP is the area where deficits have been found in alcoholics compared to controls, there was less of a pronounced difference between the controls and the alcoholics and therefore less apt to effect learning of this modality. Finally, the videotaped intervention was a 45 minute lecture given by a doctor which was presented using audio and visual modalities and the computer-based intervention used visual materials only with the text presented in written and not auditory form. It is possible that the use of visually complex materials only, instead of audio and visual content, in the computer based learning environment added cognitive load and negated what positive effects the active learning
environment might have provided when compared to the videotape. Since no learning took place in either condition it was not possible to find an effect on learning.
Research Questions and Hypotheses

Question (1): What is the difference between recovering alcoholics and normal control subjects in executive functioning?

Since recovering alcoholics have been shown to suffer from cognitive impairments, performance on cognitive tests of executive functioning should differ. Furthermore, it was shown that EF is most impaired in alcoholics in early sobriety (less than 3 months).

I hypothesize;

(H1.1): Control group participants will score significantly higher on tests of EF

(H1.2): Alcoholics in early sobriety will score significantly worse on tests of EF than those with longer term sobriety.

Question (2): What is the difference between alcoholics and controls in learning from simulations with more guidance or less guidance?

Assuming that alcoholics have been shown to have impaired EF I expect that as cognitive load is increased by offering less guidance in a learning scenario that those with impaired inhibitory functions will perform worse. If, as Mayer (2001) suggests, the prior knowledge structures that a learner brings to a learning situation can trigger a cognitive conflict between existing knowledge and novel information, then as levels of inhibition differ amongst subjects, so would the ability to resolve this conflict.

I hypothesize;

(H2): There will be a significant interaction between EF (inhibition) and instructional format (more guidance vs. less guidance) of learning materials in a population with impaired EF. More specifically:
(H2.1): Participants who score significantly better (lower) in measures of inhibition will perform better on posttests of learning from an exploratory format of instruction.

(H2.2): Participants who score worse, (higher) on tests of inhibition will perform better on posttests of learning from a guided format instruction.

and ask;

**Exploratory Question (3):** What is the relationship between cognitive functioning and learning in alcoholics?

In order to learn more about how this population learns, the relationships between EF, prior knowledge, learning environments, demographic information and covariates will be explored.
Chapter 3

Methods

A 2x2 factorial design was used with treatment on two levels (less guidance & more guidance) and group on two levels (experimental & control) in order to help predict the effect of guidance on learning from these populations. All significance was determined at a value of $p < .05$. Descriptive statistics were collected for all tests and variables in the study; Gender, Age, Pre-test of knowledge, Level of Education, Visual-Verbal, Interest, Card Rotation, Paper Folding, and Length of Sobriety. ANOVA was conducted to test for significant group differences. All group means and correlations were reported (see Table 3). Permission to conduct this study (#11-04-093-0135 – NYC) was granted by the Institutional Review Board for Human Participants of the Graduate Center, CUNY (Appendix R) and by the Institutional Review Board for Human Participants of Rutgers State University, Newark New Jersey (#Pro2013003032: Appendix S). This study in part extended the research performed by Homer and Plass (2014) to a new population. Seventy-six participants over the age of 18 years old were selected from people who attend meetings of Alcoholics Anonymous in the community as well as a control group selected from the community and randomly assigned to a treatment group.

Sample

To solicit community participants a flyer was posted in places where meetings of Alcoholics Anonymous took place, community centers and churches in the area, and in colleges (See Appendix G). The locations were selected by proximity to the Graduate Center and targeted many locations. The study was scheduled to take place evenings during the week in one-hour sessions with multiple participants per session, at the Graduate Center in a designated area. Later in the study, the Rutgers Ph.D. laboratory in the School of Nursing was recruited as a secondary
place of data collection. By adding this location, several participants could be solicited for the control group from staff at the Rutgers School of Dental Medicine. Data was also collected using portable laptops at locations where AA meetings were held at the request of the participants. Participants were randomly assigned to one of two conditions; instruction using more guidance or instruction using less guidance.

Once a participant took part in the study they were asked if they could refer another person as well. Some participants for both groups came from a group of sober motorcycle riders who started participating and also brought others significant to their recovery, some of whom became part of the control group. This aided in the completion of the study by providing ample participants but necessitated the researcher to use laptops and set up mobile data retrieving methods because participants were not willing to travel to Newark after work hours and that was the time most people were available.

The sample consisted of two groups (experimental and control) randomly assigned to two levels of instruction (less guidance, and more guidance). According to Cohen (1992), a sample size of $N = 45$ is suggested in each group for power of .80 in order to detect a medium difference ($d = .25$) with ANOVA between two independent sample means at $\alpha = .05$. In two similar studies by Homer et al. (2009) significant findings were reported with $N = 67$ participants and Homer and Plass (2010) reported significant findings with $N = 186$.

In this study I expected a larger effect size because the population was impaired. A more pronounced reaction to instructional format was expected, because there was expected a greater variance in prior knowledge and EF. In the case of a large effect ($d > .4$) $N = 72$ would yield significant results. Therefore, a total of $N = 30$ participants per group with a total of $N = 120$ participants overall should have provided sufficient results.
Inclusions and Exclusions

Inclusion criteria for the Alcoholic group was a primary diagnosis of alcohol dependence, and a minimum of 24 hours of abstinence from any mood altering substance, similar to the criteria used by Peterson, Patterson, Pillman, and Battista (2002) and by Zinn, Stein, and Swartzwelder (2004). Abstinence was measured through a self-report survey and prior alcohol usage will be measured using the SADQ-C (Appendix H). Since the goal of this research was to investigate EF, which structurally, has been linked to functioning taking place in the frontal lobes of the brain, results from dual-diagnosis patients (those with bi-polar, schizophrenia and similar disorders) were not excluded from the study. The cut-off age for the study data was 65 because validation studies have shown that after 65 there is an acute decline in EF abilities resulting from age (Houx, Jolles, & Vreeling, 1993).

Instrumentation

Participants accessed instructional material, the Cognitive Load Survey, the Stroop Test and the knowledge pre and posttests, through a web-based interface on a computer provided by the researcher. The computer was connected to the internet via a wireless LAN network or through a hotspot on the researcher’s cell phone. The website was hosted by “Godaddy” web-hosting service and it gathered data from the experimental portal and stored it in a Structered Query Language (SQL) database. Initially the experimental data was delivered in the form of Cold Fusion (.cfm) but had to be converted to .php in order to work with the current web server. The database stored data about the treatment condition and user ID, Stroop, knowledge pre-test, posttests and the cognitive load survey. The data was exported from the SQL database to the researcher’s laptop using .csv formatted excel files.
The Working Memory, Inhibition and Switching Test (WIST) was initially created and subsequently delivered using the e-prime psychology software tool package and was accessed via the researcher’s computer. All WIST information was stored in a file on the local computer. All data was imported into an excel spreadsheet and then loaded into the statistical package for the social sciences version 21 (SPSS 21) where it would was analyzed. Because of technical issues, the WIST was discontinued and no data was coded from the small batch that was collected (see discussion section for more details).

As stated, participants were randomly assigned to one of two instructional modules; more guided instruction, or less guided instruction (see Figure 1 & 2). The modules were created by the CREATE lab, NYU and demonstrate the “ideal gas laws” by providing simulated interaction of three variables; pressure, temperature and volume. Participants in the less guided group learned by manipulating the levels of each variable using a slider and watching the effects of the manipulation on the dynamic simulations which were being recorded on an adjacent graph. Participants in the more guided group learned by following a list of instructional hyperlinks which activated recordings created by recording manipulations made using the less guided module. Both treatments were presented in the online module and accessed via web browsers.
Table 2.

*Participants' Demographic Characteristics for Alcoholic and Control Groups*

<table>
<thead>
<tr>
<th></th>
<th>ALCOHOLIC GROUP</th>
<th>CONTROL GROUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants (N)</td>
<td>40</td>
<td>36</td>
</tr>
<tr>
<td>Ages (M) (range)</td>
<td>46.68 (19-65)</td>
<td>39.69 (19-65)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males (N)</td>
<td>32</td>
<td>14</td>
</tr>
<tr>
<td>Females (N)</td>
<td>8</td>
<td>22</td>
</tr>
<tr>
<td>Level of Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completed by Years (M)</td>
<td>13.97</td>
<td>15.11</td>
</tr>
<tr>
<td>Average Days Since Last</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drink (M) (in years)</td>
<td>4136.93 (11.6)</td>
<td>N/A</td>
</tr>
</tbody>
</table>
Measures

*Inhibition of Prepotent Responses.* An internet version of the Stroop color-word test (Stroop, 1935) (Appendix A) was used to measure inhibition of prepotent responses (EF). Two levels of functioning were tested, Speed (S) and Interference (I). Test-retest reliability coefficients of the standard Stroop color-word test have been found to be very high ranging from .71-.89; (Houx, Jolles, & Vreeling, 1993; Siegrist, 1997; Strauss, Allen, Jorgensen, & Cramer, 2008). The Stroop color-word test was run using a web-based version as opposed to a local computer administered version. One problem that could have arisen from using a computer-based version, rather than a paper and pencil type, is that using the internet could have caused skewed results due to transmission speed. Linnman, Carlbring, Ahman, Andersson, and Andersson (2006) validated the results of the Stroop Color-Word test for the internet by comparing it to the conventional computerized version using 20 male and 20 female participants. While in their study it was found that response times were slightly slower using the web-based version, the validity of the test was not diminished.

The Stroop consists of three separate sets; the Neutral (WNI), Congruent (WCI) and Incongruent (WII), with each set consisting of 23 trials. The scoring method used in this paper for the Stroop was first, within each of the sets remove any trial that was responded to incorrectly, then, find the mean response time for each of the 23 trials next, calculate for the standard deviation of each set using an Excel spreadsheet, and then remove outliers more than three standard deviations from the mean. In order to determine an interference score, the mean response time from trials WNI and WCI were added together then the mean of the WNI and WCI was subtracted from the WII trial mean score [(WII – (WNI+WCI)/2]. The product of that calculation was used as a continuous variable. The Stroop was also used as a categorical variable.
by calculating the mean, using that mean as a high-low cut point and then creating two groups. For the Stroop Speed test, the sum of the WNI and WCI was divided by two and that mean was used as a continuous variable \([\text{WNI} + \text{WCI}/2]\). By utilizing the mean score rather than the trials total scores of 23 responses, internet glitches in the form of outliers were controlled for. In order to use the raw interference score in SPSS analyses a \(Z\)-score was calculated using SPSS and that score was then converted to a T-score (Van der Elst, Van Boxtel, Breukelen & Jolle, 2006). It is important to note that a lower Stroop (I) score represents a higher level of EF and a lower score on the Stroop (S) represents faster processing speeds in EF.

**Instruction.** Instruction was recorded as two categorical independent measures as instructional format on two levels with version “B” being more guided and version “A” being less guided (Figures 1 & 2).

**Learning.** Learning was assessed at two levels; Comprehension (Appendix I) of basic facts and principles (10 short-answer/multiple choice questions); and Transfer (Appendix J) of concepts to related problems (6 short-answer questions). Comprehension questions were scored as pass/fail, while transfer questions were given a score of zero, one or two by graders using a rubric that was developed for the transfer test scoring. Inter-rater reliability was calculated. The raw score outcomes were used as scale items in SPSS. A rubric was used to score the test of transfer with a maximum score of 12 points possible (Appendix K).

**Prior Knowledge.** A 7 item multiple choice pre-test (Appendix L) was used to assess prior knowledge in chemistry as a continuous independent variable to be used as a between group comparison in prior knowledge. This measure assessed the learner’s knowledge of the interaction between volume, temperature, and pressure. The student’s score on this test distinguished between learners with low prior knowledge and learners with high prior knowledge.
using a linear scale of one to seven. The final single digit score was used in statistical analysis as a scale value. An inter-rater reliability of .88 has been demonstrated in past research (Plass & Homer, 2011).

**Cognitive load.** Learners answered two questions (Appendix M) on a symmetrical scale from 1 through 9, with 1 being “extremely easy” and 9 being “extremely difficult”; (1) How easy or difficult was it to work with this simulation? (2) How hard did you have to think to understand kinetic theory in this simulation? This measure was administered but was not included in the data analysis.

**Executive function.** Executive function was measured as a continuous independent variable with 3 subcomponents; *working memory* (updating), *inhibition* and *shifting* (Miyake et al., 2000) with the working memory, inhibition and shifting (WIST) task (Hallion & Jha, 2011)(Appendix C). “The WIST is a combination of several widely-used cognitive paradigms, including the N-back test of WM, a two-choice Inhibition task, and a switching task. Participants complete high- and low-demand trials for each ECP” (Hallion & Jha). The WIST has been validated (N = 80) in the above study to accurately measure EF. Results from a follow-up validation study (N = 116) (Hallion, Coutanche, & Jha, manuscript in prep) provide evidence of good convergent and discriminant validity of the WIST with existing ECP measures. This measure was eliminated from the study because of technical and logistical limitations (see chapter 5)

**Interest in chemistry.** The Chemistry Attitudes and Experiences Questionnaire (CAEQ; Dalgety, Coll, & Jones, 2003) was administered (Appendix N). This is a 4 question multiple choice assessment that uses a symmetrical 5 point likert scale ranging from “not true of me at all”, to “very true of me” when asked about how much they are interested in learning about
chemistry. The average score was calculated based on each questions response (1-5) and used in SPSS as a scale item. Though this measure was collected it was not analyzed as part of the statistics for this study.

