THE ROLE OF COMPLEXITY WITHIN INTELLIGENT DECISION AIDS ON USER RELIANCE: AN EXTENSION OF THE THEORY OF TECHNOLOGY DOMINANCE

by

Bradley Edward Williams

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Business Administration

Charlotte

2020

Approved By

____
Dr. Reginald Silver

____
Dr. Justin Webb

____
Dr. George Banks

____
Dr. Dolly King

ABSTRACT

BRADLEY EDWARD WILLIAMS. The role of complexity within intelligent decision aids on user reliance: an extension of the Theory of Technology Dominance. (Under the direction of Dr. Reginald Silver)

A decision aid only works if used within the decision-making process. As computational power and data storage become more accessible across industries, reliance on intelligent decision aids, those with embedded decision-making technologies, will become more necessary to address the increase in volume, velocity, and variety of data that is now available for use to make decisions. We analyzed over 4900 transactions provided by an organization utilizing an intelligent decision aid as part of their business processes. Using multilevel regression analysis, this dissertation evaluates whether or not differences in the complexity of embedded agents used within intelligent decision aids influence Decision Aid Reliance and whether Decision Aid Complexity moderates the relationships proposed in the Theory of Technology Dominance; Task Experience, Cognitive Fit or User Familiarity. We found that the complexity of a decision aid's embedded agent is negatively associated with Decision Aid Reliance and negatively moderates the relationship between Decision Aid Complexity and Task Experience. Our results provide additional empirical evidence to support and extend Arnold & Sutton's (1998) Theory of Technology Dominance.

DEDICATION

This paper is dedicated to Kelly, Colin, Aidan, Conor, Brennan & Robert Williams, and Kathie & James Spencer. Without your love and support, completion of this project would not have been accomplished.

ACKNOWLEDGEMENTS

I would like to thank my dissertation committee for their guidance, mentorship, encouragement, and, most of all, their patience. Dr. Reginald Silver, who served as both my dissertation chair and the Program Director, has been instrumental in creating the most rewarding journey I could have hoped for. Having the intellectual strength and support of Dr. Justin Webb, Dr. George Banks & Dr. Dolly King on this project has provided me with an outstanding foundation for further research. I would also like to express my gratitude to Dr. Ivan A. Marcotte, who supported this endeavor from its inception without whom this dissertation could not have been possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

A decision aid is a tool that houses the knowledge and decision-making process of an "expert." These expert systems are deployed by organizations to get higher quality results more frequently by systemically following or replicating the "expert's" decision-making process. These tools can be as simple as an expert in Real Estate sales providing a seller with a checklist of things to do before marketing a property, to advanced Intelligent Decision Aids (IDA) which are delivered from a technology platform, utilizing an embedded agent to execute a series of simulations in "friend or foe" identification for military aviation radar control analysts (Wang, Jamieson, & Hollands, 2009). The identifying feature, regardless of the mechanism, is that the final decision is made by the user: *to rely or not to rely*.

Businesses across many disciplines are currently working to establish Artificial Intelligence (AI) and Machine Learning (ML) solutions to automate many decision-making tasks (Chui, 2018). According to the Journal of the American Medical Association, the medical profession is anticipating that surgery, an activity including very high decision complexity and variation, could be done robotically through a machine using an autonomous decision support mechanism (Fogel & Kvedar, 2017). While we are not yet at the point where decision-making in organizations has been taken away from humans, the velocity, volume, and variety of information available (the hallmarks of Big Data) are already having an impact. Decision-makers today must absorb, synthesize, and react to events using more data points at a pace that approaches the rational bound of human cognition (Simon, 1957). Because of these factors, we can logically expect that Decision Aids will begin to incorporate more features of AI and ML in their design. How these tools are designed and implemented could influence the actions of a user, intentionally or not.

1.1 Research Objective

The objective of this research is to understand further why users of Decision Aids choose to rely on or reject its recommendations. Research on Decision Aids has waned significantly since the late 1990s and early 2000s but has seen a recent resurgence. An analysis by Sutton et al. (2014) illustrates this bimodal distribution (Figure 1) of research interest and finds that the field's nomenclature has evolved to include the advances in technology. (Sutton, Holt, & Arnold, 2016).

