THE EFFECTS OF DYNAMIC METACOGNITIVE PROMPTS ON EXPERT AUDITOR REASONING EFFICACY

by

Toby Groves

ELENI PINNOW, PhD, Faculty Mentor and Chair

PAULA FREMONT, PhD, Committee Member

JONATHAN MENDS-COLE, PhD, Committee Member

Andrea Miller, PhD, Dean of Psychology

Harold Abel School of Social and Behavioral Sciences

A Dissertation Presented in Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

Capella University

January 2018

Abstract

The purpose of this study was to compare the effects of using dynamic, inward oriented reasoning strategies versus linear, outward oriented reasoning strategies on expert auditor reasoning efficacy. The research focused on further examining whether an interaction existed between the types of prompts received and an individual's personal need for structure. The preponderance of literature on expert auditor judgment has focused on the effects of outward oriented professional skepticism on reasoning efficacy. Missing from the literature on expert auditor reasoning quality are the effects of inward oriented metacognitive strategies. This was the first experimental study using expert auditors that examined the combined effects of personal need for structure and dynamic versus linear prompts on reasoning efficacy. The research process included presenting participants with abstract and deontic hypothesis testing tasks after exposure to either a dynamic or linear reasoning prompt. The purpose of the research was to determine whether a statistically significant difference existed between the abstract and deontic selection task scores of participants based on the type of prompt received, and if an interaction existed with the trait of need for structure. Task type (abstract versus deontic) was significant in this study because abstract rule-based tasks represent many of the reasoning challenges presented to auditors. In a quantitative, two group between-participants design, 264 expert auditors participated in hypothesis testing tasks after receiving a brief 4-minute dynamic reasoning prompt or a linear reasoning prompt. The results were interpreted using two-way ANOVAs. The major finding of this research suggests that dynamic reasoning strategies, but not linear strategies enhance reasoning efficacy on abstract tasks.

Dedication

I dedicate this work to my children Jacob, Caleb, Sam, Sophie and Chloe, and to my parents, Charles and Bernice. A special dedication goes to the loving memory of Caleb who taught me the importance of embracing my errors and failures because every time I can prove myself wrong I come another step closer to the truth. Many thanks to my friends who were supportive of the efforts required to achieve the completion of this work.

Acknowledgments

My heartfelt thanks go to the expert auditors who participated in this research. It is with deep admiration that I acknowledge the noble efforts of experts in every domain that have the authenticity to question their own judgments, the courage to challenge superficial criteria, and the genuineness to build trust and improve processes in the pursuit of public health, safety, and prosperity. Thanks to my mentor Dr. Eleni Pinnow and the dissertation committee for their help and guidance. Last but certainly not least, my deepest appreciation goes to James and Patricia Cergol without whose support this research could not have been completed.

Table of Contents

List of Tables

CHAPTER 1. INTRODUCTION

Background of the Problem

An understanding of expert reasoning processes related to the way experts search for information, and how they interpret and validate evidence is crucial for public health and safety. Auditor reasoning is of special interest to researchers because of far-reaching implications of audit quality in virtually all types of organizations from environmental issues to financial markets (Koonce & Mercer, 2005; Peytcheva, 2013). Auditors are charged with having the specialized knowledge tailored for specific applications for safety, effectiveness, and security in a wide variety of domains. The overwhelming majority of policy directives and research initiatives have focused on professional skepticism of evidence and assertions of decision makers. This form of outward oriented skepticism is widely accepted within the audit industry as a determining factor in the quality of auditor reasoning and ultimately audit quality (Hurtt, 2010; Peytcheva, 2013). Successful implementation of this type of professional skepticism requires reasoners to focus on a matching frame of reasoning, or an outward orientation allowing the reasoner to assume the perspective of at least one party to the social contract that would face a loss if the contract were not honored (Dweck, Mangels, Good, Dai, & Sternberg, 2004; Gigerenzer & Hug, 1992; Tenbrunsel & Messick, 1999).

Studies, such as Cosmides (1989), Oaksford and Chater, (1994) and Peytcheva, (2014) examined reasoning processes of auditors and utilized experiments that rely on social contract theory and used hypothesis-testing tasks as tests of the quality of auditor reasoning. These

studies examined reasoning efficacy on content-independent, abstract tasks, as well as content dependent, deontic tasks, noting reasoners experience significantly higher rates of success with deontic tasks (Cosmides, 1989; Oaksford & Chater, 1994). Studies using either abstract or deontic tasks while priming with outward oriented skepticism yielded contradictory results or have produced findings inapplicable to real world reasoning contexts (Klein, 1997; Peytcheva, 2013).

Expert auditors work in highly structured environments that tend to cue a conservative, inward oriented reasoning frame and a motivation to avoid sanctions and loss of reputation (Johns, 2006; Michael, 2006; Tenbrunsel & Messick, 1999). This inward-orientation is qualitatively different from the outward orientation associated with social contract theory, and research manipulations relying on outward oriented skepticism may not match the reasoning frame of expert auditors in real world auditing contexts, yielding unreliable or misleading results (Dweck et al., 2004; Tenbrunsel & Messick, 1999). Social contract theory proposes people follow unspoken rules assumed depending on the context of the problem they face, and decisions are influenced by these assumptions (Fiddick, Brase, Cosmides, & Tooby, 2017). Gigerenzer and Hug (1992) suggested decisions become clearer for the reasoner when they are framed as a social contract where a person is able to put himself in the shoes of someone who is cheated.

Research in audit quality and auditor judgment have gained interest after waves of audit failures were reported in the news, frustrating regulators and standard-setting organizations, and spurring researchers to explore cognitive processes underlying auditor judgment (Brink, Lowe, & Victoravich, 2013; Lenard & Alam, 2009; Rezaee, 2004). The response to audit failures has been the enactment of new legislation implemented through regulatory and oversight agencies directed toward the promulgation of new procedures, decision tools, and training programs

designed to enhance outward oriented skepticism. The Sarbanes Oxley Act of 2002 authorized creation of the Public Company Accounting Oversight Board (PCAOB); the Dodd-Frank Wall Street Reform Act and the Consumer Protection Act of 2010. Numerous additions and amendments to auditing standards and guidelines also occurred (Baker, Bealing Nelson, & Blair Staley, 2006; Coffee, 2002; Lenard & Alam, 2009; Romano, 2004). Policymakers have ignored relevant behavioral research in the development of these policies and regulations, and the mismatch occurring between the reasoning frame cued by policies, and the reasoning frame that yields high quality reasoning (Romano, 2004).

While experts in all professional domains rely on taxonomies that assist in problem recognition and decision making, complex rule structures may narrow attention and result in reasoning processes based on heuristic models that could lead to judgment errors (Kahneman, 1991; Klein 1997; Michael, 2006; Tversky & Kahneman, 1975). Experts in highly structured environments are not immune to errors and biases in decision making and are prone to context specific reasoning errors related to the very tools developed to help aid in decision making (Johns, 2006; Kahneman, 1991). Linear decision processes with numerous steps can become over-burdensome, and as decision making becomes more complex, reasoners are prone to heuristic errors (Michael, 2006). Dual process theory describes two systems of cognition, system one which is fast and subconscious and system two which is slow, conscious and logical. Dual process theory provides a structure for understanding expert decision making (Evans & Over, 1996a; Kahneman, 2013; Kahneman & Frederick, 2002; Stanovich & Thompson, 2001). As a reasoner becomes more cognitively loaded and fatigued, their logical system becomes depleted and the reasoner is more prone to rely on fast and subconscious heuristics (Baumeister, Bratslavsky, Muraven, & Tice, 1998). Depleted reasoners are more likely to rely on incorrect

assumptions, stereotypes, or normative strategies that yield inappropriate outcomes without attention toward relevant evidence that does not fit into the linear model (Baumeister et al., 1998). Reasoners may benefit from a metacognitive, dynamic model of reasoning which assists reasoners in validating heuristic responses from system one with logical processes in system two (Kahneman, 2013; Mölle et al., 1996; Osman, 2007).

The audit domain has evolved to become increasingly complex with industry leaders in accounting and auditing continuously introducing new models, tools, and decision aids to assist auditors in simplifying their reasoning and decision making processes (Lenard & Alam, 2009). Standard setting agencies such as the American Institute of Certified Public Accountants (AICPA) and regulatory agencies such as the U.S. Securities and Exchange Commission (SEC) and the Financial Accounting Standards Board (FASB) have continuously issued updates to various practice standards. The standards include Generally Accepted Accounting Practices (GAAP), mandating rules and standards established for uniform financial reporting in the United States; Generally Accepted Auditing Standards (GAAS), setting standards for the quality of audits of non-public companies, and are promulgated by the Audit Standards Board (ASB). The ASB is a division of the AICPA, and Auditing Standards (AS), promulgated by the PCAOB as audit standards for public companies.

There are also numerous educational initiatives put forth from the AICPA and other leading industry organizations such as the Institute of Internal Auditors (IIA) and the Information Systems Audit and Control Association (ISACA). These well-intended actions have sought to improve the performance of auditors and audit quality, but the additive nature and evolution of regulations has resulted in an extremely precise and complex system that negatively impacts the information search process of auditors (Agoglia, Doupnik, & Tsakumis, 2011; Michael, 2006;

Peytcheva, 2013). Some of the very actions undertaken by regulatory authorities and standard setting agencies to address issues of audit quality may in fact be making the problem worse (Michael, 2006; Romano, 2004).

Statement of the Problem

The problem is a lack of research and understanding of the effects reasoning strategies employed in the audit industry have on the reasoning efficacy of expert auditors. Research on expert auditor reasoning efficacy has yielded inconsistent results and an incomplete understanding of reasoning processes that underlie expert auditor judgment (Gigerenzer & Hug, 1992; Peytcheva, 2013). Researchers have used priming mechanisms to prompt different levels of outward oriented skepticism to test reasoning performance with the assumption that outward oriented skepticism provides the best understanding of reasoning efficacy. Higher levels of outward oriented skepticism tend to produce positive results among novices more but not experts (Peytcheva, 2013). Researchers have also attempted priming participants with cheater detection prompts that yielded inconclusive results for both novices and experts (Peytcheva, 2013). The literature lacks experimental approaches that compare the effectiveness of metacognitive priming focused on inward orientations versus outward oriented skepticism of evidence and assertions. The lack of attention to prompts that cue the expert reasoner to be skeptical of their own reasoning processes could be significant because expert reasoning is distinct from novice reasoning in that outward oriented skepticism assists reasoning in novices but not experts (Gigerenzer & Hug, 1992; Kahneman, 2013; Tenbrunsel & Messick, 1999). Experts frequently work in highly structured environments which could cause cognitive rigidity, conservatism and cue expert reasoners toward an inward oriented frame of reasoning due to fear of consequences.

Consequently, a mismatch may exist in training that directs expert reasoners to use an outward oriented reasoning frame toward evidence (Dweck et al., 2004; Kahneman, 2013).

Experimental variables, which assume an inward orientation, may more closely match the expert reasoner's frame in a highly structured environment and yield more robust and reliable results on hypothesis testing tasks than manipulating outward oriented variables (Dweck et al., 2004; Johns, 2006; Kahneman, 2013; Michael, 2006; Tenbrunsel & Messick, 1999). Such inward oriented variables are focused on metacognitive strategies such as prompts spurring either dynamic or linear reasoning frames in lieu of skepticism or cheater detection prompts which are based on an outward orientation. The significance of focusing on an inward orientation rather than an outward orientation in this context is twofold: Experts, as opposed to novices, are already saturated with outward oriented skepticism, and further prompts may not achieve additional gains. Second, the rigid, linear environmental structure may cue a self-conscious frame, motivating reasoners to avoid non-compliance and related sanctions. In this environment reasoners tend to frame issues more abstractly to create mental distance and reduce the threat, and accordingly may seek the simple safety of an abstract, content independent structure and familiarity of frequently used tools and decision aids (Agoglia et al., 2011; Dweck et al., 2004; Michael, 2006).

Researchers have examined the role that auditor reasoning plays in audit quality and audit failures, and have suggested that reasoning processes are impaired by poor situational awareness leading to a syntactic focus on mechanics of search activities (Peytcheva, 2013). This lack of situational awareness is ameliorated more using inward oriented, metacognitive strategies than an even greater focus on outward oriented strategies (Agoglia et al., 2011; Dweck et al., 2004; Michael, 2006; Peytcheva, 2013). Details of auditor information search activities are considered

particularly important, including how situational context influences the information search and interpretation of evidence, how natural tendencies of individual auditors influence their judgments, and how situational context and individual tendencies interact to influence auditor motivations and evidence search behaviors (Peytcheva, 2013). Professional skepticism, widely accepted as a key factor in audit quality, is based on how auditors interpret the salience of evidence, what pattern seeking errors exist, and cognitive influences that result in common errors in reasoning such as confirmation bias (Hurtt, 2010). Researchers have conducted state and trait research, such as whether certain characteristics or personality traits of an individual auditor affect the accuracy of their judgments, and how current mood state affects reasoning processes (Peytcheva, 2013). Finally, priming effects were examined to determine how exposure to various stimuli may assist or handicap auditors in their reasoning processes, as well as how the nature of the rules themselves affect reasoning processes (Agoglia et al., 2011).

A checklist-style or linear approach to reasoning narrows decision processes by limiting the evidence that appears relevant and may lead to errors in judgment (Michael, 2006; Tenbrunsel & Messick, 1999). Peytcheva, Wright, and Majoor (2014) suggested ways auditors seek information are affected by the rules-based structure that exists, leading to a linear reasoning process that may limit auditors seeking behaviors toward certain narrow bands of information. Agoglia, Doupnik, and Tsakumis (2010) added a broader, principles-based structure leads to attention of a wider variety of options in decision making than does a rules-based structure. Michael (2006) suggested further research in linear reasoning addressing rules and decision making and noted different types of decisions may benefit from a variety of reasoning models. Likewise, a study on information search efficiency suggested there is more information needed on decision making under risk (Blay, Kadous, & Sawers, 2012).

Studies, such as Pike, Curtis, and Chui (2013); Peytcheva et al. (2014); and Knoblich et al. (1999) explored ways decision making is affected by constraints. However, there is a gap in the literature addressing the problem in the audit environment. The audit environment is a unique environment due to the volume of rules that exist in not only auditing but also the numerous specific domains in which audits are performed. Furthermore, these studies did not explore whether dynamic, metacognitive approaches to reasoning may have affected the reasoning quality of participants. Ostafin and Kassman (2012); and Gilhooly, Georgiou, and Devery (2013) examined effects of priming on reasoning, but a gap exists for this type of research in the audit environment.

Purpose of the Study

The purpose of this study was to determine whether a statistically significant difference existed between selection task scores of expert auditors exposed to a dynamic, metacognitive prompt versus a linear prompt, and if an interaction existed with personal need for structure. Individuals with higher need for structure may be more likely to make heuristic reasoning errors under a linear mindset than a dynamic mindset (Dweck et al., 2004; Kahneman, 2013; Tenbrunsel & Messick, 1999). The study investigates effects of inward oriented, metacognitive strategies versus outward oriented strategies of skepticism on expert auditor reasoning efficacy. Research in the audit domain has placed an emphasis on outward oriented skepticism toward evidence and the individual trait of skepticism (Hurtt, 2010; Hurtt, Brown-Liburd, Earley, & Krishnamoorthy, 2013). This research will fill a gap in the literature by investigating the influence of inward oriented, metacognitive strategies on reasoning processes. The abstract rule-laden structure of the audit domain may prompt expert auditors to assume an inward orientation and focus on themselves to avoid errors and possible sanctions (Tenbrunsel & Messick; 1999).

Thus, an examination of expert auditor reasoning through manipulation of variables related to inward oriented metacognitive strategies may more closely match expert auditor reasoning frames than variables related to outward oriented skepticism.

Significance of the Study

The findings from this study are significant from using dual process theory to challenge assumptions predominant in literature regarding the function of outward oriented skepticism in expert search strategies, evidence evaluation, and reasoning processes (Evans & Over, 1996a; Kahneman, 2003; Kahneman, 2013; Kahneman & Frederick, 2002; Stanovich & West, 2000). Specifically, dual process theory is used to examine reasoning processes of expert auditors by comparing reasoning efficacy differences between inward oriented versus outward oriented reasoning strategies and the interaction with personal need for structure.

Standard setting agencies have promulgated policies relying on the efficacy of outward oriented skepticism and asserted it as a key factor in the quality of auditor reasoning (Baker et al., 2006; Coffee, 2002; Hurtt, 2010; Lenard & Alam, 2009). Past research focused on outward-oriented skepticism and manipulated outward oriented, skepticism-related variables to study reasoning processes of expert auditors (Gigerenzer & Hug, 1992; Hurtt, 2010; Peytcheva, 2013). These studies experienced poor replication, inconsistent results, and lack a unifying explanation of why small changes in context have large effects on results (Johns, 2006). This study questioned whether the focus on outward oriented skepticism was misguided due to the inherent assumption in past literature auditors primarily pursue an outward oriented reasoning strategy and questioned whether outward oriented skepticism is an appropriate measure of the quality of auditor reasoning.

The success of an outward-oriented frame is tied to social contract theory that relies on the reasoner assuming the perspective of another party (Gigerenzer & Hug, 1992). This research suggests expert auditors are implicitly concerned with following the domain-specific structure, avoiding personal sanctions, resulting in a reasoning strategy on a continuum between inward and outward orientation triggered by an interaction between characteristics of the evidence, the task environment, and personal need for structure. This study included examining evidentiary characteristics by utilizing abstract or deontic tasks, manipulated the frame of the task environment by priming participants with either dynamic or linear prompts, and assessed personal need for structure using the Personal Need for Structure (PNS) inventory (Thompson, Naccarato, & Parker, 1989). The interaction of these variables alters metacognitive mechanisms related to information search and evidence interpretation strategies (Dweck, et al., 2004; Tenbrunsel & Messick, 1999). Accordingly, this study aimed to manipulate variables related to reasoners' metacognitive tendencies rather than outward-oriented skepticism processes.

This research advances scientific knowledge related to expert reasoning by filling a gap that exists in the literature on effects of dynamic metacognitive prompts, particularly dynamic versus linear reasoning frames, and the interaction of that priming with the trait of personal need for structure. The body of literature recognizes the need for research on reasoning utilizing expert reasoners instead of novices as participants and utilizing experimental manipulations related to natural environments (Klein, 1997). Research related to auditor reasoning is important to the field of psychology due to the complex environment in which auditors conduct their work (Koonce & Mercer, 2005; Ramamoorti & Nayer, 2013). Results of this research may assist policymakers and standard setting agencies as they design procedures for experts in highly structured task settings. A better understanding of cognitive performance related to information

search, evidence validation, and hypothesis-testing processes may inform policymakers and standard setting agencies on effective guidance practices in the expert audit context.

Research Questions and Hypotheses

RQ1: Is there a statistically significant difference between the combined (abstract and deontic) selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model?

$RQ1H1_{0:}$ There is not a statistically significant difference between the combined (abstract and deontic) selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model.

$RQ1H1_{a:}$ There is a statistically significant difference between the combined (abstract and deontic) selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model.

RQ1SQ1: Is there a statistically significant difference between the abstract selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model?

$RQ1SQ1H_0$: There is not a statistically significant difference between the abstract selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model.

$RQ1SQ1H_a$: There is a statistically significant difference between the abstract selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model.

RQ1SQ2: Is there a statistically significant difference between the deontic selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model?

RQ1SQ2H_0: There is not a statistically significant difference between the deontic selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model.

RQ1SQ2Ha: There is a statistically significant difference between the deontic selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model.

RQ2: Is there a statistically significant difference between the combined (abstract and deontic) selection task scores of auditors based on their scores on the personal need for structure inventory?

RQ2H1_0: There is not a statistically significant difference between the combined (abstract and deontic) selection task scores of auditors based on their scores on the personal need for structure inventory.

RQ2H1_a: There is a statistically significant difference between the combined (abstract and deontic) selection task scores of auditors based on their scores on the personal need for structure inventory.

RQ2SQ1: Is there a statistically significant difference between the abstract selection task scores of auditors based on their scores on the personal need for structure inventory?

RQ2SQ1H_0: There is not a statistically significant difference between the abstract selection task scores of auditors based on their scores on the personal need for structure inventory.

RQ2SQ1H_a: There is a statistically significant difference between the abstract selection task scores of auditors based on their scores on the personal need for structure inventory.

RQ2SQ2: Is there a statistically significant difference between the deontic selection task scores of auditors based on their scores on the personal need for structure inventory?

RQ2SQ2H_0: There is not a statistically significant difference between the deontic selection task scores of auditors based on their scores on the personal need for structure inventory.

RQ2SQ2H_a: There is a statistically significant difference between the deontic selection task scores of auditors based on their scores on the personal need for structure inventory.

RQ3: Is there a statistically significant difference between the combined (abstract and deontic) selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory?

RQ3H1$_0$: There is not a statistically significant difference between the combined (abstract and deontic) selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory.

RQ3H1$_a$: There is a statistically significant difference between the combined (abstract and deontic) selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory.

RQ3SQ1: Is there a statistically significant difference between the abstract selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning

model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory?

RQ3SQ1H_0: There is not a statistically significant difference between the abstract selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory.

RQ3SQ1H_a: There is a statistically significant difference between the abstract selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory.

RQ3SQ2: Is there a statistically significant difference between the deontic selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory?

RQ3SQ2H_0: There is not a statistically significant difference between the deontic selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory.

RQ3SQ3H_a: There is a statistically significant difference between the deontic selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory.

Definition of Terms

Reasoning is investigated from a cognitive orientation and defined for the use of this study as cognitive processes involved in seeking evidence that supports or opposes the truth of a proposition and moving from a premise to a conclusion through belief updating based on the interpretation of the evidence (Walton, 1990). This view of reasoning includes the conscious and subconscious cognitive processes described by dual process theory (Kahneman, 2003; Kahneman, 2013; Kahneman & Frederick, 2002; Stanovich, 1999; Stanovich & West, 2000) involved in deciding whether a proposition is justified and those strategies and methods employed to solve logical inconsistencies and arrive at a conclusion. The study examines evidence seeking activities, interpretation and reasoning as a goal directed behavior in the context of a logical argument.

Reasoning style was a key interest in this research. Specific interests include effects of inward versus outward orientation on reasoning efficacy and the differences between dynamic, metacognitive, and linear reasoning approaches on judgment and the shifting of neural activity from areas handling linear processing to networks known to handle dynamic metacognitive processes. Variations in reasoning style may cue attentional and interpretational differences related to an interaction between traits of the reasoner and the context of the task environment (Dweck, et al., 2004; Tenbrunsel & Messick, 1999). Linear reasoning style is used in this research as related to economic and rational theory and conceptualized by Vance, Groves, Paik, and Kindler (2007), based in Newtonian, deterministic tradition and understanding problems by viewing them as systems that are divided into independent parts that are discrete and analyzed independently.

Inward and outward orientation was used in this research in reference to attentional focus as it affects cognitive processes of evidence seeking and interpretation. Inward orientation refers to a focus on the self-regulatory and metacognitive processes when performing tasks including performance evaluation, responsibilities, concerns over competence, and anticipatory anxiety focused on adherence to procedure rather than the content of the evidence. Self-evaluative processes are a distraction and act as a constraint to task involvement, or the extent to which an individual can focus on the content of the activity (Elliot & Harackiewicz, 1994). Task-specific goals based on normative standards cue different motivational processes related to orientation than goals based on characteristics and content of the task (Saleh & Grygier, 1969).

Outward orientation focuses on the task characteristics, environmental details and content (Saleh & Grygier, 1969). An outward focus on external directives, including goals or rules assigned externally can evoke performance anxiety and interfere with task engagement (Elliot & Harackiewicz, 1994). This can reduce perceptions of autonomy and result in desire for inward orientation. Orientation is distinguished from metacognition in this research. Although orientation can be an aspect of metacognitive processing, metacognition is referenced in this research as volitionary cognition as opposed to the orientation adopted by a reasoner which is frequently immediate, automatic and involuntary (Kahneman, 2013). The orientation someone assumes is not effortful on the part of the reasoner and reasoners usually lack specific awareness of their orientation unless it is brought to their attention or they have learned to cue it volitionally.

Structure of the domain-specific task environment is a key consideration in this research (Johns, 2006). Highly evolved task environments containing multiple decision tools and compliance frameworks elicit different reasoning frames, orientations, and implicit motivations

for reasoners than less structured and less complex environments (Dweck et al., 2004; Kahneman, 2013; Michael, 2006; Tenbrunsel & Messick, 1999). Highly structured environments as used in this study refers to professional domains that have evolved taxonomies and rely on multiple decision aids with rule frameworks that carry implied or explicit sanctions to ensure compliance. Highly structured environments are subject to regulatory oversight from multiple agencies and standards from multiple standard-setting agencies. Experts in highly structured domains routinely rely on decision aids and frequently refer to procedural guidelines in the performance of their work (Michael, 2006).

This study required using a broad definition for expert auditors, which will include those auditors not considered novices or students, but were actively employed in the industry and engaged in the industry as indicated by their pursuit of continuing professional education, which is necessary in the industry to maintain professional designations, meet employment requirements, or to fulfill personal knowledge goals. This definition is used to conform to past research by (Peytcheva, 2013) that relies upon reasoning differences between novices and practicing professionals and does not conform to definitions such as Ericsson, Prietula, and Cokeley (2007) who defined expertise using a 10-year model.

The target population for this research was auditors sufficiently embedded in the professional audit domain such that they are familiar with the structure of the domain, its regulatory rules, norms, standards and guidelines. While research has indicated significant differences in reasoning processes between experts and novices (Kahneman, 2013; Klein, 1999) and as such, participants imbedded in the structure of the audit domain is of interest for the study, however, Hurtt (2010) found no meaningful demographic differences amongst experts in

relation to reasoning tasks and skepticism. Accordingly, demographics are not a factor of interest in this research.

Linear reasoning style uses a rational, analytic frame based in formal logic and relies on ordered approach to decision making using a step-process wherein each step builds from and relies on the previous step, such as decision aids, normative tools and checklists traditionally used in highly structured environments (Michael, 2006). Linear reasoning is a static style of reasoning where designated elements of an argument do not change (Walton, 1990). The linear reasoning approach assumes relationships between variables are linear and one-dimensional and the whole is understood by the sum of its parts (Lichenstein & Mendenhall, 2002). Linear, analytic reasoning detaches the content from its context, focus on attributes and assignment to categorie, and utilize rules about categories to provide explanations (Nisbett, Peng, Choi, & Norenzayan, 2001)

Dynamic reasoning style was conceptualized for this study as using a divergent approach to reasoning known to spur a shift in thought processes from system one brain areas to system two brain areas with greater dynamic complexity as opposed to convergent reasoning (Kahneman, 2013; Mölle et al., 1996). A convergent reasoning approach uses linear and analytical cognitive processes, while divergent reasoning focuses on creating alternative ideas and perspectives. Dynamic reasoning uses a divergent approach which affects attentional control by lessening controls over self-regulatory processes. Dynamic reasoning refers to the complexity in cortical activity correlated with a divergent approach to reasoning caused by the reasoner searching to make new associations between cues and shifting between unique thoughts. Dynamic cortical activity is reduced during convergent reasoning because the reasoner

is attempting to suppress information perceived as irrelevant and less activity between competing ideas that never arise compared to a dynamic style of reasoning (Mölle et al., 1996).

Dynamic reasoning style was further conceptualized through definitions by Nisbett, Peng, Choi, and Norenzayan (2001) and Vance et al. (2007) regarding nonlinear dynamical systems and holistic reasoning in an Eastern cultural tradition; and Richmond (1997), who asserted dynamic reasoning should allow problems to frame in a scale of time, recognizing the historical state, the current state, and a future state, allowing dynamic simulations and feedback loops (Maani & Maharaj, 2004). The dynamic reasoning frame means systems are interrelated and interdependent parts that are viewed holistically, and content as inseparable from its context. Dynamic reasoning is a type of reasoning that can seek new lines of approach as it progresses, and is not tied to fixed, succeeding steps (Walton, 1990). Dynamic reasoning does not rely on a chronological or step process, and supports counterintuitive or counterfactual reasoning for decision making.

Engaging in counterintuitive or counterfactual reasoning and recognition of assumptions are core components of dynamic reasoning styles. Dynamic reasoning is considered as a counterpoint to static reasoning, which works from a current perspective and assumes a linear relationship between the current state and future goals (Richmond, 1997). Dynamic reasoning utilizes a historical perspective and a non-linear path to the conclusion. To engage in dynamic reasoning, the reasoner should recognize and question the validity of assumptions already made. The reasoner would mentally simulate a decision and update beliefs in a feedback loop that creates an evolution of simulations until a conclusion is reached. The dynamic reasoning approach perceives organizations as an organism that both adapts and resists change. To achieve improved reasoning, an investment in effortful reasoning is required, questioning reasoning

processes, exploring perspective shifts, and examining non-obvious, underlying systemic relationships (Richmond, 1997).

Need for structure was assessed using the Personal Need for Structure (PNS) scale (Thompson et al., 1989). The PNS scale helps examine how individuals deal with ambiguous information within their environment. Individuals differ in their discomfort with uncertainty, cognitive inconsistencies, and ambiguities and thus their desire for structure (Neuberg & Newsom, 1993). Individual need for structure affects cognitive associations related to reasoning processes. The Personal Need for Structure scale was introduced by Neuberg and Newsom (1993) as a self-report scale where the subject agrees or disagrees with statements about how they feel in relation to their environment on a six-point scale. The theoretical foundation of the PNS is rooted in an assumption that coping mechanisms related to cognitive load are related to the ability to reduce ambiguity, and that an individual's reaction to ambiguity influences individual reasoning processes. Higher scores on the PNS are related to desire for simple structure and the tendency to rely on heuristics, assumptions, and stereotypes (Neuberg & Newsom, 1993).

Research Design

This quantitative study used an experimental approach with random assignment and a 2x3 factorial design. The dependent variables in the study were participant scores on five variations of hypothesis testing tasks based on the Wason Selection Task (WST) (Wason; 1966; Wason, 1968) as seen in Appendix B. The Wason Selection Tasks included one abstract task and four deontic tasks. There were two independent variables: Participant scores on the Personal Need for Structure scale (PNS) (Thompson et al., 1989), and priming type that was either a dynamic or a linear reasoning prompt in the form of a short, written narrative as seen in

Appendix C. Two-way ANOVAs were used to assess differences between the dependent variable and the interaction between independent variables. Descriptive statistics were utilized as a lens to organize results (Leedy & Ormrod, 2013). The research design used probability sampling and the sampling method was purposive.