**Learner preference.** A modified version of the Visualizer Verbalizer Questionnaire (VVQ; Appendix E) (Richardson, 1977) was administered. According to Leutner and Plass (1998) visual and verbal learning preferences should have an influence on learners' cognitive processes in test situations. This data can help answer questions regarding the effect of individual differences on learning. Learners answered 4 questions which assessed their learning preferences when approaching the subject of chemistry. An inter-item consistency was reported at .86.

**Visuo-Spatial abilities.** Spatial abilities were measured using the Card Rotation test S-1 (ETS, 1962, 1975; Appendix B), and the Paper Folding test VZ-2 (ETS, 1962; Appendix F) and visual abilities were tested using the matching pictures test P-3 (ETS, 1962, 1975; Appendix D) via pencil-and-paper (Ekstrom, French, & Harman, 1976). In the S-1, participants determined whether a shape was the rotated version of a target shape or the flipped version. If it was rotated they checked the letter “s” and if it was flipped they checked the letter “d”. In the VZ-2 participants imagined folding a piece of paper into different shapes, punching a hole all the way through the paper and then unfolding it. To the right there were 5 possible choices of what the paper would look like based on where the holes are in the paper. In the P-3 participants matched one picture on the left to a matching picture on the right, but this measure was discontinued during the study because of time restraints. In a factor analysis the S-1 test measured .694 and the VZ-2 test measured .744 with the factor “spatial ability” suggesting high construct validity (Mayer & Massa, 2003). The P-3 was eliminated from the study after a few participants responded in order to shorten the length of the study.
Exclusionary and inclusionary measures

Demographic survey. (Appendix O) This survey was used to collect gender (male, female), age (in years), education level, how many chemistry classes the participant had taken, present mental health conditions and “clean or sober” date of last usage of alcohol. This survey was also used to randomly assign participants to their treatment and to a four digit ID number. Descriptive statistics were calculated for each of the variables. An age above 65 excluded a participant from the study.

Severity of alcohol dependence questionnaire (SADQ-C). A twenty item survey design to test the severity of alcohol dependence within one year of administration (Appendix H). This test has achieved very high test-retest reliability (0.85) on construct, content and criterion reliabilities (Stockwell, Murphy & Hodgson, 1983). A score of less than 16 points on this test qualifies someone as not having a problem with alcohol and does not exclude the participant from the control group. If the participant scores over 16 on this test they are excluded from the study as control participants.

Mini International Neuropsychiatric Interview, Parts I and J (MINI 6.0.0). A short survey used to determine the severity of Alcohol and Drug Abuse (Appendix P). This test scored very high (8.8 to 1.0) on test-retest reliability measures (Lecrubier et al., 1997). A score of two or less on I2, and J2 questions on this test qualifies someone as not having a problem with alcohol or drugs and does not disqualify them as a participant in the control group. Aside from very rare cases, all alcoholics who attend AA meetings and identify themselves as recovering or recovered alcoholics will honestly report their sober date, therefore self-report is sufficient (W, B., 1976, pp. xxiv).
Data Collection

This study initially took 90 minutes to complete and was later altered to take 60 minutes to complete. The timing was strictly adhered to for each section. A manila envelope of paper-and-pencil materials was distributed to each learner containing an information form, a demographic survey, the interest in chemistry survey, the visualizer – verbalizer, spatial tasks, the SADQ and the MINI. Each demographic sheet and envelope had a preselected identification number written on it and a group designation (A or B). To randomize the treatment, the envelopes were randomly placed on the table next to a computer and a participant was instructed to take a seat at whichever computer they wanted to. If there was only one participant, the next available envelope would be administered with the envelopes ordered in alternate fashion with the letters A and B. To begin, an informational sheet (Appendix Q) with a description of the study was read explaining the general purpose and rights of the participants. All participants were given the informational form to keep (signature was waived to protect their anonymity) and asked to fill out the demographic survey.

Upon completion of the demographic survey the learners were instructed to fill out the interest survey and the visualizer – verbalizer. Upon completion of that portion of the paper-and-pencil materials participants were instructed to click on a link that brought them to a login screen on the research website which was the start page of the study. The learners then entered the identification number that was assigned to them on their demographic form and selected the group letter that was assigned to them on the demographic survey. This brought them to the next screen where they were given directions on how to take the Stroop test. This was not timed and participants could move forward at their own pace. After completing the Stroop test, they had Six minutes to complete the seven question pre-test on Chemistry knowledge. After submitting the
pretest they were prompted to wait for the proctor to tell them to press next to get to the next section.

On the next page was the introduction to the learning simulation. This exercise directed the student’s attention to what type of information they would be looking for during the simulation. Once the learners felt they were ready, they entered the simulator. The learners in either of the two conditions learned for 12 minutes. The learning was self-guided and the learners could ask questions about how the simulation worked, but not content based questions. This allowed for any participants who had less computer experience to manipulate the simulations and the website. Once the participants finished the learning portion they moved to the next screen. On the following screen was the cognitive load survey. This asked learners to reflect on the difficulty of using the simulator and how hard it was to learn either theory.

The next screen brought the learner to the posttests which measured their knowledge of gas laws. For reasons that will be explained in the discussion section, during the initial stages of the study data from the posttests were being collected in a Structured Query Language (SQL) database on a “Godaddy” hosted site, but later in the study participants were asked to enter their answers on a word document and save the document to the desktop after completion. Learners were allowed six minutes per section of the test and were prompted before each section to stop when they reached the end of that section. Learners could not go backward into any section. Upon completion a screen appeared thanking the learners for their participation in the study.

Once the computer-based learning was complete, the paper-and-pencil version of the S-1 was administered in six minutes. Next the learners were instructed on how to take the paper-and-pencil version of the VZ-2. This test allotted 3 minutes per section, for a total of 6 minutes. Next
the learners were instructed on how to take the paper-and-pencil version of the P-3. This test allotted 1.5 minutes per section, for a total of 3 minutes.

Next, the computerized WIST test was administered. The computer controls the administration of the test, and prompts the student to continue after each part. The speed and duration of the test is controlled by the participant. This test took approximately 25 minutes to administer and needed specialized software installed on a designated computer in order to run it.
Chapter 4

Results

A quantitative analysis was conducted. The first research question asked, what is the difference between recovering alcoholics and normal control subjects in executive functioning? Hypothesis H1.1 stated; control group participants will score significantly higher on tests of EF. To test for this, ANCOVA was conducted to compare the main effects of EF between groups, Stroop (I) and (S) as continuous dependent variables, group on two levels (alcoholics and control) as a categorical independent variable, and age, gender and education as covariates and the results were reported (see Table 5). To confirm or reject H1.1, a significant between groups difference on the ANCOVA test using interference as the dependent variable, and a mean score of Stroop interference by the alcoholic group would be needed. This test would need to show significant differences between the two groups and that the experimental group scored a higher response time signifying lower levels of EF. Furthermore, this test measured the amount of variance in the Stroop that could be attributed to education, gender and age.

In order to confirm or reject H1.2, that alcoholics with more time sober would score significantly better on tests of EF; first, the experimental group was coded into three groups based on differences using length of sobriety with a cut point of 180 days and another at 5110 days because according to some literature long term abstinence is considered 6 months to 14 years (Fein, Shimotsu, Chu & Barakos, 2009). Then, within the two groups, effects of Stroop (I) and length of sobriety were tested for using one-way ANOVA.

As a secondary measure, a correlation between length of sobriety as a continuous variable and performance on the Stroop (I) test as a continuous variable using a listwise exclusion of the control group was examined. A significant finding on both of the tests, with the regression
showing that there was a positive linear relationship between better performance on the Stroop (I) and length of time sober were needed to confirm H1.2.

Multivariate General Linear Modeling was used to test for main effects and interactions in order to confirm or reject the hypotheses generated from question 2 (H2.1, H2.2). First a two-way ANOVA was conducted. Next, a linear regression was performed to identify variables that had significant contribution to the outcome of the dependent variable and to test for multicollinearity. In order to further understand how learning outcomes were affected by instructional format (more guided instruction vs. less guided instruction) and group (experimental vs. control) MANCOVA ($p < .05$) was conducted with two levels of learning as dependent variables (comprehension and transfer), two groups as independent variables (experimental and control) and significant covariates. A series of ANCOVAs were conducted testing each covariate separately. In order to accept Hypothesis 2.1, that the alcoholic group would perform better on posttests in the more guidance group, the interaction effect of group and treatment on posttests would need to be significant and the mean score of the alcoholic group would have to be higher on posttests when they learned from treatment B then on treatment A. The opposite effect would be needed to confirm Hypothesis H2.2. For that hypothesis, the control group needed to score higher on posttests when learning from treatment A then from treatment B. Data from this analysis was used to answer exploratory questions.
Quantitative Results

Comparison of conditions on demographic and pretest variables. All demographic condition mean scores are presented in Table 2 of the methods section. Descriptive statistics and means were collected for all pre-test variables and reported in Table 3. One-way ANOVA was used to compare all group differences on pretest variables. Missing data fields were removed from the analysis.

A significant difference between the experimental group and the control group in the level of education completed was reported in Table 3. The experimental group averaged a little more than a year less schooling than the control group, but all participants in the study reported at least 12 years of schooling. Although the alcoholic group reported a lower mean score in level of education, that group averaged the level of an associate’s degree (14 years) and neither the control group (\( M = 1.31, SD = 1.11 \)) or the Experimental group (\( M = 1.00, SD = 1.24 \)) showed any significant prior amount of chemistry classes, \( F(1, 74) = 1.26, p = .265 \). As seen in Table 3, results of a one-way ANOVA showed that there was no significant difference in performance on the pretest of chemistry knowledge between the experimental and control groups (\( p > .05 \)). Participants in both groups entered the experiment with a similar degree of comprehension in chemistry, with the experimental group (\( M = 3.77, SD = 1.44 \)) scoring slightly higher than the control group (\( M = 3.26, SD = 1.65 \)). There was a significant difference between the groups in Gender. In the experimental group 20% were female and of the control group 60 % were female. This could have potentially confounded the Stroop test because there have been differences reported in performance between men and women so a possible effect for gender was further investigated in the next section (Elst et al., 2006).
Table 3

*Comparison of Groups on Demographic and Pretest Measures*

<table>
<thead>
<tr>
<th>PRETEST MEASURE</th>
<th>CONDITION</th>
<th>MEAN</th>
<th>STD. DEV.</th>
<th>F</th>
<th>SIG.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prior Knowledge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemistry Pretest (max. 7)</td>
<td>Alcohol</td>
<td>3.77</td>
<td>1.44</td>
<td>2.07</td>
<td>.155</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>3.26</td>
<td>1.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level of Education</td>
<td>Alcohol</td>
<td>13.97</td>
<td>2.29</td>
<td>5.34</td>
<td>.024*</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>15.11</td>
<td>1.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemistry Classes Taken</td>
<td>Alcohol</td>
<td>1</td>
<td>1.24</td>
<td>1.30</td>
<td>.258</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>1.31</td>
<td>1.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning Preferences</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(VVQ)</td>
<td>Alcohol</td>
<td>3.89</td>
<td>.826</td>
<td>2.96</td>
<td>.090</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>3.58</td>
<td>.692</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interest in Chemistry</td>
<td>Alcohol</td>
<td>2.57</td>
<td>1.040</td>
<td></td>
<td>.869</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>2.61</td>
<td>1.237</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive Tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paper Folding Test</td>
<td>Alcohol</td>
<td>8.86</td>
<td>3.24</td>
<td>.187</td>
<td>.667</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>8.5</td>
<td>3.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Card Rotation Test</td>
<td>Alcohol</td>
<td>100.22</td>
<td>33.48</td>
<td></td>
<td>.930</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>99.5</td>
<td>34.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stroop Speed</td>
<td>Alcohol</td>
<td>1096.56</td>
<td>319.6</td>
<td>5.95</td>
<td>.017*</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>937.67</td>
<td>226.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stroop Interference</td>
<td>Alcohol</td>
<td>48.748</td>
<td>10.50</td>
<td></td>
<td>.243</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>51.48</td>
<td>9.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ages (M) (range)</td>
<td>Alcohol</td>
<td>46.68</td>
<td>10.28</td>
<td>5.98</td>
<td>.017*</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>39.63</td>
<td>14.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>Alcohol</td>
<td>.20</td>
<td>.492</td>
<td>15.84</td>
<td>.000*</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>.61</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05

Results of Hypothesis H1.1. Hypothesis 1.1 posited that control group participants would score lower on tests of EF than the experimental alcoholic group. The Stroop (I) test was normally distributed across participants (skewness = 0.78, $SE = 0.28$)(Figure 3.1),
Figure 3.1. Histogram of the Stroop Interference Test Distribution with Normal Curve

and the Stroop (S) test had a positive skewness of 1.23 ($SE = 0.28$) and kurtosis of 2.245 ($SE = 0.55$) (Figure 3.2).
Figure 3.2. Histogram of the Stroop Speed Test Distribution with Normal Curve

There were a few outliers recorded in the long-time sobriety group. ANOVA was conducted with the outliers removed, but that had no results on the significance of the test, therefore, the outliers were left in the analysis because they were accurate measurements of the participant’s performance and not shown to be caused by technological error. Between group differences tested with ANOVA (see Table 3) uncovered a significant main effect for Stroop (S) ($F = 5.95$, $p = .017$), but not for Stroop (I) ($F = 1.38$, $p = .243$). Alcoholics scored significantly slower (worse) on the Stroop (S) test when the means of the two groups were compared, but conversely,
slightly slower (better) on the Stroop (I). Though there was no significant difference between groups on the Stroop (I), the findings were unexpected and therefore further explored in the discussion section.

In this study the mean age between groups was significantly different with the control group approximately six years on average younger than the experimental group (approx. 40 & 47 respectively) (Table 3), but had no participants over the age of 65. One confounding variable commonly found when using the Stroop measurement is age. The population between 40 and 65 is most susceptible to Stroop age effects so a possible effect for age was further investigated in the next section (Elst et al., 2006).

*Examination of Hypothesis H1.1 Results.* Hypothesis 1.1 posited that control group participants would score lower on tests of EF than the experimental alcoholic group. This was partially confirmed in a one-way ANOVA reported in Table 3, as displayed in Table 3 there was a significant effect for Speed, but not for Interference. The experimental group took a significantly longer amount of time to complete the Stroop (S) test than the control group suggesting some impairment, but the Stroop (I) score, while slightly lower (better) for alcoholics was not significantly different. Since Interference was the variable of interest to this study, a correlational analysis between potential confounding variables was conducted and the results presented in Table 4.
Using Pearson’s *r*, a significant correlation was found between age and Speed and a significant negative correlation was found between age and Interference (Table 4). A significant correlation was also found (see Table 4) between age and time sober, but this is of little interest based on the linear nature of the two variables. In order to control for age, and other possible confounding variables ANCOVA was conducted between groups using Stroop (I) as the dependent variable, and age, gender and level of education as covariates. As seen in Table 5, there was a
### Table 5.

*Analysis of Covariance with Stroop (I) as the Dependent Variable by Group*

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>DF</th>
<th>Mean Square</th>
<th>F-Stat</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>224.953</td>
<td>1</td>
<td>214.903</td>
<td>2.486</td>
<td>.119</td>
</tr>
<tr>
<td>Sex</td>
<td>0.091</td>
<td>1</td>
<td>1.490</td>
<td>.017</td>
<td>.896</td>
</tr>
<tr>
<td>Age</td>
<td>573.655</td>
<td>1</td>
<td>808.893</td>
<td>9.359</td>
<td>.003*</td>
</tr>
<tr>
<td>Group</td>
<td>0.342</td>
<td>1</td>
<td>0.342</td>
<td>.110</td>
<td>.741</td>
</tr>
<tr>
<td>Error</td>
<td>5877.194</td>
<td>68</td>
<td>89.889</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>190608.457</td>
<td>73</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. R Squared = .154 (Adjusted R Squared = .105)
b. Significant at the $p < .05$ level

significant main effect for age ($F = 9.359, p = .003$) but there was no significant main effect for group, gender or level of education. Based on the findings of the Stroop (S) test Hypothesis H1.1 was partially accepted in that there was a significant difference between groups in processing speed, but not in the variable of interest, level of interference.

**Results of Hypotheses H2.1.** Hypothesis H1.2 stated that the longer alcoholics were sober the lower they would perform on the Stroop (I). To test for Hypothesis 1.2, a within-group investigation was conducted for both Stroop (S) and Stroop (I). This test was conducted on the results of the Stroop tests both as categorically data and continuous data. Figure 3.3 demonstrates the results for Stroop (I) by Length of Sobriety as categorical data divided into three groups.
When displayed as categorical data in a Box-plot, Figure 3.3, does not demonstrate a linear progression between length of sobriety and performance on the Stroop (I), instead there appears to be a performance curve, with the Long-term group (6 months to 14 years) appearing to have done a bit worse than the Beginners and Old-timers. The boxplot in Figure 3.4 shows an opposite performance curve for the Stroop (S) test where the midrange group scored better.
than the other two groups by taking less time to respond to the exercise. In the Scatterplot in Figure 3.5, which uses continuous data to test whether alcoholics score better on the Stroop (I) the longer they are sober, there does not seem to be a significant effect, but it does appear as if there is a trend in that direction.
The Scatterplot in Figure 3.6 demonstrates a slight linear relationship between time sober and performance on the Stroop (S) with speed slowing over time but also demonstrates that there was no significant change in performance on the Stroop (S) based on length of sobriety. There was no significant correlation in a listwise analysis using continuous data between amount of days sober and Interference \( (r = -0.235, p > .05) \) or Speed tests \( (r = 0.056, p > .05) \) but there was a significant negative correlation between Stroop (I) and Stroop (S) \( (r = -0.294, p < 0.05) \). Therefore, Hypotheses 1.2 was partially rejected because there was some indication that as length of sobriety increased performance on the Stroop (I) got better.
Examination of hypothesis two: Hypothesis H2 was stated in two parts, the first part (H2.1) stated that non-alcoholics would perform better on posttests of learning from a format of instruction that offered less guidance and second, that alcoholics would perform better on posttests of learning from instruction that offered more guidance. Table 6 contains the mean scores of the participants on the pretest and two posttests,
Table 6

*Group by Treatment Means of Pretest and Posttests (Comprehension and Transfer)*

<table>
<thead>
<tr>
<th>Group</th>
<th>Treatment</th>
<th>Pre-test</th>
<th>Comprehension</th>
<th>Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>Treatment A</td>
<td>Mean</td>
<td>3.95</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Std. Deviation</td>
<td>1.56</td>
<td>2.449</td>
</tr>
<tr>
<td>Treatment B</td>
<td>Mean</td>
<td>3.56</td>
<td>5.32</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>18</td>
<td>19</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>1.60</td>
<td>2.583</td>
<td>2.806</td>
</tr>
<tr>
<td>Control</td>
<td>Treatment A</td>
<td>Mean</td>
<td>3.10</td>
<td>4.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>1.56</td>
<td>2.447</td>
<td>2.953</td>
</tr>
<tr>
<td>Treatment B</td>
<td>Mean</td>
<td>3.42</td>
<td>5.06</td>
<td>4.67</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>17</td>
<td>17</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>1.77</td>
<td>2.015</td>
<td>3.039</td>
</tr>
</tbody>
</table>

organized by group and treatment. Inter-rater reliability on the posttest of transfer was ($\alpha = .920$). The groups were fairly well distributed, though the control group that learned from treatment B had a smaller number ($n = 15$) of participants than the other 3 groups. In order to control for group differences in prior knowledge that could have confounded the results, a pretest of chemistry knowledge was given. There was no significant difference between groups in prior knowledge so the groups were considered to have similar prior knowledge. (There was a significant difference between groups in level of education as seen in table 3. It was predicted then, that differences in education would contribute to the variance found in learning outcomes between the two groups since there was little difference in EF and this was tested later in the study.)

**Hypothesis 2.1.** MANCOVA (see Table 7) was then conducted to test hypothesis H2.1, for main effects and interactions, with two levels of learning as dependent variables
### Table 7

*Multivariate Analysis of Covariance for All Main Effects and Interactions*

<table>
<thead>
<tr>
<th>Source</th>
<th>Dependent Variable</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F-Stat</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Card Rotation</td>
<td>Comprehension</td>
<td>4.545</td>
<td>1</td>
<td>4.545</td>
<td>1.223</td>
<td>.275</td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
<td>42.050</td>
<td>1</td>
<td>42.050</td>
<td>6.598</td>
<td>.014*</td>
</tr>
<tr>
<td>Paper Folding</td>
<td>Comprehension</td>
<td>.184</td>
<td>1</td>
<td>.184</td>
<td>.049</td>
<td>.825</td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
<td>3.087</td>
<td>1</td>
<td>3.087</td>
<td>.484</td>
<td>.490</td>
</tr>
<tr>
<td>Visualizer – Verbalizer</td>
<td>Comprehension</td>
<td>4.473</td>
<td>1</td>
<td>4.473</td>
<td>1.203</td>
<td>.279</td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
<td>8.016</td>
<td>1</td>
<td>8.016</td>
<td>1.258</td>
<td>.268</td>
</tr>
<tr>
<td>Pretest of Knowledge</td>
<td>Comprehension</td>
<td>19.319</td>
<td>1</td>
<td>19.319</td>
<td>5.197</td>
<td>.027*</td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
<td>22.505</td>
<td>1</td>
<td>22.505</td>
<td>3.531</td>
<td>.067</td>
</tr>
<tr>
<td>Stroop</td>
<td>Comprehension</td>
<td>2.549</td>
<td>1</td>
<td>2.549</td>
<td>.686</td>
<td>.412</td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
<td>2.340</td>
<td>1</td>
<td>2.340</td>
<td>.367</td>
<td>.548</td>
</tr>
<tr>
<td>Chemistry</td>
<td>Comprehension</td>
<td>17.389</td>
<td>1</td>
<td>17.389</td>
<td>4.677</td>
<td>.036*</td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
<td>20.137</td>
<td>1</td>
<td>20.137</td>
<td>3.160</td>
<td>.082</td>
</tr>
<tr>
<td>Level of Education</td>
<td>Comprehension</td>
<td>22.552</td>
<td>1</td>
<td>22.552</td>
<td>6.066</td>
<td>.018*</td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
<td>26.849</td>
<td>1</td>
<td>26.849</td>
<td>4.213</td>
<td>.046*</td>
</tr>
<tr>
<td>Stroop</td>
<td>Comprehension</td>
<td>8.886</td>
<td>1</td>
<td>8.886</td>
<td>2.390</td>
<td>.129</td>
</tr>
<tr>
<td>Interference</td>
<td>Transfer</td>
<td>.915</td>
<td>1</td>
<td>.915</td>
<td>.144</td>
<td>.707</td>
</tr>
<tr>
<td>Treatment</td>
<td>Comprehension</td>
<td>.738</td>
<td>1</td>
<td>.738</td>
<td>.198</td>
<td>.658</td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
<td>5.831</td>
<td>1</td>
<td>5.831</td>
<td>.915</td>
<td>.344</td>
</tr>
<tr>
<td>Group</td>
<td>Comprehension</td>
<td>20.554</td>
<td>1</td>
<td>20.554</td>
<td>5.529</td>
<td>.023*</td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
<td>15.051</td>
<td>1</td>
<td>15.051</td>
<td>2.362</td>
<td>.131</td>
</tr>
<tr>
<td>Treatment * Group</td>
<td>Comprehension</td>
<td>.204</td>
<td>1</td>
<td>.204</td>
<td>.055</td>
<td>.816</td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
<td>18.450</td>
<td>1</td>
<td>18.450</td>
<td>2.895</td>
<td>.096</td>
</tr>
<tr>
<td>Error</td>
<td>Comprehension</td>
<td>167.298</td>
<td>45</td>
<td>3.718</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
<td>286.788</td>
<td>45</td>
<td>6.373</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a. *p < .05
(comprehension and transfer) and instructional format and group as categorical independent measures with all independent variables included in the model. Multiple analysis of covariance also provided evidence that there was no interaction between group and treatment on the posttests for comprehension or transfer ($F = .055, p = .816$; $F = 2.895, p = .096$, respectively). Based on these findings hypothesis H2.1 was rejected.

Levene’s test of equality for error variances was not significant for Comprehension ($F = 1.003, p = .399$) or for Transfer ($F = .561, p = .643$). The results of the MANCOVA in Table 6 demonstrate that there was a main effect for group on the comprehension posttest ($F = 5.529, p = .023$) but not on the transfer posttest ($F = 2.362, p = .131$) There was also a main effect for level of education on the comprehension posttest ($F = 6.066, p = .018$) and on the transfer posttest ($F = 4.213, p = .046$). In this model the level of education took up a significant amount of the model’s explained variance making it a very good predictor of performance on the posttest of comprehension and transfer posttests.