The population sample was drawn from professional auditors attending continuing education conferences. Participants completed the PNS assessment followed by exposure to a dynamic or linear prompt that consisted of a two or three paragraph narrative, and then completed five selection tasks. The five selection tasks were in the form of one abstract Wason Selection Task and four deontic Wason Selection Tasks that inferred a social contract. The sample size for this study was 264. Sample size was determined by an *a-priori* power analysis using a standard alpha level of .05 and power of .9 with a medium effect size. Based on these assumptions the minimum required sample size was 252 however the actual sample size yielded was 264. s= Sample calculations were obtained by using G*Power (Faul, Erdfelder, Lang, & Buchner, 1996).

The study employed a survey method of data collection with participants responding on keypads to research questions administered in-person by the researcher at educational conferences in various locations around the United States. The researcher collected data according to the following steps: Participants were first provided with informed consent and responded electronically by affirming or declining their participation in the study. All potential participants were informed they could decline to answer any of the questions, may stop their participation at any time, or could alternatively choose to participate in the study's activities but still decline to have their responses saved for use in the research. A yes response was accepted as affirmation by the participant, and any other response or lack of response was considered as a

declination and that individual's data were not saved for use in the study. Care was taken to assure participants' anonymous keypad responses were unknown to the researcher and other participants.

Participants began by responding via keypad to the 12 item PNS assessment as seen in Appendix A. Each of the 12 items was presented individually and participants were given 10 seconds to respond to each question. The PNS assessment was followed by randomly assigned exposure to either a dynamic or a linear prime. The priming mechanism took place in the form of a written narrative either based on standardized technical language used in the audit industry regarding sampling procedures (linear prime) or a narrative prompting auditors to recognize their intuitive assumptions and explore alternatives (dynamic prime). After exposure to the priming mechanism, participants responded electronically to five separate hypothesis validation tasks based on the Wason Selection Task (Wason, 1966; Wason 1968). The first selection task was an abstract task in the form of *if p then q*, and was followed by four deontic versions of the selection task framed in a form related to social contracts.

Data were collected electronically, reviewed for accuracy and missing data, then transformed to a format for analysis using SPSS. Raw data contained no individually identifiable information of participants, and was encrypted and stored on a flash drive kept in a secure and locked drawer at the researchers' home office. Descriptive statistics were performed via SPSS to obtain variable mean scores and independent *t*-tests were run on the average (mean) score of the PNS assessment, two levels of the priming mechanism, and interaction between independent variables with the dependent variable. Two-way ANOVAs were performed to explore relationships between all levels of independent variables including five levels of the PNS assessment and two levels of the priming variable.

Assumptions and Limitations

Assumptions

The researcher made certain assumptions in conducting this research including that the specific content of work performed by expert auditors varies between specific domains, but the content independent nature of the work is comparable and should prompt similar reasoning strategies. Participants were auditors from a wide variety of environments including financial, environmental, and process auditors working in organizations that range from airlines to investments and it was assumed any operational differences in specific domains do not alter reasoning processes investigated by this study. It was assumed other than as related to the personal need for structure, the specific workplace environment or personal experiences of participants was evenly distributed to match the population and not significantly affecting average (mean) responses to study experiments. It was assumed all participants were familiar with normative sampling language used in the study that is commonly used within the auditing industry. Finally, it was assumed varying levels of experience amongst participants are similar to the population of interest.

Limitations

The goal of this research was to discover effects of dynamic versus linear priming effects on expert auditor reasoning using a hypothesis testing task. Therefore, the study was limited to reasoning of expert auditors in the audit environment. The research was also limited to reasoning related to hypothesis testing and evidence selection activities. The research is not intended to address novice reasoning, reasoning in non-audit or less structured environments or mechanics of general cognitive processes.

Organization of the Remainder of the Study

The remainder of the study is organized into chapters. This description is the conclusion of Chapter 1 that provided a background of the research problem, a statement of the research problem, the purpose of the study, the significance of the study, research questions, definition of terms, research design and assumptions and limitations. Chapter 2 will provide a literature review including methods of searching, theoretical orientation of the study, review of the literature specific to the topic, synthesis of research findings, findings, critique of previous research methods, and a summary. Chapter 3 will provide the research methodology including research questions and hypotheses, the research design, the target population and sample, research procedures, research instruments, ethical considerations, and summary. Chapter 4 will provide research results, background, sample description, hypothesis testing and summary. Finally, Chapter 5 will provide a summary of results, discussion of results, conclusions, limitations, implications for practice, recommendations for future research and conclusions. The dissertation will conclude with the reference section.

CHAPTER 2. LITERATURE REVIEW

Methods of Searching

The literature in this study was acquired via electronic databases accessed through the university library. Peer reviewed articles were accessed in electronic journals via ProQuest Psychology Journals and PsycArticles. The majority of research used in the study was sourced by following reference paths through the body of literature on the subject. Literature was acquired from disciplines of psychology, cognitive science, anthropology, and neuroscience. Most effective keywords included *expert reasoning*, *dual process theory*, *Wason selection task*, *hypothesis testing*, *professional skepticism*, and *evidence validation*. Searches were also completed related to relevant theories such as naturalistic decision making, belief updating, and recognition primed decision making.

Theoretical Orientation for the Study

Theoretical foundations of the research are based in cognitive psychology with a focus on two competing theories of expert reasoning: dual process theory (Kahneman, 2013) and naturalistic decision making theory (Lipshitz, Klein, Orasanu, & Salas, 2001). Dual process theory (Kahneman, 2013) assists as a framework to dissect auditor reasoning on the continuum between explicit conscious and implicit subconscious processes. The related theory of heuristics and biases (Tversky & Kahneman, 1975) coupled with dual process theory supports explanations for reasoning errors and understanding the role of personality traits and affective states related to auditor reasoning processes. The study draws from naturalistic decision making theory and the accompanying recognition-primed decision model (Lipshitz et al., 2001) to explore pattern matching and cue relevancy in expert decision making in domain specific settings. While Kahneman and Tversky's heuristics and biases model were traditionally thought to conflict with

the naturalistic decision making framework (Kahneman & Klein, 2009; Lipshitz, et al., 2001), the research will address this conflict and explain how the theories provide complementary frameworks for an examination of auditor reasoning processes.

This research relied on Domain Specific Reasoning (Gigerenzer & Hug, 1992) to assist in applying dual process theory (Kahneman, 2013) to explain reasoning processes unique to the audit context. A common theme amongst these theories is each theory refutes rational economic theories such as expected utility theory as described by Mongin (1997), which purports decision makers make choices using linear, formal logic by comparing calculations of expected utility values of potential outcomes. The Belief Adjustment model (Hogarth, & Einhorn, 1992) was relied upon to supplement an understanding of reasoning processes of auditors as associated with heuristics and biases paradigm (Tversky & Kahneman, 1975) in specific relation to belief updating processes. The study critically examines reliance on social contract theory (Cosmides, 1989; Cosmides & Tooby, 1989) and skepticism models (Hurtt, 2010) in the expert audit domain.

The philosophy aligned with the research relied on an ontological approach, assuming decision processes of the population of United States auditors were observable using tools from this research; by employing an epistemological view that the reasoning under examination in this study were measurable using selected tools and instruments validated and accepted in the literature in the field of expert reasoning. The axiological viewpoint of the study was that a variety of priming effects influence decision processes, and it is important to policymakers and researchers in the field when constructing future policies and procedures. The methodological view of this research was that of a quantitative, experimental approach based on a positivist interpretation.

Review of the Literature

Beginning in the mid 1960's, researchers began to challenge the validity of rational decision making models based in economic theory which dominated accounting and auditing research up to that point (Arnold, & Sutton, 1997). Subsequently, interest grew for examining psychological theories that offer a more thorough understanding of auditor reasoning processes, considered an important area of research due to consequences to the public of faulty auditor judgment (Koonce & Mercer, 2005). Auditors are experts charged with a responsibility as gatekeepers of the validity and reliability of processes and information in critical fields from governance and policy to public safety to finance (Coffee, 2002; Cravens, Oliver, & Ramamoorti, 2003; se, 2006).

Nearly all aspects of accounting and auditing include some aspect of decision making or behavior that is not well explained by economic theory (Koonce & Mercer, 2005). Psychological theories such as those in cognitive and social psychology are informative in the audit context because they do not rely on assumptions of strict rationality, and produce different predictions about decision making and behavior than allowed by economic theories. A more thorough understanding of auditor reasoning and behavior is gained from psychological theories that examine decision making, belief updating, heuristics and biases (Koonce & Mercer, 2005). Auditors and standard setting agencies must understand auditor reasoning processes and unconscious errors if they are to achieve gains in audit quality (Bazerman, Loewenstein, & Moore, 2002).

Researchers point to the importance of examining reasoning processes of auditors in natural decision environments related to heuristics and biases paradigm of Kahneman and Tversky (1975), understanding mechanisms behind how auditors make decisions under

uncertainty and risk, and how auditors seek, collect and update their beliefs about evidence (Gigerenzer & Hug, 1992; Klein, 1997; Peytcheva, 2013; Blay et al., 2012). Researchers noted the importance of examining specific issues affecting reasoning in the audit context such as attention and framing related to evidence salience (Kida, 1984; Tenbrunsel & Messick, 1999; Peytcheva, 2013).

This review will examine the literature in cognitive psychology related to auditor reasoning, which has a rich background in the examination of reasoning processes under uncertainty (Kahneman, 2013), as well as an investigation of cognitive processes that explain differences in attention and salience of evidence (Peytcheva, 2013). Influences of affect and personality traits, referred to in the literature as state and trait respectively, are also surveyed (Peytcheva, 2013). This review will include an examination of the literature on professional skepticism which is assumed by many researchers and audit industry experts to positively correlate with auditor reasoning (Blay et al., 2012; Peytcheva, 2013). The review will conclude with a synthesis of the literature on expert reasoning in the audit context.

The Role of Professional Skepticism in the Literature

Auditing standards adopted by the PCAOB and approved by the SEC emphasize professional skepticism as integral to audit quality and require auditors to exercise due professional care in the performance of their work and specifically require auditors to exercise professional skepticism (SAS 99, PCAOB AU §230.07, §316.13). PCAOB auditing standards recognize three elements of auditor skepticism including auditor attributes, auditor mindset and auditor actions. Standards identify attributes as auditor knowledge, skills, and abilities while auditor mindset directs auditors to assume neither honesty nor dishonesty. Auditor actions, as

related to skepticism, are identified as diligence in the gathering of evidence and a critical assessment of that evidence (PCAOB AU §230.07, §230.08, §230.09).

According to Hammersley (2011), auditors' cognitive performances are correlated with personality characteristics related to their motivations to understand situations measured through a professional skepticism scale developed by Hurtt (2010). The skepticism scale is widely used in the literature. The scale is a 30 item, Likert scale self-report measure that asks participants to agree or disagree with statements such as *I take my time when making decisions* on a scale of 1 to 6. Hurtt (2010) identified six major characteristics of professional skepticism. The characteristics include: having a questioning mind which is defined as approaching information with suspicion or doubt, seeking meaning and clarification of definitions, validating data, suspending judgment which is defined as waiting to gather all evidence prior to making a judgment. Suspending judgment was noted as being the inverse of need for cognitive closure. The third characteristic identified was search for knowledge, which Hurtt distinguished from having a questioning mind by defining it as a general curiosity and lacking the element of disbelief and doubt inherent in the questioning mind component.

Hurtt (2010) collectively described the first three elements as related to the way an auditor evaluates evidence. The fourth characteristic of professional skepticism is interpersonal understanding, which addresses the auditor's perceptual abilities regarding the integrity and motivations of those who provide evidence to the auditor. Interpersonal understanding would include opportunities and incentives that might exist that could cause an individual to commit fraud. The fifth characteristic is autonomy, which addresses the auditor's objectivity and independence. The sixth and final characteristic is self-esteem, or having confidence in one's abilities.

Hurtt (2010) asserted auditors who score higher on the skepticism scale will engage in more skeptical behaviors identified in the scale, focus greater attention toward the assessment of evidence, and produce a greater number of alternative problem explanations to explore. A higher level of professional skepticism was purported to correlate with superior logical reasoning, more effective evidence selection procedures, and improved performance on hypothesis testing tasks (Dawson, Gilovich, & Regan, 2002; Hammersley, 2011). Peytcheva (2014) suggested however, there is a lack of empirical evidence to support assertions that higher levels of skepticism improve auditor-reasoning efficacy. Although research indicates higher levels of professional skepticism leads auditors to perform more audit work, they also expand their search for evidence to include substantial amounts of irrelevant evidence that impairs logical reasoning and ultimately reduces audit quality (Peytcheva, 2013; Popova, 2012). Auditors high in trait skepticism may over-audit, which further reduces audit efficiency and quality (Hurtt, Eining, & Plumlee, 2012).

While professional skepticism received significant emphasis in the literature, regulatory, and professional standards, it was understood and defined poorly (Nelson, 2009). An understanding of how professional skepticism affects auditor reasoning is incomplete and the assumption that levels of professional skepticism explain the quality of auditor reasoning is challenged (Peytcheva, 2013). Professional skepticism includes possessing an attitude of a questioning mind and to critically assess the appropriateness and sufficiency of evidence (AICPA, 2007). The foundation of strong professional skepticism is the ability to attend to appropriate and relevant evidence, which is also one of the most important aspects of auditor reasoning (Peytcheva, 2013).

Appropriate professional skepticism includes sensitivity to relevant evidence, being critical of audit evidence, having an appropriate level of suspicion, and being alert to errors in the evidence (Peytcheva, 2013). The selection of relevant evidence entails efficient information gathering and belief updating processes, which is affected by skepticism (Kahneman, 2013). Nelson (2009) further defined auditor judgment as it relates to professional skepticism, saying auditors should display a heightened sensitivity to risks associated with incorrect assertions and assume a presumptive doubt view versus a neutral view of evidence. Auditors high in professional skepticism should require greater quantities and a higher quality of persuasive evidence before accepting an assertion is correct. Accordingly, PCAOB AU §230.09 states an auditor should not measure evidence based on a belief that management is honest.

Assessing and measuring professional skepticism is difficult, and while regulators may point to lower levels of skepticism as the cause of audit deficiencies, the reason why skepticism varies between and within individuals is unclear (Hurtt et al., 2013). Skepticism is complex and multidimensional, influenced by individual characteristics and environmental factors such as standards, regulations and the control environment (Hurtt, 2010). Skepticism is both a trait and a state, simultaneously influenced by stable and enduring individual characteristics and temporary arousal due to situational qualities. Hurtt et al. (2013) suggested the complex interaction of individual auditor characteristics, evidentiary characteristics, client characteristics and environmental characteristics are antecedents to professional skepticism in auditing, and influence the way skeptical judgment translates to behavior.

Individual personality characteristics of auditors are a prominent feature in the literature examining skepticism, including an individual's predisposition to have a questioning mind, referred to in the literature as trait skepticism, as well as training, expertise level, experience and

affect, referred to in the literature as state skepticism (Hurtt et al., 2013. Environmental characteristics influencing skepticism include incentives or pressures and levels of accountability that exist during an audit. Client characteristics include perceived knowledge and integrity of management, the complexity of organizational structure, and the existence and strength of internal audit staff. Evidentiary characteristics include the amount of evidence, the availability of evidence and the confirmatory or disconfirming nature of the evidence.

The best way to demonstrate that we understand a behavior is to be able to reverse that behavior through the manipulation of research variables (Kahneman, 2013); however, research on auditor reasoning related to skepticism and hypothesis-testing tasks yielded inconsistent and contradictory results and consistently failed to support investigated biases in expert auditor reasoning (Peytcheva, 2013; Smith & Kida, 1991). The way skeptical judgments translate into behavior is complex, and while an individual auditor may possess the knowledge, skills, and experience necessary for appropriate levels of skepticism, they may lack the confidence or courage to act on those beliefs (Hurtt et al., 2013). Research has found experienced auditors who are asked to assume an outward orientation and be skeptical of evidence rarely change their judgments about a situation. When those same auditors are primed with a dynamic perspective to be skeptical of their own judgments, thus cueing an inward orientation, they are more likely to produce explanations related to fraud than are less experienced auditors (Grenier, 2011).

Grenier (2011) also found industry experts were more likely to exhibit sound reasoning when they were prompted dynamically to examine their inward orientation by questioning their own judgments. It is important to note experienced auditors are sometimes conditioned to be less skeptical depending on whether they were personally exposed to irregularities (Hurtt et al. 2013). Inward versus outward orientation appears to be an important factor in cue sensitivity and

goal framing in auditor reasoning processes. Research found asking novices to question management's assertions (outward orientation) created more skeptical judgments than when those auditors were asked to assume an inward orientation and question their own judgment (Harding & Trotman, 2011). More skeptical judgments do not necessarily equate to better reasoning efficacy however (Peytcheva, 2013). Differences between findings is possibly due to complex differences in how inward versus outward orientation frames are cued in experienced versus non-experienced auditors and the extent to which their reasoning is focused on justification (Hurtt et al., 2013; Tenbrunsel & Messick, 1999; Michael, 2006).

The ability to engage in a dynamic versus linear reasoning was identified as having a significant influence on auditor reasoning. Rather than requiring auditors to focus on a linear process, reasoning may benefit from using a dynamic reasoning approach; using multiple different approaches and cueing different unconscious processes (Hurtt et al., 2013). Trotman, Simnett, and Khalifa (2009) found asking auditors to use a dynamic, backward reasoning approach improved reasoning. In their study, participants were primed to make a hypothetical assumption by imagining they were at some point in the future and a certain unexpected event had occurred. Participants were directed to work backwards and imagine why and how that event occurred, and participants generated a greater number of quality ideas related to the problem than participants who were not cued to use the backward approach.

In a study that asked accounting students to engage in reasoning tasks, researchers found those who had received training in law engaged in less biased reasoning than those who were only trained in accounting processes (Pinsker, Pennington, & Schafer, 2009). Authors suggested this was because the students adopted a different perspective and noted law students are trained to be dynamic thinkers, to explore contrary positions, recognize disconfirming as well as

confirming evidence, and view issues from multiple points of view. When auditors are trained to use both convergent and divergent reasoning, they engage in better reasoning than do auditors only trained to use a single perspective (Plumlee, Rixom, & Rosman, 2012). Peytcheva (2014), using a selection task experiment and a skepticism related priming mechanism found a professional skepticism prompt that increased suspicion, and a cheater detection form of an experimental manipulation did not improve cognitive performance of auditors. This research suggested auditor reasoning might benefit more from auditors questioning their own reasoning processes than simply directing them to be more skeptical of evidence (Hurtt et al., 2013).

Auditors are motivated to seek evidence corresponding to the problem perceived to be the most important and attention is drawn more to evidence that matches intrinsic motivations or arouses emotions (Kahneman, 2002). Auditors are rewarded more on outcomes than the quality of the decision process, and regulators are more aggressive in punishing bad outcomes even if the decision process was made with good intentions, than good outcomes that were the result of faulty reasoning (Bazerman & Tenbrunsel, 2011). These incentives and sanctions may mean outcomes and following standardized procedures are more salient than well-crafted decision processes. Additionally, rules change the frame of the decision maker to narrow attention and develop an inward focus motivated to follow rules and avoid possible sanctions. The result is an inward focus with attention motivated more toward evidence corresponding to auditor standards than externally motivated social evidence related to skepticism (Tenbrunsel & Messick, 1999). The PCAOB, as an example, only addresses negative outcomes from inspections, reinforcing auditor's beliefs that outcomes are more important than the quality of the reasoning process and the safest choice would be to protect one's self by carefully following proscribed procedures (Peecher, Solomon, & Trotman, 2010).

Skepticism of assertions benefits from the activation of a social contract frame in which the auditor recognizes one party has agreed to meet some requirement or pay some cost in associated with receiving a corresponding benefit (Cosmides, 1989; Gigerenzer & Hug, 1992). Skepticism of evidence and assertions requires a perspective of an outward orientation wherein the auditor is cued to the perspective of a party who is cheated in relation to a social contract (Gigerenzer & Hug, 1992). The rule and incentive structure that surrounds experienced auditors may cue instead, an inward orientation (Tenbrunsel & Messick, 1999). Cheater-detection prompts similarly require an outward orientation to affect reasoning.

Domain Specific Context and Task Environment

Experts are expected to have superior judgment and decision making skills associated with their experience and mastery of a professional framework of knowledge (Shanteau, 2015). Diverse professional domains from auditing and medicine to law have underlying science that evolves as the profession matures, developing increasingly complex taxonomies and formal decision models that assist professionals in recognizing problems, their likely etiology (Kahneman, 2013). Formal decision models include a variety of decision aids such as checklists and algorithms designed to assist experts in difficult decision situations. It is assumed experts within a domain will develop similar and increasingly precise approaches to problems, will use similar decision making tools, and will have common backgrounds and experiences. In domain contexts that are quantitative in nature such as auditing, it is assumed there is only one correct answer and any expert who does not arrive at that one correct answer is biased in choice of evidence or interpretation of the evidence (Shanteau, 2015). Specialized training in specific methods increases the likelihood a reasoner will regularly employ that approach to problem solving (Smith & Kida, 1991).

The assertion there is only one correct answer to a problem, and experts will follow proscribed taxonomies in a predictable manner is based in economic theory and expected utility theory. This theory leads us to expect experts within the same domain will consistently reach the same conclusions, although research has consistently demonstrated this is untrue (Shanteau, 2015). Consensus between experts and reliability of conclusions in the same domain is viewed as a requisite part of expertise (Einhorn, 1974). It is expected when the environment of a domain involves uniform standards the domain will experience higher levels of agreement and higher levels of compliance within that domain (Michael, 2006). Shanteau (2015) suggested two important aspects of expert reasoning are the ability to discriminate between stimuli that are similar but not identical, and internal consistency, or the ability to be consistent by agreeing with themselves over time. An important part of distinguishing expert judgment from non-expert judgment is the ability to determine what evidence is relevant versus what is irrelevant. Understanding influences that can help experts make these discriminations is important to policymakers and researchers.

Despite the expectation of homogeneity among expert responses, research shows significant differences occur amongst conclusions of experts even in situations where experts have the same data and follow identical procedures (Johns, 2006). Disagreements are frequently perceived between experts because they see different approaches to problem solutions (Shanteau, 2015). Experts who use rules and guidelines as decision making aids routinely fail to apply rules uniformly, and differences in context that spur reasoning variances can be so small as to be nearly undetectable (Johns, 2006). Researchers have customarily viewed disagreement amongst experts within the same domain as being a sign of incompetence of some of the experts. It is difficult however, for the public to ascertain which experts may be wrong so the disagreement

amongst those experts leads to a distrust of the entire group (Shanteau, 2015). Shanteau (2015) suggested a conclusion termed either right or wrong is overly simplistic in many expert domains, and past investigations into expert decision making were performed with an insufficient understanding of complexities facing expert decision makers and without an appropriate appreciation for variances between domains. Varying amounts of disagreement between experts are accepted, and expected because of these contextual and individual differences (Shanteau, 2015).

Biases in decision making do not occur evenly between contexts, making the predictability of their effects difficult (Johns, 2006; Kahneman & Tversky, 1982). Care is necessary in using conclusions of research from one audit context in another context as heuristics and biases that affect reasoning are highly differential (Smith & Kida, 1991). In some domains, experts show very slight signs of bias while other domains show significant biases for reasons that are poorly understood (Johns, 2006; Shanteau, 2015). Additionally, structures designed to assist decision makers and diminish decision errors and biases can have unexpected effects. As an example, in at least some domains, the more complex the rule structure becomes, the more likely decision makers will fall prey to certain biases. Redelmeier and Shafir (1995) found just one extra option in a decision model increased complexity to the extent that it caused reasoning difficulty for physicians and dramatically influenced their professional judgment. The research suggested even small changes in choice complexity can result in decision makers being more likely to favor the default option. Research in the accounting and auditing domain indicates more similar conclusions are reached by professionals following a less precise principles based system rather than a more precise rule based system, surprising industry experts (Agoglia, Doupnik & Tsakumis, 2011). This suggests relevancy among choice options is difficult to

interpret, particularly for experts in certain domains. Decision researchers found experts make some of the same errors made by novices (Kahneman, 1991).

Per Shanteau (2015), variances among consensus levels of expert conclusions in different domains are correlated with the contextual and cultural characteristics in which the expert is embedded, such as error tolerance and the nature of the stimuli examined. Professional domains that experience higher levels of consensus are more likely to be dealing with stimuli that are static and predictable such as physical systems rather than human behavior, which is dynamic and less predictable. Higher levels of consensus are also more likely amongst repetitive tasks than when under unique and changing conditions. The routine employment of normative decision aids such as checklists or algorithms also supports more uniform agreement amongst experts. Consensus is also more likely in cultures that provide feedback over those that do not, and perhaps surprisingly, where errors in decision making are tolerated versus cultures where errors are not tolerated (Shanteau, 2015). Agoglia (2011) findings explain, at least in part, results of that principles based accounting procedures result in more agreement amongst experts than rules based procedures due to the advantageous interpretation of rules and the triggering of a moral reflection frame when rules do not exist as opposed to a rule that is interpreted as approval of an otherwise unacceptable decision.

Rules spur linear reasoning, narrow attention and lead to the recognition of commonly recognized patterns and cause difficulty in the recognition of unique or unusual connections (Michael, 2006). Precise laws and rule-based structures tend to make people self-focused to protect themselves from inferred sanctions and are more likely to follow checklists or guidelines that may not be the most relevant to given circumstances. Rules and sanctions change the frame in which decision makers view their decision, making them more likely to perceive their choices

as a business decision rather than an ethical decision and make decision makers blind to certain options in the decision process (Tenbrunsel & Messick, 1999). Decision makers tend to focus on the letter of a rule rather than its sprit especially when sanctions are present, and sanctioning systems tend to reduce cooperation (Michael, 2006: Tenbrunsel & Messick, 1999).

Decision makers can face inner conflict and confusion between what they want to do and what they think they should do in relation to rules and guidelines (Bazerman, Tenbrunsel, & Wade-Benzoni, 1998). This conflict can result in anxiety and pressure on decision makers to fall back on automatic default judgments that feel safe (Baumeister et al., 1998). Michael (2006) suggested this conflict may be part of the reason the most highly regulated domains are those which also experience the highest rates of non-compliance. Research indicates reasoning benefits from a culture that supports mental freedom to recognize patterns in changing and dynamic situations, not when experts are obsessed with avoiding disagreements (Johns, 2006; Michael, 2006;; Shanteau, 2015). Researchers also suggested the goal of experts in decision making is not to simply arrive at a single answer but to understand complexities and triangulate by arriving at answers through different means and using a variety of approaches (Agoglia et al., 2011; Michael, 2006; Shanteau, 2015).

It is generally accepted within the audit domain that professional skepticism is required for effective reasoning processes (Hurtt, 2010). Auditors however, must simultaneously contend with a formal domain structure and content independent logic as well as content specific algorithms and domain specific analysis that can confuse skeptical views and squelch skeptical behaviors (Michael, 2006). Reasoning does not occur in a dichotomous fashion between structure and content, but rather on a continuum and understanding auditor reasoning requires understanding the interaction between objective, formal rules and subjective skepticism

(Gigerenzer & Hug, 1992). A significant amount of research suggests the reasoner must adopt a social contract, or conditional permission or obligation frame in relation to a rule to adopt the skepticism that standard setting agencies require (Cheng & Holyoak, 1985; Cosmides, 1989; Gigerenzer & Hug, 1992). In other words, to apply skepticism and test assertions and evidence properly, it is important for an auditor to be able to adopt the viewpoint of a party that may lose if the other party to a social contract fails to hold up to their side of the contract (Gigerenzer & Hug, 1992).

Social contract theory is a domain specific and evolutionary perspective of human reasoning relating to the need for cooperation and perceived costs and benefits of a proposition. An example of such a social contract would be one party granting access to information in return for another party promising to give something in return. Gigerenzer and Hug (1992) purported simply adopting a social contract view is not sufficient to achieve effective hypothesis testing. They asserted the critical determinant is pragmatic, not semantic, and the reasoner must be able to move beyond the meaning of the rule and adopt the perspective of the party who is cheated. Gigerenzer and Hug (1992) suggested human reasoning does not occur based on rational, content independent logic as demonstrated by Wason and Johnson-Laird (1970), and reasoning is governed by task content and not its formal structure. However, the audit domain is embedded with content independent rules that the expert auditors must navigate.

Implicit Theories and Expert Reasoning

Problem framing and implicit beliefs can significantly affect reasoning in the audit context (Dweck et al., 2004; Tenbrunsel & Messick, 1999). While standard setting agencies provide guidance to auditors regarding skepticism, auditors are also required to follow formal, complex protocol. Audit is based in rules and standards which can lead auditors to engage in a compliance mindset, an inward orientation focused primarily on complying with auditing standards, not a social contract frame (Dweck et al., 2004). This inward orientation may be at the cost of an external orientation focused on the formal meaning of rules instead of content, the reasonableness of evidence and managements assertions (Dweck et al., 2004; Michael, 2006). Personal beliefs such as whether an individual believes implicitly in a growth versus a fixed mindset of intelligence is generally shown to correlate with success in learning and motivation, however Good and Dweck (2006) found environment can undermine personal beliefs. A compliance mindset and climate of fixed ability causes a reasoner to engage less with content, and be more sensitive to negative feedback and personal negative outcomes. This has evolved as a self-protective measure and motivational effects the compliance mindset has on cognitive strategies is underestimated (Good & Dweck, 2006). Good & Dweck (2006) found a fixed mindset is related to lower levels of effort on challenging tasks, but Bame-Aldred (2011) found no support for a relationship between implicit beliefs of intelligence and task effort or performance for auditors.