There was a main effect of the card rotation test on Transfer. The card rotation test was a better predictor of performance on the posttest than level of education and therefore was of great interest, because transfer is a higher level of learning than comprehension (Anderson & Krahtwal, 2001; Bloom, Engelhart, Furst, Hill & Krathwohl, 1956). The relationship of the card rotation test to learning was also explored later in this analysis. There was a main effect for prior knowledge as measured by the pretest for comprehension but not for transfer, and this was to be expected since transfer is a higher level of learning and thus the ability to transfer learning could not be accomplished by prior knowledge alone. There were also main effects on performance on the test of comprehension for the VVQ, chemistry classes taken, and interest in chemistry.
Hypothesis 2.2. Hypothesis 2.2 stated that the control group, who were participants expected to score higher (worse) on tests of inhibition, would perform better on posttests of comprehension and transfer when learning from a more guided format of instruction. The ANCOVA in Table 7 confirms that although there was some evidence that the groups varied in cognitive function, there was little effect if any on learning outcomes from two different types of learning materials. Therefore, Hypothesis H2.2 was rejected.

One problem with H.2 was that it was predicated on hypothesis H1 being confirmed. Since H.1 was not confirmed, the analysis conducted in H.2 could only provide data by “group” differences and not by “Stroop” differences. In order to further explore the role of Stroop on level of guidance, a second ANCOVA was conducted that tested for an interaction effect between Stroop (I) and Treatment (level of guidance). As seen in Table 8, there was no interaction effect between Stroop (I) and treatment. This confirmed that regardless of the results of Hypothesis H.1, Hypothesis H.2 would have been rejected in this study.
Table 8

Multiple Analysis of Variance With Stroop Interference as a Categorical Dependent Variable

<table>
<thead>
<tr>
<th>Source</th>
<th>Dependent Variable</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F-Stat</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>Comprehension</td>
<td>.507</td>
<td>1</td>
<td>.507</td>
<td>.100</td>
<td>.753</td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
<td>.017</td>
<td>1</td>
<td>.017</td>
<td>.002</td>
<td>.964</td>
</tr>
<tr>
<td>Stroop (I)</td>
<td>Comprehension</td>
<td>19.302</td>
<td>1</td>
<td>19.302</td>
<td>3.792</td>
<td>.056</td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
<td>70.293</td>
<td>1</td>
<td>70.293</td>
<td>8.512</td>
<td>.005</td>
</tr>
<tr>
<td>Treatment * Stroop (I)</td>
<td>Comprehension</td>
<td>.612</td>
<td>1</td>
<td>.612</td>
<td>.120</td>
<td>.730</td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
<td>1.470</td>
<td>1</td>
<td>1.470</td>
<td>.178</td>
<td>.674</td>
</tr>
<tr>
<td>Error</td>
<td>Comprehension</td>
<td>320.656</td>
<td>63</td>
<td>5.090</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
<td>520.267</td>
<td>63</td>
<td>8.258</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < .05

Exploratory Questions on Learning, Alcoholics, and Visually Complex Simulations. In order to understand more about learning, alcoholics and levels of guidance in visually complex simulations a series of exploratory analysis were conducted.

Variables that Contributed to Learning. A stepwise linear regression was run to find out which, if any, variables contributed most to learning regardless of group (see Table 9. Included in the regression model were all independent variables except for Interest, VVQ, and pretest. These variables were not included because they were not objective direct measurements. Interest survey and VVQ are subjective reporting methods and while the data...
Table 9

*Stepwise Regression Analysis For Variables Predicting Posttest Comprehension Scores*

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>F-Stat</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SE B</td>
<td>Beta</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Card Rotation</td>
<td>.020</td>
<td>.009</td>
<td>.275*</td>
</tr>
<tr>
<td>2</td>
<td>Card Rotation</td>
<td>.017</td>
<td>.009</td>
<td>.240*</td>
</tr>
<tr>
<td></td>
<td>Chemistry Classes</td>
<td>.504</td>
<td>.250</td>
<td>.239*</td>
</tr>
</tbody>
</table>

*a *p < .05
b. Dependent Variable: Comprehension

was collected from all participants on these measures, it was of more interest to this study to see how direct measurements of performance such as the card rotation task and level of education, would contribute to the model. The data from the self-report measures will be used in a later analysis. As seen in Table 9, the card rotation test related most to performance on the comprehension posttest, but the amount of chemistry classes taken was also was significantly linked to performance. As demonstrated by the stepwise regression the most likely factor related to EF to have contributed to comprehension was the card rotation test (Table 9). A second linear regression model was assembled, this time using transfer as a dependent variable (Table 10).
Table 10

*Stepwise Regression Analysis For Variables Predicting Posttest Transfer Scores*

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>F-Stat</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$B$</td>
<td>$SE B$</td>
<td>Beta</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Card Rotation</td>
<td>0.045</td>
<td>0.011</td>
<td>0.467**</td>
</tr>
<tr>
<td>2</td>
<td>Card Rotation</td>
<td>0.039</td>
<td>0.011</td>
<td>0.407**</td>
</tr>
<tr>
<td></td>
<td>Level of Education</td>
<td>0.371</td>
<td>0.161</td>
<td>0.269*</td>
</tr>
</tbody>
</table>

*a *p < .05, **p < .01
b Dependent Variable: Transfer

In the stepwise regression shown in table 10, the card rotation and the level of education were the best predictors of performance. Card rotation was clearly the best cognitive predictor of knowledge transfer. There was no predictive ability for Stroop (I) or Stroop (S) on either of the posttest measures ($p > .05$) and age played no significant role in predicting dependent measures either.

**Level of Education and Performance.** In a further exploration of the role of education on performance a MANOVA was run using education on three levels (High School, Associates - Bachelors, and Masters) and treatment on two levels (A & B) as between-subjects factors, and comprehension and transfer as dependent variables. This analysis uncovered a main effect for education on transfer and an interaction effect between level of education and treatment on transfer (see Table 11).
<table>
<thead>
<tr>
<th>Source</th>
<th>Dependent Variable</th>
<th>df</th>
<th>Mean Square</th>
<th>F-Stat</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>Comprehension</td>
<td>1</td>
<td>0.092</td>
<td>0.018</td>
<td>.892</td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
<td>1</td>
<td>0.022</td>
<td>0.003</td>
<td>.888</td>
</tr>
<tr>
<td>Education</td>
<td>Comprehension</td>
<td>2</td>
<td>11.651</td>
<td>2.331</td>
<td>.106</td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
<td>2</td>
<td>26.977</td>
<td>3.503</td>
<td>.044*</td>
</tr>
<tr>
<td>Treatment * Education</td>
<td>Comprehension</td>
<td>2</td>
<td>3.615</td>
<td>0.723</td>
<td>.489</td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
<td>2</td>
<td>31.296</td>
<td>4.064</td>
<td>.020*</td>
</tr>
<tr>
<td>Error</td>
<td>Comprehension</td>
<td>62</td>
<td>4.997</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Transfer</td>
<td>68</td>
<td>7.702</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < .05

An analysis of mean scores (see Table 12) showed that a linear progression of performance when learning from treatment A occurred, where on tests of comprehension and transfer those students with more education scored higher. When learning from treatment B the performance was not as linear. Those students with some college did not perform as well as high school level participants when learning from treatment B and they also scored quite a bit lower than their counterparts in the treatment A group. This shows an expertise reversal effect for education, whereas the load reducing measures used in treatment B inhibited learning from those learners with higher prior education and those with higher education learned better from less guidance.
Table 12

*Means of Posttest Scores By Level of Education*

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Level of Education</th>
<th>Comprehension</th>
<th>Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less Guidance</td>
<td>HS</td>
<td>Mean</td>
<td>4.42</td>
</tr>
<tr>
<td>(A)</td>
<td>N</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>2.314</td>
<td>3.239</td>
</tr>
<tr>
<td></td>
<td>Bachelors</td>
<td>Mean</td>
<td>5.7</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>23</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>2.566</td>
<td>2.743</td>
</tr>
<tr>
<td></td>
<td>Grad</td>
<td>Mean</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>1.581</td>
<td>2.236</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>Mean</td>
<td>5.48</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>40</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>2.481</td>
<td>3.15</td>
</tr>
<tr>
<td>More Guidance</td>
<td>HS</td>
<td>Mean</td>
<td>5</td>
</tr>
<tr>
<td>(B)</td>
<td>N</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>2.357</td>
<td>2.134</td>
</tr>
<tr>
<td></td>
<td>Bachelors</td>
<td>Mean</td>
<td>5.21</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>19</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>2.07</td>
<td>3.008</td>
</tr>
<tr>
<td></td>
<td>Grad</td>
<td>Mean</td>
<td>5.43</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>3.101</td>
<td>2.191</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>Mean</td>
<td>5.19</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>36</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>2.303</td>
<td>2.875</td>
</tr>
</tbody>
</table>

As shown in Table 12, learners with higher levels of education (masters level) scored higher than learners with lower levels of education (high school level) on both posttests when learning from treatment A rather than from treatment B. Those with a high school level education scored better on tests of transfer when they learned from treatment B. Students who had completed some
college scored better from treatment A on tests of transfer. These results are consistent with the idea that expert learners learn better from less guidance and novice learners learn better from more guidance
Chapter 5

Discussion

Relevance

This study investigated the role of level of guidance and EF in the design of visually complex science simulations with a sample of recovering alcoholics. It tested whether learners with different levels of EF, primarily inhibition of prepotent responses, would be able to learn better from simulations that offered more guidance or less guidance during learning. Based on the results of Homer and Plass (2014), a further investigation of EF was warranted, because these study showed that junior high school students were not affected by expertise reversal when learning from a visually complex science simulation, contrary to the expected outcome and that there was an interaction between EF and level of guidance. This is of interest because the expertise reversal effect for representation has consistently been reported in the literature for this population when learning from science simulations (Lee, Plass & Homer, 2006). In order to explore these results the researchers investigated whether inhibition of pre-potent responses, as measured by a Stroop Color-Word test, was the mediating factor. This study sought to further investigate the role of EF in learning by studying a normal adult learner population and adult learners who might be expected to suffer from impaired cognitive function in the area of inhibition. Since there is some evidence in the literature that people who suffer from alcoholism as a group might have cognitive deficiencies in EF, specifically in the area of inhibition of prepotent responses, it was decided to use a sample from this population.

The experimental design used in this study and prior studies incorporated a simulation that models the ideal gas laws; which, when the subject is properly understood, enables a person to conceptualize, for example, how tires are inflated, how hot air balloons rise and fall, or why
air bubbles change size underwater. Understanding chemistry is important as general knowledge for adults, and for school age learners building knowledge in the sciences, but also for alcoholics in a rehabilitation setting where being able to build new complex conceptual models allows them to learn important information about alcohol and drug abuse. For example, Zinn et al., (2004) suggested that alcoholics need to learn the difference between the alcohol content in different drinks, and how alcohol level in the blood affects decision making. In this paper’s attempt to discover learner-centered factors that contributed to performance in learning from simulations, some lessons were learned.

Consistent with the literature, there was evidence that prior knowledge, level of education and card rotation skills were significant predictors of learning, as indicated by performance on posttests. The performance on the card rotation test was the best direct predictor of posttest performance in comprehension, and the card rotation test was the best direct predictor of performance on the posttest of transfer. According to Wu and Shah (2004) card rotation tests are good predictors of the ability to learn chemistry, and this study showed that even in a diverse population of adult learners it was a significant predictor of comprehension and higher level transfer of knowledge. Although this was a diverse set of adult learners there was only a significant difference between the two groups on tests of learning based on the amount of education they had, and not based on cognitive factors.

The goal of CLT is to reduce the load in instructional design so that learners can create schema. Schemata are a mental model containing both depictive and descriptive representations of knowledge. According to Schnotz (2005) mental model creation takes place in the visuospatial sketchpad and that sketchpad is fractionated into visual and spatial functions. Spatial functions are more dependent upon the CE in order to manipulate objects while visual ability relies less on
EF and is associated with speed of processing images. If learning chemistry requires higher levels of spatial ability, then the card rotation should be a better predictor than speed of processing. This is consistent with the findings that the Stroop (S) was not a factor in higher level learning while the card rotation was. Lee (2007) used the identical pictures test and the card rotation test as spatial ability tests, but according to this research speed of processing and spatial manipulation might be two different constructs, or at least use different levels of EF in their control. Though it is logical that increased processing speed might contribute to rotational abilities, there needs to be more research on the contribution that speed adds to the equation.

To add to this model, there were significant predictions on learning from the card rotation test but not from the paper folding test, even though there was a significant correlation between the two tests, suggesting that limitations in spatial abilities might have existed that were not important to this learning model. This is evident by the poor performance on the paper folding test for both groups. This paper would then suggest that Mohr and Linden (2005) were correct in their assumption that the subsystems of the VSSP have separate functionality and that one system is responsible for identifying images and the other for rotation and manipulation of visual information. Furthermore, there might be different levels of spatial abilities needed to perform well on varied topics or subject matters. Reconsidering Miyake et al. (2000), if the most important part of learning chemistry is spatial ability then it is likely that there is a different effect on learning between two-dimensional and three-dimensional rotational abilities. In this case levels of inhibition might be related to spatial abilities because it correlated with the card rotation test.