The failure of the research to find this relationship is potentially because of a mismatch between the mindset or reasoning frame of the participant and the environmental frame of the task (Fischoff, 1982). Framing related to a task produces alternate descriptions of that task, not by altering the content, but by spurring a shift in the way a reasoner thinks about that task (Jamal,

Johnson, & Berryman, 1995). The framing may influence the decision goal, which could explain inconsistent decision outcomes on identical task content between and within reasoners (Mueller & Anderson, 2002). There may be a basic misunderstanding in that researchers do not design the research task in the frame the reasoner typically uses when performing that task, and standards may fail to produce desired results because standard setting agencies may design policies without considering complex reasoning frames auditors use in the real world (Fischoff, 1982; Smith & Kida, 1991).

Researchers should be careful to examine individuals by having them perform tasks in the frame that they normally perform them and be cautious in generalizing the behavior as realworld audit environments are complex (Kleinmuntz, 1985; Smith & Kida, 1991). Biases are likely to reduce with a match between the task and the participant (Bame-Aldred, 2011). How the task environment is framed is key to the validity of research results and understanding the reasoning process (Fischoff, 1982). If the task environment is more abstract, such as a compliance and a rules-based environment, and the reasoner is primed with an inward oriented, compliance mindset then changes in judgment and reasoning are more readily manipulated and observed by changes that are related to the environment and frame, such as linear versus non-linear reasoning frames and using tasks that are more abstract than based in social contract theory (Fischoff, 1982).

Dual ProcessTheory

Dual process theory (Evans & Over, 1996a; Kahneman & Frederick, 2002; Stanovich & Thompson, 2001) provides a structure to examine expert auditor reasoning (Baumeister et al., 1998). The dual process cognitive framework is described as coexisting systems of implicit and explicit reasoning; however, some researchers believe it is better to focus on functionality rather

than consciousness. Researchers describe a dual cognitive framework involving two systems, known as system one and system two in the literature, that interact in the decision-making process (Kahneman, 2003; Kahneman, 2013; Stanovich & West, 2000). The dual system is thought of as two operating systems that communicate like a software environment (Morewedge & Kahneman 2010). System one is generally considered instinctive, fast, intuitive, emotional and subconscious. System one relies on associative learning processes and is more likely to be subject to bias (Evans, 2003). System two is slow, logical and conscious, and monitors heuristic operations of system one. System two functionality is described as checking the validity of automatic system one intuitions (Kahneman, 2013). System two requires effort which has important implications in the reasoning process as this effort, as well as the stress it causes, varies amongst reasoners (Baumeister et al., 1998).

Decision making errors occur when system one generates an intuitive error that is not detected by system two or when system two falsely detects an error where none exists in system one intuition (Morewedge & Kahneman 2010). These cognitive errors occur due a lazy system two, or an oversensitive and overactive system one (Baumeister et al., 1998). Intuitive judgments are involved with many systematic biases related to automated associations identified by system one. Recognized stimuli spur self-reinforcing patterns of associative memory that may affect reasoning frames and attention that result in biases. System one is perceptual; it can interpret and create complex representations but is incapable of rule-governed operations (Morewedge & Kahneman 2010), hence system two is capable of abstraction and is required to effectively process abstract problems (Evans & Curtis-Holmes, 2005).

An understanding of reasoning inconsistencies is gained from the dual process perspective between our intuition and our logical reasoning processes (Kahneman, 2002).

Tversky and Kahneman (1971) found sophisticated statisticians routinely committed obvious reasoning errors related to sample size despite their expertise in the field. Experts intuitively accepted small sample sizes as having much more power than their statistical knowledge should allow them to believe. The failure to recognize errors of expert intuition continues to be a stubborn problem as the operation of system one is difficult to control and relies on the effortful invocation of system two to monitor intuitive errors.

A central feature affecting intuitive judgment is accessibility; why some thoughts come to mind with greater ease than others (Higgins, 1996). Categories, labels, descriptions, and dimensions employed to make associations are variable in their accessibility depending on the context of the situation and the mood state of the reasoner at the time of the judgment. Accessibility operates on a continuum from those associations that are learned, implicit, and automatic, to cognitive operations that require effortful logic. Reasoners typically require explicit recognition of a purpose to undertake effortful logic. Accessibility of thoughts relies on attentional attributes such as perceived salience of stimuli, situational context, and priming effects (Higgins, 1996; Kahneman, 2013).

Automatic attentional operations are overridden by deliberate attention such as intentional attention to a stimulus (Kahneman, 2103). Attention is also drawn more to those stimuli that match motivations or arouse emotions. During times of high motivational or emotional arousal, accessibility will be easier for those stimuli that relate to the arousal (Kahneman, 2002). Framing effects also impact accessibility, as some aspects of a stimuli occur automatically while others require conscious computation. Kunreuther and Slavic (2002) introduced evidence for the affect heuristic proposing every stimulus spurs an affective

evaluation which occur automatically and are substituted in place of stimulus attributes that require more complex evaluation.

Depletion: Effortful Reasoning as a Limited Resource

Mental fatigue plays an important role in expert auditor reasoning processes and cortical activity related to dual process theory. Baumeister, Bratslavsky, Muraven, and Tice (1998) asserted system two processes including self-regulation, persistence at complicated tasks, making complex choices, and the inhibition of automatic responses represents expenditure of a limited cognitive resource not dissimilar to expending stores of glucose in skeletal muscles. Mental energy becomes depleted when burdened and the depletion of self-regulatory energy is detrimental to cognitive performance. Fatigued reasoners face a persistence obstacle as they engage in problem solving. They tend to give up more quickly and rely on normative decision tools. When reasoners are primed with an initial energy-depleting act, subsequent task performance is impaired (Baumeister et al., 1998). Schmeichel, Vohs, and Baumeister (2003) found participants asked to follow attentional rules reported increased task difficulty and exercised more regulatory exertion as opposed to participants who were not given rules. Researchers reported the depletion negatively affected performance on complex tasks and reading comprehension, but not simple tasks. Mental energy is replenished however, through mental rest or prompts that raise the mood of the reasoner (Tice, Baumeister, Shmueli, & Muraven, 2007).

Kahneman (2013) suggested this depletion results in the reasoner's inability to escape system one heuristics and have heightened susceptibility to stereotypic and superficial reasoning strategies. Stanovich and West (2000) found that reasoners vary in their ability to activate and sustain system two reasoning and describe system two as containing two distinct sections,

algorithmic reasoning, and rational reasoning. Algorithmic reasoning is complex and deals with computations and slower operations. People better at algorithmic reasoning usually do well on standard IQ tests, but nevertheless are susceptible to reasoning biases. Stanovich and West (2000) asserted that judgment biases are failures of rationality and related to superficial reasoning strategies and a failure of the reflective mind. The researchers say that many reasoning errors are not due to lack of intelligence or working memory issues but rather a tendency to engage in superficial reasoning (Stanovich & West, 2000). Superficial reasoning occurs when reasoners fail to invest the necessary energy to double-check default responses produced by system one (Kahneman & Frederick, 2002). Per Frederick (2005), reasoning errors related to superficial reasoning are not simply the fault of individual differences in standard intelligence, but are related to a reasoner's tendency to engage in reflective thought processes.

Heuristics and Biases, Prospect Theory, and Ambiguity Theory

A significant body of research indicates reasoners are susceptible to heuristics and biases and reasoning heuristics are used as an adaptive response to situational variables (Klayman & Ha, 1987). Studies indicated auditor judgment is susceptible to many of those same biases including anchoring and belief adjustment, confirmatory behaviors, and poor inferences regarding probabilities and a tendency to implicitly trust smaller sample sizes than is appropriate (Smith & Kida, 1991). Context is an important modifier however, and generalizability is a problem related to research in this area. Auditor judgment research indicates that expertise and task realism is correlated with better reasoning performance; novices are more susceptible to biases than are experienced auditors. Experience level and task familiarity tend to mitigate bias but results are inconsistent across environments and individuals.

An important facet of auditor reasoning is the development of expectations regarding the reasonableness of evidence in regard to what it is purported to indicate. During this process, auditors are prone to collecting significant amounts of evidence that also includes irrelevant evidence as part of the decision making process (Peytcheva, 2013). The consideration of irrelevant evidence raises audit risk, the chance that the financial statement contains a material misstatement when the audit opinion states financial statements are free of material misstatement (Pike, Curtis, & Chui, 2013). Pike et al. (2013) found auditors rely on irrelevant information in the development of analytical procedures that can result in the incorrect acceptance of material misstatements by gathering information that supports their expectations rather than contradicts them.

A factor influencing the evidence gathering process is availability, or ease of retrieval of information that is a system one phenomenon (Kahneman, 2013). Standardized procedures are most easily retrieved from a process perspective. From a content perspective, auditors are drawn to evidence depending on the ease of recall of applicable audit rules or recall of past interactions with similar data or situations that may cue an emotional response, such as receiving reprimand for an error related to the evidence. Ease of recall of evidence and ease of retrieval of applicable rules is more important than the amount or quality of the content involved in the recall. As auditors are pressed to recall a significant amount of information, the ease of recall naturally diminishes, resulting in the auditor feeling less effective. The auditor then may feel a need to collect more evidence, which frequently includes less relevant information (Kahneman, 2013).

Counter-intuitively, the diminishing ease of retrieval can result in individuals with a greater amount of recall feeling less secure, cueing an inward orientation. Auditors are more likely to over-emphasize the importance of ease of retrieval subconsciously in lieu of content

when they are required to engage in tasks requiring significant effort. The over-emphasis is stronger when the reasoner has powerful decision aids and tools at their disposal because they will attempt to escape the discomfort of uncertainty by seeking the safety of familiar structure and tools. Fatigue from the complicated structure of rules used by auditors may cause an individual to align with system one (Baumeister et al., 1998). Individuals guided by system one are more likely to fall prey to availability bias (Kahneman, 2013).

When auditors become aware of unaudited values prior to performing their substantive work, they are anchored to this value instead of updating their beliefs appropriately. The anchor influences auditors' evidence seeking behaviors who tend to gather information that is more likely to support the anchor value even though this information should be irrelevant to the formation of beliefs and expectations (Pike et al., 2013). In these cases, auditors are participating in a logical fallacy of circular reasoning, where the reasoner begins by assuming what they are attempting to prove. Auditors do not update their beliefs appropriately because they fail to consider significant competing or alternative information (Pike et al., 2013).

Confirmation bias is widely recognized as a problem related to auditor reasoning and while a significant amount of research has occurred on confirmation bias in auditing (Peytcheva, 2013), some studies regarding auditor judgment and evidence searching strategies do not support the assertion experienced auditors are biased toward confirmatory evidence (Smith & Kida, 1991). A study by Butt and Campbell (1989) that examined hypothesis testing strategies of auditors manipulated beliefs regarding the quality of an internal control system, were presented with positive and negative evidence, and then explicitly directed to seek confirmatory or disconfirming evidence or received no instruction. Results suggested experienced auditors do

not engage in confirmatory evidence search strategies within familiar contexts unless specifically directed to do so.

Other studies that examined the presence of confirmatory strategies in hypothesis testing found changing the direction of the hypothesis or posing it in either a positive or negative frame did not cause auditors to recall more confirmatory information (Anderson, 1989). Smith and Kida (1991) and Peytcheva (2014) suggested this is because experienced auditors are already cued to be more sensitive to negative data that may nullify the confirmation bias in their information search strategies. They further suggested auditors revise their sequential beliefs more through exposure to negative data than positive data, which may is caused by conservatism due to serious consequences of auditor judgments.

The risk of consequences that may come from failure to follow standards and guidelines act to frame auditor judgment and behaviors culminating in conservatism, adopting a cautious, suspicious view and maintaining established traditions. Conservatism may result in auditors habitually posing hypotheses in a negative frame (Smith & Kida, 1991). These observations may explain differences between behaviors of novices and experienced auditors and why characteristics of the audit environment are important in information search and reasoning processes. Klayman and Ha (1987) purported much of what is investigated as confirmation bias is looked at rather as a positive testing strategy. It is not the case that individuals are necessarily trying to verify their current beliefs, as much as they are naturally drawn to test cases that include the property of greatest interest, rather than those that lack that property.

An important factor determining the preference of a reasoning strategy is when the greatest concern of the reasoner is judgment error (Klayman & Ha, 1987). Smith and Kida (1991) reported that serious risk of consequences from judgment errors results in conservatism

and great concern regarding negative information. They further found the use of skepticism prompts did not give strong support for confirmatory biases, they did find positive or negative framing did affect data search preferences and suggested the choice of strategy depended on the cost structure associated with task elements. While Smith and Kida (1991) suggested this conservatism may cause auditors to habitually frame problems in negative terms, Klayman and Ha (1987) reported this concern leads to a positive test strategy directed toward elements of greatest importance or concern. In a conservative frame, the default approach would be a tendency toward normative strategies perceived as a quicker and safer way to avoid costly errors. In a conservative environment, an abstract task may frame problems in a less threatening format than a social contract frame, indicating focusing on metacognitive aspects of reasoning may have more power in affecting reasoning performance than a focus on evidentiary skepticism. Einhorn and Hogarth (1978) stated an important conflict exists in the validation process between strategies that assist in belief updating related to long term success and heuristic strategies perceived to maximize short term success.

As a critique of expected utility theory, Kahneman and Tversky (1979) offered prospect theory. Prospect theory assigns values to gains and losses associated with a decision instead of assigning values to final outcomes. Authors noted reasoners display inconsistent decision processes when the same problem is posed in different forms, indicating people respond differently to gains and losses, not in a uniform manner, as rational theories, such as expected utility propose. Prospect theory purports reasoners give more weight to losses and evidence framed negatively than gains or evidence framed positively. Reasoners also tend to overweight options perceived as certain over those that contain probabilities.

Additional opposition to expected utility theory is offered by Einhorn and Hogarth's (1988) ambiguity theory, a formal decision model that suggests as prospect theory does, reasoners follow a quantitative series of calculations in decision making. The authors asserted expected utility theory fails to address how reasoners deal with uncertainty, contextual differences, and effects of probabilities. Ambiguity theory suggests people deal with uncertainty in different ways that results in lack of uniformity in decision making, and asserts people use an anchoring and adjustment approach. The formal model purports reasoners begin with a starting point that acts as an anchor, and reasoners adjust for ambiguity. Reasoners continue the process by using available information or probabilities, updating until a decision is made (Einhorn & Hogarth, 1988; Kahneman, 2013).

Belief Updating and Expert Reasoning Processes

The belief-adjustment model (Hogarth, & Einhorn, 1992) is a belief updating theory of decision making that has provided theoretical support for much research in auditing and accounting and provides additional understanding of reasoning processes related to dual process theory (Arnold, Collier, Leech, & Sutton 2000; Kahneman, 2013). The belief-adjustment model addresses the interaction between task characteristics and strategies of information processing and particularly primacy, recency, and order effects. Order bias occurs when the conclusion of the decision-maker is affected by the order in which information is presented. Recency bias occurs when there is an over-reliance on information presented last and primacy bias is an over-reliance on information that is presented first (Arnold et al., 2000). The model incorporates two types of processing, either a step-by-step approach, where the decision maker updates beliefs with each piece of information, or an end-of-sequence approach where the updating is completed after all information is gathered. The model predicts decision makers will over-rely on

information given last when the problem is complex. Task complexity is defined in relation to task familiarity and information load. The model suggests the decision-maker's attitude regarding the evidence is an important consideration as are prior beliefs held by the decision-maker and how new information is interpreted in relation to those beliefs.

Previously held beliefs are anchors that are updated based on the attitude toward new information. Auditor attitudes are likely to affect many decision processes, including information search strategies, their beliefs regarding new evidence, and recall of previous evidence, and evaluating the strength of evidence (Bamber, Ramsay, & Tubbs, 1997). Hogarth and Einhorn (1992) suggested an understanding of contrasting sensitivities, particularly to confirming or disconfirming evidence would be important in understanding decision behaviors. These beliefs, attitudes and affect or mood of the auditor are influenced by not only their own trait characteristics, but also the state of the environment, and specific context in which they are immersed (Johns, 2006).

Studies have shown order bias is mitigated by experience, but typically, those decision-makers have used end-of-sequence approach (Messier & Tubbs, 1994; Trotman & Wright, 1996). When using a step-by-step approach such as a checklist, studies indicated recency bias occurs even with experienced participants (Asare & Wright, 1995; Krull ,1993). Many decision aids employed by auditors are used in a step-by-step fashion which may make them more susceptible to order bias. Kahneman (2013) suggested decision-makers would naturally and implicitly tend toward the simplest models of decision making, expending the least amount of energy necessary to draw a conclusion. Decision makers who desire the most certainty are the most likely to draw on quick and simple models of decision making; auditors and accountants typically rank high on need for certainty, suggesting auditors may be more likely to have a

stronger attraction to checklist-style reasoning tools and are thus more likely to fall prey to biases, such as the recency bias. Per Bamber, Ramsay, and Tubbs (1997), this makes auditors more prone to confirmation bias, or the tendency to seek information that confirms rather than contradicts their beliefs.

Naturalistic Decision Making and the Recognition Primed Decision Model

Dual process theory and the heuristics and biases model (Tversky & Kahneman, 1975) have received some criticism as not effectively representing decisions faced in real world situations, and naturalistic decision making (NDM) (Lipshitz et al., 2001) was offered as a contrasting model (Kahneman & Klein, 2009). Real world, expert decision making shares common themes across a wide variety of professional disciplines that include facing decisions under uncertainty, severe time limitations, high personal risk, and a dynamic environment (Lipshitz et al., 2001). Examining this type of decision making defies experimentation in controlled laboratory settings, or using novices as study participants, because the complexity of the real-world environment and cognitive processes of trained and experienced experts are the main features of the discipline of study which are not captured outside of the real-world environment or by using non-experts. Naturalistic decision making (NDM) is a popular approach for examining real world expert decision making (Bryant, 2002). NDM avoids formal decision models, tools, and attempts to examine decision making in real world contexts where the situation is familiar and situational cues are meaningful based on experience and learning. Reasoners routinely violate rules of rational choice; normative and classical models of decision making that assert decision makers compare competing choice options against one another in a quantitative weighting manner.

Among those models considered formal choice models by NDM are prospect theory (Kahneman & Tversky, 1979) and the Ambiguity model (Einhorn & Hogarth, 1986). These models suppose different models of comparison in choosing amongst a set of competing alternatives. Naturalistic decision making theory asserts instead, reasoners compare characteristics of a response to characteristics of the situation to determine a match, and the expert reasoner is not comparing alternatives against each other alternative and calculating the best result. NDM suggests reasoners rely on cues of orientation and matching characteristics of the situation to appropriate actions in response. This infers novices will respond differently than experts because situational cues will cue different meanings. Because of this, NDM emphasizes the study of expert behavior in real world settings (Pruitt, Cannon-Bowers, & Salas, 1997).

NDM does not focus on predicting process options that are implemented, but rather seeks to be descriptive of processes expert decision makers employ. NDM emphasizes the importance of the information sought by decision makers as well as how they interpret it and the decision process employed, not predictions based on formal decision models, or expected rational decision processes. The model relies on explanations of recognition and pattern matching, not details behind reasoning processes (Bryant, 2002). NDM theory suggests many different cognitive processes occur simultaneously, which are highly context dependent, and people do not consciously weigh their options and make choices based upon a purely logical process, as utilitarian theories suggest (Kaempf, Klein, Thordsen, & Wolf, 1996).

The NDM model suggests decision makers use a matching model to form choices by comparing choice options to a standard, not by comparing options to each other. NDM suggests expert decision making is highly context dependent and domain specific (Lipshitz et al., 2001). Bryant (2002) explained NDM is a holistic approach rather than a fragmented comparison of

alternatives, and it relies on recognition of environmental characteristics and matching them to known patterns of response. NDM suggests since a rapid response is frequently necessary, reasoners attach to a workable solution quickly rather than spending time evaluating a large group of options amongst each other. NDM rejects classical decision models such as the economically based rational choice model, suggesting people do not engage in a formal, conscious reasoning process that scores and compares options against one another.

NDM is more thoroughly understood by the Recognition-Primed Decision Making model (RPD) (Klein, Calderwood, & Macgregor, 1989) which is an archetypic model within the NDM framework (Bryant, 2002; Klein, 1997). RPD suggests reasoners begin the decision making process by assessing the familiarity of a situation and matching features of a current situation to characteristics of prior events. Where characteristics fail to cue a match, the reasoner creates a story that helps to explain the situation, and then builds mental simulations to form expectations and will update and adjust beliefs and expectations as inconsistencies are encountered. Accordingly, the familiarity with the situational context surrounding a decision is a significant factor influencing the success of decision making (Lipschitz et al., 2001). Higher levels of cognitive load mean there is less energy available for comparative reasoning, and quick recognition becomes important, leaving no room or time for carrying out formal, computational approaches to decision making in the real-world environment (Bryant, 2002). There are three approaches to RPD. In the simplest form, experts assess the typical situation and can usually generate a successful response on the first try. In a second form, when uncertainty exists, the reasoner will build a mental simulation of the situation and the possible response and decide whether the response is appropriate. In the third approach, experts simulate the situation and

response, but go a further step by assessing possible unintended consequences (Lipschitz et al., 2001).

The RPD model purports due to the quick orientation and interpretation of situational characteristics, time pressure effects on expert reasoners is minimal compared to novices and more time to make the decision is unlikely to improve decision efficacy among experts. When errors do occur, the RPD model says they should not immediately imply heuristic errors as suggested by heuristics and biases paradigm of Kahneman and Tversky (1979). According to Bryant (2002), heuristics and biases paradigm focuses on errors as a negative even though errors are unavoidable to some extent in unstable environments. Efforts to attain an error free environment are maladaptive because errors serve an important source of learning in dynamic environments in which experts operate.

Paradigms of heuristics and biases (Kahneman & Tversky, 1979) and NDM (Lipschitz et al., 2001) are often viewed as conflicting because heuristics and biases are skeptical of expert intuition and typically focus on judgment errors and NDM focuses on judgment successes (Kahneman & Klein, 2009). Authors noted however, the two paradigms have several important points of agreement. First, both models agree expert reasoning can be good in some circumstances, and faulty in others. Both models also demonstrate the importance of making a distinction between intuitive skill and overconfident assumptions. Both models agree system one and system two are a central feature of expert decision making, and system one represents intuitive judgment. Both models support a stable task environment yields the best expert intuition, when cues and characteristics of the environment have a stable connection to subsequent events. Outcomes are difficult to predict even for experts, in environments of low

stability and validity. The models also merge on the idea that experts are typically unaware of cues that form and guide their actions.

The models further align with the idea experts are usually aware when they do not know something, while novices are usually unaware. The authors agree the quality of intuitive reasoning is tied to the predictability of the task environment and opportunities afforded the reasoner to learn norms and characteristics of the environment (Kahneman & Klein, 2009). The most prominent difference between the two approaches is the focus on intuitive successes versus failures. Experiments are approached differently as well, as heuristics and biases approach typically employs controlled experiments with non-experts frequently used as participants. The NDM approach prefers to use active practitioners in real world environments to study decision making. Results from differing approaches can provide useful comparisons that inform both models and highlight needs for future research (Kahneman & Klein, 2009).

Personal Need for Structure

Reasoning relies upon constructing meaning through a coherent structure of associations with which to understand an environment and build expectations (Kahneman, 2013; Thompson et al., 1989). Environments that challenge one's associational structure can produce an unsettled state of uncertainty in which people are motivated to reduce the discomfort associated with that uncertainty. Individuals have different cognitive styles and preferences in relation to information gathering and decision making. Individuals also experience varying levels of cognitive dissonance and discomfort when faced with cognitive inconsistencies and ambiguities that result in different individual needs for cognitive closure (Webster & Kruglanski, 1996). According to Slijkhuis, Rietzschel, and Van Yperen (2013), there is significant overlap between need for structure and need for cognitive closure with people higher in need for structure and closure

tending to more quickly accept evidence that matches expectations, reduces dissonance and allows closure.

Expert decision makers in audit domains operate in an environment that is complex and data intensive, frequently pushing people beyond their attentional limits and cognitive capacities. In response, reasoners attempt to simplify the structure they face to obtain a cohesive understanding of the world. Standardized approaches and norms constructed by the professional domain help decision makers simplify complicated situations. In an automatic effort to minimize the expenditure of cognitive assets, reasoners fall prey to heuristic responses learned through the environment that can deviate from logical best practices (Rossi, Cassotti, Moutier, Delcroix, & Houdé, 2015). The rapid evolution of professional life makes it more important than ever to examine the ability of domain experts to be adaptable to changing situations that are negatively impacted by standardized approaches developed to assist in the efficiency of decision making (Hamtiaux & Houssemand, 2012).

The cognitive load many decision makers experience can cause uncertainty and discomfort and should reduce to a level that is processed efficiently. Individuals vary in levels of discomfort with uncertainty and their desire for simple structure, which has important connections to the way they process information as well as fundamental aspects of behavior and psychological functioning. Reasoners reduce cognitive load through a variety of avoidance strategies or cognitive structuring and restructuring strategies (Neuberg & Newsom, 1993). Avoidance strategies limit exposure to the stimuli in the environment that is achieved by either ignoring information or implementing barriers that stop stimuli from intruding on attention. The reduction of information is also achieved by employing standardized scripts or formalized schemas to reduce complexity and quantity of information. Cognitive structuring utilizes

constructs, or abstract mental representations such as checklists or standardized methods to simplify complex situations and allow generalizations which maximizes cognitive efficiency and minimizes the expenditure of cognitive resources. These generalizations may assist with efficiency in cognition, but may not apply accurately or translate evenly between contexts (Neuberg & Newsom, 1993).

The Personal Need for Structure (PNS) scale developed by Thompson, Naccarato, and Parker (1989) and introduced by Neuberg and Newsom (1993) is a self-report scale where the subject agrees or disagrees with statements about their views of the environment on a six-point scale. The PNS scale focuses on how individuals process ambiguous information within their environment. The theoretical foundation of the PNS is rooted in an assumption that coping mechanisms related to cognitive load or overload is related to the ability to reduce ambiguity, and that an individual's reaction to ambiguity or lack of structure significantly influences the way people understand and experience the world. A negative correlation is generally assumed between creativity, flexibility, and a higher personal need for structure. A positive correlation is expected between rigidity, anxiety, neuroticism, and a higher personal need for structure (Ruiselová, Prokopcáková, & Kresánek, 2012). These researchers found there was a negative response when individuals with a high Personal Need for Structure were exposed to counterfactual reasoning models, but they did not find that the need for structure influenced the frequency of counterfactual reasoning. These findings suggested problem solving and critical reasoning processes may not have simple frequency correlations to the need for simple structure, but there may be complex interactions in cognition related to the consideration of hypothetical alternatives in the problem-solving process.

Individuals with a higher need for structure and certainty are threatened by uncertainty and have strong emotional responses to anything that disrupts standardization, prefer the simple organization of data, are less likely to change their attitudes, and are more likely to rely on stereotypes. These individuals prefer simple heuristics, pre-formed categories and methodologies, and are likely to respond to failure or unforeseeable consequences with learned helplessness (Ruiselová et al., 2012). Uncertainty and awareness of personal inconsistency in reasoning can be a self-threat that results in compensation through defensiveness, authoritarianism and extreme, unfounded conviction to a position to defend against threats to self-integrity (McGregor, Zanna, Holmes, & Spencer, 2001).

Certainty is desired more in states of ambiguity, anxiety, and fear, including when there are rule structures that infer sanctions that are perceived as self-threats (Michael, 2006; Tenbrunsel & Messick, 1999). Categorization provides structure and high PNS individuals have a higher desire to categorize than low PNS individuals (Moskowitz, 1993). However, evaluative aspects of the PNS may have negative effects on motivation and creativity when they are perceived as controlling as opposed to informational. Informational evaluation is associated with higher levels of motivation and better creative performance, but this finding was only identified among individuals low in PNS (Slijkhuis, Rietzschel, & Van Yperen, 2013). Furthermore, high PNS individuals are more likely to accept weakly validated information and feedback because of the possibility that it may reduce ambiguity and be less critical in the evaluation of information and the acceptance of irrelevant information throughout the decision process (Slijkhuis et al., 2013). The increasing need for structure is reflected in higher levels of conscientiousness and a desire to be thorough but is also related to a fear of personal invalidity and higher levels of neuroticism and a tendency to focus on errors and to be self-conscious and overly self- critical,

resulting in a narrowing of attention toward expected patterns and standardized responses (Michael, 2006; Thompson, 1998).

Neuberg and Newsom (1993) and Personal Need for Structure scale developers, Thompson et al. (1989) said the need for simple structure is based in part on a motivation to avoid ambiguity and reduce complexity. Webster and Kruglanski (1994) suggested these motivations are related to personal disposition and are separated into different variables including an individual's preferences for structure and order, levels of desire for predictability, and response to ambiguous situations, and closed-mindedness.

Higher Personal Need for Structure is associated with a variety of reasoning errors. Schaller, Boyd, Yohannes, and O'Brien (1995) found need for structure significantly influenced inferential reasoning processes and high PNS participants used overly simplistic reasoning strategies in comparison to low PNS participants, and participants with a high PNS were more likely than those with a low PNS to make attributional errors and to make erroneous group stereotypes. High PNS individuals are more likely to dismiss information that is inconsistent with standardized schemas (Hess, Follett, & McGee, 1998), and are less likely to update their knowledge with new information (Okun & Rice, 1997; Hess, 2001). Anxiety, including math anxiety reduces working memory capacity and interrupts efficient cognitive processing (Ashcraft & Kirk, 2001). Elovainio and Kivimaki (2001) suggested job characteristics are associated with role ambiguity and PNS has a moderating effect on role ambiguity. Higher PNS is associated with higher role ambiguity and the lower the feeling of control the higher the feeling of role ambiguity that is associated with occupational strain. Occupational strain and ambiguity result in a higher need for structure.

Part of natural heuristics of reducing overwhelming amounts of data is the automatic human preference to seek validating information for our beliefs over invalidating information (Hallihan & Shu, 2013; Kahneman, 1979). Confirmation bias, or the tendency to seek and interpret new information in a way that confirms one's preexisting beliefs, is a bias that limits reasoning effectiveness amongst auditors in a wide variety of contexts (Peytcheva, 2013). While the phenomenon is widely studied, there has been a lack of action on the part of those who design investigative procedures, and criteria to make changes that limit undesirable outcomes (Hallihan & Shu, 2013). Hallihan and Shu (2013) said confirmation bias may distort an individual's concepts related to data evaluation processes not only in the analysis of data gathered in the investigational process but also in the design of investigational procedures and leads to the misinterpretation of evidence. Confirmation bias results in design fixation and the restriction of ideas as concepts are developed and restricting further examination and proper validation of ideas (Hallihan & Shu, 2013; Kahneman, 1979). Per Jermias (2006), confirmatory reasoning contributes to overconfidence and a resistance to change that creates a feedback cycle that resists invalidating evidence.