Understanding the role of EF as a control mechanism for spatial abilities is important. According to Homer, Plass and Hayward (2009, pp. 3) the effectiveness of learning materials
“depends on whether learners have sufficient cognitive resources to perceive and process the essential information in dynamic visualizations.” In this study allowing the learner to control or not control the level of interaction between the elements had no effect when regressed against cognitive resources as judged by performance, only prior education and chemistry classes had an effect. Although there was a correlation between performance on the card rotation test and the paper folding test, the paper folding test was beyond the ability of most participants. Only a few participants in either group did well. The paper folding test asks learners to perform three dimensional rotations that necessitates several iterations of an object be held in WM at a single time. Even learners with good spatial manipulation skills did not have a high enough level of skill to perform well on the paper folding. Bridging the gap between the two dimensional and three dimensional models which need to be internally constructed in a mental model should be a design consideration in reducing cognitive load. One way to improve upon the simulation used in this study, then, would be to use 3D computer modeling.

Comparing the present study to the dual-task employed by Gutkin (2010) -- in the 2010 study a secondary visual task that measured the speed of the response did not inhibit learning. Though learning was not significant, it was equal across groups with or without the secondary task. This means that responding to a secondary visual task, such as the screen changing colors, did not interfere with the spatial resources that were being used at the same time. This would suggest that spatial and visual functions are separable functions of the VSSP. If spatial abilities and visual abilities are two separate subsystems that rely differently on the control of EF for their performance, then working memory should not be measured on a capacity level without understanding the contribution of EF. The additivity principle does not apply unless the assumption is that the capacity of the VSSP is limited to the sum of the spatial and visual
subparts. In that case, measuring each of the subparts and using a dual-task to delineate the two by altering tasks would give an additive measure. Future studies of cognitive will be designed that consider the difference relationships between speed and interference. One potential way to do this is to create a secondary, or several different tasks that tap the spatial abilities of a learner during learning. This would isolate the measurement to the spatial processing function of the VSSP.

There were significant cognitive differences between alcoholics and non-alcoholics. The alcoholic group scored significantly slower on the Stroop (S) parts WCI and WNI, but the Stroop (I) WII score did not reflect a significant group difference. The results of the Stroop Speed (S) are consistent with most of the literature in addiction studies in that alcoholics are deficient in processing speed. Of interest is that the alcoholics scored better than the controls in the Interference test. A significant difference in age in the sample could have confounded the Stroop test results, but age correction for the Stroop is applied to participants over the age of 65 and there were no participants over 65 in this study (Elst et al., 2006). Alcoholics were expected to perform better on the Stroop as they gained longer sobriety, but not better than controls. One factor that needs to be considered is that the average age of the alcoholics was 46 and the average amount of sobriety time was 11.6 years, demonstrating that most of the participants stopped at a very young age, many in their middle 20’s to early 30’s. In a study with a similar population, the average age of the alcoholic participant was 46, but the average length of sobriety was only 6 years (Fein et al., 2006). In that study 6 years was considered very long-term sobriety and they found minor spatial differences between controls and experimental groups. There were no studies found that reported lengths of sobriety of this many participants having longer than 20 years.
Urben, Van der Linden, and Barisnikov (2011) suggest that working memory speed and inhibition are two separate constructs. The findings of that study support this discussion. If inhibition of prepotent responses is a separate construct than processing speed, another possible explanation for why there was no significant difference between the groups on Stroop (I), but was on the Stroop (S), is that there might be a relationship between being able to inhibit prepotent responses and being able to stay sober. According to Noel, et al., (2012) the ability to inhibit prepotent responses to control alcohol consumption is a necessity. The thought of drinking is an automatic impulse, and inhibition of prepotent responses controls behaviors. If a thought arises within an alcoholic that makes them think of drinking, without the ability to suppress that thought they will pick up a drink. Therefore, it is possible that only those participants who were able to successfully inhibit responses took part in the study (the rest were out drinking). To begin, one has to make it through the first year without a drink in order to benefit from cognitive recovery and that is not an easy task (Stavro, Pelletier & Potvin, 2003).

In my study all participants stayed sober exclusively by attending meetings of Alcoholics Anonymous. Is it possible that cognitive changes occurred as a result of attending AA meetings? From a social cognitive perspective it is possible that interaction with other people can change cognitive structures (Bandura, 1986). According to Galantar (2014) stories told about one’s past during an AA meeting (known as “qualifying”) are drawn from self-schemas. These self-schemas shape a person’s reaction to social stimuli, such as being conditioned to pick up a drink to cope. In order to stay sober then, one’s self-concept needs to change from thinking they have to drink, to believing they don’t have to drink. One way they do this is by reconstructing their own self-schemas, or alcoholic stories, by listening to other people’s stories and identifying parts that are similar to their own.
The historical stories are made up of episodic self-schemas, which primarily consist of imagery and emotional content (Boyer, 2008). These episodic thoughts come involuntarily and tend to be negative by default, conflicting greatly with current goals and moods. When applying the knowledge of self-schema to alcoholics in recovery, it becomes clear that as alcoholics experience involuntary recall of negative thoughts (schemata), where their only previously known response was to drink, if they can’t inhibit these thoughts they will drink again. Over time, learning how to cope with these thought allows them the skill to refrain from acting in a previously automatic way. One of the problems is that when alcoholics are looking toward the future, they are activating the same neural pathways as when they remember past experiences.

One goal, then, is to retrain the thought process so that previous experiences are now integrated with the proper attributions ingrained through interaction with the AA program. In order for this change to spread across other self-schemas, AA’s practice these principles in all of their affairs, as it states in “Step 12” of the AA program. By learning how to cope with involuntary self-schemas it is possible that recovering alcoholics improve in their ability to inhibit prepotent thoughts as a skill set for survival, or make them more aware of discrepancies causing them to think more about their reaction. This is consistent with the literature. Houben, Nederkoorn, Wiers, and Jansen (2010) attempted to train drinkers how to increase their ability to inhibit responses to the urge to drink. They found that teaching drinkers how to stop a response inhibition for alcohol-related cues decreased alcohol intake.

From an educational perspective, there are circumstances in which alcoholics in recovery take part in learning other than attending meetings and listening to speakers. According to Galantar (2014), “sponsors” guide “newcomers” in the program through empathetic mutuality and mirroring. One way is through the study of their basic text (and the 12 Steps found within),
the book for which the program was named for, “Alcoholics Anonymous.” One question then is how newcomers should be introduced to the 12 steps, whether higher levels of guidance should be offered or a more exploratory method. Based on my research it seems that those with higher levels of education, might benefit from a more exploratory approach while those with less ability to read and comprehend the text, would benefit from more guidance. This is based on the idea that those with higher level of education performed better on tests of transfer. Furthermore, it could be suggested that an expertise reversal effect can occur if those with more experience are not allowed to explore their own meanings of the book. More research is needed on this area.

On the other side of the spectrum, there is a possibility that the positive performance on the Stroop (I) in relation to the length of sobriety many of the participants achieved is due to an underlying, pre-existing ability to inhibit prepotent responses in relation to stopping drinking. This was not measured in this study because in order to do so a longitudinal analysis would have had to be conducted beginning with participants who are in their first week of alcohol rehabilitation and then tracked and compared to those who have succeeded in achieving substantial time sober. One question that would need to be answered is why, if alcoholics have this capacity, why they are not able to use it while they are still drinking. Conversely to the idea that inhibition is acquired through retraining and social interaction, this approach would suggest that there are certain situations which can bring out the effective use of inhibition that already existed. Understanding the dynamic relationship of Stroop (I) in this situation would be beneficial to several areas of cognition and addiction treatment.

Another finding in this paper was that level of education interacted with level of guidance. It is interesting because the participants that reported each level of education were not actively at that level and several had completed their degrees as much as 20 years prior. This
suggests that participating in higher education could possibly predict one’s ability to learn new content and navigate new areas of learning outside of the major area. There were no chemistry majors reported in the participant pool, yet the level of general education achieved predicted performance on chemistry and interacted with level of guidance regardless of EF or whether they were alcoholic. For instance, of all participants, those who had a high school level education scored $M = 3.09$ on transfer when learning with less guidance, but scored $M = 5.38$, respectively when learning with more guidance. When given more guidance these learners achieved $M = 5.38$ on tests of transfer. Those participants who reported having college level education in the past scored $M = 6.05$ on transfer tests when given less guidance and $M = 3.89$ when given more guidance. These findings represent an expertise reversal for level of guidance based solely on past level of education. For those reporting masters level education there was no difference in performance based on guidance with both groups scoring $M = 7$.

What is also interesting about the role of level of education on the ability to transfer knowledge is that prior education did not predict scores on the pretest of chemistry knowledge. If it had it could be assumed that the chemistry was learned in college. It is possible that it can be explained by the fact that learners who are able to succeed in college have learned how to learn better, or apply critical thinking skills to new situations. But why did the expertise reversal effect occur? The expertise reversal effect has been associated with prior knowledge (Kalyuga, Ayres, Chandler, & Sweller, 2003), EF (Homer & Plass, 2009) and representation (Homer & Plass, 2010). It would be of interest to find a common denominator that explains why this is so. A future study might conduct a more in depth investigation in this area. It would have to consider why card rotation, a measurement of spatial ability, clearly predicted performance regardless of any other measure or treatment, but level of education was sensitive to levels of guidance.
Limitations Within the Study

The original design of this study was to take place in a state run alcohol rehabilitation setting where all participants would have had less than 30 days of sobriety. Data would have been collected from participants on a two week rotation (15 entered the rehab and 15 left every two weeks). A repeated measures design would have measured cognition and learning upon entry and exit. The first 30 days is when cognitive deficiency is at its worst according to most research. This might have offered higher variance to the Stroop test. Unfortunately the director of the center had second thoughts about applying for IRB to conduct the study, and the opportunity was lost, so the alternative group was used. Future studies that can capture early measures of alcoholics obtaining sobriety is important for future knowledge on this population. It is also important to note that chemistry was used as a subject matter in this study but future studies should use materials that are contextual to learning about how to stop drinking.

After identifying a new sample of participants, using flyers to gather participants was not very successful for either group, and because there was no reward for participating there was a very slow response. Furthermore, during the study the researcher began working at Rutgers in Newark, New Jersey, making it very difficult to schedule research at the Graduate Center in NYC. It was decided to apply for IRB approval at Rutgers and collect data there as well using the PhD Nursing Lab. One problem with using the Newark site was that many participants did not want to travel to the research location after hours. Using Rutgers during work hours had limitations as well. The first problem was that the study was 90 minutes long and everyone in the school had a one hour lunch break. The second problem was that because the computer lab in the Nursing School is controlled by a centralized IT department, the software for the WIST could not be installed. In order to move forward with the study the WIST test needed to be eliminated. The
time of the test also needed to be less than an hour so the identical pictures test was eliminated as well. This brought the length of the study to 50 minutes and allowed it to be run on any computer. The elimination of these tests did not alter the effectiveness of the study as they were not variables central to confirming hypotheses, though the data from the identical pictures test and the WIST would have been beneficial to answer other questions.

In order to complete the data collection the researcher solicited staff and students from the School of Dental Medicine. This resulted in an increase of participants. Another addition to the study was that the researcher used a portable hotspot and laptops to travel to meet participants in public areas or buildings where AA meetings take place. The laptops worked very well except for a few exceptions. During the study a few technical conditions arose that changed the way posttest data was collected. The posttest data was supposed to be collected via the web-based portal and worked for several participants until at some time during the study and for reasons never understood the posttest data randomly stopped registering in the database and some posttest data was lost. In order to continue collecting data a word document template was created for the posttests. The procedure changed for administering the test so that the researcher stopped the participants as they were ready to take the posttest and had them switch to the word document and fill the answers in there. The word document answers were saved and copied into a google drive folder which held all research materials making it accessible from a central computer.

Limitations and Future Directions

The cross-section of participants that responded were diverse, from different professions and backgrounds with varied education and history, but were drawn from a convenient group of participants. This could diminish the generalizability of the study. Regardless, the picture of an
alcoholic who is laying in the street with a bottle certainly did not apply to the 40 experimental participants who took part in this study. As shown earlier, many had Masters and Bachelors degrees and hadn’t drank alcohol in over 20 years, one or two with over 30 years. The overall sample size of 76 had a strong enough effect to report significant findings on spatial abilities and education.

The results pointed to a test of spatial measurement that was highly reliable, card rotation and level of education as a predictor for level of guidance. One question that still needs to be addressed is how an assessment of a learner can aid in the ad-hoc presentation of the proper learning environment for alcoholics and controls. The potential of computer-based learning environments is that instruction can be differentiated between different type learners. The goal is to develop a measurement of cognition that could predict which learning scenario a learner would perform best from, using a cognitive measurement that has a high predictive value, and then present the learner with materials that would maximize learning.

One of the primary goals of this paper was to inform researchers on how to create learning materials that can help alcoholics get and stay sober. There is a dearth of literature that discusses instructional design for this population. Most literature is geared toward patients in their first 30 days, because this is when they are in rehabilitation centers. Based on the findings in this paper there is little evidence that changing levels of guidance or decreasing visual complexity will help designers create effective materials for alcoholics in their first 30 days of recovery without special consideration for their cognitive deficits, but these findings do suggest alcoholics in recovery that have sustained sobriety for a year or for several years do not need special considerations. Therefore the focus of future research in this area should be conducted in
rehabilitation settings and in Intensive Outpatient Programs (IOP) and potentially, cognitive remediation.

It would be of importance to develop multimedia training that can promote learning in areas specific to alcohol treatment or that can retrain cognitive functions in the area of inhibition. The simulations worked very well in an experimental setting because the subject matter was unique to the learners, but presenting instruction geared toward recovery would be more beneficial. There are several ways that instruction can be designed. One way is by using video game-based solutions. They are desirable because human-like characters can be used in order to activate mirror neurons. This form of Social Cognition is useful because it can help patients change self-schemas by modeling positive behaviors and ideals in a more life-like setting. Game-based training also allows for players to succeed on multiple levels. This is not a replacement for actual attendance at AA meetings, but can provide initial support during early sobriety in rehabilitation centers. Success of the cognitive retraining can be measured via a simple test such as the Stroop, but can also measure responses to involuntary thoughts as those thoughts are reshaped into healthier ones. It is possible that these changes can only occur after many years of sobriety. Unfortunately the literature suggests that many newcomers can’t afford to wait because they pick up a drink within the first year. Therefore instruction can be delivered with varied levels of guidance, with more direct guidance to novice learners and a removal of scaffolding as the learner progresses. Clearly there is much work needed in this area but this paper has been able to add some validity to recent findings in EF and alcoholics.

The inclusion of this population to multimedia learning literature has proven somewhat beneficial. It is the first study to explore the interaction between recovering alcoholics, EF and visually complex science simulations. Even thought the findings did not confirm the hypotheses
the study was a success. On a theoretical level the paper has helped to confirm ideas that
cognitive processing speed is independent of inhibition of prepotent responses as measured by a
Stroop test and that there is potential for being able to improve EF skills through social
interaction. This was demonstrated when alcoholics scored higher on speed tests but lower on
inhibition tests. On this same theoretical level it corroborated the idea that the ability to rotate
two dimensional objects has a direct relationship to learning chemistry. On a practical level it
was found that the level of learning that is achieved affects a person’s ability later in life, even
many years after that education was obtained, to learn from novel learning environments. This
was demonstrated when the level of guidance interacted with level of education.

These findings have an impact on the development of Cognitive Load Theory. Cognitive
load theory is concerned with answering questions pertaining to level of guidance in learning and
this paper has demonstrated that there are still more questions to answer, but that the theory is
vital to instructional design research based on the findings. Primarily, is that it offers a possible
explanation for why the card rotation test is a good predictor of learning across platforms and
that has to do with intrinsic complexity and the ability to rotate two dimensional objects. If the
subject matter is intrinsically complex, the ability to rotate objects becomes more important. This
might be independent of processing speed and more dependent on the ability to focus on the
subject matter in order to create and hold a mental model in working memory while building
schema. This simulation used high complexity because that is when load reducing designs
become most crucial to learning. Clearly choosing the proper load reducing measures are highly
important when teaching to different levels of students.
APPENDIXES

Appendix A: Stroop Color-Word Test

STROOP TEST

The following material consists of three parts.

Please read instructions for each part carefully.

To begin, click the START button below.

Start

Version 1.4a  stroop_upload.cfm  j
Part I

You will be shown a word and two colored boxes.

As Quickly As Possible

press the RIGHT ARROW or LEFT ARROW to indicate which BOX best indicates the MEANING of the word

When you are ready, press the RIGHT ARROW to continue
Part II

You will be shown a word and two colored boxes

As Quickly As Possible

press the RIGHT ARROW or LEFT ARROW to indicate which BOX best indicates the MEANING of the word

When you are ready, press the RIGHT ARROW to continue
Part III

You will be shown a word and two colored boxes

As Quickly As Possible

press the RIGHT ARROW or LEFT ARROW to indicate which BOX best indicates the MEANING of the word

When you are ready, press the RIGHT ARROW to continue
Appendix B: Card Rotation Test (S-1)

CARD ROTATIONS TEST — S-1 (Rev.)

This is a test of your ability to see differences in figures. Look at the 5 triangle-shaped cards drawn below.

All of these drawings are of the same card, which has been slid around into different positions on the page.

Now look at the 2 cards below:

These two cards are not alike. The first cannot be made to look like the second by sliding it around on the page. It would have to be flipped over or made differently.
Appendix C: Working Memory, Inhibition and Switching Test (WIST)
Appendix D: Identical Pictures Test

Name ___________________________

IDENTICAL PICTURES TEST -- P-3

How fast can you match a given object? This is a test of your ability to pick the correct object quickly. At the left of each row is an object. To the right are five test objects, one of which matches the object at the left. Look at the example below:

[Diagram of objects]

The third test object has been marked by blackening the space below it, because it is the same as the object at the left.
Appendix E: Visualizer – Verbalizer Survey (VVQ)

1. In a learning situation sometimes information is presented verbally (e.g., with printed or spoken words) and sometimes information is presented visually (e.g., with labeled illustrations, graphs, or animations). Which type of information do you usually prefer?

<table>
<thead>
<tr>
<th>Strongly more verbal than visual</th>
<th>Moderately more verbal than visual</th>
<th>Slightly more verbal than visual</th>
<th>Equally verbal and visual</th>
<th>Slightly more visual than verbal</th>
<th>Moderately more visual than verbal</th>
<th>Strongly more visual than verbal</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

2. I am good at learning chemistry from labeled pictures, illustration, graphs, maps and animations

<table>
<thead>
<tr>
<th>Strongly agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

3. I am good at learning chemistry from printed text.

<table>
<thead>
<tr>
<th>Strongly agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

4. I can solve chemistry problems better when they are presented as

<table>
<thead>
<tr>
<th>Textual information only</th>
<th>Mostly text with some pictures</th>
<th>Equal amount of text and pictures</th>
<th>Mostly pictures with some text</th>
<th>Pictures only</th>
</tr>
</thead>
</table>
Appendix F: Paper Folding Test (VZ-2)

Name

PAPER FOLDING TEST — VZ-2

In this test you are to imagine the folding and unfolding of pieces of paper. In each problem in the test there are some figures drawn at the left of a vertical line and there are others drawn at the right of the line. The figures at the left represent a square piece of paper being folded, and the last of these figures has one or two small circles drawn on it to show where the paper has been punched. Each hole is punched through all the thicknesses of paper at that point. One of the five figures at the right of the vertical line shows where the holes will be when the paper is completely unfolded. You are to decide which one of these figures is correct and draw an X through that figure.

Now try the sample problem below. (In this problem only one hole was punched in the folded paper.)

```plaintext
A  B  C  D  E

```

The correct answer to the sample problem above is C and so it should have been marked with an X. The figures below show how the paper was folded and why C is the correct answer.
Appendix G: Solicitation Posters

Solicitation Flyer Control Group Version

PARTICIPANTS WANTED

My name is Jeff and I am a doctoral participant in Educational Psychology. I am recruiting participants to take part in a study on memory and learning.

To participate, you must be 18 years old and less than 65. The study will take approximately 90 minutes to complete. In that time, you will be asked to complete a survey assessing memory ability, as well as participate in a learning experiment on the computer. Your participation will be anonymous and your results will be strictly confidential. Information gathered in this study will help researchers to better understand how people learn and how to create resources to aid learning.

When: Weekday Evenings

Where: The Graduate Center, CUNY. 365 5th avenue

Who: Any person between the ages of 18 and 65

Researcher: Jeff Gutkin, Graduate Center - CUNY; CREATE lab, NYU

To sign up or for more information, email Jeff at jgtn1@aol.com; or call 917-560-5656.

Thank You! Your participation in this study could help us to help develop computer-based learning materials.

This flyer has been approved for posting by the IRB board at the Graduate Center – CUNY
PARTICIPANTS WANTED

Take part in a study on memory and learning!

To participate, you must be a recovering alcoholic and 18 years old or older. The study will take approximately 90 minutes to complete. In that time, you will be asked to complete a survey assessing memory ability, as well as participate in a learning experiment on the computer. Your participation will be anonymous and your results will be strictly confidential. Information gathered in this study will help researchers to better understand how alcoholics learn and how to create resources to aid learning.

**When:** Tuesday Evenings, 7:00 PM

**Where:** The Graduate Center, CUNY. 365 5th avenue

**Who:** Alcoholics who are abstaining or sober

**Researcher:** Jeff Gutkin, PhD Candidate in Educational Psychology

The Graduate Center, CUNY

To sign up or for more information, email Jeff at jgtnl1@aol.com; or call 917-560-5656.

Thank You! Your participation in this study could help us to help other alcoholics.
Appendix F

SEVERITY OF ALCOHOL DEPENDENCE QUESTIONNAIRE (SADQ-C)

NAME________________________AGE______No.________

DATE:

Please recall a typical period of heavy drinking in the last 6 months.

When was this? Month:____________________Year:________________________

Please answer all the following questions about your drinking by circling your most appropriate response.

During that period of heavy drinking

1. The day after drinking alcohol, I woke up feeling sweaty.
   ALMOST NEVER SOMETIMES OFTEN NEARLY ALWAYS

2. The day after drinking alcohol, my hands shook first thing in the morning.
   ALMOST NEVER SOMETIMES OFTEN NEARLY ALWAYS

3. The day after drinking alcohol, my whole body shook violently first thing in the morning if I didn't have a drink.
   ALMOST NEVER SOMETIMES OFTEN NEARLY ALWAYS

4. The day after drinking alcohol, I woke up absolutely drenched in sweat.
   ALMOST NEVER SOMETIMES OFTEN NEARLY ALWAYS

5. The day after drinking alcohol, I dread waking up in the morning.
   ALMOST NEVER SOMETIMES OFTEN NEARLY ALWAYS

6. The day after drinking alcohol, I was frightened of meeting people first thing in the morning.
   ALMOST NEVER SOMETIMES OFTEN NEARLY ALWAYS

7. The day after drinking alcohol, I felt at the edge of despair when I awoke.
   ALMOST NEVER SOMETIMES OFTEN NEARLY ALWAYS

8. The day after drinking alcohol, I felt very frightened when I awoke.
   ALMOST NEVER SOMETIMES OFTEN NEARLY ALWAYS

9. The day after drinking alcohol, I liked to have an alcoholic drink in the morning.
   ALMOST NEVER SOMETIMES OFTEN NEARLY ALWAYS

10. The day after drinking alcohol, I always gulped my first few alcoholic drinks down as quickly as possible.
    ALMOST NEVER SOMETIMES OFTEN NEARLY ALWAYS

11. The day after drinking alcohol, I drank more alcohol to get rid of the shakes.
    ALMOST NEVER SOMETIMES OFTEN NEARLY ALWAYS

---

NOTES ON THE USE OF THE SADQ

The Severity of Alcohol Dependence Questionnaire was developed by the Addiction Research Unit at the Maudsley Hospital. It is a measure of the severity of dependence. The AUDIT questionnaire, by contrast, is used to assess whether or not there is a problem with dependence.

The SADQ questions cover the following aspects of dependency syndrome:
- physical withdrawal symptoms
- affective withdrawal symptoms
- relief drinking
- frequency of alcohol consumption
- speed of onset of withdrawal symptoms.

Scoring
Answers to each question are rated on a four-point scale:

Almost never - 0
Sometimes 1
Often 2
Nearly always 3

A score of 31 or higher indicates "severe alcohol dependence".
A score of 16 - 30 indicates "moderate dependence"
A score of below 16 usually indicates only a mild physical dependency.
A chlordiazepoxide detoxification regime is usually indicated for someone who scores 16 or over.

It is essential to take account of the amount of alcohol that the patient reports drinking prior to admission as well as the result of the SADQ.

There is no correlation between the SADQ and such parameters as the MCV or GGT.
Appendix I: Knowledge of Chemistry Posttest for Comprehension

Post-Test

1. If the temperature of a gas is increased, what happens to the gas particles?
   - The speed of the particles increases
   - The speed of the particles decreases
   - The number of particles increases
   - The size of the particles expands
   - Don't know

2. What causes the pressure produced by a gas in a container?
   - The gas particles collide with each other within the container
   - The gas particles collide with the walls of the container
   - The gas particles expand their size within the container
   - The gas particles fall to the bottom of the container and press on it
   - Don't know

3. If the temperature is held constant, what happens to the pressure of a gas sample if the volume of its container is decreased?
   - The gas pressure increases
   - The gas pressure decreases
   - The gas pressure does not change
   - The gas pressure increases first, then slowly decreases to the previous level
   - Don't know

4. If the pressure remains constant, what happens to the volume of a gas sample if its temperature is decreased?
   - The gas volume increases
   - The gas volume decreases
   - The gas volume does not change
   - The gas volume decreases first, but then bounces back to the previous level
   - Don't know

Submit
5. If you are given a gas sample in a container with a fixed volume, what would you change in order to increase the gas pressure?