High PNS is correlated with a higher sensitivity to confirming evidence than disconfirming evidence (Bamber et al., 1997). Peytcheva (2014) purported evidence seeking activities of auditors and the cognitive processes involved in determining the relevance of evidence is a critical part of the auditor's reasoning process, and confirmation bias is a common error amongst auditors. Cosmides (1989) suggested individuals might perform better on tasks prone to confirmation bias if they perceive the task in the frame of a social contract rather than as an abstract problem. Comparatively, Michael (2006) suggested rules-based structures, such as those existing in the auditing domain create a more abstract environment that focuses more on

the letter of the rule, instead of its underlying spirit, damaging some aspects of the reasoning process.

Hypothesis Testing and the Wason Selection Task

Hypothesis testing is a routine duty for auditors, and thus an auditor's hypothesis validation strategy is important and using hypothesis testing in research is an effective way to examine auditor reasoning (Peytcheva, 2013). According to Evans (2003), and Rossi, Cassotti, Moutier, Delcroix, and Houdé (2015), the Wason Selection Task (WST) is recognized as an effective tool for the exploration of the dual system reasoning model and heuristics and biases framework introduced by the seminal work of Tversky and Kahneman (1975) and more recently by Stanovich & Thompson (2001), Stanovich and West (2000), Kahneman (2003) and Morewedge and Kahneman (2010). Reasoning research examining the hypothesis testing effectiveness of auditors however, has yielded notably inconsistent and contradictory results (Gigerenzer & Hug, 1992). Little progress was made in the understanding of reasoning processes or mitigating reasoning errors due in part to the presupposition that a simple dichotomy exists between structure and content, but no simple delineation exists between the content independent rules and the content that they address, or between which evidence is relevant or irrelevant in rational reasoning processes (Gigerenzer & Hug, 1992).

Past research focused on rational and formal logic and purported content independent logic provided the best reasoning (Arnold & Sutton, 1997). Content-independent formal rules however, are unable to describe human reasoning fully (Kahneman & Tversky, 1982). Some research suggested deontic frame is required (Oaksford & Chater, 1994), or that a social contract frame is important (Cosmides, 1989), while other research suggested domain specific or pragmatic framing is what influences auditor reasoning (Cheng & Holyoak, 1985). Still other

research found reasoning is best when the frame of the reasoner matches the frame of the hypothesis (Fischoff, 1982; Smith & Kida, 1991). Researchers examined how content-independent formal rules and content attentive framing affects reasoning processes and found reasoners experience dissonance between the symbolic, logical structures of rules and the content of the evidence (Gigerenzer & Hug, 1992).

An effective tool for examining dual process reasoning in auditor hypothesis testing processes is the Wason Selection Task (Evans, 2007; Peytcheva, 2013; Wason; 1966; Wason, 1968). The Wason Selection Task (WST) is especially informative when investigating reasoning processes because participants respond to slight variations in framing and content (Dawson et al., 2002). The WST is particularly sensitive to evidence seeking biases and is an excellent experimental tool to examine reasoning in auditors because it models similar choices to those auditors must make as a typical part of their work (Peytcheva, 2013). The WST has been one of the most investigated and intriguing problems in reasoning psychology (Evans & Over, 1996; George, 1991). The WST is a robust challenge to economic theories of rational reasoning due to the ability of the simple task to draw an overwhelming percentage of participants to make logically irrational choices (Oberauer & Wilhelm, 1999).

Peter Wason introduced the famous reasoning assessment (Wason; 1966; Wason, 1968) known as the Wason four-card selection task that has become one of the most popular experimental tools for studying human reasoning (Almor & Sloman, 1996; Klauer, Stahl, & Erdfelder, 2007). The WST has several versions, but one of the more popular versions is an abstract task that uses four cards presented to participants that have a letter on one side of the card, and a number on the opposite side. The task is known to follow what is referred to in the literature as the if *p* then *q* format yielding possible selections of *p*, *not p*, *q* and *not q*. Two of

the four cards showing have a number exposed while the remaining two cards have a letter exposed. The cards read in order A, K, 8 and 5. Participants are given a conditional rule associated with the cards in an if-then format*, if there is a vowel on one side of a card there will be an even number on the other side of the card* (Wason, 1966; Wason, 1968).

Participants are informed the conditional rule may be either true or false. Participants are asked to turn over the minimum number of cards that best help them determine whether the rule is true or false. Logically the participant should choose the A and 5 or *p* and *not q* cards representing affirmation of the antecedent and falsification of the consequent respectively as this is the only combination of cards that have the power to falsify the rule (Rossi et al., 2015). The antecedent is the first part of a conditional rule and the consequent is the last part of the conditional rule. As an explanation of logical choice in the above example, the A card would be a logical choice because it has the power to affirm the antecedent, the K card, would have no power and is irrelevant because the rule does not involve consonants; the 8 card is a less logical choice, as it could affirm the consequent but also may reveal a consonant which is irrelevant; and the 5 card is the most logical second choice because it has the power to falsify the consequent. The proper way to solve the Wason Selection Task is to seek disconfirmation. People who are motivated to reject a task rule should generally display better reasoning than those who are motivated to confirm it. A skeptical mindset is most helpful when the skepticism is leveled toward reasoners reasoning processes and not only the evidence (Dawson et al., 2002).

The abstract version is content independent and based more in the formal rule structure. A social contract version of the WST might propose a rule such as *if a person is drinking alcohol, then that person is at least 21 years old*. The four cards would show *drinking alcohol*, *not drinking alcohol*, *at least 21*, and *not 21*, representing *p*, *not p*, *q* and *not q* respectively.

Some research suggests participants in this social contract version will focus on the content and are more likely to successfully recognize the logically correct choices of drinking alcohol and not 21, or *p* and *not q*. Researchers are intrigued by the low success rate on the task, which hovers around 20%, and frustrated while trying to find ways to improve performance regardless of intelligence levels of reasoners (Cosmides, 1989; Dawson et al., 2002; Griggs & Cox, 1982). Some success in boosting performance was found when using familiar tasks and altering instructions to frame the task in familiar, social contract format (Evans, 1996a). Research suggested a cheater detection prompt or skepticism prompt increases reasoning effectiveness (Mueller & Anderson, 2002), while other studies found this type of priming is effective in novices, but not expert auditors (Peytcheva, 2013).

A common bias experienced by expert auditors is the tendency to verify hypotheses by seeking confirmatory evidence that supports a given proposition (Peytcheva, 2013). Even those reasoners with no motivation tied to a proposition are biased toward verification of that proposition (Dawson et al., 2002). In addition to the tendency toward confirmatory evidence, people typically engage in motivated reasoning and seek evidence that provides the best outcome for them personally, an effect that is pronounced in outcome-oriented environments (Dweck et al., 2004; Tenbrunsel & Messick, 1999). People are more sensitive to evidence that correlates with a presented personal threat. While people wish to be objective in their reasoning, they also have an implicit tendency toward self-protection and evidence that is associated with the self (Dawson et al., 2002). A key issue in reasoning is the motivation behind evidence seeking activities. Reasoners who are motivated to support a given hypothesis are more likely to seek confirmatory evidence and to accept it quickly to avoid conflicting evidence that causes discomfort. Reasoners who are motivated to reject a given hypothesis are more likely to seek

evidence that disconfirms that hypothesis (Dawson et al., 2002). Accordingly, depending on their motivational orientation, reasoners may ignore disconfirming evidence as irrelevant or heavily scrutinize disconfirming evidence. Hurtt et al. (2013) found auditors tend toward seeking confirming evidence but this tendency reverses when the hypothesis is stated negatively.

Cheng and Holyoak (1989) reported reasoners typically frame abstract knowledge structures pragmatically rather than syntactically, straying from standard rational and logical frames, and participants tended to reason in a normatively logic frame when prompted with social contract related schemas. Researchers noted the irony of this pattern since standard logic does not contain a deontic frame using permissions and obligations. Their research found that training in standard logic did not assist reasoners except for a slight reduction in the error of affirming the consequent, and that formal logical training was unhelpful with errors of affirming the consequent or denying the antecedent. They concluded purely abstract training in formal logic is almost totally ineffective at reducing errors in reasoning. A critique was offered; however, in that reasoners may fall prey to bi-conditional bias cued by the if-then mode of the selection task in the abstract mode where there is a tendency to interpret the *if p then q* proposition as *p if and only if q*, which is typically not cued in the permission and obligation framing mode of the problem. Additionally, while authors noted participants were most successful in reasoning in permission and obligation schemas, those schemas were proposed with direction toward checking for a unilateral violation of rules rather than checking bilaterally, whether they were true or false, which may have altered reasoning outcomes. Cheng and Holyoak (1985) suggested their findings supporting the success of pragmatic schemas in reasoning are partly due to the difficulty of inducing formal deductive rules. They suggested psychologists such as Piaget might have been wrong about reasoners naturally having formal

deductive logic but correct about them having an inductive rule system, and that formal logic training is thus unlikely to improve deductive reasoning.

If researchers understand a behavior, that behavior must be reversible through experimentation (Kahneman, 2013). Per Gigerenzer and Hug (1992), a perspective change on the part of the reasoner from an inward orientation to the perspective of the parties in a social contract problem with bilateral cheating options resulted in an almost perfect reversal of *p and not q* responses to *not p and q* responses. Unilateral cheating options, where only one party can cheat the other party produced only partial success versus bilateral cheating options where both parties can cheat one another, presumably due to the reasoner receiving cues to the perspective of only one party to the contract instead of both. Cheng and Holyoak (1989) critiqued conclusions of Cosmides (1985) that social contract problems prompt a high p and not q responses over non-social contract problems because they asserted Cosmides used problems that lacked a symmetrical social exchange and are thus pseudo-exchanges that are more of a straight permission which is not a social contract. Rules that were considered social contracts by Cosmides (1989) were considered permissions and obligations by Cheng and Holyoak (1989). Gigerenzer and Hug (1992) argued concepts between Cheng and Holyoak and Cosmides are nearly identical and the key theoretical argument is whether a rule is a social contract, a permission, or some combination.

Gigerenzer and Hug (1992) concluded reasoning occurs on a continuum and not a dichotomy between content-dependent and content-independent theories and in regard to the extent to which a reasoner's adaptation to specific structures of an environment exists, an analysis of adapted reasoning mechanisms in that environment would be indispensable. Accordingly, expert auditors do not frame reasoning problems dichotomously, strictly dividing

between content-dependent and content-independent views, or strictly between inward and outward orientation but frame on a continuum that has adapted to the audit environment. It is important to note the audit environment has consistently saturated expert auditors with outward orientation via skepticism prompts while imbedding the environment with rules and sanctions auditors can avoid by following those rules. Auditors may tend to frame problems with an inward orientation due to personal effects involved (Dweck et al., 2004; Tenbrunsel & Mesick, 1999).

Considering reasoners view problems on a continuum between inward oriented, content independent, abstract rule structure, and an outward oriented, content-dependent frame (Gigerenzer & Hug, 1992), and reasoners implicitly tend toward evidence that holds the most personal importance (Dweck et al., 2004; Kahneman, 2013), auditors are likely to place significant importance and implicit attention on the abstract rule structure that dominates their environment (Bame-Aldred, 2011; Dweck et al., 2004; Gigerenzer & Hug, 1992; Kahneman, 2013). Furthermore, considering Cheng and Holyoak's (1989) finding that reasoners tend to frame abstract problems pragmatically, and use standard logic when prompted with social contract related schemas, expert auditors may respond more to prompts aimed at manipulating inward oriented reasoning frames such as priming dynamic or linear reasoning frames as opposed to outward oriented reasoning frames such as skepticism or cheater detection prompts when completing hypothesis testing tasks. Additionally, since the task structure can cue a frame of reasoning (Dweck et al., 2004; Kahneman, 2013), these inward oriented prompts may be significantly more effective at altering reasoning outcomes on abstract tasks than social contract tasks, which is important since expert auditors do not reason solely in an outward oriented, social contract frame. In such respect, variances in effectiveness of inward oriented priming on

reasoning outcomes between abstract and social contract based tasks could be indicative of where reasoners lie on the continuum between inward and outward orientation related to domain specific tasks.

Some researchers criticized the use of the WST in research, and particularly abstract versions of the task as unhelpful in understanding expert reasoning in real world contexts, arguing they do not represent a realistic task performed in real-world audit environments (Almor & Sloman, 1996; Klein, 1997). Inferred in this view is the assumption abstract frames are entirely separated from real world environments, but real-world environments embedded in an abstract environment can cue abstract reasoning frames (Tenbrunsel & Messick, 1999). Expert auditor reasoning occurs on a continuum between inward oriented, content independent and outward oriented, content-dependent frames, not dichotomously (Gigerenzer & Hug, 1992). Tasks in the style of the Wason Selection Task are effective tools for examining auditor reasoning mechanisms (Peytcheva, 2013). Reasoners make logical errors by using heuristics that result in deviations from logical norms and the WST is an effective tool to explore these mechanisms through system one and system two responses (Rossi et al., 2015).

Success on the WST entails an inhibition of heuristic responses of system one as well as activation of analytical procedures of system two that check the validity of automatic operations of system one (Rossi et al., 2015). System one reflects automatic operations spurred by domain specific learning, with conscious access only allowed to the product of those operations. System two represents cognitive networks supporting effortful, sequential processing required for hypothetical reasoning. Dual process theory (Kahneman, 2013), described previously, proposes mechanics with which two systems interact in reasoning processes. The operations of system one, born out of domain specific learning, knowledge and beliefs may produce correct responses

in most instances, but also create biases in complex situations where the analytical system needs to override the heuristic system to avoid errors in reasoning and inhibit the acceptance of irrelevant evidence or strategies (Rossi et al., 2015). This inhibition must not only inhibit the acceptance of irrelevant evidence but also of ineffective, automatic strategies that are triggered by complex or unfamiliar situations.

Rapid Response and Reasoning Processes

Expert decision makers in real world situations rely on rapid assessments of the situation and quick decisions based on heuristics (Lipshitz et al., 2001). Constraining the response time of reasoners while they solve problems will inhibit system two operations, which are necessary for effective abstract thinking, and make it more likely reasoners will fall prey to matching bias (Roberts & Newton, 2002). Superficial assessments of a situation and rapid heuristic decisions can lead to reasoning errors and biased decision processes (Kahneman & Tversky, 1979). According to Evans and Curtis-Holmes (2005), speed of processing is a critical aspect of dual process operation and a rapid, heuristic response increases the tendency for belief bias, or the tendency to evaluate the validity an argument not based on merits of the logic but whether the reasoner agrees with the conclusion. A study by these researchers found a condition requiring participants to respond within 10 seconds to a syllogism procedure like the Wason Selection Task experienced poorer reasoning efficacy than a free time version of the task, which supports predictions about reasoning processes based on dual process theory (Evans & Curtis-Holmes, 2005). The study found higher incidence of belief bias, which authors equated to matching bias on the Wason Selection Task where reasoners tend to select cards that match conditional statements regarding the rule. While results showed a clear tendency toward belief bias, they also indicated a logical component was still involved due to differing results among tasks of

varying believability. The conflict between belief and logic was of interest in the study, which is when experimental manipulations hold the most power to influence reasoning processes. The rapid response version of the WST clearly affects reasoners' processes, which tend toward heuristic processing (Evans & Curtis-Holmes, 2005). Information obtained prior to the task however, could change the frame of the reasoner and influence the cognitive processing of the problem (Dweck et al., 2004; Kahneman, 2013; Tenbrunsel & Messick, 1999).

Heuristic response errors are caused by attending to superficial aspects of the problem solving, minimizing cognitive efforts and avoiding conflicting responses between automatic processes of system one and the effortful and controlled processes of system two (Kahneman, 2013). Greater levels of cognitive self-control are needed not just to override heuristic errors of system one, but to go further and activate system two processes. It is possible to inhibit system one activities without spurring activation of system two (Rossi et al., 2015). The self-regulation required to inhibit heuristic errors in system one and spur the effortful logic of system two is metacognitive and an inward-oriented mechanism (Dweck et al., 2004; Kahneman, 2013).

System one operates at much faster speeds than system two (Evans, 2007). The difference in speed at which the dual process system works has brought about interest in researching the speed of decisions in real world environments and timing involved in solving reasoning problems (Klein, 1999; Rossi et al., 2013). Expert decision making in real world contexts relies on quick and efficient recognition processes drawn from experience and knowledge (Reyna, 2004). Real world decisions frequently occur without significant conscious deliberation (Klein, 1999). Crucial decision processes may occur in the very early stages of the reasoning process and superficial aspects of problems, including the relevancy and believability of evidence choices can direct reasoning processes beginning with unconscious system one

processes (Evans, 2007). Accordingly, rapid response tasks that require reasoners to respond in very short periods of time are particularly informative in expert domains.

Rossi et al. (2015) varied the amount of time participants had to provide answers on a selection task and found reasoners' responses in their rapid response time task did not significantly differ from the free response time version; although, they did note an increase in the selection of a matching card in the consequent of the task that represented a less logical choice in the rapid response treatment. The rapid response treatment seemed to activate system one but the free response time version did not seem to activate system two (Rossi et al., 2015). Individual differences in cognitive style are an important consideration in this regard (Evans, 2007). Mismatching errors in the selection task occur due to the interpretation of non-matches as irrelevant and reasoners not connecting to logical demands of the task. Reasoners must overcome heuristic tendencies inherent in system one and also engage system two by redirecting attention toward logically relevant evidence**.**

Effects of Priming on Reasoning Processes in Selection Tasks

Rossi et al., (2015) tested effects of priming reasoners with pre-test training in executive learning and a rule falsification exercise on reasoning performance on the WST. They found the training affected performance, and the order in which participants completed the training and exercise further affected performance. The executive learning training cued reasoners to use a nonlinear, imaginative reasoning frame and to question their intuitive reasoning processes. The training also included warnings to avoid intuitive pitfalls, intended to evoke an emotional response involving a dynamic interaction of cognitive networks. A rule falsification task was performed by participants to support learning directing them toward the falsification of a conditional rule, qualitatively similar to directions of skepticism toward evidence.

Researchers found executive learning reduced matching errors but did not improve logical performance; participants still chose other illogical options indicating reasoners may need further prompting to engage in relevancy testing, not simply any choice that is non-matching. The study further noted a metacognitive learning effect was indicated when the rule falsification task was performed prior to the WST, participants were too strict in their transfer of knowledge which reduced matching but did not improve logical reasoning, emphasizing the complex nature of learning transfer. The rule falsification task was more formal logical reasoning and required absolute rejection of matching, while the WST contains two components, the interpretation of conditionals, and the selection process and decision.

Osman (2007) found tutoring based on dynamic, metacognitive strategies by having reasoners focus on their misinterpretations was effective, particularly under circumstances of limited cognitive resources such as a rapid-response version of a selection task. Researchers suggested reasoners were no longer trapped by system one influences. Prado and Noveck (2007) also suggested executive learning is helpful in inhibiting the matching heuristic, prompting reasoners to shift from system one reasoning to system two reasoning. Finally, Houdé et al. (2000) found inhibition training was effective at reducing matching bias and heuristic errors while normative logical training or repetition was not helpful. They suggested a shift away from perceptual processing regions associated with system one and toward higher order brain processing regions associated with system two was crucial.

Normative performance on the WST is adversely affected when a perceptual match is expected but the correct response requires cognitively negotiating a mismatch (Prado & Noveck, 2007). According to dual process theory, a bias erupts immediately because mismatches are sensed as perceptually irrelevant to the conditional rule. When a conditional rule presents two

elements to the reasoner, and two elements in the choice set do not correspond with elements presented, mismatches are considered as irrelevant. The reasoner must not only overcome the system one tendency to only perceive matching elements as relevant, but also must engage system two to redirect their attention toward logically relevant elements, which in the case of the WST would indicate the true antecedent and the false consequent.

Synthesis of the Research Findings

Professional skepticism is assumed to correlate with the quality of auditor judgment; higher levels of professional skepticism were long purported to improve auditor reasoning processes (Hurtt, 2010). Engaging a skeptical mindset requires activation of an outward orientation where the auditor is cued to the perspective of a party who is cheated in relation to a social contract (Cosmides, 1989; Gigerenzer & Hug, 1992). Cueing an outward oriented frame is less complicated with novices, but rules and sanctions based environment of the audit domain may predispose expert auditors toward an inward orientation (Tenbrunsel & Messick, 1999).

A significant amount of experimental research on auditor reasoning was conducted using students as participants; however, there are marked differences between ways that novices and experts respond to experimental manipulations (Kahneman, 2013; Klein, 1999). One of the primary differences between novices and experts is the complex way orientation frames are cued in expert auditors versus novices and the extent to which expert auditor reasoning is focused on justification, which affects information search motivations (Hurtt et al., 2013; Michael, 2006; Tenbrunsel & Messick, 1999). Research found while the reasoning quality of novices seem to benefit from prompts of outward oriented skepticism, findings were not consistently reproduced using expert auditors as participants (Grenier, 2011; Harding & Trotman, 2011; Peytcheva, 2013).

An explanation for findings may be the audit domain already over-saturates expert auditors with conservative skepticism and because the frame used in many experimental manipulations does not match the orientation of expert auditors. Reasoning is best when the frame of the reasoner matches the frame of the hypothesis (Fischoff, 1982; Smith & Kida, 1991), but many experiments are based on an outward orientation while expert auditors may be adopting an inward orientation (Agoglia et al., 2011; Dweck et al., 2004; Michael, 2006; Peytcheva, 2013). Additionally, experienced auditors are conditioned to varying levels of skepticism depending on whether they were personally exposed to irregularities; an influence not inherent in novices (Hurtt et al. 2013).

As opposed to outward oriented skeptical prompts, Grenier (2011) found experts were more likely to exhibit sound reasoning when they were prompted with dynamic, metacognitive strategies using an inward orientation aimed at questioning their own judgments. Trotman et al. (2009) similarly found prompting auditors to use a dynamic metacognitive strategy utilizing a backward reasoning approach resulted in improved reasoning. Dynamic metacognitive strategies may assist expert reasoners by reducing heuristic matching errors on experimental selection tasks. It is unclear however, to what extent these strategies may engage logical processes to inhibit the selection of irrelevant evidence and further assist the reasoner in identifying the most relevant elements to select on these tasks and why those elements are logically appropriate (Rossi et al., 2015). Heuristic errors occur quickly as crucial decision processes occur in the very early stages of the reasoning process and superficial aspects of problem solving direct reasoning processes (Evans, 2007). Rossi et al. (2015) found reasoners' responses in rapid response time tasks did not significantly differ from free response time versions, indicating expert reasoners rely heavily on quick problem assessments. Accordingly, experimental tasks

that require reasoners to respond in very short periods of time are particularly informative in expert domains.

A particularly effective experimental tool for assessing auditor reasoning efficacy is the hypothesis testing task (Dawson et al., 2002; Evans, 2007). Hypothesis testing tasks are effective in examining auditor reasoning because hypothesis testing is a routine duty for auditors (Peytcheva, 2013). One of the most popular experimental methods for researching hypothesis testing is the Wason Selection Task (Wason, 1966). The WST has several versions, but one of the more popular versions is an abstract version that is content independent and based on a formal rule structure.

The abstract WST uses four cards presented to participants that have a letter on one side of the card, and a number on the opposite side. Participants are given a conditional rule associated with cards in an if-then format such as *if there is a vowel on one side of a card there will be an even number on the other side of the card* (Wason, 1966; Wason, 1968). Participants are informed the conditional rule may be either true or false. Participants are asked to turn over the minimum number of cards that best help them determine whether the rule is true or false. A social contract version of the WST might propose a rule such as *if a person is drinking alcohol, then they are at least 21 years old*. The four cards would show *drinking alcohol*, *not drinking alcohol*, *at least 21*, and *not 21*, representing *p*, *not p*, *q* and *not q* respectively.

While higher levels of skepticism have been purported to improve performance on reasoning tasks such as the WST (Dawson et al., 2002; Hammersley, 2011; Hurtt, 2010), Peytcheva (2014) asserted there is lack of empirical evidence to support this presumption. In fact, when used in research on auditor reasoning, the WST yielded results that contradicted expectations about effects of skepticism prompts; frequently failing to help expert reasoners

achieve better results on hypothesis testing tasks (Gigerenzer & Hug, 1992; Peytcheva, 2013). While asserting reasoners have more success solving tasks that are framed as social contracts because the context is more familiar, the literature contains criticisms of using abstract, content independent experimental tasks in the study of expert reasoning because some researchers consider abstract tasks irrelevant to real world decision making contexts (Cosmides, 1989; Klein, 1999).

However, as previously noted, the real-world task domain for expert auditors is one imbedded in an abstract rule structure and social contract problems framed with skepticism prompts may not match the frame of the expert auditor (Dweck et al., 2004; Fischoff, 1982; Klauer et al., 2007; Michael, 2006; Tenbrunsel & Messick, 1999). Auditors are likely to frame problems abstractly because they implicitly place significant importance on the abstract rule structure that dominates their environment (Bame-Aldred, 2011; Dweck et al., 2004; Gigerenzer & Hug, 1992; Kahneman, 2013; Klayman & Ha, 1987; Smith & Kida, 1991). Reasoners may employ a positive test strategy toward elements perceived to be of greatest importance on tasks such as the WST because the rule based environment may stress the reasoner, pushing them toward familiar rules and easy matches (Kahneman, 2013; Klayman & Ha, 1987).

The rule structure inherent in the audit domain may change the frame in which expert auditors make decisions, and change their underlying information search motivations related to those decisions (Dweck, et al., 2004; Johns, 2006; Neuberg & Newsome, 1993). Reasoners implicitly tend toward evidence that holds the most personal importance, and in the audit domain that may mean auditor attention tends toward normative models to avoid personal sanctions (Dweck et al., 2004; Kahneman, 2013). When the reasoner is concerned with judgment error this represents an inward orientation and an abstract task frame is perceived as less threatening

(Smith & Kida, 1991). Reasoners may rely on abstract mental representations such as checklists or other normative methods to simplify problems (Neuberg & Newsome, 1993). Since reasoners vary in their reaction to complexity, uncertainty, and their desire for simple structure, they will differ in their avoidance and coping strategies.

A measurement related to need for certainty is achieved through the Personal Need for Structure (PNS) scale developed by Thompson et al. (1989) and introduced by Neuberg and Newsom (1993). The PNS is a self-report scale where the subject agrees or disagrees with statements about their views of the environment on a six-point scale. The PNS scale focuses on how individuals process ambiguous information within their environment. Certainty is desired more in states of ambiguity, anxiety and fear, including when there are rule structures that infer sanctions that are perceived as a self-threat, such as in the audit environment (Michael, 2006; Tenbrunsel & Messick, 1999). Individuals who have a higher personal need for structure tend to rely on normative models and are more susceptible to matching errors on selection tasks such as the WST (Kahneman, 2013; Neuberg & Newsome, 1993). High PNS individuals are more likely to dismiss information inconsistent with standardized schemas, making them more susceptible to matching errors (Hess et al., 1998), and are less likely to update their knowledge with new information (Hess, 2001; Okun & Rice, 1997).

Rossi et al. (2015) tested effects of priming on the WST and found executive learning training that used a dynamic reasoning frame affected performance. The executive learning training cued reasoners to use a nonlinear, imaginative reasoning frame and to question their intuitive reasoning processes. The training also included warnings to avoid intuitive pitfalls, intended to evoke an emotional response involving an interaction of cognitive networks. Osman (2007) found tutoring based on dynamic, metacognitive strategies by having reasoners focus on

their misinterpretations was effective, particularly under circumstances of limited cognitive resources such as a rapid-response version of a selection task.

Priming with executive learning involving efforts to inhibit the matching heuristic in selection tasks was substantiated by neuroimaging studies exhibiting a dramatic shift from posterior brain regions handling perceptual stimuli related to matching bias toward prefrontal networks required for hypothetical reasoning (Houdé et al., 2000). Houdé et al. (2000) reported inhibition training that encouraged dynamic reasoning was effective in assisting reasoners in escaping heuristic traps and lessening reasoning errors as opposed to convergent, normative logical training.

Critique of Previous Research Methods

Rational decision making models dominated literature until the 1960's when researchers began to challenge the validity of formal decision approaches (Arnold & Sutton, 1997). Researchers repeatedly demonstrated reasoners do not treat all situations equally, and apply decision making models differently depending on the context (Johns, 2006 Kahneman, 2013). Nearly all aspects of accounting and auditing include some aspect of decision making or behavior that is not well explained by economic theory (Koonce & Mercer, 2005). Dual process theory (Kahneman, 2013) was a fundamental model offered as an alternative to rational models in decision making research which assists this study as a framework to dissect auditor reasoning on the continuum between explicit conscious and implicit subconscious processes. Heuristics and biases (Tversky & Kahneman, 1975) coupled with dual process theory supports explanations for reasoning errors and understanding the role of personality traits and affective states related to auditor reasoning processes. Dual process theory explains decision making errors occur when the intuitive, automatic processes of system one generate an error which is not detected by

logical, conscious operations of system two or when system two falsely detects an error in a correct system one intuition (Morewedge & Kahneman 2010).

The naturalistic decision making model (Lipschitz et al., 2001), criticizes the heuristics and biases model by suggesting imprecise judgments should not immediately imply heuristic errors as suggested by the heuristics and biases paradigm. The recognition primed decision model explores pattern matching and cue relevancy in expert decision making and in domain specific settings. NDM suggests experts in real world environments do not have the time to consider all possible option choices and compare them against each other, but instead make associations between past learning and experience to conjure a possible solution and mentally test that solution against a construct of acceptability under given circumstances.

NDM and its counterpart, the recognition primed decision model purported errors are unavoidable in a complex environment and are needed for learning, and the goal should not be an environment free of errors. Error avoidance, authors suggested, leads to slower learning. NDM further suggests the heuristics and biases model, as it relates to dual process theory (Kahneman, 2013), does not appropriately address real world decision making because experiments are typically highly controlled, uses unrealistic, abstract tasks, and employs the use of non-experts used as participants. The NDM approach prefers to use active practitioners in real world environments to study decision making. While the perspective that using realistic experimental methods and experts in real world environments for research that is not always possible and dismisses valuable information that is garnered from controlled experimental methods. NDM research criticized the use of abstract tasks, suggesting these experimental manipulations do not represent real world decision characteristics, but in doing so NDM ignores experts who work in abstract environments. The perspective dismisses the functionality of

abstract experimental framing, as abstract framing may more appropriately describe decision making frames in highly structured, rule-based environments such as the audit environment.

Authors of both NDM and heuristics, and biases paradigm note models have several important points of agreement. First, both models agree expert reasoning can be successful in some circumstances while unsuccessful in others. Both models also believe it is important to distinguish between intuitive skill and assumptions. Both models agree system one and system two are central features of expert decision making, and system one represents intuitive judgment, while system two represents logical, conscious processes. Both models support that a stable environment yields the best expert intuition. Outcomes are difficult to predict in unstable environments, even for experts. The models also merge on the idea experts are typically unaware of cues that form and guide their actions (Kahneman & Klein, 2009).

Hurtt (2010) developed a skepticism scale that is widely employed in research examining auditor reasoning efficacy. The skepticism scale addresses six key characteristics of skepticism include having a questioning mind, or approaching information with suspicion or doubt, seeking meaning and clarification of definitions, and validating data. A higher level of professional skepticism is assumed to correlate with superior logical reasoning, more effective evidence selection procedures, and improved performance on hypothesis testing tasks (Dawson et al., 2002; Hammersley, 2011). The focus of Hurtt (2010) on outward oriented skepticism assumed it is a key determinant of reasoning efficacy and manipulating auditor skepticism will alter reasoning efficacy in expert auditors.

Hurtt's (2010) scale is considered to have validity in its measure of outward oriented skepticism, although it is not evidence of sound auditor reasoning more generally. Peytcheva (2014) purported there is a lack of empirical evidence to support the assumption that higher

levels of skepticism improve reasoning. In a study using both novices and experts as participants, Peytcheva (2014) found skepticism prompts affected reasoning processes of novices but not experts. The study is part of a body of literature focusing on outward oriented skepticism, which produces results that contradict expectations and fail to support hypotheses regarding effects of skepticism on auditor reasoning. The body of literature in this field also fails to produce a coherent explanation for why experimental manipulations of skepticism yield contradictory results or a model that explains the disagreement amongst researchers.

Smith and Kida (1991) analyzed heuristics and biases literature by specifically addressing expert auditor reasoning. The authors specified the body of literature identifies different experimental results from expert auditors versus novices. Of the studies reviewed, little to no support was found for the use of confirmatory strategies amongst experienced auditors, which contradicts findings and expectations of other researchers (Kahneman, 2013). Similar to later criticisms by Klein (1997) and Lipschitz (2001), Smith and Kida (1991) criticized the body of research for failing to use expert participants and familiar tasks that they say would be the most relevant context. They particularly criticized the use of tasks that are unfamiliar to participants; failing to appreciate there is an abstract context to many rules inherent in the audit environment. While context is a crucial part of learning and decision making, researchers should not ignore the context includes content-independent, abstract rules (Johns, 2006; Michael, 2006). Research that is more recent continues to support the use of outward oriented skepticism prompts and realistic and familiar tasks while ignoring the value of inward oriented metacognitive strategies and abstract framing that is a common part of the audit domain (Hurtt, 2010; Klein, 1999; Lipschitz, et al., 2001; Peytcheva, 2013).

Smith and Kida (1991) found strong support for anchoring bias amongst expert auditors. Anchoring bias amongst auditors was supported through research conducted by Pike et al. (2013) who found auditors who were aware of unaudited balances prior to conducting their work tended to give more weight to evidence that led to confirming those balances than auditors who performed their work without knowledge of balances. Anchoring bias is tied to the use of confirmatory strategies (Tversky & Kahneman, 1975).

Summary

The taxonomy and research surrounding the audit domain and expert auditor reasoning has evolved from rational reasoning models based in economic theory to psychological theories exploring biases and heuristics and realistic task experiments (Kahneman & Klein, 2009; Tversky & Kahneman, 1975; Klein, 1997; Koonce & Mercer, 2005; Peytcheva, 2013). There has been a focus on outward oriented reasoning strategies in the literature with significant attention toward how varying levels of skepticism affect auditor reasoning processes (Hurtt, 2010). A continuing course of research outcomes using realistic tasks and skepticism prompts revealed these outward oriented primes affect novice reasoning but their effects on expert reasoning is limited and remains unclear. The literature has largely ignored the influence the abstract rule structure of the audit domain may play in auditor reasoning processes and the potential usefulness of abstract task experiments. Researchers also focused on outward oriented skepticism and largely ignored the role inward oriented metacognitive strategies play in auditor reasoning (Dweck et al., 2004; Hurtt, 2010; Tenbrunsel & Messick, 1999). Research showed experts rely heavily on quick, mental associations, matching current situational characteristics to past knowledge, with additional time rarely affecting reasoning outcomes (Lipschitz et al., 2001). Research further indicated framing effects involving inward versus outward orientation are

significant influences on expert reasoning (Gigerenzer & Hug, 1992; Grenier, 2011; Harding & Trotman, 2011; Peytcheva, 2013). Accordingly, the present research focuses on the influence of inward oriented metacognitive states and traits on expert auditor reasoning

CHAPTER 3. METHODOLOGY

Purpose of the Study

The purpose of this study was to determine whether a statistically significant difference existed between the selection task scores of expert auditors exposed to a dynamic prompt versus a linear prompt, and if an interaction existed between those prompts and an individual's personal need for structure. An inherent interest exists in this study regarding the influence of metacognitive strategies in reasoning efficacy, particularly in expert domains where reasoners are imbedded in a linear, abstract rule structure. The audit domain, in accord with other investigative domains, places great emphasis on outward oriented frames of reasoning when assessing evidence, including levels of skepticism employed toward the analysis of evidence (Hurtt, 2010). Experts in such domains frequently rely on decision tools, such as linear step-by-step checklists and norms to simplify their environment, and may face reputation risk or sanctions if they commit errors, which may cause an inward oriented frame of self-concern and an increased tendency to rely on norms and tools to simplify the environment (Dweck, et al., 2004; Michael, 2006; Tenbrunsel & Messick, 1999).

Dynamic prompts addressing metacognitive strategies should spur an inward orientation, more closely matching the reasoning frame experts use when they employ abstract strategies or face abstract problems, and cue the reasoner to be aware of and question superficial assumptions (Dweck et al., 2004; Gigerenzer & Hug, 1992; Grenier, 2011; Harding & Trotman, 2011). The study is also interested in the interaction between the linear or dynamic state cued by the prompt and the trait of personal need for structure. While the dynamic or linear prompt addresses a current state of cognition related to temporal frame or affect, personal need for structure addresses an enduring set of traits of the reasoner related to discomfort with ambiguity and desire

for closure. Individuals experience different levels of discomfort when they confront cognitive inconsistencies, which result in different individual needs for cognitive closure; people higher in need for structure and closure tend to more quickly accept evidence that matches expectations (Slijkhuis et al., 2013; Webster & Kruglanski, 1996).

Research Questions and Hypotheses

RQ1: Is there a statistically significant difference between the combined (abstract and deontic) selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model?

$RQ1H1_{0:}$ There is not a statistically significant difference between the combined (abstract and deontic) selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model.

$RQ1H1_{a:}$ There is a statistically significant difference between the combined (abstract and deontic) selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model.

RQ1SQ1: Is there a statistically significant difference between the abstract selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model?

$RQ1SQ1H_0$: There is not a statistically significant difference between the abstract selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model.

$RQ1SQ1H_a$: There is a statistically significant difference between the abstract selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model.

RQ1SQ2: Is there a statistically significant difference between the deontic selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model?

RQ1SQ2H_0: There is not a statistically significant difference between the deontic selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model.

RQ1SQ2Ha: There is a statistically significant difference between the deontic selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model.

RQ2: Is there a statistically significant difference between the combined (abstract and deontic) selection task scores of auditors based on their scores on the personal need for structure inventory?

RQ2H1_0: There is not a statistically significant difference between the combined (abstract and deontic) selection task scores of auditors based on their scores on the personal need for structure inventory.

RQ2H1_a: There is a statistically significant difference between the combined (abstract and deontic) selection task scores of auditors based on their scores on the personal need for structure inventory.

RQ2SQ1: Is there a statistically significant difference between the abstract selection task scores of auditors based on their scores on the personal need for structure inventory?

RQ2SQ1H_0: There is not a statistically significant difference between the abstract selection task scores of auditors based on their scores on the personal need for structure inventory.

RQ2SQ1H_0: There is a statistically significant difference between the abstract selection task scores of auditors based on their scores on the personal need for structure inventory.

RQ2SQ2: Is there a statistically significant difference between the deontic selection task scores of auditors based on their scores on the personal need for structure inventory?

RQ2SQ2H_0: There is not a statistically significant difference between the deontic selection task scores of auditors based on their scores on the personal need for structure inventory.

RQ2SQ2H_a: There is a statistically significant difference between the deontic selection task scores of auditors based on their scores on the personal need for structure inventory.

RQ3: Is there a statistically significant difference between the combined (abstract and deontic) selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory?

RQ3H1$_0$: There is not a statistically significant difference between the combined (abstract and deontic) selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory.

RQ3H1$_a$: There is a statistically significant difference between the combined (abstract and deontic) selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory.

RQ3SQ1: Is there a statistically significant difference between the abstract selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning

model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory?

RQ3SQ1H_0: There is not a statistically significant difference between the abstract selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory.

RQ3SQ1H_a: There is a statistically significant difference between the abstract selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory.

RQ3SQ2: Is there a statistically significant difference between the deontic selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory?

RQ3SQ2H_0: There is not a statistically significant difference between the deontic selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory.

RQ3SQ3H_a: There is a statistically significant difference between the deontic selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory.

Research Design

The process for conducting this study was an experimental approach with random assignment and a 2x3 factorial design. Dependent variables were abstract task scores, deontic task scores and combined abstract and deontic task scores of participants on five hypothesis-testing tasks based on the Wason Selection Task (WST) (Wason; 1966; Wason, 1968). The WST tasks included five successive tasks; one abstract task presented first which was content independent, followed by four deontic tasks framed as social contracts. The abstract task was presented first to avoid learning order effects, followed by four deontic tasks. Tasks differ in that the abstract task is rule based and content independent, matching the rule based structure of the audit environment, and deontic tasks were based on a social contract frame and were content dependent. There were two independent variables in this study: Participant scores on the Personal Need for Structure scale (PNS) (Thompson et al., 1989), and either a dynamic or linear prompt in the form of a short, written narrative. Two-way ANOVAs were conducted for abstract selection task scores, deontic selection task scores, and combined selection task scores. Descriptive statistics were utilized to organize results (Leedy & Ormrod, 2013).

Target Population and Sample

Population

The population for this study was practicing auditors active in professional auditing within the recent 12-month period, and engaged in the audit domain as indicated by their seeking continuing professional education in auditing. A key feature of the audit population indicative of experts engaged in the industry is active involvement in a professional audit or accounting organization that awards professional licenses or designations, both requiring continuing professional education. Such continuing education is not typically attended by novices.

Accordingly, a core characteristic of the target population would be attendance at continuing professional education courses or conferences.

Sample

The sample of 264 expert auditors was drawn from auditors attending continuing professional education conferences offered through recognized professional auditing and accounting organizations around the United States. Recognized professional entities include any of the five major professional organizations in the auditing and accounting industry that are part of the Committee of Sponsoring Organizations of the Treadway Commission (COSO) originally organized in 1985 to sponsor the National Commission on Fraudulent Financial Reporting. These include the following organizations, or any official chapter of any of these organizations: The AICPA or any State CPA Society, the Institute of Internal Auditors (IIA), the Institute of Management Accountants (IMA), Financial Executives International (FEI), or the American Accounting Association (AAA). Criteria for participant inclusion in this study included active employment in the audit industry at any point during the most recent 12 months, and industry engagement as demonstrated by attending at least one professional continuing education course in the most recent year. Adherence to the inclusion element of attending a continuing education course was accepted by attendance at the current event in which the study was conducted. There were no exclusion criteria.

Power Analysis

Sample size was determined by an *a-priori* power analysis using a standard alpha level of .05 and power of .9 with a medium effect size. Based on these assumptions the minimum required sample size was 252 however the actual sample size yielded was 264. Sample calculations were obtained by using G*Power (Faul, Erdfelder, Lang, & Buchner, 1996). The

post-hoc calculation with 264 participants (two levels of priming and three levels of PNS) was .9154.

Procedures

Participant Selection

The researcher contacted leaders of 42 continuing professional education conferences for auditors and obtained agreement from 12 conference leaders to recruit participants for this study. The study recruited participants from six of those conferences. A recruitment document was distributed amongst prospective participants followed by an informed consent document. Prospective participants were informed of inclusion criteria that included attendance at continuing professional education conference for auditors and participants responded via electronic keypad as to their choice regarding inclusion in the study.

Protection of Participants

The study was designed to produce minimal risk to participants. The study protected participants by explaining the purpose of the study and expected benefits and possible harms before obtaining informed consent from all participants. The researcher took steps to minimize the risk of harm to participants by avoiding deceptive practices and protecting anonymity and confidentiality of participants by collecting anonymous data via electronic keypads that did not collect any personally identifiable information of individual participants. In order to further protect participant anonymity, data were transformed to spreadsheets immediately after collection, and data that included device identifiers that was necessary to collect data were deleted to prevent device identifiers used in the collection of data from saving such information. Participants were informed they had the right to withdraw from the research at any time.

Data Collection

The study employed a live polling method of data collection with participants responding electronically to research questions administered directly by the researcher at live educational conferences in locations around the United States. The researcher collected data according to the following steps: Participants were first provided with informed consent and responded electronically by affirming or declining their participation in the study. All potential participants were informed they could choose to answer, or not to answer, any of the questions; may stop their participation at any time; or alternatively, could choose to participate in study activities but still decline to have their responses collected for use in the research; or they could fully participate and allow their data to contribute to the research. A yes response was accepted as affirmation by the participant, and any other response or lack of response was considered as a declination and that individual's data were not saved for use in the study. Failure to answer any questions excluded the participant's data from use in the study. The anonymous keypad response regarding inclusion in the study, and other responses, or lack of response during the study was blind to the researcher and other participants.

After informed consent and choice of participation, participants were directed to respond via the keypad to the 12 item PNS assessment. The Personal Need for Structure (PNS) scale developed by Thompson et al. (1989) and introduced by Neuberg and Newsom (1993) is a self-report scale where the subject agrees or disagrees with statements about their views of the environment on a six-point scale. Participants were led through the PNS assessment, viewing items one at a time on a screen and responding on their device. The researcher read the item aloud and participants were allowed 10 seconds to respond to each item. The display showed

participants response options marked alphabetically instead of numerically to avoid the possibility of respondents responding to the potential influence of numerical weighting.

Following the PNS assessment, participants were exposed to a dynamic or linear priming mechanism as seen in Appendix C. Documents were distributed in random order using a random number generator, face down in front of participants with instructions to only turn over and read the document upon notification. Once participants were notified to read the priming document, they were given 4 minutes to complete their review. The priming mechanism took place in the form of a written narrative either based on standardized sampling language used in the audit industry (linear prime) or a narrative prompting auditors to use a metacognitive approach, to be skeptical of their own reasoning processes, and consider alternatives (dynamic prime).

After exposure to the priming mechanism, participants responded electronically to five separate hypothesis validation tasks based on the Wason Selection Task (Wason, 1966; Wason 1968). The first selection task was an abstract task in the form of *if p then q*, followed by four deontic/social contract versions of the selection task. The order of tasks was chosen to reduce learning bias; learning is reduced by presenting abstract tasks prior to deontic tasks (Arnold et al., 2000). Additionally, the abstract version was presented first to minimize the chance of deontic versions cueing a social contract frame that could affect participant responses on the abstract version, and research has shown little to no effect of learning from the abstract version to the deontic versions in this regard (Cheng & Holyoak, 1985; Osman, 2007; Rossi et al., 2015). Multiple versions of the four deontic tasks, including numeric and non-numeric tasks were used to gather sufficient data for analysis and to examine for learning effect and order bias.

The first three versions of the deontic task (tasks 2, 3 and 4) had an identical correct answer pattern, with the fourth version of the task having a different correct answer pattern. In

each version of the task, the researcher displayed four cards on the screen; read aloud a conditional rule associated with the cards in an if-then format. Participants were informed the conditional rule may be either true or false and then participants are asked to select two cards that best helped them determine whether the rule is true or false. To cue a rapid response, participants were allowed 10 seconds to make their selections. Rapid response tasks were utilized because experts assess situations quickly, and do not typically change their responses after longer exposure to problem cues (Klein, 1997; Osman, 2007; Rossi et al., 2015). Only first answers were accepted, and participants could not change their selections. Completion of the fifth WST concluded the data collection portion of the study.

Data Analysis

Data were collected using electronic polling and converted to a format for analysis using SPSS. Data contained no individually identifiable information of participants, and was encrypted and stored on a flash drive kept in a secure and locked safe. Descriptive statistics were performed via SPSS to obtain variable mean scores and independent *t*-tests were run on the average (mean) score of the PNS assessment, two levels of the priming mechanism, and interaction between independent variables with the dependent variable. Factorial ANOVAs and two-way MANOVAs were conducted for each of the dependent variables on abstract selection task scores, deontic selection task scores, and combined selection task scores.

Instruments

The study utilized two instruments, the PNS scale (Thompson et al., 1989; Neuberg & Newsom, 1993) as seen in Appendix A, and the Wason Selection Task (Wason; 1966; Wason, 1968) as seen in Appendix B. The PNS scale was used as an independent variable to examine

the trait of desire for simple structure. The WST was employed as the dependent variable to examine expert auditor reasoning via hypothesis testing behaviors.

Personal Need for Structure Inventory

The Personal Need for Structure scale (PNS) was presented by Thompson et al. (1989) at the Annual Meeting of the Canadian Psychological Association in 1986. The Personal Need for Structure inventory (PNS) is a 12 question Likert scale inventory that asks participants to respond to each of the 12 statements on a six-point scale from *strongly agree* to *strongly disagree*. Scores are then used to place participants in one of five groups: Low, Moderately Low, Medium, Moderately High, and High. According to Neuberg and Newsom (1993), the PNS scale began with a larger set of items based on face validity but was reduced through psychometric analysis to 12 items. The PNS scale is not copyrighted or listed in the *Mental Measurements Yearbook* as a formal psychometric tool. The scale and psychometric properties were unpublished until Neuberg and Newsom (1993) completed psychometrics including reliability and convergent and discriminant validity.

Validity. The Personal Need for Structure scale addresses two distinct factors: the desire for structure, and how individuals respond to lack of structure. Confirmatory factor analyses completed by Neuberg and Newsom (1993) compared the two-factor model against a single factor model and on six subject populations; results supported the two-factor model provided best fit for data. Interscale correlations with other measures as well as big five personality traits were reviewed. Confirmatory factor analysis showed PNS factors are negatively correlated with the big five factors of openness and desire for structure, while positively correlating with conscientiousness. A correlation was indicated between higher PNS scores and the tendency toward behavioral routinization. Confirmatory analysis showed scales such as Authoritarianism,

Dogmatism, and Intolerance of Ambiguity were weakly correlated with PNS, indicating scale measures contained orthogonal components. Neuberg and Newsom (1993) asserted based on their psychometric analysis, the PNS possesses adequate convergent and discriminant validity.

Reliability. In tests of internal reliability by Neuberg and Newsom (1993) across six subject populations, the two-factor scale returned a Cronbach alpha= .77 and test-retest validity was significant at r=.84. Neuberg and Newsom (1993) suggested the PNS possesses adequate reliability for use in scientific inquiries. Cronbach's alpha was performed for this study yielding Cronbach alpha= .74. Further information regarding the reliability of the PNS was unavailable.

Wason Selection Task

One of the most popular tests for examining hypothesis testing processes is the Wason Selection Task (Evans, 2007; Peytcheva, 2013; Wason; 1966; Wason, 1968). The Wason Selection Task (WST) (Wason; 1966; Wason, 1968) was introduced by Peter Wason and is commonly known as the Wason four-card selection task. The WST examines propositional logic and information seeking behaviors. The WST is considered an effective tool for examining dual process reasoning in auditors (Peytcheva, 2013). The WST has become one of the most popular experimental tools for studying human reasoning (Almore & Sloman, 1996; Klauer et al., 2007), but it is not copyrighted or listed in the *Mental Measurements Yearbook* and the task and psychometric properties remain unpublished as a formal psychometric tool.

The WST is especially informative when investigating reasoning processes because participants respond to slight variations in framing and content (Dawson et al., 2002). The WST is particularly sensitive to evidence seeking biases and is an excellent experimental tool to examine reasoning in auditors because it models similar choices to those that auditors must make as a typical part of their work (Peytcheva, 2013). Designed to test propositional logic, the WST

is known to follow what is referred to in the literature as the *if p then q* format yielding possible selections of *p*, *not p*, *q* and *not q*. The participant is presented four cards. Two of the four cards showing have a number exposed while the remaining two cards have a letter exposed. Cards read in order A, K, 8 and 5. Participants are given a conditional rule associated with cards in an if-then format, *if there is a vowel on one side of a card there will be an even number on the other side of the card* (Wason, 1966; Wason, 1968).

Participants are informed the conditional rule may be either true or false. Participants are asked to turn over the minimum number of cards that best help them determine whether the rule is true or false. Logically the participant should choose the A and 5 or *p* and *not q* cards representing affirmation of the antecedent and falsification of the consequent respectively as this is the only combination of cards that have the power to falsify the rule (Rossi et al., 2015). The antecedent is the first part of a conditional rule and the consequent is the last part of the conditional rule. There are up to two possible correct selections, the *p* card and *not q* card, yielding a score of zero, one or two points per task.

Validity. According to Evans (2003), and Rossi et al. (2015), the Wason Selection Task (WST) is recognized as a valid and effective tool for the exploration of the dual system reasoning model and the heuristics and biases framework introduced by the seminal work of Tversky and Kahneman (1975) and more recently by Stanovich and Thompson (2001), Stanovich and West (2000), Kahneman (2003) and Morewedge and Kahneman (2010). Peytcheva (2014) reported the WST is particularly valid for the examination of auditor reasoning as auditors routinely face decisions of the if-then variety as part of their work. Validity information is limited, as the WST is not a published instrument and further information regarding the reliability of the WST was unavailable.

Reliability. Reliability data for the WST are unpublished as the WST is not a formally published or copyrighted tool. Cronbach's alpha was performed for this study yielding Cronbach alpha= .707. Numerous versions of the WST are used which diminish the ability to determine reliability, however it is one of the most widely used tools in the study of logical reasoning. Further information regarding the reliability of the WST was unavailable. Reliability information is limited, as the WST is not a published instrument and further information regarding the reliability of the WST was unavailable.

Ethical Considerations

This study adhered to the Belmont Principles of respect for persons, justice, and beneficence. This study did not expose participants to greater than minimal risk, as no significant risks to participants were identified. Individually identifying information was not collected as part of the study. Informed consent was obtained electronically. No vulnerable populations were used in this study. Individual identifiers were assigned to each participant, and data were immediately transferred to a spreadsheet without identifying information for cleaning.

Protection of participant privacy and anonymity was paramount, and the study utilized electronic polling technology to gather data from participants. The study attempted to enhance participant autonomy and reduce the possibility of coercion by supporting anonymous participation and privacy through electronic polling and allowing participants to participate in exercises of the study without the saving of data for use in the study. Participants were informed they could refuse to answer any individual questions, or cease participation at any point in the study.

Summary

Chapter 3 described the methodological approach for the study, and the purpose of the study, explaining the relevance of dynamic versus linear prompts in the examination of hypothesis testing and information seeking behaviors of expert auditors. The chapter outlined the research design, research questions, and hypotheses and provided information about the target population, described how participants were selected, and described protection of participants. A description of data collection procedures and data analysis was provided. The chapter continued with a review of instruments used in the study which included the Wason Selection Task and the Personal Need for Structure scale; following with validity and reliability information for instruments, and finally ethical considerations specific to the study.

CHAPTER 4. RESULTS

Introduction

This chapter presents results of the quantitative, experimental study designed to answer the three research questions and six subquestions to determine whether a statistically significant difference exists between selection task scores of expert auditors exposed to a dynamic, metacognitive prompt versus a linear prompt, and if an interaction exists with personal need for structure. The study investigated effects of inward oriented, metacognitive strategies versus outward oriented strategies of skepticism on expert auditor reasoning efficacy. The design relied on research indicating experts tend to make rapid assessments of a situation and develop reasoning approaches reliant on heuristic responses based on past experiences and learned responses based on tools representing best practices for their domain (Kahneman & Klein, 2009; Lipshitz et al., 2001). These rapid decisions may result in heuristic errors based in dual process theory, and a failure of the reasoner to examine the validity of their judgment, particularly in highly structured environments (Kahneman & Klein, 2009; Michael, 2006). Expert training focusing solely on outward oriented skepticism while ignoring the value of metacognitive strategies may result in continuing judgment errors (Kahneman & Klein, 2009). Professional skepticism commonly refers to having a questioning mind; there may be value in turning that lens back toward reasoners in the form of questioning their assumptions and examining the justification for their own beliefs, particularly in the form of confirmation bias (Hurtt, 2010; Peytcheva, 2004).

The preponderance of literature on expert auditor judgment has investigated auditor reasoning with a focus on outward oriented skepticism. Missing from the literature is a comparison of effects of inward oriented, metacognitive strategies versus outward oriented

strategies on the judgment quality of experts. This was the first experimental study on expert auditors that examined effects of personal need for structure and dynamic, metacognitive versus linear prompts on reasoning efficacy, comparing both abstract and deontic hypothesis testing task outcomes.

Description of the Sample

The population for this study was practicing auditors, active in professional auditing within the recent 12-month period in the United States, and engaged in the audit domain, as indicated by their membership in a recognized professional organization and seeking continuing professional education in auditing. A key feature of the audit population indicative of engagement in the industry is active involvement in a professional audit or accounting organization that awards professional licenses or designations, both requiring continuing professional education. Accordingly, a core characteristic of the target population would be attendance at continuing professional education courses or conferences.

The sample of 264 expert auditors was drawn from auditors attending continuing professional education conferences offered through recognized professional auditing and accounting organizations around the United States. Recognized professional entities include any of the five major professional organizations in the auditing and accounting industry that are part of the Committee of Sponsoring Organizations of the Treadway Commission (COSO) originally organized in 1985 to sponsor the National Commission on Fraudulent Financial Reporting. These include the following organizations, or any official chapter of any of these organizations: The AICPA or any State CPA Society, the Institute of Internal Auditors (IIA), the Institute of Management Accountants (IMA), Financial Executives International (FEI), or the American Accounting Association (AAA). Criteria for participant inclusion in this study included active

employment in the audit industry at any point in the previous 12 months, and industry engagement as demonstrated by attending at least one professional continuing education course in the most recent year. Adherence to the inclusion element of attending a continuing education course was accepted by attendance at the current event in which the study was conducted. There were no exclusion criteria.

Table 1 displays the frequency distribution for three outcome scores. The total correct ranged from zero to six points (M = 2.14, SD = 1.05). The total abstract correct ranged from zero to two points (M = 1.11, SD = 0.47). The total deontic correct ranged from zero to four points (M = 1.03, SD = 0.90). Table 2 displays frequency counts for two independent variables. The two priming type groups were of equal size. The personal need for structure level ranged from low (1.5%) to high (11.4%) with the median category being *moderately high*. For ANOVA models in this study, the five PNS levels were collapsed into three categories.

Assumption of normality. Univariate normality was assessed using the skewness and kurtosis indices of the variables. Analysis of the distribution showed that the distribution was symmetric. The value for combined scores was skew=0.50, SE=0.15 and kurtosis= - 0.34, *SE*=0.30. The value for abstract scores was skew=0.35, SE=0.15 and kurtosis= - 1.12, *SE*=0.30. The value for deontic scores was skew=0.35, *SE*=0.15 and kurtosis= - 0.77, SE=0.30. Both the skew and the kurtosis of the distribution did not deviate by more than two in absolute value. Additionally, due to the large sample size, histograms were visually inspected and revealed normal distributions. In the case of larger sample sizes, the assumptions of normality can be further assessed through visual interpretation of Normal Q-Q Plots. As displayed in appendix D, inspection of Normal Q-Q Plots showed only minor deviations from normality and is deemed to not negatively affect the results (Shadish, Cook & Campbell, 2002).

Homogeneity of variance. Levene's test was used to test the assumption of homogeneity of variance for the interaction (Howell, 2017). Results for combined scores showed $F(2, 261) = 1.104$, $p = .333$. Results for abstract scores showed $F(2, 261) = 1.444$, $p = .238$. Results for deontic scores showed $F(2, 261) = 0.764$, $p = .467$. The Levene's test for each dependent variable was not significant, therefore, the assumption was met.

Missing data and outliers. Missing data was analyzed via frequency counts and there were no cases with missing data. In a sample size of 264, only six outliers were identified. The outliers are being considered in the analysis because there were relatively few outliers and the results were not materially affected as determined by comparing the results of the analysis excluding the outliers (Shadish, Cook & Campbell, 2002).

Table 1

Frequency Distribution for Outcome Scores (N = 264)

Score	Category	*n*	%
Total Correct [a]			
	0	2	0.8
	1	86	32.6
	2	81	30.7
	3	65	24.6
	4	28	10.6
	5	1	0.4
	6	1	0.4
Total Abstract Correct [b]			
	0	16	6.1
	1	202	76.5
	2	46	17.4
Total Deontic Correct [c]			
	0	91	34.5
	1	85	32.2
	2	78	29.5
	3	9	3.4
	4	1	0.4

[a] Correct: $M = 2.14$, $SD = 1.05$.
[b] Abstract: $M = 1.11$, $SD = 0.47$.
[c] Deontic: $M = 1.03$, $SD = 0.90$.