- Decrease the gas temperature
- Do not change the temperature
- The gas pressure cannot be changed
- Increase the gas temperature
- Don't know

6. If the temperature is held constant, what change in the volume of a gas container will decrease the gas pressure from 10 atm to 5 atm?

- Double the volume of the container
- Halve the volume of the container
- Increase the volume of the container by three times
- Decrease the volume of the container to a quarter of its original amount
- Don't know

7. Consider a gas sample that occupies a volume of 10 liters at a temperature of 300 K. What change would increase the gas volume to 15 liters? Assume that the pressure is constant.

- Increase the temperature to 500 K
- Increase the temperature to 450 K
- Decrease the temperature to 150 K
- Increase the temperature to 400 K
- Don't know

8. Consider a gas sample with a volume of 12 liters and a pressure of 0.5 atmospheres. If the pressure was increased to 2 atmospheres, how would that affect the volume of the container? Assume that the gas remains at constant temperature.

- Decrease the volume to 2 liters
- Decrease the volume to 6 liters
- Decrease the volume to 4 liters
- Decrease the volume to 3 liters
- Don't know
9. If the temperature is constant, which of the following graphs represents the relationship between pressure and volume of a gas sample?
10. If the pressure is constant, which of the following graphs represents the relationship between temperature and volume of a gas sample?
Appendix J: Knowledge of Chemistry Posttest of Transfer

11. Why does a toy balloon burst when it rises in the air?

12. Explain how gas laws apply to a person's breathing.

13. Suppose you inflate your car tires to a recommended pressure. What would happen to the pressure if you travel to a much colder place? Explain your answer.

14. How do the bubbles of scuba divers change in size as they travel from deep water to the surface? Explain your answer.

15. Imagine you are flying a hot air balloon and the height above the ground becomes too low. What would you do? Explain the reasons behind your actions.

16. Mountain hikers carry oxygen tanks when hiking at high altitudes. Imagine that the pressure in the tank becomes too low. Suggest two different ways by which you could increase the oxygen pressure in the tank?
Appendix K: Scoring Rubric for Transfer Posttest

QUESTION 11: Why does a toy balloon burst when it rises in the air?

Answer: As the balloon rises the external pressure pushing against the balloon decreases because of the thinner air in the atmosphere. The particles in the balloon are at a specific internal pressure that pushes against the balloon. As the pressure of the atmosphere decreases the balloon will get bigger until it bursts.

*0 points

as volume increases pressure decreases (incorrect causality)

change in pressure

*examples of 1 point:

Correct mention of volume (but not pressure)

the higher it goes the larger it gets (no mention of pressure)

Balloons get bigger as they rise

pressure is lower, higher (in the sky)

Wrong 11, BUT correct mention of pressure

pressure increases if lower OR pressure decreases if higher (but if includes an incorrect conclusion, 0 pts, eg: less pressure in sky so balloons get smaller)

*examples of 2 points:

Correct, complete explanation:

when they are lower in the sky there is so much pressure on the balloons they will not burst and when they start getting higher its opposite there won't be as much pressure and volume will expand making it burst.

pressure of the air will lessen so that the volume of the balloon will increase

less pressure higher up, thus bigger volume and balloon bursts.

Correct 11 AND correct statement about effect of pressure on volume, even if no reference to direction in the water

pressure of the air will lessen so that the volume of the balloon will increase

the volume of the size of the balloon will increase as pressure decreases
QUESTION 12: Explain how gas laws apply to a person’s breathing

*Answer:* While inhaling, the lungs enlarge due to muscular movement and so internal pressure decreases and the volume increases. Since product of pressure and volume is constant as pressure in lungs decreases volume increases and air from the outside gushes in. When exhaling, external pressure acts on the lungs, decreases volume and pushes air out.

*0 points

No correct causality

As volume increases pressure decreases (incorrect causality)

Reference to temperature

*examples of 1 point:

Correct mention of volume (but not pressure)

Correct mention of pressure (but not volume)

Pressure is lower or higher

Pressure in the lungs is greater

Volume increases

*examples of 2 points:

Correct, complete explanation:

Correct statement about effect of pressure on volume

Pressure of the lungs will decrease so that the volume of the lungs will increase
QUESTION 13: Suppose you inflate your car tires to a recommended pressure. What would happen to the pressure if you travel to a much colder place? 13a: Explain your answer.

Answer: The pressure would decrease. Why? The effect of the cold on the air molecules inside the tire would cause them to slow down creating fewer collisions with the tire wall causing the pressure to drop.

*0 points
Wrong answer
Volume decreases
Pressure increases

*examples of 1 point:
Correct answer to 13:
Pressure decreases

*examples of 2 points:
Correct 13 and:
Correct mention of pressure
A drop in temperature would have a direct cause on the pressure to decrease because the molecules would slow down and collide less with the tire wall

:
QUESTION 14: How do the bubbles of scuba divers change in size as they travel from deep water to the surface? 14a: Explain your answer

*Answer*: The bubbles would get larger as they near the surface. Why? The lower pressure of the water would allow the molecules inside the bubbles to raise the internal pressure and increase the volume

*0 points*

Get smaller

*examples of 1 point:*

Correct answer to 14:
Get bigger (but no reference to pressure or volume)

*examples of 2 points:*

Correct 14 and;
Correct statement about effect of pressure
Less pressure from the water above
statement about volume (no mention of pressure)
Get bigger, pressure drops
As the bubbles rise the volume increases because of less pressure
QUESTION 15. Imagine you are flying a hot air balloon and the height above the ground becomes too low. What would you do? Explain the reasons behind your actions.

*Answer:* Turn up the heat. Why? Heating the air on the balloon will cause the molecules to move faster causing more collisions with the balloon. This will increase the pressure and the volume of the balloon and cause the air inside to be less dense and it will become lighter and move further away from the ground.

*0 points*

Lower the heat

*examples of 1 point:*

Correct answer to 15:
make it hotter, turn up flame (but no reference to pressure or volume)

*examples of 2 points:*

Correct 15 and;
Correct statement about effect of pressure
Correct statement about volume
Turn up the heat to create more pressure in the balloon
Turn up the heat to create more volume in the balloon
As the heat rises the volume increases because of more internal pressure
QUESTION 16: Mountain hikers carry oxygen tanks when hiking at high altitudes. Imagine that the pressure in the tank becomes too low. Suggest two different ways by which you could increase the oxygen pressure in the tank

Two possible answers: Heat the tank, crush the tank,

*0 points

Go back down

*examples of 1 point:

One correct answer

*examples of 2 points:

Two correct answers
Appendix L: Knowledge of Chemistry Pretest

Knowledge Pre-test

1. When a gas sample in a closed container is heated, what happens to the speed of the gas particles?
   - The speed remains the same
   - The speed increases
   - The speed decreases
   - Gas particles don't move
   - Don't know

2. A gas in a closed container has an internal pressure, which we can measure. What feature of kinetic theory is used to explain this internal pressure?
   - Gas particles moving around all over the container
   - Gas particles colliding with the walls of the container
   - Gas particles colliding with each other
   - Gas particles colliding with container walls and each other
   - Don't know

Submit
3. A balloon filled with helium rises when released in the air. Which of the following diagrams best represents the distribution of helium particles inside the balloon?

[Diagram of balloons with varying distributions of dots, indicating helium particles]

Don't know

Submit
The diagram below shows three containers with the same volume and at the same temperature. However, each container has a different number of gas particles.

4. In which container are the particles moving with the fastest speed?
   - A
   - B
   - C
   - All the same
   - Don't know

5. Which container has most collisions between the particles and the walls?
   - A
   - B
   - C
   - All the same
   - Don't know

6. Which container has the lowest internal pressure?
   - A
   - B
   - C
   - All the same
   - Don't know
7. When you climb to higher altitude, like the top of a mountain, we say that the air becomes "thinner." This means that there are fewer air particles in a particular volume (e.g. 1 liter). What effect does this have on the pressure exerted by the air?

- The pressure remains the same
- The pressure is higher
- The pressure is lower
- It is not possible to determine the pressure
- Don't know
# Appendix M: Cognitive Load Survey

## Cognitive Load Survey

1. How easy or difficult was it to work with this simulation?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Extremely easy</strong></td>
<td><strong>Neither easy nor difficult</strong></td>
<td><strong>Extremely difficult</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. How hard did you have to think to understand the Gas Laws in this simulation?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Extremely easy</strong></td>
<td><strong>Neither easy nor difficult</strong></td>
<td><strong>Extremely difficult</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix N: Interest in Chemistry Survey

Interest Survey

1. I have a general interest in science.  
   Not true of me at all  
   Sometimes true  
   Very true of me

2. I think chemistry is an interesting subject.  
   Not true of me at all  
   Sometimes true  
   Very true of me

3. I would like to learn more about chemistry.  
   Not true of me at all  
   Sometimes true  
   Very true of me

4. I find chemistry easy.  
   Not true of me at all  
   Sometimes true  
   Very true of me

Submit
Appendix O: Demographics Survey

Demographics Survey

Please circle, or fill in the answers to the questions below:

Male or Female (circle one)  M  F

How old are you? ______

Highest level of education completed: ________________

How many days/months/years has it been since your last drink or substance use? ______

Are you an alcoholic in recovery?  Y  N.  If “Yes”, At what age did you first start drinking? ___

What age did you first try to stop? _____ What age did you stop drinking? _____

Have you ever been diagnosed with a mental illness such as bipolar, depression, or schizophrenia?  Y  N

If “Yes”, (1) what is your diagnosis and __________(2) what, if any, medications you are taking to treat that disorder? list__________________________________________________________

Have you ever taken chemistry classes in high school or college?  Y  N  If yes, how many? __________

The letter circled will be the group you are assigned to.  A  B

ID Number __________
## Appendix P: MINI 6.0.0

### I. ALCOHOL DEPENDENCE / ABUSE

(♦ means: go to diagnostic boxes, circle NO in both and move to the next module)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td><strong>In the past 12 months</strong>: have you had 3 or more alcoholic drinks, - within a 3-hour period, - on 3 or more occasions?</td>
</tr>
</tbody>
</table>

12 | **In the past 12 months:**  
|   | a. Did you need to drink a lot more in order to get the same effect that you got when you first started drinking or did you get much less effect with continued use of the same amount? | NO | YES |
|   | b. When you cut down on drinking did your hands shake, did you sweat or feel agitated? Did you drink to avoid these symptoms (for example, “the shakes”, sweating or agitation) or to avoid being hungover? If yes to any, code YES. | NO | YES |
|   | c. During the times when you drank alcohol, did you end up drinking more than you planned when you started? | NO | YES |
|   | d. Have you tried to reduce or stop drinking alcohol but failed? | NO | YES |
|   | e. On the days that you drank, did you spend substantial time in obtaining alcohol, drinking, or in recovering from the effects of alcohol? | NO | YES |
|   | f. Did you spend less time working, enjoying hobbies, or being with others because of your drinking? | NO | YES |
|   | g. If your drinking caused you health or mental problems, did you still keep on drinking? | NO | YES |

Are 3 or more 12 answers coded YES?  
* If YES, skip 13 questions and go to next module. "Dependence preempts abuse" in DSM IV TR.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
</table>
| 13 | **In the past 12 months:**  
|   | a. Have you been intoxicated, high, or hungover more than once when you had other responsibilities at school, at work, or at home? Did this cause any problems?  
(If yes only if this caused problems.) | NO | YES |
|   | b. Were you intoxicated more than once in any situation where you were physically at risk, for example, driving a car, riding a motorbike, using machinery, boating, etc.? | NO | YES |
|   | c. Did you have legal problems more than once because of your drinking, for example, an arrest or disorderly conduct? | NO | YES |
|   | d. If your drinking caused problems with your family or other people, did you still keep on drinking? | NO | YES |
ARE 1 OR MORE 13 ANSWERS CODED YES?