Hypothesis Testing

Research Question 1 was, is there a statistically significant difference between the combined (abstract and deontic) selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model? The related null hypothesis was, H_0 There is not a statistically significant difference between the combined (abstract and deontic) selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model. To answer this question, Table 5 displays the relevant two-way ANOVA model. Table 3 displays relevant

mean and standard deviation data. There was not a significant interaction between the combined (abstract and deontic) selection task scores of expert auditors and type of priming, F (1, 258) = 0.00, p = .75. Thus, the null hypothesis was retained.

Research Question 1- Subquestion 1 was, is there a statistically significant difference between the abstract selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model? The related null hypothesis was, H_0 There is not a statistically significant difference between the abstract selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model. To answer this question, Table 6 displays the relevant two-way ANOVA model. Table 3 displays relevant mean and standard deviation data. There was a significant main effect for priming on total abstract selection task scores of expert auditors, F (1, 258) = 3.80 p =.05, partial eta-squared =.015. Mean and standard deviation data is displayed in Table 3. Participants in the dynamic prime group (M=1.19, SD =.48) had a significantly higher total abstract score than those in the linear prime group (M=1.04, SD =.45). Thus, the null hypothesis was rejected.

Table 2

Frequency Counts for Independent Variables (N = 264)

Variable	Category	*n*	*%*
Priming type			
	Linear prime	132	50
	Dynamic prime	132	50
Personal Need for Structure Level			
	Low	4	1.5
	Moderately low	27	10.2
	Medium	89	33.7
	Moderately high	114	43.2
	High	30	11.4
Personal Need for Structure Recorded			
	Low	31	11.7
	Medium	89	33.7
	High	144	54.5

Research Question 1- Subquestion 2 was, is there a statistically significant difference between the deontic selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model? The related null hypothesis was, H_0 There is not a statistically significant difference between the deontic selection task scores of expert auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model. To answer this question, Table 7 displays the relevant two-way ANOVA model. Table 3 displays relevant mean and standard deviation data. There was not a significant interaction between the combined (abstract and deontic) selection task scores of expert auditors and type of priming, F (2, 258) = .1.17, p = .28. Thus, the null hypothesis was retained.

Table 3

Means and Standard Deviations by Priming Type

Outcome	Priming Type	*M*	*SD*
Total Correct			
	Linear prime	2.17	1.11
	Dynamic prime	2.12	1.00
Total Abstract Correct			
	Linear prime	1.04	0.45
	Dynamic prime	1.19	0.48
Total Deontic Correct			
	Linear prime	1.13	0.95
	Dynamic prime	0.93	0.84

Note. For Linear Prime n=132, for Dynamic Prime n=132.

Research Question 2 was, is there a statistically significant difference between the combined (abstract and deontic) selection task scores of auditors based on their scores on the personal need for structure inventory? The related null hypothesis was, H_0 There is not a statistically significant difference between the combined (abstract and deontic) selection task scores of auditors based on their scores on the personal need for structure inventory.

To answer this question, Table 5 displays the relevant two-way ANOVA model. Table 4 displays relevant mean and standard deviation data. There was not a significant interaction between the combined (abstract and deontic) selection task scores of expert auditors and type of priming, $F(2, 258) = .96$, $p = .38$. Thus, the null hypothesis was retained.

Research Question 2-Subquestion 1 was, is there a statistically significant difference between the abstract selection task scores of auditors based on their scores on the personal need for structure inventory? The related null hypothesis was, H_0 There is not a statistically

significant difference between the abstract selection task scores of auditors based on their scores on the personal need for structure inventory. To answer this question, Table 6 displays the relevant two-way ANOVA model. Table 4 displays relevant mean and standard deviation data. There was not a significant interaction between the abstract selection task scores of expert auditors and type of priming, $F(2, 258) = 1.25$, $p = .29$. Thus, the null hypothesis was retained.

Research Question 2-Subquestion 2 was, is there a statistically significant difference between the deontic selection task scores of auditors based on their scores on the personal need for structure inventory? The related null hypothesis was, H_0 There is not a statistically significant difference between the deontic selection task scores of auditors based on their scores on the personal need for structure inventory.

To answer this question, Table 7 displays the relevant two-way ANOVA model. Table 4 displays relevant mean and standard deviation data. There was not a significant interaction between the deontic selection task scores of expert auditors and type of priming, $F(2, 258) = .51$, $p = .60$. Thus, the null hypothesis was retained.

Table 4

Means and Standard Deviations by PNS Level

Outcome	Structure Level	*M*	*SD*
Total Correct			
	Low	2.39	1.12
	Medium	2.13	1.04
	High	2.10	1.05
Total Abstract Correct			
	Low	1.23	0.43
	Medium	1.09	0.54
	High	1.10	0.44
Total Deontic Correct			
	Low	1.16	0.97
	Medium	1.04	0.86
	High	0.99	0.91

Note. For Low PNS level n=31, for Medium PNS level n=89, for High PNS level n=144.

Research Question 3 was, is there a statistically significant difference between the combined (abstract and deontic) selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory? The related null hypothesis was, H_0 There is not a statistically significant difference between the combined (abstract and deontic) selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory. To answer this question, Table 5 displays results of the two-way ANOVA model for total correct score based on priming type and PNS level. There was not a significant interaction between the combined (abstract and deontic) selection task scores of expert auditors based on the interaction between 1: Being primed with a

dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory, F (5, 258) = .54, p = .75. Hence, we focus on the main effects. Inferential tests for main effect for combined (abstract and deontic) selection task scores show there was not a significant main effect for priming, F (1, 258) = 0.00, p = .95. Additionally, there was not a significant main effect for PNS, F (2, 258) = 0.96, p = .38. Thus, the null hypothesis was retained.

Table 5

Two-Way ANOVA for Total Correct Score Based on Priming Type and Personal Need for Structure (PNS) Level (N = 264)

Source	*SS*	*df*	*MS*	*F*	*p*	Partial Eta Squared
Full Model	3.00	5	0.60	0.54	.75	.010
Priming Type	0.01	1	0.01	0.00	.95	.000
PNS	2.15	2	1.07	0.96	.38	.007
Type X PNS	0.76	2	0.38	0.34	.71	.003
Error	287.53	258	1.11			
Total	290.53	263				

Research Question 3-Subquestion 1 was, is there a statistically significant difference between the abstract selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory? The related null hypothesis was, H_0 There is not a statistically significant difference between the abstract selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model

compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory. To answer this question, Table 6 displays results of the two-way ANOVA model for total abstract score based on priming type and PNS level. There was a significant interaction between the abstract selection task scores of expert auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory, F (5, 258) = 2.58., *p* = .03, *partial eta-squared* = .048. Thus, the null hypothesis was rejected.

Table 6

Two-Way ANOVA for Total Abstract Score Based on Priming Type and Personal Need for Structure (PNS) Level (N = 264)

Source	*SS*	*df*	*MS*	*F*	*p*	Partial Eta Squared
Full Model	2.79	5	0.56	2.58	.03	.048
Priming Type	0.82	1	0.82	3.80	.05	.015
PNS	0.54	2	0.27	1.25	.29	.010
Type X PNS	0.69	2	0.34	1.59	.21	.012
Error	55.80	258	0.22			
Total	58.59	263				

Research Question 3-Subquestion 2 was, is there a statistically significant difference between the deontic selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory? The related null hypothesis was, H_0 There is not a statistically significant difference between the deontic selection task scores of

auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory. To answer this question, Table 7 displays results of the two-way ANOVA model for total deontic score based on priming type and PNS level. There was not a significant interaction between the deontic selection task scores of expert auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory, F (5, 258) = .88, p = .49. Hence, we focus on the main effects. Inferential tests for main effect for deontic selection task scores show there was not a significant main effect for priming, F (1, 258) = 1.17, p = .49. Additionally, there was not a significant main effect for PNS, F (2, 258) = 0.19, p = .82. Thus, the null hypothesis was retained.

Table 7

Two-Way ANOVA for Total Deontic Score Based on Priming Type and Personal Need for Structure (PNS) Level (N = 264)

Source	*SS*	*df*	*MS*	*F*	*p*	Partial Eta Squared
Full Model	3.60	5	0.72	0.88	.49	.017
Priming Type	0.96	1	0.96	1.17	.28	.005
PNS	0.83	2	0.41	0.51	.60	.004
Type X PNS	0.31	2	0.16	0.19	.82	.001
Error	210.16	258	0.81			
Total	213.76	263				

Supplemental analysis was conducted to determine whether there were differences between types of reasoning errors committed between different experimental conditions. This analysis is considered of significant practical interest to researchers in the field (Houdè & Moutier, 1999). Specifically, data were analyzed to determine if participants fell prey to matching bias or irrelevant selection error based on the type of prime (dynamic or linear) and type of task (abstract or deontic). Matching bias would be the selection of cards in the WST that attempt to affirm the antecedent and the consequent of the conditional if-then statement instead of correctly seeking to affirm the antecedent and disaffirm the consequent. Irrelevant selection error would be the selection of a card in the WST that was not addressed in the conditional if-then statement. The analysis revealed that matching errors increased under the linear prime condition (.64) from the dynamic prime condition (.43). Additionally, while results related to the PNS scores were not statistically significant, this may be due to skewed PNS scores in the population sample because of the likelihood of individuals high in need for structure to be drawn to the audit industry (Reardon & Lenz, 1999). A review of scores between low and high PNS scores revealed that of the participants in the study, low PNS scorers were much more likely to correctly answer the abstract selection task (.75 versus .17).

Summary

This chapter presented results of the analysis designed to answer the three main research questions and the six sub-questions. These research questions had a goal of determining whether a statistically significant difference existed between selection task scores of expert auditors exposed to a dynamic, metacognitive prime versus a linear prime, and if an interaction existed with personal need for structure.

The first research question asked if there was a statistically significant difference between combined (abstract and deontic) selection task scores of expert auditors who are primed with a dynamic reasoning model compared with those primed with a linear reasoning model. No statistically significant difference was found. The first sub-question to the first research question asked if there was a statistically significant difference between the abstract selection task scores of expert auditors who are primed with a dynamic reasoning model compared with those primed with a linear reasoning model. A statistically significant difference was found and the null hypothesis was rejected. The second sub-question to the first research question asked if there was a statistically significant difference between the deontic selection task scores of expert auditors who are primed with a dynamic reasoning model compared with those primed with a linear reasoning model. No statistically significant difference was found.

The second research question asked is there a statistically significant difference between combined (abstract and deontic) selection task scores of auditors based on their scores on the personal need for structure inventory. No statistically significant difference was found. The first sub-question of the second research question asked if there was a statistically significant difference between abstract selection task scores of auditors based on their scores on the personal need for structure inventory. No statistically significant difference was found.

The second sub-question of the second research question asked if there was a statistically significant difference between deontic selection task scores of auditors based on their scores on the personal need for structure inventory. No statistically significant difference was found.

The third research question asked if there was a statistically significant difference between combined (abstract and deontic) selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those

primed with linear reasoning and 2: Scores on the personal need for structure inventory. While the main interaction effect was not statistically significant, the priming effect was statistically significant. The interaction was not statistically significant.

The first sub-question of the third research question asked if there was a statistically significant difference between abstract selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory. Results showed a statistically significant main interaction. The second sub-question of the third research question asked if there was a statistically significant difference between deontic selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory. No statistically significant difference was found.

CHAPTER 5. DISCUSSION, IMPLICATIONS, RECOMMENDATIONS

Summary of the Results

The purpose of this study was to determine whether a statistically significant difference existed between selection task scores of expert auditors exposed to a dynamic, metacognitive prompt versus a linear prompt, and if an interaction existed with an individual's personal need for structure. This study differed in the approach to examining reasoning efficacy, as research in the audit domain has traditionally emphasized an examination of outward oriented skepticism toward evidence and possessing the individual trait of a sufficiently skeptical mindset (Hurtt, 2010; Hurtt et al., 2013). Inherent in this traditional emphasis is the assumption higher levels of outward oriented skepticism are associated with higher levels of reasoning efficacy. However, research indicates while higher levels of outward oriented skepticism have shown to enhance reasoning effectiveness for novices, these benefits do not extend to, and may even diminish reasoning efficacy in experts (Gigerenzer & Hug, 1992; Kahneman, 2013; Michael, 2006; Peytcheva, 2013; Tenbrunsel & Messick, 1999).

The body of literature, which emphasizes outward oriented skepticism, lacks experimental approaches that compare the effectiveness of inward oriented, metacognitive strategies versus outward oriented strategies of skepticism of evidence on reasoning efficacy. The lack of attention to inward oriented skepticism may be detrimental to the understanding of expert reasoning processes due to inherent differences existing between expert and novice reasoning. The significance of examining effects of an inward orientation as opposed to an outward orientation of skepticism in this context is twofold: First, experts as opposed to novices are already saturated with outward oriented skepticism, and further prompts may not achieve additional gains. Second, the rigid, abstract, linear, and rule-laden structure of the audit

environment may cue a self-conscious and conservative reasoning frame, motivating reasoners to avoid non-compliance and related sanctions. In this environment reasoners tend to frame issues more abstractly to create mental distance and reduce the threat and accordingly may seek the simple safety of an abstract, content independent structure and familiarity of frequently used tools and decision aids (Agoglia et al., 2011; Dweck et al., 2004; Michael, 2006). Consequently, a mismatch may exist in the belief experts need training to use continually higher levels of outward oriented skepticism, because while the outward orientation may match the reasoning frame of novices, it may not match the inward orientation assumed by experts in highly structured environments. Additionally, experts are already oversaturated and fatigued with outward oriented skepticism (Dweck et al., 2004; Kahneman, 2013).

The body of literature is consistent in that it also emphasizes the trait of skeptical mindset, largely ignoring an examination of effects metacognitive traits such as a need for structure play in reasoning efficacy. The literature lacks a critical examination of the role a skeptical mindset plays in reasoning efficacy, assuming the higher one scores on skepticism; the more effective are his reasoning processes (Hurtt, 2010). Expert reasoning efficacy however, relies on more than a skeptical mindset, it relies on the reasoner constructing meaning through a coherent structure of associations with which to understand an environment and build expectations (Kahneman, 2013; Thompson et al., 1989). While building this coherent structure, individuals have different cognitive styles and preferences that influence information gathering and decision making processes. Individuals experience varying levels of cognitive dissonance and discomfort when faced with cognitive inconsistencies and ambiguities that result in different individual needs for structure (Webster & Kruglanski, 1996). The Personal Need for Structure (PNS) scale developed by Thompson et al. (1989) and introduced by Neuberg and Newsom

(1993) is a self-report scale focusing on how individuals process ambiguous information within their environment.

Expert decision makers in audit domains operate in an environment that is complex and data intensive, frequently pushing people beyond their attentional limits and cognitive capacities. The cognitive load many decision makers experience can cause uncertainty and discomfort, and needs reducing to a level that is more efficiently processed. Individuals vary in levels of discomfort with this uncertainty and their desire for simple structure that has important connections to reasoning processes. Reasoners reduce cognitive load through a variety of avoidance strategies or cognitive structuring and restructuring strategies (Neuberg & Newsom, 1993). Higher Personal Need for Structure is associated with a variety of reasoning errors. Individuals with a higher need for structure and certainty are threatened by uncertainty and have strong emotional responses to anything that disrupts standardization, they prefer the simple organization of data, are less likely to change their attitudes, and are more likely to rely on stereotypes. These individuals prefer simple heuristics, pre-formed categories and methodologies, and are likely to respond to failure or unforeseeable consequences with learned helplessness, and individuals high in PNS are drawn to professional domains such as auditing due to the structure it offers (Ruiselová et al., 2012).

Schaller et al. (1995) found that need for structure significantly influenced inferential reasoning processes, and high PNS participants used overly simplistic reasoning strategies in comparison to low PNS participants and participants with high PNS were more likely than those with a low PNS to make attributional errors, and to make erroneous group stereotypes. High PNS individuals are more likely to dismiss information that is inconsistent with standardized schemas (Hess et al., 1998), and are less likely to update their knowledge with new information

(Hess, 2001; Okun & Rice, 1997). High PNS is also correlated with a higher sensitivity to confirming evidence than disconfirming evidence (Bamber et al., 1997). Peytcheva (2014) purported evidence seeking activities of auditors and cognitive processes involved in determining the relevance of evidence is a critical part of the auditor's reasoning process and confirmation bias is a common error amongst auditors. Accordingly, this research examined the need for structure and its effects on reasoning efficacy.

This research fills a gap in the literature by examining expert auditor reasoning through the manipulation of variables related to inward oriented metacognitive strategies that may match expert auditor reasoning frames more closely than variables related to outward oriented skepticism. This research also fills a gap in the literature by critically examining the trait of skepticism and the role that need for structure plays in reasoning efficacy. Filling this gap in the literature provides researchers and policymakers with a deeper understanding of mechanisms that underlie expert auditor reasoning processes, informing them about the complexity of behavioral reactions that are considered when developing policy and training initiatives.

The study utilized a quantitative, two group, between-participants design. Participants were 264 expert auditors who completed a Personal Need for Structure assessment and five hypothesis-testing tasks after receiving a 4-minute dynamic or linear prime. The five hypothesis testing tasks included one abstract task presented first, followed by four deontic tasks. Results were interpreted using two-way ANOVAs.

The first research question and its two sub-questions examined whether there were statistically significant differences on participants' selection task scores based on their exposure to a dynamic or linear prompt. The first main research question asked if there was a statistically significant difference between combined (abstract and deontic) selection task scores of expert

auditors that are primed with a dynamic reasoning model compared with those primed with a linear reasoning model. No statistically significant difference was found.

The first sub-question to the first research was narrowed to examine only abstract selection task scores of expert auditors primed with a dynamic reasoning model compared with those primed with a linear reasoning model. In this case, a statistically significant difference was found and the null hypothesis was rejected. The second sub-question to the first research was narrowed to examine only deontic selection task scores of expert auditors primed with a dynamic reasoning model compared with those primed with a linear reasoning model. No statistically significant difference was found.

The second research question and two sub questions examined whether there was a statistically significant difference between selection task scores of expert auditors based upon their scores on the Personal Need for Structure inventory. The second main research question asked if there was a statistically significant difference between combined (abstract and deontic) selection task scores of auditors based on their scores on the personal need for structure inventory. No statistically significant difference was found.

The first sub-question of the second research question narrowed the analysis and asked if there was a statistically significant difference between abstract selection task scores of auditors based on their scores on the personal need for structure inventory. Results were not statistically significant. The second sub-question of the second research question narrowed the analysis to ask if there was a statistically significant difference between deontic selection task scores of auditors based on their scores on the personal need for structure inventory. Results showed no statistically significant difference.

The third research question addressed the interaction between the type of prompt received and scores on the Personal Need for Structure Inventory. The third main question asked if there was a statistically significant difference between combined (abstract and deontic) selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory. In this case, the main interaction effect was not statistically significant.

The first sub-question of the third research question narrowed the analysis to ask if there was a statistically significant difference between abstract selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory. The results showed a statistically significant interaction effect and the null hypothesis was rejected.

The second sub-question of the third research question narrowed the analysis to ask if there was a statistically significant difference between deontic selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory. No statistically significant difference was found. Results showed the interaction effect between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory on deontic selection task scores of expert auditors was not statistically significant. There was also no statistically significant main effect of priming on deontic selection task scores of expert auditors.

Discussion of the Results

The first main research question focused on effects of priming mechanisms, specifically asking if dynamic versus linear priming would result in different scores on the Wason Selection Task. The main research question examined overall scores, combining selection task scores of both the single abstract task and four deontic tasks. Cosmides (1989) and Cosmides and Tooby (1989) predicted reasoners would have more success on deontic tasks that have a social contract frame than on abstract tasks that are content independent. The social contract frame of the question may cue relevant contextual decision elements that, for example, allow the reasoner to clearly understand when a party to the social contract is cheated.

Abstract, rule-based tasks are content independent. It is important to examine performance on these abstract tasks because they are similar in structure of abstract rules in the audit environment (Peytcheva, 2013). According to Agoglia et al. (2011) and Dweck et al. (2004), reasoners in such a rule-laden environment tend to frame issues more abstractly to create mental distance. The abstract version was presented first to minimize the chance of deontic versions cueing a social contract frame that could affect participant responses on the abstract version, and research has shown little to no effect of learning when participants move from the abstract version of the task to the deontic versions (Cheng & Holyoak, 1985; Osman, 2007; Rossi et al., 2015). Rapid response was cued by allowing participants 10 seconds to respond with an answer. Rapid response was used to more closely match decision making in real world contexts. Expert decision makers in real world situations rely on rapid assessments of the situation and quick decisions based on heuristics (Lipshitz et al., 2001). According to Evans and Curtis-Holmes (2005), speed of processing is a critical aspect of dual process operation and a rapid, heuristic response increases the tendency for belief bias, or the tendency to evaluate the

validity an argument not based on merits of the logic but whether the reasoner agrees with the conclusion.

Based on past research, it is expected metacognitive priming mechanisms may not be as effective with deontic tasks as with abstract tasks (Fiddick et al., 2017; Peytcheva, 2013). Deontic tasks comprised 80 % of overall task scores examined in the first research question. These expectations played out further when sub-questions separated the analysis between abstract and deontic tasks. The first sub-question to the first research examined only abstract selection task scores, and in this case, results showed statistically significant differences in scores on the abstract Wason Selection Task which increased from the linear prime condition to the dynamic prime condition. When the analysis was narrowed to examine priming effects only on deontic selection task scores, no statistically significant difference was found. Task scores increased very slightly from the linear prime condition to the dynamic prime condition, but these differences were not statistically significant.

The second main research question focused on effects the trait of personal need for structure may have on reasoning efficacy by examining participant scores from the Personal Need for Structure inventory (Thompson et al., 1989). The PNS is a self-report scale focusing on how individuals process ambiguous information within their environment. Higher Personal Need for Structure scores are associated with a variety of reasoning errors, including being more likely to dismiss information that is inconsistent with standardized schemas (Hess et al., 1998), being less likely to update their knowledge with new information (Okun & Rice, 1997; Hess, 2001), and having a higher sensitivity to confirming evidence than disconfirming evidence (Bamber et al., 1997).

The second main research question asked if a statistically significant difference existed between combined (abstract and deontic) selection task scores of auditors based on their scores on the personal need for structure inventory. Results showed the average total combined scores on the Wason Selection Task decreased notably from the Low PNS condition, however differences between groups was not statistically significant. Findings are at least in part due to the distribution of participant PNS scores. An extremely low number of the 264 participants fell in the original low PNS category (n = 4), while the greatest number fell in the moderately high category (n = 114) and nearly eight times as many participants were in the high PNS category (n = 30) as in the low PNS category. This is due to the nature of the sample population, being that individuals who score higher on the PNS scale are drawn to the audit profession (Ruiselová et al., 2012).

Sub-questions to the second research question examined effects of PNS level separately on abstract and deontic selection tasks. The first sub-question to the second research question examined selection task scores on the abstract task that failed to show a statistically significant difference. The second sub-question to the second research question examined selection task scores on deontic versions of task. Differences between groups was not statistically significant, which is due to both the nature of reasoning processes on deontic tasks which are less affected by PNS level, and negatively skewed PNS scores of the sample population.

The third main research question examined the interaction between the type of prompt received and scores on the Personal Need for Structure inventory. The third main question asked if there was a statistically significant difference between combined (abstract and deontic) selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on

the personal need for structure inventory. The analysis did not indicate a statistically significant interaction effect.

Sub-questions to the third research question examined the interaction of priming type and PNS level separately on abstract and deontic selection tasks. The first sub-question of the third research question narrowed the analysis to ask if there was a statistically significant difference between abstract selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory. Results showed a statistically significant difference in the scores. The results point to the complexity of influences affecting reasoning between abstract versus deontic task types.

The second sub-question of the third research question narrowed the analysis to ask if there was a statistically significant difference between the deontic selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory. Again, the analysis showed no statistically significant interaction effect.

Conclusions Based on the Results

This section will focus on examining findings in relation to the previous literature and theoretical framework of the study. The section will begin by comparing findings with those of the previous literature and continue with an interpretation of findings in light of the theoretical foundation of the research.

Comparison of the Findings with the Theoretical Framework and Previous Literature

Dual process theory served as a theoretical foundation for the study. Dual process theory describes two cognitive systems, system one and system two, that interact in the decision-making

process (Kahneman, 2003; Kahneman, 2013; Kahneman & Frederick, 2002; Stanovich & West, 2000). System one is instinctive, fast, intuitive, emotional and subconscious and relies on associative learning processes, and is more likely to be subject to bias (Evans, 2003). System two is slow, logical and conscious, and monitors heuristic operations of system one. System two functionality is described as checking the validity of automatic system one intuitions (Kahneman, 2013). System two requires effort and causes stress which may fatigue reasoners causing depletion. Stress levels vary amongst reasoners depending on their need for structure (Baumeister et al., 1998; Neuberg & Newsom, 1993). Individuals experience varying levels of cognitive dissonance and discomfort when faced with cognitive inconsistencies and ambiguities that result in different individual needs for structure (Webster & Kruglanski, 1996).

Need for structure was measured using the Personal Need for Structure (PNS) scale developed by Thompson et al. (1989). The PNS was introduced by Neuberg and Newsom (1993) as a self-report scale focusing on how individuals process ambiguous information within their environment. High PNS is correlated with a variety of reasoning errors that include a higher sensitivity to confirming evidence than disconfirming evidence (Bamber et al., 1997). Based on this research, it is expected individuals scoring higher on the PNS are more likely to make heuristic errors than those who score lower.

Decision making errors occur when system one generates an intuitive error, which is not, detected by system two, or when system two falsely detects an error where none exists in system one intuition (Morewedge & Kahneman 2010). Reasoning errors can occur where the analytical system needs to override the heuristic system to avoid errors in reasoning and inhibit the acceptance of irrelevant evidence or strategies (Rossi et al., 2015). This inhibition must not only

inhibit the acceptance of irrelevant evidence but also of ineffective, automatic strategies that are triggered by complex or unfamiliar situations.

The Wason Selection Task (WST) was used to examine expert auditor reasoning related to dual process theory. The WST is a hypothesis-testing task that proposes a reasoning frame auditors must routinely navigate as part of their work (Peytcheva, 2013). There are two types of the WST, a deontic version that is context dependent and is based in a social contract frame, and an abstract version that is rule based and content independent. The abstract version of the task is criticized for not being similar to real-world decision making environments; however, the real-world task domain for expert auditors is one imbedded in an abstract rule structure and social contract problems framed with skepticism prompts may not match the frame of the expert auditor (Dweck et al., 2004; Fischoff, 1982; Klauer et al., 2007; Michael, 2006; Tenbrunsel & Messick, 1999). Auditors are likely to frame problems abstractly because they implicitly place significant importance on the abstract rule structure that dominates their environment (Bame-Aldred, 2011; Dweck et al., 2004; Gigerenzer & Hug, 1992; Kahneman, 2013; Klayman & Ha, 1987; Smith & Kida, 1991).

Successful reasoning on the WST requires the reasoner to inhibit heuristic responses of system one and to activate analytical procedures of system two that check the validity of system one's automatic operations (Rossi et al., 2015). Heuristic response errors are caused by attending to superficial aspects of problem solving, minimizing cognitive efforts and avoiding conflicting responses between automatic processes of system one and effortful and controlled processes of system two (Kahneman, 2013). Greater levels of cognitive self-control are needed not just to override heuristic errors of system one, but to go further and activate system two processes. It is possible to inhibit system one activities without spurring activation of system

two (Rossi et al., 2015). The self-regulation required to inhibit heuristic errors in system one and spur the effortful logic of system two is metacognitive and an inward-oriented mechanism (Dweck et al., 2004; Kahneman, 2013).

System one reflects automatic operations spurred by domain specific learning, with conscious access only allowed to the product of those operations. System two represents cognitive networks supporting effortful, sequential processing required for hypothetical reasoning. Normative performance on the WST is adversely affected when a perceptual match is expected but the correct response requires cognitively negotiating a mismatch (Prado & Noveck, 2007). According to dual process theory, a bias erupts immediately because mismatches are sensed as perceptually irrelevant to the conditional rule. When a conditional rule presents two elements to the reasoner, and two of the elements in the choice set do not correspond with elements presented, mismatches are considered as irrelevant. The reasoner must not only overcome the system one tendency to only perceive matching elements as relevant, but also must engage system two to redirect their attention toward logically relevant elements, which in the case of the WST would indicate the true antecedent and the false consequent.

A common bias experienced by auditors on WST types of tasks is the tendency to verify hypotheses by seeking confirmatory evidence supporting a given proposition (Peytcheva, 2013). Reasoners may employ a positive test strategy toward elements perceived to be of greatest importance on tasks such as the WST because the rule based environment may stress the reasoner, pushing him toward familiar rules and easy matches (Kahneman, 2013; Klayman & Ha, 1987). Cosmides (1989) suggested individuals may perform better on tasks prone to confirmation bias if they perceive the task in the outward oriented frame of a social contract rather than as an abstract problem, but auditors work in highly structured environments which

tend to cue a conservative, inward oriented reasoning frame and a motivation to avoid sanctions and loss of reputation (Johns, 2006; Michael, 2006; Tenbrunsel & Messick, 1999). Higher levels of outward oriented skepticism tend to produce positive results more among novices but not experts (Peytcheva, 2013).

Higher levels of outward oriented skepticism were purported to improve performance on reasoning tasks such as the WST (Dawson et al., 2002; Hammersley, 2011; Hurtt, 2010), but Peytcheva (2014) asserted there is lack of empirical evidence to support this presumption. Auditors are motivated to seek evidence corresponding to the problem perceived to be the most important and attention is drawn more to evidence that matches intrinsic motivations or arouses emotions (Kahneman, 2002). Auditors are rewarded more on outcomes than on the quality of the decision process, and regulators are more aggressive in punishing bad outcomes even if the decision process was made with good intentions, than good outcomes that were the result of faulty reasoning (Bazerman & Tenbrunsel, 2011). In these situations, auditors are likely to place significant importance and implicit attention on the abstract rule structure that dominates their environment and frames issues conservatively (Bame-Aldred, 2011; Dweck et al., 2004; Gigerenzer & Hug, 1992; Kahneman, 2013).

In a conservative frame, the default approach would be a tendency toward normative strategies perceived as a quicker and safer way to avoid costly errors. In a conservative environment, an abstract task may frame problems in a less threatening format than a social contract frame, indicating focusing on metacognitive aspects of reasoning may have more power in affecting reasoning performance than a focus on evidentiary skepticism. Inward oriented frames that utilize metacognitive strategies may be more effective at enhancing reasoning than outward oriented strategies because outward oriented skepticism may not match the reasoning

frame of expert auditors in real world auditing contexts (Dweck et al., 2004; Tenbrunsel & Messick, 1999).

When reasoners are primed with an energy-depleting act on the WST, subsequent task performance is impaired (Baumeister et al., 1998). Schmeichel et al. (2003) found participants asked to follow attentional rules reported increased task difficulty and exercised more regulatory exertion as opposed to participants who were not given rules. Kahneman (2013) suggested this depletion results in a reasoner's inability to escape system one heuristics and has heightened susceptibility to stereotypic and superficial reasoning strategies.

Osman (2007) found tutoring based on dynamic, metacognitive strategies by having reasoners focus on their misinterpretations was effective, particularly under circumstances of limited cognitive resources such as a rapid-response version of a selection task. Researchers suggested reasoners were no longer trapped by system one influences. Prado and Noveck (2007) also suggested executive learning is helpful in inhibiting the matching heuristic, prompting reasoners to shift from system one reasoning to system two reasoning. Houdé et al. (2000) found inhibition training was effective at reducing matching bias and heuristic errors while normative logical training or repetition was not helpful. The training may be more effective on abstract tasks than on deontic tasks where the reasoner is already cued to the relevant content (Cosmides, 1989).

Grenier (2011) found experts were more likely to exhibit sound reasoning when they were prompted with dynamic, metacognitive strategies using an inward orientation aimed at questioning their own judgments rather than outward oriented skeptical strategies. Trotman et al. (2009) similarly found prompting auditors to use a dynamic metacognitive strategy utilizing a backward reasoning approach resulted in improved reasoning. Dynamic metacognitive strategies

may assist expert reasoners by reducing heuristic matching errors on experimental selection tasks. Rossi et al. (2015) found priming reasoners with pre-test training in executive learning and rule falsification assisted reasoning performance on the WST. The executive learning training cued reasoners to use a nonlinear, imaginative reasoning frame and to question their intuitive reasoning processes. The training also included warnings to avoid intuitive pitfalls, intended to evoke an emotional response involving a dynamic interaction of cognitive networks.

Research grounded in dual process theory suggests expert auditors would be more likely to make heuristic errors in the form of matching bias on abstract tasks when primed with a linear reasoning strategy, and they would make fewer matching errors on the WST when primed with a dynamic or metacognitive reasoning strategy. These effects are not expected on deontic tasks because according to Cosmides (1989), deontic tasks on the WST framed, as social contracts are not affected as much by framing because the task is content dependent and already cues the reasoner toward relevant stimuli. The linear strategy is considered outward oriented skepticism of evidence in concordance with the audit environment. The dynamic strategy is considered in part, to cue an inward oriented frame of reasoning.

Results of this study corresponded with expectations from research on priming effects related to abstract versus deontic tasks (Cosmides, 1989; Peytcheva, 2013). Findings related to the first research question and sub-questions found statistically significant effects only in the first sub-question that examined abstract selection task scores. Findings related to deontic tasks were not statistically significant.

Research related to the Personal Need for Structure inventory as reviewed above suggested auditors scoring higher on the PNS scale would be more likely to make heuristic errors, especially matching errors because higher scorers prefer simple structures and normative

practices. Research question two and sub-questions examined this by asking if a statistically significant difference existed between combined (abstract and deontic) selection task scores of auditors based on their scores on the personal need for structure inventory. No statistically significant differences were found. The distribution of scores on the PNS were heavily skewed towards higher scores in the population sample. A sample that included more scores in the lower ranges may yield a statistically significant result.

Research related to priming and the Personal Need for Structure inventory as reviewed above suggested there would be an interaction between PNS levels and the type of priming because higher PNS scorers are expected to fall prey to reasoning biases more than low PNS scorers when primed with linear reasoning versus dynamic reasoning frames because they would have a stronger preference for the simple structure of the linear prime than low PNS scorers. This was addressed in the third research question and sub-questions which asked if there was a statistically significant difference between combined (abstract and deontic) selection task scores of auditors based on the interaction between 1: Being primed with a dynamic reasoning model compared with those primed with linear reasoning and 2: Scores on the personal need for structure inventory. Results, while not statistically significant, showed selection task scores increased under the dynamic prime condition from the linear prime condition in every PNS category. While data points toward concurrence with expectations, results not statistically significant, which may be due in part to the severity of the negative skew in participant PNS scores.

Interpretation of the Findings

Results of the study supported expectations that dynamic primes would assist reasoning efficacy on abstract tasks. Similar to findings from Grenier (2011), Osman (2007), Rossi et al.,

(2015), and Trotman et al.'s (2009), results from research questions one and three showed reasoning was enhanced on abstract tasks when reasoners were prompted with dynamic, metacognitive strategies which prompted participants to use an inward orientation aimed at questioning their own judgments rather than outward oriented skeptical strategies. Although results indicated dynamic priming enhanced reasoning efficacy on the abstract task, in accordance with Cosmide's (1989) predictions, dynamic priming did not benefit reasoning processes related to deontic tasks. Linear priming did not assist either abstract or deontic types of tasks, and supplemental analysis showed participants experienced significantly more matching errors under the linear prime condition than the dynamic prime condition on the abstract task.

Results related to priming suggest the state of mind of the reasoner is important to the efficacy of the reasoning process. Findings suggest it would be important to use reasoning strategies that match the frame of the reasoner, as it exists in a given context. Auditors work in cognitively rigid and abstract rule-based environments that may cue reasoners to frame problems in abstract terms, or to rely on normative tools for decision making because of a motivation to avoid errors. In this context, reasoners may benefit more from inward oriented priming on tasks that are abstract in nature, or to be skeptical of their own reasoning processes, which would prompt reasoners to use system two processes to examine system one for heuristic errors of reasoning. In the case of problems framed as social contracts, an outward oriented reasoning strategy may work as inward-based orientation, but reasoners may consider using both outward and inward oriented strategies to change perspective when appropriate. Results suggest the reasoner should use judgment in considering the extent to which he or she is motivated to solve a problem quickly, possibly falling prey to heuristic biases as a result.

Results related to need for structure were not statistically significant under any experimental conditions. Data pointed to support for expectations regarding effects of higher scores on the PNS, such as the tendency to prefer simple rules and superficial reasoning, but findings were not statistically significant due to the skewed distribution of participant scores on the PNS scale. Results indicated participants with low scores on the PNS scale generally had higher scores on the Wason Selection Task, but the distribution of participants amongst different levels of the PNS were negatively skewed, with only four participants in the low PNS category. While data could not allow for the drawing of conclusions regarding effects of the PNS on reasoning efficacy, it may be indicative of preferences of individuals who score higher on the PNS scale toward jobs in the audit field.

Interaction effects between PNS scores and type of priming were addressed in the third research question, where results were not statistically significant. However, tests for main effects again indicated statistically significant effects for priming. Priming effects played a consistent role in affecting reasoning processes throughout the study, while effects of PNS scores were inconclusive due to the skewness of PNS scores with participants heavily weighted toward high PNS scores.

Limitations

The study recruited participants who were auditors attending continuing professional education conferences offered through professional associations that provide professional designations in the auditing industry. While this source of recruitment is known to include a wide variety of professional auditors from government and corporate environments, some public audit firms offer in-house continuing professional education to their staff auditors, and other auditors access their education through online webinar formats or self-study courses. The

population of auditors who seek continuing professional education solely through these sources would not have been accessible to participate in this study, possibly limiting the study's generalizability.

Instruments used in the study included the Personal Need for Structure inventory (Thompson et al., 1989) and the Wason Selection Task (Wason; 1966; Wason, 1968), each of which have limited validity and reliability data available and are unpublished tests.

The study did not find statistically significant effects related to PNS scores that is in part due to the extremely small number of participants in the low PNS category. While some research indicated this issue is unavoidable due to preferences of high PNS individuals toward jobs in auditing, a larger sample size may have allowed results that were statistically significant (Reardon & Lenz, 1999). A different recruitment strategy that could have yielded participants with a greater dispersion of PNS scores may also have solved this issue.

The study was limited to the study of expert auditor reasoning and not students or novice's due to the nature of the inquiry investigating cognitive influences of reasoning in experts as opposed to non-experts. Due to this research focus, results are not intended to provide information related to non-experts. The study was also limited to the audit domain, and results are not intended to provide information in relation to other expert domains.

Implications for Practice

Results of this study showed dynamic primes, as opposed to linear primes, enhanced the reasoning efficacy of expert auditors on abstract tasks, but not deontic tasks. Findings highlight the complexity of expert reasoning processes and the influence task characteristics and cognitive framing have on expert reasoning processes. Findings are a challenge to long-held assumptions regarding the value of traditional skepticism, being higher levels of skepticism equate to higher

efficacy in reasoning and higher quality judgment (Peytcheva, 2013). Results indicate dynamic and metacognitive strategies, including skepticism toward one's own reasoning processes, are more effective at enhancing reasoning efficacy amongst expert auditors than traditional skepticism on abstract rule based tasks that auditors frequently encounter (Osman, 2007; Prado & Noveck, 2007).

Implications of these results in practice are the need for a fundamental shift in professional training that teaches auditors to recognize important differences amongst task types and reasoning frames so they can more effectively adjust their reasoning frame to enhance their reasoning efficacy under different contexts (Gigerenzer & Hug, 1992; Grenier, 2011; Harding & Trotman, 2011). Results also point to the need for regulators and policymakers to consider characteristics which behavioral research have shown influence reasoning processes of expert auditors when drafting policy. Policy initiatives should focus on real world application and implementation in the audit context, avoid oversaturating auditors with linear strategies that are not supported in behavioral literature, and move toward policy that supports dynamic reasoning through challenging assumptions and discouraging blindly following rules and procedures (Kahneman, 2003; Kahneman, 2013; Klein, 1999; Michael, 2006; Stanovich & West, 2000).

Recommendations for Further Research

Supplemental analysis was conducted to determine whether there were differences between types of reasoning errors committed between different experimental conditions which is considered of significant practical interest to researchers in the field (Houdè & Moutier, 1999). The analysis revealed that matching errors increased under the linear prime condition (.64) from the dynamic prime condition (.43). Further research illuminating details that influence reasoning efficacy in this regard would be useful. Additionally, while results related to the PNS scores

were not statistically significant, this may be due to skewed PNS scores in the population sample because of the likelihood of individuals high in need for structure to be drawn to the audit industry (Reardon & Lenz, 1999). A review of scores between low and high PNS scores revealed that of the participants in the study, low PNS scorers were much more likely to correctly answer the abstract selection task (.75 versus .17). Further research that employs sampling strategies that yield a better distribution of PNS scores would be informative.

Further research is recommended in other professional domains outside of auditing including those of physicians, pilots, lawyers and other experts to determine if similar task characteristics and reasoning frames influence reasoning efficacy in those environments. Findings could have important implications for training and policymaking in various expert domains.

Conclusion

This experimental study used a quantitative, 2x3 factorial design to compare effects of dynamic versus linear reasoning strategies on experts' reasoning efficacy on hypothesis testing tasks. The research also examined whether an interaction existed between the type of prompt received and an individual's personal need for structure. The purpose of the research was to determine whether a statistically significant difference existed between the abstract and deontic selection task scores of participants based on receiving either a dynamic or linear reasoning prompt, or whether an interaction existed with the trait of personal need for structure. This was the first experimental study using expert auditors that examined combined effects of personal need for structure and dynamic versus linear prompts on reasoning efficacy. The research presented participants, who were 264 expert auditors, with abstract and deontic hypothesis

testing tasks after exposure to either a dynamic or linear prompt. Results were interpreted using two-way ANOVAs.

The research used dual process theory (Kahneman, 2003; Kahneman, 2013; Stanovich & West, 2000) as a theoretical foundation to examine reasoning processes of expert auditors by comparing differences in reasoning efficacy on hypothesis testing tasks between inward oriented metacognitive versus outward oriented linear reasoning types through the lens of the dual process cognitive framework. Participants were primed through a brief 4-minute exposure to priming document designed to cue either dynamic or linear reasoning strategies.

There were three research questions, each with two sub-questions. The first research question and sub-questions addressed effects of dynamic or linear priming on abstract and deontic selection task scores. Results showed priming significantly affected performance on the abstract task. Specifically, dynamic priming, but not linear priming, enhanced performance on the abstract task, which is notable because auditors routinely face abstract, rule-based reasoning tasks. Linear, outward oriented priming, similar to the traditional skepticism auditors are directed to use, did not assist in reasoning on any form of hypothesis testing tasks. Additionally, neither dynamic nor linear priming assisted performance on deontic tasks, which may be due to the nature of these tasks already cueing reasoners toward relevant evidence (Cosmides, 1989; Cosmides & Tooby, 1989).

The second research question and sub-questions addressed differences in selection task scores based on scores on the Personal Need for Structure scale (Thompson et al., 1989). The PNS is a self-report scale focusing on how individuals process ambiguous information within their environment. Higher scorers on the PNS scale are more likely to make superficial assessments, rely on standardized tools, and fill in missing information with their assumptions

(Neuberg & Newsom, 1993). It is expected that individuals scoring higher in need for structure on the PNS would be more likely to make heuristic errors than those who score lower (Bamber et al., 1997). While results pointed in this direction, results were not statistically significant due to the negative skew of the original five levels of participant scores on the PNS scale. Only four participants out of 264 expert auditors fell in the original low PNS category, which did not allow a statistically significant result. The third research question addressed the interaction between the type of priming and the PNS scores. Due to the negatively skewed distribution of participant scores on the PNS scale, no statistically significant results could be obtained. The few scores that were obtained indicate the possibility of a strong interaction effect.

Supplemental analysis revealed details related to auditor reasoning processes. The linear prime condition that was designed to be similar to the guidelines auditors employ to make decisions, increased the likelihood of making matching errors in reasoning. Additionally, while results related to the PNS scores were not statistically significant, this may be due to skewed PNS scores in the population sample because of the likelihood of individuals high in need for structure to be drawn to the audit industry (Reardon & Lenz, 1999). Supplemental review of scores between low and high PNS scores revealed that of the participants in the study, low PNS scorers were much more likely to correctly answer the abstract selection task (.75 versus .17). These results suggest that further research is needed to explore whether the audit domain is attracting those individuals most likely to make these reasoning errors, and the training that is designed to ameliorate these reasoning errors may make them more and not less likely.

Results of this study challenge assumptions about the role of traditional definitions of professional skepticism in auditor reasoning and audit quality. Findings suggest dynamic metacognitive reasoning strategies are more effective at enhancing reasoning efficacy than

higher levels of traditional skepticism on abstract tasks auditors most routinely face (Osman, 2007; Prado & Noveck, 2007). Expert auditors as opposed to novices, are already saturated with outward oriented skepticism, and are cued to an inward orientation, motivated to protect themselves by strictly following industry rules and guidelines to avoid errors that may bring sanctions or loss of reputation (Tenbrunsel & Messick, 1999). Auditor reasoning is enhanced by matching the reasoning approach to the frame cued by the task, suggesting auditors may enhance their reasoning and judgment by turning the lens of skepticism toward themselves by questioning not just the validity of the evidence, but their own reasoning processes (Dweck et al., 2004; Tenbrunsel & Messick, 1999).

The findings illustrate the need for policies, regulations, and professional education to be designed with a consideration of the real-world influences that affect reasoning processes of auditors. Results of this study suggest fundamental changes in professional training should be considered to educate auditors about cognitive influences of task characteristics and cognitive framing. This type of training would teach auditors to recognize and adjust their cognitive approach via metacognitive strategies to enhance the efficacy of their reasoning processes in order to achieve higher quality judgment and ultimately, audit quality.

REFERENCES

Agoglia, C. P., Doupnik, T. S., & Tsakumis, G. T. (2011). Principles-based versus rules-based accounting standards: The influence of standard precision and audit committee strength on financial reporting decisions. *The Accounting Review*, *86*(3), 747-767. doi.org/10.2308/accr.00000045

Almor, A., & Sloman, S. A. (1996). Is deontic reasoning special? *Psychological Review*, *103*(2), 374-380. doi.org/10.1037/0033-295X.103.2.374

American Institute of Certified Public Accountants (AICPA). (2002). *Consideration of fraud in a financial statement audit. Statement on Auditing Standards No. 99.* New York, NY: Author.

American Institute of Certified Public Accountants (AICPA). (2007). *Consideration of fraud in a financial statement audit. AU Section 316.* New York, NY: Author.

American Institute of Certified Public Accountants (AICPA). (2012). *Analytical procedures: AICPA audit guide.* New York, NY: Author.

Anderson, B. H. (1989). *An examination of factors affecting auditors' hypothesis testing strategies* (Doctoral dissertation). Retrieved from Proquest. (AAI8917326.)

Arnold, V., & Sutton, S. G. (Eds.). (1997). *Behavioral accounting research: Foundations and frontiers.* American Accounting Association. Lakewood Ranch, FL.

Asare, S. K., and A. Wright. 1995. Normative and substantive expertise in multiple hypotheses evaluation. *Organizational Behavior and Human Decision Processes* 64 (3): 171–184. doi.org/10.1006/obhd.1995.1098

Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General*, *130*(2), 224. doi.org/10.1037/0096-3445.130.2.224

Baker, R. L., Bealing, W. E., Jr, Nelson, D. A., & Blair Staley, A. (2006). An institutional perspective of the Sarbanes-Oxley Act. *Managerial Auditing Journal*, *21*(1), 23-33. doi.org/10.1108/02686900610634739

Bamber, E. M. (1983). Expert judgment in the audit team: A source reliability approach. *Journal of Accounting Research*, 396-412. doi.org/10.1016/S0361-3682(96)00029-3

Bamber, E. M., Ramsay, R. J., & Tubbs, R. M. (1997). An examination of the descriptive validity of the belief-adjustment model and alternative attitudes to evidence in auditing. *Accounting, Organizations and Society*, *22*(3), 249-268. doi:10.1016/S0361-3682(96)00029-3

Bame-Aldred, C. (2011). The impact of implicit theories of intelligence on professional decision making. *Journal of Behavioral Studies in Business*, *4*, 1.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*(5), 1252. doi.org/10.1037/0022-3514.74.5.1252

Bazerman, M. H., & Tenbrunsel, A. E. (2011). Ethical breakdowns. *Harvard Business Review*, *89*(4), 58-65.

Bazerman, M. H., Loewenstein, G., & Moore, D. A. (2002). Why good accountants do bad audits. *Harvard Business Review*, *80*(11), 96-103.

Bazerman, M. H., Tenbrunsel, A. E., & Wade-Benzoni, K. (1998). Negotiating with yourself and losing: Making decisions with competing internal preferences. *Academy of Management Review*, *23*(2), 225-241. doi.org/10.5465/AMR.1998.533224

Blay, A. D., Kadous, K., & Sawers, K. (2012). The impact of risk and affect on information search efficiency. *Organizational Behavior and Human Decision Processes*, *117*(1), 80-87. doi.org/10.1016/j.obhdp.2011.09.003

Brink, A. G., Lowe, D. J., & Victoravich, L. M. (2013). The effect of evidence strength and internal rewards on intentions to report fraud in the Dodd-Frank regulatory environment. *Auditing: A Journal of Practice & Theory*, *32*(3), 87-104. doi.org/10.2308/ajpt-50449

Brown, J. R. (2006). Criticizing the critics: Sarbanes Oxley and quack corporate governance. *Marquette Law Review, 90*, 309.

Bryant, D. J. (2002). *Making naturalistic decision making "fast and frugal." Defence Research Development*. Canada. Unpublished manuscript.

Butt, J. L., & Campbell, T. L. (1989). The effects of information order and hypothesis-testing strategies on auditors' judgments. *Accounting, Organizations and Society*, *14*(5), 471-479. doi.org/10.1016/0361-3682(89)90012-3

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, *17*(4), 391-416. doi.org/10.1016/0010-0285(85)90014-3

Coffee, J. C., Jr. (2002). Understanding Enron: "It's about the gatekeepers, stupid." *The Business Lawyer*, 1403-1420. doi.org/10.2139/ssrn.325240

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, *31*(3), 187-276. doi.org/10.1016/0010-0277(89)90023-1

Cosmides, L., & Tooby, J. (1989). Evolutionary psychology and the generation of culture, part II: Case study: A computational theory of social exchange. *Ethology and Sociobiology*, *10*(1), 51-97. doi:10.1016/0162-3095(89)90013-7

Cox, J. R., & Griggs, R. A. (1982). The effects of experience on performance in Wason's selection task. *Memory & Cognition*, *10*(5), 496-502.

Cravens, K., Oliver, E. G., & Ramamoorti, S. (2003). The reputation index: Measuring and managing corporate reputation. *European Management Journal*, *21*(2), 201-212. doi:10.1016/S0263-2373(03)00015-X

Dawson, E., Gilovich, T., & Regan, D. T. (2002). Motivated reasoning and performance on the Wason selection task. *Personality and Social Psychology Bulletin*, *28*(10), 1379-1387. doi:10.1177/014616702236869

Dweck, C. S., Mangels, J. A., Good, C., Dai, D. Y., & Sternberg, R. J. (2004). Motivational effects on attention, cognition, and performance. *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development*, *2*, 41-55.

Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review*, *85*(5), 395. doi: 10.1037/0033-295X.85.5.395

Einhorn, J. (1974). Expert judgment: Some necessary conditions and an example. *Journal of Applied Psychology*, *59*, 562–571. doi:10.1037/h0037164

Einhorn, H. J., & Hogarth, R. M. (1988). Decision making under ambiguity: A note. In *Risk, decision and rationality* (pp. 327-336). Springer Netherlands. doi:10.1007/978-94-009-4019-2_19

Elovainio, M., & Kivimäki, M. (2001). The effects of personal need for structure and occupational identity in the role stress process. *The Journal of Social Psychology*, *141*(3), 365-378. doi.org/10.1080/00224540109600558

Elliot, A. J., & Harackiewicz, J. M. (1994). Goal setting, achievement orientation, and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology*, *66*(5), 968-980. doi:10.1037/0022-3514.66.5.968

Ericsson, K. A., Prietula, M. J., & Cokely, E. T. (2007). The making of an expert. *Harvard business review*, *85*(7/8), 114.

Evans, J. S. B. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, *7*(10), 454-459. doi:10.1016/j.tics.2003.08.012

Evans, J. S. B. (2013). The psychology of deductive reasoning (psychology revivals). *Psychology Press*. New York, NY.

Evans, J. S. B., & Curtis-Holmes, J. (2005). Rapid responding increases belief bias: Evidence for the dual-process theory of reasoning. *Reasoning & Reasoning, 11*(4), 382-389. doi:10.1080/13546780542000005

Evans, J. S. B., & Over, D. E. (1996a). *Rationality and reasoning*. Hove, UK: Psychology Press.

Evans, J. S. B., & Over, D. E. (1996b). *Rationality in the selection task: Epistemic utility versus uncertainty reduction.* Hove, UK: Psychology Press.

Evans, J. T. (2007). On the resolution of conflict in dual process theories of reasoning. *Reasoning & Reasoning*, *13*(4), 321-339. doi:10.1080/13546780601008825

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191. doi:10.3758/BRM.41.4.1149

Fiddick, L., Brase, G. L., Cosmides, L., & Tooby, J. (2017). Rethinking relevance: Repetition priming reveals the psychological reality of adaptive specializations for reasoning. *Evolution and Human Behavior*, *38*(3), 366-375. doi.org/10.1016/j.evolhumbehav.2016.11.008

Fischoff, B. (1982). Debiasing. In D.Kahneman, P. Slovic, &A.Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 422- 444). New York: Cambridge University Press. doi.org/10.1017/CBO9780511809477.032

Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives*, *19*(4), 25-42. doi.org/10.1257/089533005775196732

George, C. (1991). Facilitation in the Wason selection task with a consequent referring to an unsatisfactory outcome. *British Journal of Psychology*, *82*(4), 463-472. doi.org/10.1111/j.2044-8295.1991.tb02413.x

Gigerenzer, G., & Hug, K. (1992). Domain-specific reasoning: Social contracts, cheating, and perspective change. *Cognition*, *43*(2), 127-171. doi.org/10.1016/0010-0277(92)90060-U

Gilhooly, K. J., Georgiou, G., & Devery, U. (2013). Incubation and creativity: Do something different. Thinking & Reasoning, 19(2), 137-149. doi.org/10.1080/13546783.2012.749812

Good, C., & Dweck, C. S. (2006). A motivational approach to reasoning, resilience, and responsibility. Optimizing student success with the other three Rs, *reasoning, resilience, and responsibility*, 39-56.

Grenier, J. H. (2011). Encouraging professional skepticism in the industry specialization era. *Journal of Business Ethics*, 1-16. doi.org/10.2139/ssrn.1533622

Griggs, R. A., & Cox, J. R. (1982). The elusive thematic- materials effect in Wason's selection task. *British Journal of Psychology*, *73*(3), 407-420. doi.org/10.1111/j.2044-8295.1982.tb01823.x

Hallihan, G. M., & Shu, L. H. (2013). Considering confirmation bias in design and design research. *Journal of Integrated Design and Process Science*, *17*(4), 19-35.

Hammersley, J. S. (2011). A review and model of auditor judgments in fraud-related planning tasks. *Auditing: A Journal of Practice & Theory*, *30*(4), 101-128. doi.org/10.2308/ajpt-10145

Hamtiaux, A., & Houssemand, C. (2012). Adaptability, cognitive flexibility, personal need for structure, and rigidity. *Psychology Research*, *2*(10), 563.

Harding, N., & Trotman, K. T. (2011). Enhancing professional skepticism via the fraud brainstorming discussion outcomes. School of Accounting Australian School of Business The University of New South Wales Sydney NSW, 2052.

Hess, T. M. (2001). Ageing-related influences on personal need for structure. *International Journal of Behavioral Development*, *25*(6), 482-490. doi.org/10.1080/01650250042000429

Hess, T. M., Follett, K. J., & McGee, K. A. (1998). Aging and impression formation: The impact of processing skills and goals. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *53*(3), P175-P187. doi.org/10.1093/geronb/53B.3.P175

Higgins, E. T. (1996). The “self digest": self-knowledge serving self-regulatory functions. Journal of Personality and Social Psychology, 71(6), 1062. doi.org/10.1037/0022-3514.71.6.1062

Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology*, *24*(1), 1-55. doi.org/10.1016/0010-0285(92)90002-J

Houdé, O., & Moutier, S. (1999). Deductive reasoning and experimental inhibition training: The case of the matching bias. New data and reply to Girotto. *Cahiers de Psychologie Cognitive*, *18*(1), 75-85.

Houdé, O., Zago, L., Crivello, F., Moutier, S., Pineau, A., & Mazoyer, B. (2001). Access to deductive logic depends on a right ventromedial prefrontal area devoted to emotion and feeling: Evidence from a training paradigm. *Neuroimage*, *14*(6), 1486-1492. doi.org/10.1006/nimg.2001.0930

Houdé, O., Zago, L., Mellet, E., Moutier, S., Pineau, A., Mazoyer, B., & Tzourio-Mazoyer, N. (2000). Shifting from the perceptual brain to the logical brain: The neural impact of cognitive inhibition training. *Journal of cognitive neuroscience*, *12*(5), 721-728. doi.org/10.1162/089892900562525

Howell, D. C. (2017). *Fundamental statistics for the behavioral sciences*. (9th ed.) Boston, MA: Cengage Learning.

Hurtt, R. K. (2010). Development of a scale to measure professional skepticism. *Auditing: A Journal of Practice & Theory*, *29*(1), 149-171. doi.org/10.2308/aud.2010.29.1.149

Hurtt, R. K., Brown-Liburd, H., Earley, C. E., & Krishnamoorthy, G. (2013). Research on auditor professional skepticism: Literature synthesis and opportunities for future research. *Auditing: A Journal of Practice & Theory*, *32*(sp1), 45-97. doi.org/10.2308/ajpt-50361

Hurtt, R. K., Eining, M., & Plumlee, D. (2012). *Linking professional skepticism to auditors' behaviors* (Working Paper). Baylor University, and The University of Utah.

Jaccard, J. (1998). *Interaction effects in factorial analysis of variance* (No. 118). Sage. doi.org/10.4135/9781412984508

Jamal, K., Johnson, P. E., & Berryman, R. (1995). Detecting framing effects in financial statements. *Contemporary Accounting Research*, *12*(1), 85-105. doi.org/10.1111/j.1911-3846.1995.tb00482.x

Jermias, J. (2006). The Influence of Accountability on Overconfidence and Resistance to Change: A Research Framework and Experimental Evidence. Management Accounting Research, 17(4), 370-388. doi.org/10.1016/j.mar.2006.03.003

Johns, G. (2006). The essential impact of context on organizational behavior. *Academy of management review*, *31*(2), 398-408. doi.org/10.5465/AMR.2006.20208687

Kaempf, G. L., Klein, G., Thordsen, M. L., & Wolf, S. (1996). Decision making in complex naval command-and-control environments. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *38*(2), 220-231. doi.org/10.1177/001872089606380204

Kahneman, D. (1991). Judgment and decision making: A personal view. *Psychological science*, *2*(3), 142-145. doi.org/10.1111/j.1467-9280.1991.tb00121.x

Kahneman, D. (2002). Maps of bounded rationality: A perspective on intuitive judgment and choice. [Nobel prize lecture]. *8*, 351-401.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review*, *93*(5), 1449-1475. doi.org/10.1257/000282803322655392

Kahneman, D. (2013). Think, fast and slow. *Farrar, Straus and Giroux, New York*.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgement. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgement* (pp. 49 – 81). Cambridge, MA: Cambridge University Press. doi.org/10.1017/CBO9780511808098.004

Kahneman, D., & Frederick, S. (2007). Frames and brains: Elicitation and control of response tendencies. *Trends in cognitive sciences*, *11*(2), 45-46. doi.org/10.1017/CBO9780511808098.004

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist*, *64*(6), 515. doi.org/10.1037/a0016755

Kida, T. (1984). The impact of hypothesis-testing strategies on auditors' use of judgment data. *Journal of Accounting Research*, 332-340. doi.org/10.2307/2490714

Klauer, K. C., Stahl, C., & Erdfelder, E. (2007). The abstract selection task: New data and an almost comprehensive model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 680. doi.org/10.1037/0278-7393.33.4.680

Klayman, J., & Ha, Y. W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*(2), 211. doi.org/10.1037/0033-295X.94.2.211

Klein, G. (1997). The recognition-primed decision (RPD) model: Looking back, looking forward. *Naturalistic Decision Making*, 285-292.