ALCOHOL ABUSE
CURRENT

NO YES
### J. SUBSTANCE DEPENDENCE / ABUSE (NON-ALCOHOL)

(⇒ MEANS: GO TO THE DIAGNOSTIC BOXES, CIRCLE NO IN ALL DIAGNOSTIC BOXES, AND MOVE TO THE NEXT MODULE)

<table>
<thead>
<tr>
<th>J1</th>
<th>In the past 12 months, did you take any of these drugs more than once, to get high, to feel elated, to get &quot;a buzz&quot; or to change your mood?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NO</td>
</tr>
</tbody>
</table>

**CIRCLE EACH DRUG TAKEN:**

- **Stimulants:** amphetamines, "speed", crystal meth, "crank", "rush", Dextedrine, Ritalin, diet pills.
- **Cocaine:** snorting, IV, freebase, crack, "speedball".
- **Narcotics:** heroin, morphine, Dilaudid, opium, Demerol, methadone, Darvon, codeine, Percodan, Vicodin, OxyContin.
- **Hallucinogens:** LSD ("acid"), mescaline, peyote, psilocybin, STP, "shrooms", "ecstasy", MDA, MDMA.
- **Phencyclidine:** PCP ("Angel Dust", "PeeCe Pill", "Tranq", "Hog"), or ketamine ("Special K").
- **Inhalants:** "glue", ethyl chloride, "rush", nitrous oxide ("laughing gas"), amyl or butyl nitrate ("poppers").
- **Cannabis:** marijuana, hashish ("hash"), THC, "pot", "grass", "weed", "reefer".
- **Tranquilizers:** Quaalude, Seconal ("reds"), Valium, Xanax, Librium, Ativan, Dalmane, Halcion, barbiturates, Miltown, GHB, Roofinol, "Roofies".
- **Miscellaneous:** steroids, nonprescription sleep or diet pills. Cough Medicine? Any others?

**SPECIFY THE MOST USED DRUGS:** __________________________

**WHICH DRUG(S) CAUSE THE BIGGEST PROBLEMS:** __________________________

**FIRST EXAMINE THE DRUG(S) CAUSING THE BIGGEST PROBLEMS AND MOST LIKELY TO MEET DEPENDENCE / ABUSE CRITERIA.**

**IF MEETS CRITERIA FOR ABUSE OR DEPENDENCE, SKIP TO THE NEXT MODULE. OTHERWISE, EXAMINE THE NEXT MOST PROBLEMATIC DRUG.

### J2 Considering your use of [NAME THE DRUG / DRUG CLASS SELECTED], in the past 12 months:

- **a** Have you found that you needed to use much more [NAME OF DRUG / DRUG CLASS SELECTED] to get the same effect that you did when you first started taking it?  
  NO YES

- **b** When you reduced or stopped using [NAME OF DRUG / DRUG CLASS SELECTED], did you have withdrawal symptoms (aches, shaking, fever, weakness, diarrhea, nausea, sweating, heart pounding, difficulty sleeping, or feeling agitated, anxious, irritable, or depressed)? Did you use any drug(s) to keep yourself from getting sick (withdrawal symptoms) or so that you would feel better?  
  IF YES TO EITHER, CODE YES.

- **c** Have you often found that when you used [NAME OF DRUG / DRUG CLASS SELECTED], you ended up taking more than you thought you would?  
  NO YES

- **d** Have you tried to reduce or stop taking [NAME OF DRUG / DRUG CLASS SELECTED] but failed?  
  NO YES

- **e** On the days that you used [NAME OF DRUG / DRUG CLASS SELECTED], did you spend substantial time (>2 hours), obtaining, using or in recovering from the drug, or thinking about the drug?  
  NO YES

- **f** Did you spend less time working, enjoying hobbies, or being with family or friends because of your drug use?  
  NO YES

- **g** If [NAME OF DRUG / DRUG CLASS SELECTED] caused you health or mental problems, did you still keep on using it?  
  NO YES

M.I.N.I. 6.0.0 (October 1, 2009) 18
Are 3 or more 12 answers coded yes?
Specify drug(s): __________________________

* if yes, skip J3 questions, move to next disorder.
"dependence preempts abuse" in DSM IV TR.

Considering your use of (name the drug class selected), in the past 12 months:

<table>
<thead>
<tr>
<th></th>
<th>NO</th>
<th>YES</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NO</th>
<th>YES</th>
</tr>
</thead>
</table>

Are 1 or more J3 answers coded yes?
Specify drug(s): __________________________
Appendix Q: Informational Sheet

Informational Sheet

You are being invited to take part in a research study about the effectiveness of various multimedia elements on learning. The study is being conducted by Jeff Gutkin, a doctoral student in Educational Psychology at the Graduate Center, City University of New York. You are eligible to participate in the study if you are over 18 years of age. Approximately 120 participants will take part in this study.

If you agree to participate in the study, you will be asked to fill out a brief survey asking about your past drinking history, and assessing your spatial and verbal abilities. You will also participate in a computer based learning module about a concept in science, such as kinetic theory. The study session will last approximately 90 minutes.

Participation in this study is voluntary. You may refuse to participate or withdraw from the study at any time without penalty. Additionally, you have the right to skip or not answer any questions you prefer not to answer.

There are no known risks associated with your participation in this research beyond those of everyday life. The direct benefits to you are minimal, but information gathered in this study will help researchers to better understand how alcoholics learn and how to create resources to aid learning.

Your participation in this study will be completely anonymous. Participant names will not be written on tests or surveys given during the study. Instead you will be assigned a random identity in the form of a four-digit number. There is no key to link participant names with the random numbers they have been assigned. We may publish results of the study, but names of people or any identifying characteristics will not be used in any of the publications.

Confidentiality of your research records will be strictly maintained. Participant performance in this study will be kept confidential and dissemination of results will occur in an aggregate format with removal of all information that could lead to the identification of any participant. Data will be archived for 3 years.

If there is anything about the study or your participation that is unclear or that you do not understand, or if you have questions or wish to report a research-related problem, you may contact Jeff Gutkin at (917) 560-9411, jgutkin@gc.cuny.edu, 365 5th Ave., New York, NY, Department of Educational Psychology.

For questions about your rights as a research participant, you may contact Kay Powell (IRB Administrator), CUNY Graduate Center, 365 Fifth Avenue, Room 8309, New York, NY 10016-4309, 212-817-7525, kpowell1@gc.cuny.edu.

Thank you for considering participation in our study. This Information Sheet is yours to keep.
Appendix R: Graduate School & University Center (CUNY) IRB Approval

DATE: November 7, 2013
TO: Jeff Gutkin, PhD
FROM: Graduate School & University Center (CUNY) HRPP Office
PROJECT TITLE: [352969.3] Designing Instruction for Recovering Alcoholics: The Role of Executive Function and Level of Guidance in Learning from Visually Complex Simulations
SUBMISSION TYPE: Amendment/Modification
ACTION: APPROVED
APPROVAL DATE: November 7, 2013
EXPIRATION DATE: November 6, 2014
RISK LEVEL: Minimal Risk
REVIEW TYPE: Expedited Review
REVIEW CATEGORY: Expedited review category # 7

Thank you for your submission of Amendment/Modification materials for this project. The University Integrated IRB has APPROVED your research. This approval is based on an appropriate risk/benefit ratio and a project design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

PLEASE NOTE THAT NO RESEARCH PROCEDURES MAY BEGIN AT THE NEW FIELDSITE UNTIL DOCUMENTATION OF IRB APPROVAL FROM THE SITE HAS BEEN SUBMITTED TO OUR OFFICE.

Please note that any modifications/changes to the approved materials must be approved by this IRB prior to implementation. Please use the appropriate modification submission form for this request.

All UNANTICIPATED PROBLEMS (UPS) involving risks to subjects or others, NON-COMPLIANCE issues, and SUBJECT COMPLAINTS must be reported promptly to this office. All sponsor reporting requirements must also be followed. Please use the appropriate submission form for this report.

This research must receive continuing review and final IRB approval before the expiration date of November 6, 2014. Your documentation for continuing review must be received with sufficient time for the IRB to conduct its review and obtain final IRB approval by that expiration date. Please use the appropriate continuation submission forms for this procedure. PLEASE NOTE: The regulations do not allow for any grace period or extension of approvals.

If you have any questions, please contact Claire Panetta at (212) 817-7527 or gc-irb@gc.cuny.edu. Please include your project title and reference number in all correspondence with this committee.
Appendix S: Rutgers State University of New Jersey, IRB Approval

https://eirb.rutgers.edu/eIRB/Doc/0/UDNVUCA7A4V3QSUIDG7...
### Submission Details

<table>
<thead>
<tr>
<th>Submission Type:</th>
<th>Research Protocol/Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approval Date:</td>
<td>12/9/2013</td>
</tr>
<tr>
<td>Expiration Date:</td>
<td>12/8/2014</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pregnancy Code:</th>
<th>No Pregnant Women as Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pediatric Code:</td>
<td>No Children As Subjects</td>
</tr>
<tr>
<td>Prisoner Code:</td>
<td>No Prisoners As Subjects</td>
</tr>
</tbody>
</table>

**Protocol Other Materials:**
- Card Rotations Test v6.1
- Control Group Flyer.pdf
- Demographic Survey.pdf
- Experimental group flyer.pdf
- Familiar Pictures Test
- Information Sheet.pdf
- Interest Survey.pdf
- MINI.pdf
- Paper Folding Test v2.2
- Pre and post test, Cognitive Load
- Stroop Test and treatments.pdf
- Protocol.pdf
- SADQ
- Visualizer Verbalizer Questionnaire.pdf
- WST.pdf

*Study Performance Sites:*

- Rutgers School of Dental Medicine 110 Bergen Street, Newark NJ 07101
- CUNY Graduate Center 395 Fifth Avenue New York NY 10016

### ALL APPROVED INVESTIGATOR(S) MUST COMPLY WITH THE FOLLOWING:

1. Conduct the research in accordance with the protocol, applicable laws and regulations, and the principles of research ethics as set forth in the Belmont Report.
2. Continuing Review: Approval is valid until the protocol expiration date shown above. To avoid lapses in approval, submit a continuation application at least eight weeks before the study expiration date.
3.Expiration of IRB Approval: If IRB approval expires, effective the date of expiration and until the continuing review approval is issued. All research activities must stop unless the IRB finds that it is in the best interest of individual subjects to continue. (This determination shall be based on a separate written request from the PI to the IRB.) No new subjects may be enrolled and no samples/charts/surveys may be collected, reviewed, and/or analyzed.
4. Amendments/Modifications/Revisions: If you wish to change any aspect of this study, including but not limited to, study procedures, consent form(s), investigators, advertisements, the protocol document, investigator drug brochure, or accrual goals, you are required to obtain IRB review and approval prior to implementation of these changes unless necessary to eliminate apparent immediate hazards to subjects.
5. Unanticipated Problems: Unanticipated problems involving risk to subjects or others must be reported to the IRB Office (45 CFR 46, 21 CFR 312, 812) as required. In the appropriate time as specified in the attachment online at: http://irb.rutgers.edu/hasweb
6. Protocol Deviations and Violations: Deviations from Warnings of the approved study protocol must be reported to the IRB Office (45 CFR 46, 21 CFR 312, 812) as required. In the appropriate time as specified in the attachment online at: http://irb.rutgers.edu/hasweb
7. Consent/Assent: The IRB has reviewed and approved the waiver and/or alteration described in this protocol as required by 45 CFR 46 and 21 CFR 50, 56. (If FDA regulated research, only the versions of the documents included in the approved process may be used to document informed consent and/or assent of study subjects; each subject must receive a copy of the approved form(s), and a copy of each signed form must be filed in a secure place in the subject's medical/patient/research record.)
8. Completion of Study: Notify the IRB when your study has been stopped for any reason. Neither study closure by the sponsor nor the investigator removes the obligation for submission of timely continuing review application or final report.

9. The investigator(s) did not participate in the review, discussion, or vote of this protocol.

CONFIDENTIALITY NOTICE: This email communication may contain private, confidential, or legally privileged information intended for the sole use of the designated and/or duly authorized recipient(s). If you are not the intended recipient or have received this email in error, please notify the sender immediately by email and permanently delete all copies of this email including all attachments without reading them. If you are the intended recipient, secure the contents in a manner that conforms to all applicable state and/or federal requirements related to privacy and confidentiality of such information.
References


Hallion, L.S., Coutanch, M., & Jha, A.P., (manuscript in prep)


observeration of preferential choice behavior in authentic learning situations: The 
Visualizer/Verbalizer Behavior Observation Scale (VV-BOS). *Computers in Human 


Associates, Inc.

(Ed.), *The Cambridge handbook of multimedia learning* (pp. 147-158). New York, NY 
US: Cambridge University Press.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and 


Mayer, R. E. (2004). Should There Be a Three-Strikes Rule Against Pure Discovery Learning? 
*American Psychologist, 59*(1), 14-19.

Cambridge University Press.

Press.

Ability, Cognitive Style, and Learning Preference. *Journal of Educational Psychology, 
95*(4), 833-841.


visual presentation modes. *Journal of Educational Psychology*, 87(2), 319-334.

Murphy MK1, Bijur PE, Rosenbloom D, Bernstein SL, Gallagher EJ. (2013). Feasibility of a computer-assisted alcohol SBIRT program in an urban emergency department: patient and research staff perspectives. *Addiction Science Clinical Practices*, 16 (8), 2-


Plass & Homer (manuscript under review). Efficacy of a Computer-Based Learning Environment to Teach Chemistry.