Klein, G. A., Calderwood, R., & Macgregor, D. (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man, and Cybernetics*, *19*(3), 462-472. doi.org/10.1109/21.31053

Kleinmuntz, D. N. (1985). Cognitive heuristics and feedback in a dynamic decision environment. *Management Science*, *31*(6), 680-702. doi.org/10.1287/mnsc.31.6.680

Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal Of Experimental Psychology: Learning, Memory, And Cognition*, *25*(6), 1534-1555. doi.org/10.1037/0278-7393.25.6.1534

Koonce, L., & Mercer, M. (2005). *Using psychology theories in archival financial accounting research* (McCombs Research Paper Series No. ACC-01-05).

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: “Seizing" and “freezing." *Psychological Review*, *103*(2), 263. doi.org/10.1037/0033-295X.103.2.263

Krull, D. S. (1993). Does the grist change the mill? The effect of the perceiver's inferential goal on the process of social inference. *Personality and Social Psychology Bulletin*, *19*(3), 340-348. doi.org/10.1177/0146167293193011

Kunreuther, H., & Slavic, P. (2002). 12 The affect heuristic: Implications for understanding and managing risk-induced stigma. *Judgments, Decisions, and Public Policy*, 303.

Leedy, P., D., & Ormrod, J. E. (2013*). Practical research: Planning and design* (10th ed.). New York, NY: Pearson.

Lenard, M. J., & Alam, P. (2009). An historical perspective on fraud detection: From bankruptcy models to most effective indicators of fraud in recent incidents. *Journal of Forensic & Investigative Accounting*, *1*(1), 1-27.

Lichtenstein, B. M. B., & Mendenhall, M. (2002). Non-linearity and response-ability: Emergent order in 21st-century careers. *Human Relations*, *55*(1), 5-32. doi.org/10.1177/0018726702055001604

Lipshitz, R., Klein, G., Orasanu, J., & Salas, E. (2001). Taking stock of naturalistic decision making. *Journal of Behavioral Decision Making*, *14*(5), 331-352. doi.org/10.1002/bdm.381

Maani, K. E., & Maharaj, V. (2004). Links between systems reasoning and complex decision making. *System Dynamics Review*, *20*(1), 21-48. doi.org/10.1002/sdr.281

McGregor, I., Zanna, M. P., Holmes, J. G., & Spencer, S. J. (2001). Compensatory conviction in the face of personal uncertainty: Going to extremes and being oneself. *Journal of Personality and Social Psychology*, *80*(3), 472. doi.org/10.1037/0022-3514.80.3.472

Michael, M. L. (2006). Business ethics: The law of rules. *Business Ethics Quarterly*, 475-504. doi.org/10.5840/beq200616445

Messier. W.F. & Tubbs R.M. (1994) Recency effects in belief revision: the impact of audit experience and the review process. *Auditing: A Journal of Practice and Theory* 13, 57-72.

Mölle, M., Marshall, L., Lutzenberger, W., Pietrowsky, R., Fehm, H., & Born, J. (1996). Enhanced dynamic complexity in the human EEG during creative reasoning. *Neuroscience Letters*, *208*(1), 61-64. doi.org/10.1016/0304-3940(96)12539-8

Mongin, P. (1997). Expected utility theory. *Handbook of Economic Methodology*, 342-350.

Morewedge, C. K., & Kahneman, D. (2010). Associative processes in intuitive judgment. *Trends in Cognitive Sciences*, *14*(10), 435-440. doi.org/10.1016/j.tics.2010.07.004

Moskowitz, G. B. (1993). Individual differences in social categorization: The influence of personal need for structure on spontaneous trait inferences. *Journal of Personality and Social Psychology*, *65*(1), 132. doi.org/10.1037/0022-3514.65.1.132

Mueller, J. M., & Anderson, J. C. (2002). Decision aids for generating analytical review alternatives: The impact of goal framing and audit-risk.level. *Behavioral Research in Accounting*, *14*(1), 157-177. doi.org/10.2308/bria.2002.14.1.157

Nelson, M. W. (2009). A model and literature review of professional skepticism in auditing. Auditing: *A Journal of Practice & Theory*, *28*(2), 1-34. doi.org/10.2308/aud.2009.28.2.1

Neuberg, S. L., & Newsom, J. T. (1993). Personal need for structure: Individual differences in the desire for simpler structure. *Journal of Personality and Social Psychology*, *65*(1), 113. doi.org/10.1037/0022-3514.65.1.113

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review*, *108*(2), 291. doi.org/10.1037/0033-295X.108.2.291

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, *101*(4), 608. doi.org/10.1037/0033-295X.101.4.608

Oberauer, K., Wilhelm, O., & Diaz, R. R. (1999). Bayesian Rationality for the Wason Selection Task? A Test of Optimal Data Selection Theory. *Thinking & Reasoning*, *5*(2), 115-144. doi.org/10.1080/135467899394020

Okun, M. A., & Rice, G. E. (1997). Overcoming elders' misconceptions about accurate written medical information. *Journal of Applied Gerontology*, *16*(1), 51-70. doi.org/10.1177/073346489701600103

Osman, M. (2007). Can tutoring improve performance on a reasoning task under deadline conditions? *Memory & Cognition*, *35*(2), 342-351. doi.org/10.3758/BF03193455

Ostafin, B. D., & Kassman, K. T. (2012). Stepping out of history: mindfulness improves insight problem solving. *Consciousness and cognition*, *21*(2), 1031-1036. doi.org/10.1016/j.concog.2012.02.014

Palmrose, Z. V. (2006). Maintaining the value and viability of independent auditors as gatekeepers under SOX: An auditing master proposal. *Financial Gatekeepers: Can They Protect Investors,* 103-135.

Peecher, M. E., Solomon, I., & Trotman, K. T. (2013). An accountability framework for financial statement auditors and related research questions. *Accounting, Organizations and Society*, *38*(8), 596-620. DOI: 10.1016/j.aos.2013.07.002

Peytcheva, M. (2013). Professional skepticism and auditor cognitive performance in a hypothesis-testing task. *Managerial Auditing Journal*, *29*(1), 27-49. doi.org/10.1108/MAJ-04-2013-0852

Peytcheva, M., Wright, A. M., & Majoor, B. (2014). The impact of principles-based versus rules-based accounting standards on auditors' motivations and evidence demands. *Behavioral Research in Accounting*, *26*(2), 51-72. doi.org/10.2308/bria-50707

Pike, B. J., Curtis, M. B., & Chui, L. (2013). How does an initial expectation bias influence auditors' application and performance of analytical procedures? *The Accounting Review*, *88*(4), 1413-1431. doi.org/10.2308/accr-50426

Pinsker, R., Pennington, R., & Schafer, J. K. (2009). The influence of roles, advocacy, and adaptation to the accounting decision environment. *Behavioral Research in Accounting*, *21*(2), 91-111. doi.org/10.2308/bria.2009.21.2.91

Pituch, K. A., & Stevens, J. P. (2016). *Applied multivariate statistics for the social sciences* (6th ed.). New York, NY: Routledge.

Plumlee, D., Rixom, B. A., & Rosman, A. J. (2012). *Training auditors to think skeptically.* (Working paper). The University of Utah, and University of Connecticut.

Popova, V. (2012). Exploration of skepticism, client-specific experiences, and audit judgments. *Managerial Auditing Journal*, *28*(2), 140-160. doi.org/10.1108/02686901311284540

Prado, J., & Noveck, I. A. (2007). Overcoming perceptual features in logical reasoning: A parametric functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience*, *19*(4), 642-657. doi.org/10.1162/jocn.2007.19.4.642

Pruitt, J. S., Cannon-Bowers, J. A., & Salas, E. (1997). In search of naturalistic decisions. *Decision making under stress- Emerging themes and applications (A 99-12526 01-53), Aldershot, United Kingdom, Ashgate, 1997,* 29-42.

Public Company Accounting Oversight Board (PCAOB). (2003). Audit documentation. Auditing standard §230.07, 230.08, 230.09, 316.13. Washington, DC: Author.

Public Company Accounting Oversight Board (PCAOB). (2004). *An audit of internal control over financial reporting performed in conjunction with an audit of financial statements*. PCAOB Auditing Standard No. 2. Washington, DC: Author. Available at: https://pcaobus.org/Standards/Archived/Pages/Auditing_Standard_2.aspx

Public Company Accounting Oversight Board (PCAOB). (2007). *An audit of internal control over financial reporting that is integrated with an audit of financial statements*. PCAOB Auditing Standard No. 5. Washington, DC: Author. Available at: https://pcaobus.org/Standards/Auditing/pages/auditing_standard_5.aspx

Ramamoorti, S., & Nayar, M. K. (2013). The importance of information integrity. Internal Auditor, 70(1), 29-31.

Reardon, R. C., & Lenz, J. G. (1999). Holland's theory and career assessment. *Journal of Vocational Behavior*, *55*(1), 102-113. doi.org/10.1006/jvbe.1999.1700

Redelmeier, D. A., & Shafir, E. (1995). Medical decision making in situations that offer multiple alternatives. Journal of the American Medical Association (JAMA), *273*(4), 302-305. doi.org/10.1001/jama.1995.03520280048038

Reyna, V. F. (2004). How people make decisions that involve risk a dual-processes approach. *Current Directions in Psychological Science*, *13*(2), 60-66. doi.org/10.1111/j.0963-7214.2004.00275.x

Rezaee, Z. (2004). Restoring public trust in the accounting profession by developing anti-fraud education, programs, and auditing. *Managerial Auditing Journal*, *19*(1), 134-148. doi.org/10.1108/02686900410509857

Richmond, B. (1997). Dynamic reasoning: A behavioral context. *The Systems Thinker*, *8*(6), 6-7.

Roberts, M. J., & Newton, E. J. (2002). Inspection times, the change task, and the rapid-response selection task. *Quarterly Journal of Experimental Psychology, 54A*, 1031 – 1048.

Romano, R. (2004). The Sarbanes-Oxley Act and the making of quack corporate governance. *Yale Law and Econ Research Paper,* 297. doi.org/10.2139/ssrn.596101

Rossi, S., Cassotti, M., Moutier, S., Delcroix, N., & Houdé, O. (2015). Helping reasoners succeed in the Wason selection task: When executive learning discourages heuristic response, but does not necessarily encourage logic. *PloS One*, *10*(4). doi.org/10.1371/journal.pone.0123024

Ruiselová, Z., Prokopcáková, A., & Kresánek, J. (2012). Personal need for structure related to counterfactual reasoning and wisdom. *Studia Psychologica*, *54*(4), 287.

Saleh, S. D., & Grygier, T. G. (1969). Psychodynamics of intrinsic and extrinsic job orientation. *Journal of Applied Psychology*, *53*(6), 446-450. doi.org/10.1037/h0028661

Schaller, M., Boyd, C., Yohannes, J., & O'Brien, M. (1995). The prejudiced personality revisited: Personal need for structure and formation of erroneous group stereotypes. *Journal of Personality and Social Psychology*, *68*(3), 544. doi.org/10.1037/0022-3514.68.3.544

Schmeichel, B. J., Vohs, K. D., & Baumeister, R. F. (2003). Intellectual performance and ego depletion: Role of the self in logical reasoning and other information processing. *Journal of Personality and Social Psychology*, *85*(1), 33. doi.org/10.1037/0022-3514.85.1.33

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Wadsworth Cengage learning.

Shanteau, J. (2015). Why task domains (still) matter for understanding expertise. *Journal of Applied Research in Memory and Cognition*, *4*(3), 169-175. doi.org/10.1016/j.jarmac.2015.07.003

Slijkhuis, J. M., Rietzschel, E. F., & Van Yperen, N. W. (2013). How evaluation and need for structure affect motivation and creativity. *European Journal of Work and Organizational Psychology*, *22*(1), 15-25. doi.org/10.1080/1359432X.2011.626244

Smith, J. F., & Kida, T. (1991). Heuristics and biases: Expertise and task realism in auditing. *Psychological Bulletin*, *109*(3), 472. doi.org/10.1037/0033-2909.109.3.472

Stanovich, K. E., & Thompson, V. A. (2001). Who is rational: Studies of individual differences in reasoning. *Canadian Journal of Experimental Psychology*, *55*(1), 77.

Stanovich, K. E., & West, R. F. (2000). Advancing the rationality debate. *Behavioral and Brain Sciences, 23*(05), 701-717. doi.org/10.1017/S0140525X00623439

Tenbrunsel, A. E., & Messick, D. M. (1999). Sanctioning systems, decision frames, and cooperation. *Administrative Science Quarterly, 44*(4), 684-707. doi.org/10.2307/2667052

Thompson, M. M. (1998). Individual differences in decision making: Establishing cognitive style normative values for a military sample. Downsview, Ontario: Defence and Civil Institute of Environmental Medicine.

Thompson, M. M., Naccarato, M. E., & Parker, K. E. (1989). *Assessing cognitive need: The development of the personal need for structure and personal fear of invalidity scales.* Paper presented at the annual meeting of the Canadian Psychological Association, Halifax, Nova Scotia.

Tice, D. M., Baumeister, R. F., Shmueli, D., & Muraven, M. (2007). Restoring the self: Positive affect helps improve self-regulation following ego depletion. *Journal of Experimental Social Psychology*, *43*(3), 379-384. doi.org/10.1016/j.jesp.2006.05.007

Trotman, K. T., Simnett, R., & Khalifa, A. (2009). Impact of the type of audit team discussions on auditors' generation of material frauds. *Contemporary Accounting Research*, *26*(4), 1115-1142. doi.org/10.1506/car.26.4.5

Trotman, K.T. & Wright, A. (1996). Recency effects: Task complexity, decision mode, and task-specific experience. *Behavioral Research in Accounting* 8, 175-193.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, *76*(2), 105. doi.org/10.1037/h0031322

Tversky, A., & Kahneman, D. (1975). Judgment under uncertainty: Heuristics and biases. In *Utility, probability, and human decision making* (pp. 141-162). Netherlands: Springer. doi.org/10.1007/978-94-010-1834-0_8

Vance, C. M., Groves, K. S., Paik, Y., & Kindler, H. (2007). Understanding and measuring linear–nonlinear thinking style for enhanced management education and professional practice. *Academy of Management Learning & Education*, *6*(2), 167-185. doi.org/10.5465/AMLE.2007.25223457

Walton, D. N. (1990). What is reasoning? What is an argument? *The Journal of Philosophy*, *87*(8), 399-419.

Wason, P.C. (1966). Reasoning. In B. Foss (Ed.), *New horizons in psychology*. Harmondsworth, UK: Penguin.

Wason, P. C. (1968). Reasoning about a rule. *The Quarterly Journal of Experimental Psychology*, *20*(3), 273-281. doi.org/10.1080/14640746808400161

Wason, P.C. (1969). Regression in reasoning? *British Journal of Psychology*, *60*, 471-480. doi.org/10.1111/j.2044-8295.1969.tb01221.x

Wason, P.C. (1977). Self-contradictions. In P.N. Johnson-Laird & P.C. Wason (Eds.), *Reasoning: Readings in cognitive science.* Cambridge, UK: Cambridge University Press.

Wason, EC., & Green, D. (1984). Reasoning and mental representation. *Quarterly Journal of Experimental Psychology*, 36A, 597-610. doi.org/10.1080/14640748408402181

Wason, P.C., & Johnson-Laird, EN. (1970). A conflict between selecting and evaluating information in an inferential task. *British Journal of Psychology*, *61*, 509-515. doi.org/10.1111/j.2044-8295.1970.tb01270.x

Wason, EC., & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology*, *23*, 63-71. doi.org/10.1080/00335557143000068

Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology*, *67*(6), 1049. doi.org/10.1037/0022-3514.67.6.1049

STATEMENT OF ORIGINAL WORK

Academic Honesty Policy

Capella University's Academic Honesty Policy (3.01.01) holds learners accountable for the integrity of work they submit, which includes but is not limited to discussion postings, assignments, comprehensive exams, and the dissertation or capstone project.

Established in the Policy are the expectations for original work, rationale for the policy, definition of terms that pertain to academic honesty and original work, and disciplinary consequences of academic dishonesty. Also stated in the Policy is the expectation that learners will follow APA rules for citing another person's ideas or works.

The following standards for original work and definition of *plagiarism* are discussed in the Policy:

> Learners are expected to be the sole authors of their work and to acknowledge the authorship of others' work through proper citation and reference. Use of another person's ideas, including another learner's, without proper reference or citation constitutes plagiarism and academic dishonesty and is prohibited conduct. (p. 1)
>
> Plagiarism is one example of academic dishonesty. Plagiarism is presenting someone else's ideas or work as your own. Plagiarism also includes copying verbatim or rephrasing ideas without properly acknowledging the source by author, date, and publication medium. (p. 2)

Capella University's Research Misconduct Policy (3.03.06) holds learners accountable for research integrity. What constitutes research misconduct is discussed in the Policy:

> Research misconduct includes but is not limited to falsification, fabrication, plagiarism, misappropriation, or other practices that seriously deviate from those that are commonly accepted within the academic community for proposing, conducting, or reviewing research, or in reporting research results. (p. 1)

Learners failing to abide by these policies are subject to consequences, including but not limited to dismissal or revocation of the degree.

Statement of Original Work and Signature

I have read, understood, and abided by Capella University's Academic Honesty Policy (3.01.01) and Research Misconduct Policy (3.03.06), including Policy Statements, Rationale, and Definitions.

I attest that this dissertation or capstone project is my own work. Where I have used the ideas or words of others, I have paraphrased, summarized, or used direct quotes following the guidelines set forth in the APA *Publication Manual*.

Learner name and date  11-21-2017

APPENDIX A. PERSONAL NEED FOR STRUCTURE SCALE

Read the following statements and decide how much you agree with each according to your attitudes, beliefs, and experiences. It is important for you to realize that there are no "right" or "wrong" answers to these questions.

1 = *strongly disagree* 2 = *moderately disagree* 3 = *slightly disagree* 4 = *slightly agree*

5 = *moderately agree* 6 = *strongly agree*

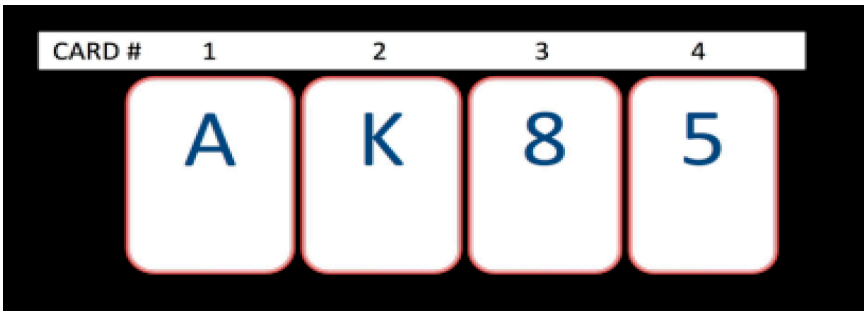
1. It upsets me to go into a situation without knowing what I can expect from it
2. I'm not bothered by things that interrupt my daily routine.
3. I enjoy having a clear and structured mode of life.
4. I like to have a place for everything and everything in its place.
5. I enjoy being spontaneous
6. I find that a well-ordered life with regular hours makes my life tedious
7. I don't like situations that are uncertain
8. I hate to change my plans at the last minute
9. I hate to be with people who are unpredictable
10. I find that a consistent routine enables me to enjoy life more
11. I enjoy the exhilaration of being in unpredictable situations
12. I become uncomfortable when the rules in a situation are not clear.

Items 2, 5, 6 and 11 are reverse scored

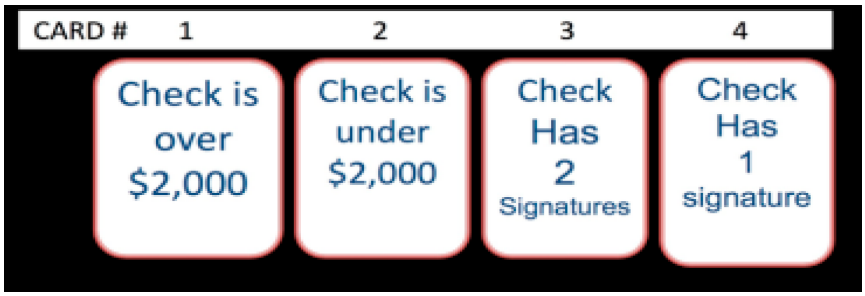
Scores are separated into 5 groups

Low	12-23
Moderately Low	24-35
Medium	36-47
Moderately High	48-59
High	60-72

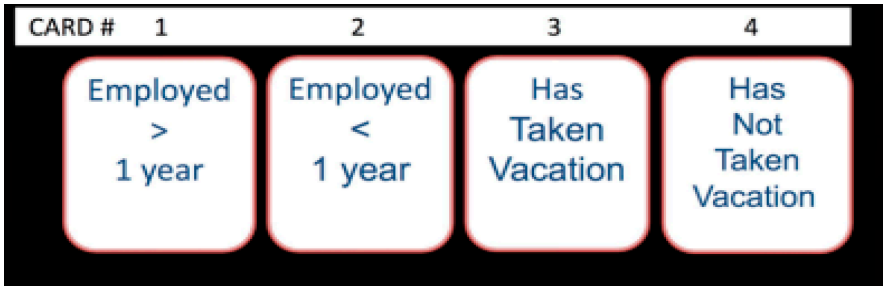
APPENDIX B. WASON SELECTION TASKS



Each card has a number on one side and a letter on the other side. Indicate which two cards you would choose to validate the following rule: If there is a vowel on one side of the card, there is an even number on the other side of the card.

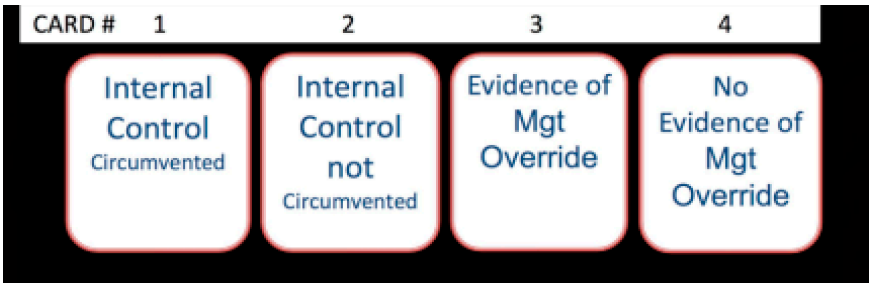


Each card indicates whether a check is over or under $2,000 on one side, and whether the check has 1 signature or two signatures on the other side. Indicate which two cards you would choose to validate the following rule: If a check is over $2,000 then it has two signatures.

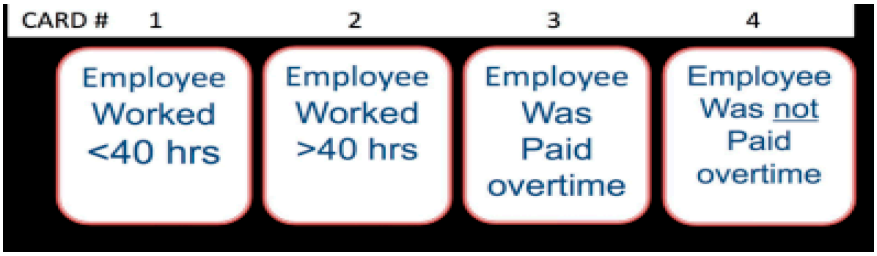


Each card indicates whether an employee has been employed more than one year or less than one year on one side and whether the employee has taken vacation or not on the other side.

Indicate which two cards you would choose to validate the following rule: If an employee has been employed for more than a year then that employee has taken vacation.



Each card indicates whether an internal control has been circumvented or not on one side, and whether there is evidence of management override or not on the other side. Indicate which two cards you would choose to validate the following rule: If an internal control was circumvented, there is evidence of management override.



Each card indicates whether an employee has worked more than 40 hours or less than 40 hours on one side and whether the employee was paid overtime or not on the other side. Indicate which two cards you would choose to validate the following rule: If an employee worked more than 40 hours then they were paid overtime.

APPENDIX C: DYNAMIC AND LINEAR PRIMING DOCUMENTS

N N

PLEASE READ CAREFULLY

When seeking to validate information, auditors often fall prey to something psychologists call "confirmation bias," which is giving attention to information that supports what you were told or what you already believe.

As an example of this type of reasoning, consider that you would not be able to find omitted items that cause an understatement of an account by sampling the recorded items. You would have to seek ways to find information that might seem counterintuitive.

Research indicates that auditors may achieve better critical reasoning by asking themselves why they are paying attention to certain criteria, and challenging themselves to create context by looking at contrasting information that is "opposite" to which their attention was first drawn. Many times this contrasting information holds more power to properly validate what you have been told. This is not always true though – sometimes the information is irrelevant. You can determine if information is relevant by determining if the information has the power to both confirm and disconfirm (validate as well as invalidate) what you have been told. Again, you will be naturally drawn to those things that confirm what you have been told – so fight that urge and look toward contrasting information.

You will now be verbally presented with a series of scenarios related to what you see on the screen and you will be asked to make choices

N N

L L

PLEASE READ CAREFULLY

There are two general approaches to audit sampling: non-statistical and statistical. Both approaches require that the auditor use professional judgment in planning, performing, and evaluating a sample and in relating the audit evidence produced by the sample to other audit evidence when forming a conclusion about the related account balance or class of transactions. Evaluating the appropriateness of audit evidence is solely a matter of auditing judgment and is not determined by the design and evaluation of an audit sample.

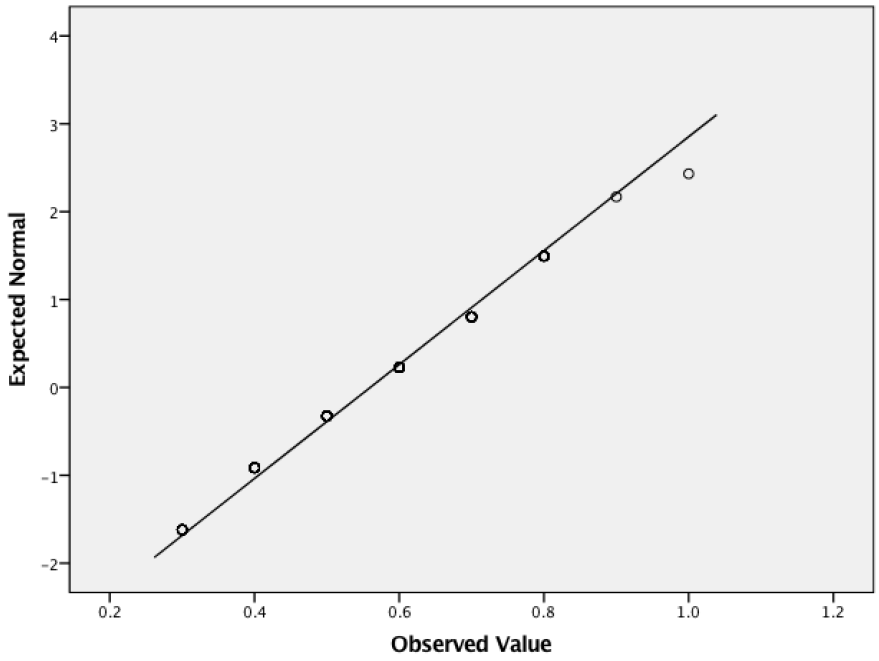
When planning a sample, the auditor should consider the specific audit objective to be achieved. Sample results for tests of controls are evaluated by comparing the sample deviation rate to the tolerable deviation rate and calculating an allowance for the sampling risk. The sample deviation rate is calculated by dividing the number of sample deviations by the sample size. The allowance for sampling risk is calculated by subtracting the sample deviation rate from the tolerable deviation rate. If the allowance for sampling risk is large and positive the auditor would most likely conclude that the design and operation of an internal control is effective. However, if the allowance for sampling risk were small or negative the auditor would conclude that the design and operation of an internal control is not effective. What constitutes a large enough difference is a matter for professional judgment? Sample size for a non-statistical sample is at the auditor's discretion.

You will now be verbally presented with a series of scenarios related to what you see on the screen and you will be asked to make choices

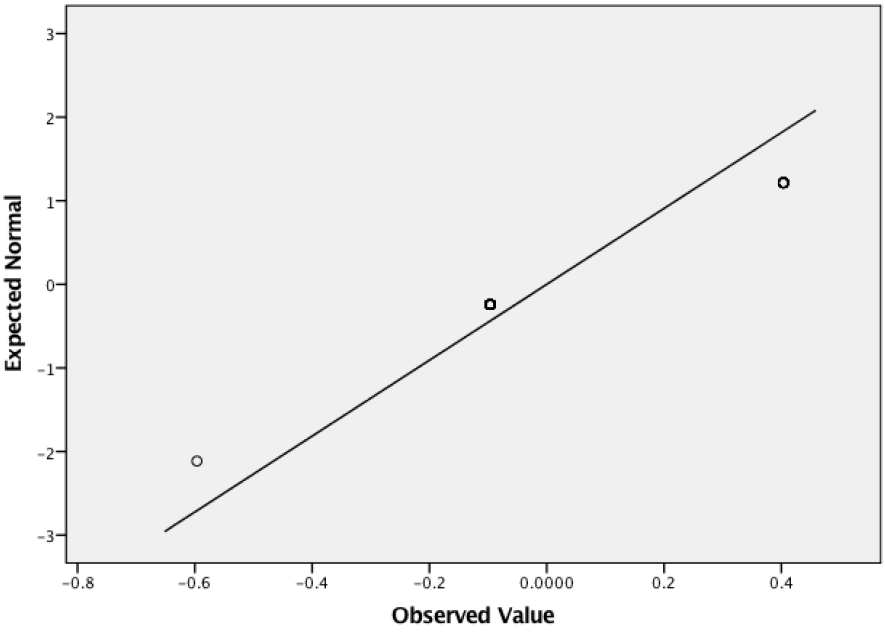
L L

Appendix D: Normal Q-Q Plots

Normal Q-Q Plot for Total Scores



Normal Q-Q Plot for Abstract scores



Normal Q-Q Plot for Deontic Scores