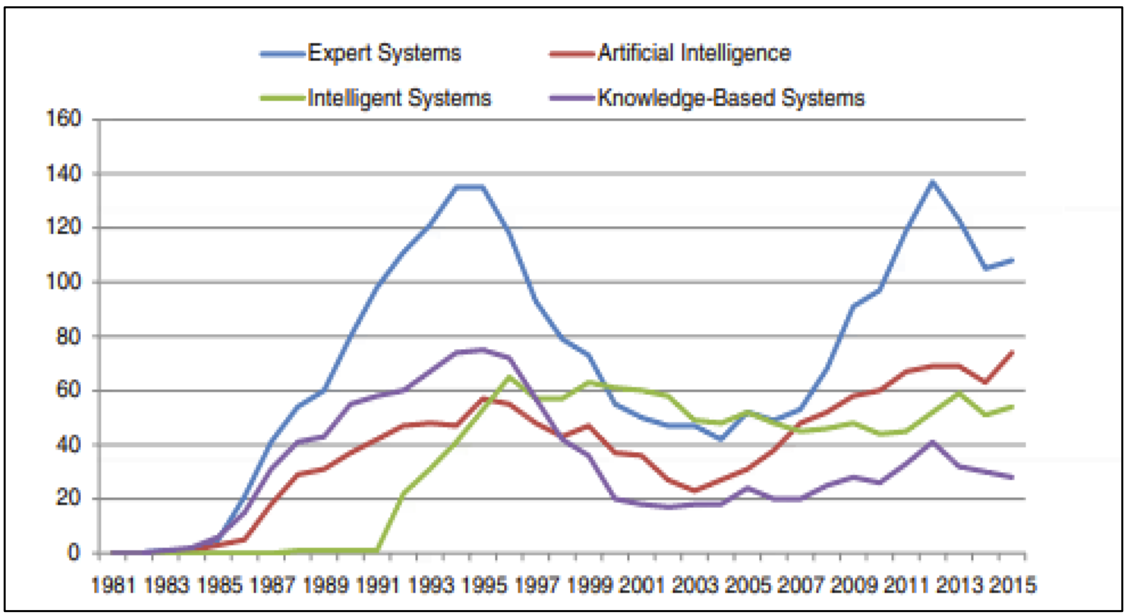

Figure 1: Research Trends

This resurgence is due in part to the democratization of computing power and lower cost of storage, which causes AI and ML technology to be more economically feasible than they have previously been. We believe that where this research left off with knowledge-based and expert systems literature can now be extended due to these recent macro-level technological advances and that we should investigate whether earlier findings remain currently relevant or provide boundary conditions not previously considered. Our research motivation is to expand what is known about the relationship between the technical attributes of Decision Aids and their impacts on the user.

To explain the phenomenon of Decision Aid Reliance, Arnold & Sutton (1998) proposed the Theory of Technology Dominance as a conceptual framework to evaluate the conditions that influence why users accept or reject the recommendation of these expert systems. Arnold & Sutton (1998) suggest that Task Experience is the primary influencing component of Reliance, which is then affected by equivalent levels of Decision Aid Familiarity, Cognitive Fit, and Task Complexity (Arnold, Collier, Leech, & Sutton, 2004; Arnold & Sutton, 1998). The proposition of TTD suggests that lower experienced users would rely upon a Decision Aid more than an experienced user unless the high level of Task Experience was also accompanied by high levels of Decision Aid Familiarity, Cognitive Fit, and Task Complexity (Arnold & Sutton, 1998; Hampton, 2005).

The current literature on decision aid reliance behavior has found that feedback and guidance provided by an Intelligent Decision Aid influence whether a user will rely on or reject the recommendation of the tool. These feedback and guidance mechanisms designed into a Decision Aid still assume that the user has the cognitive capacity to process all of the information necessary to make a choice to rely upon or reject the recommendation. With the volume of data available to decision-makers increasing, powerful computing platforms becoming more readily accessible and the cost of storage becoming more affordable, feedback or guidance designed into an Intelligent Decision Aid could easily become insufficient in the future to serve their purpose and result in more of the decision-making processes being done away from user interaction. We hypothesized that where these guidance or interaction mechanisms are not apparent to the user, there can still be impacts in the design of the Decision Aid, which may cause a user to trend towards or away from the IDA's recommendation.

While there can be many potential areas of differentiation among decision aid design elements, we chose to differentiate on the complexity of the IDA's embedded agent as it is a common decision-making mechanism within today's Artificial Intelligence and Machine Learning algorithm solutions which is hidden from the user.

Given the advances in computational power and affordability of data storage, we expect that Decision Aids will become both more complex in their design and that more of the decision-making process will be done through AI and ML methods that remain hidden from the end-user. Expanding on TTD, we will evaluate an additional construct, Decision Aid Complexity, as an influencing construct of Decision Aid Reliance. This dissertation will analyze the direct impact of Decision Aid Complexity on Decision Aid Reliance and as a moderating influence on Task Experience, Decision Aid Familiarity, and Cognitive Fit.

1.2 Dissertation Organization

This dissertation is organized into seven chapters, including this introduction (Chapter 1). We then present the extant literature on the theories underpinning our research, develop the hypothesis we intend to evaluate and present those through illustrating a conceptual framework (Chapter 2). We then discuss our research methods to include the data collection and sample design, the analysis approach, and construct definitions (Chapter 3). We then present the results of our analysis (Chapter 4), discuss the results and their meanings as it relates to accepting or rejecting our research hypothesis (Chapter 5). We acknowledge and discuss limitations we identified in our research design and results as well as considerations for future areas of research (Chapter 6). We conclude with an overall summarization of the project and its contribution to the field of decision aid research (Chapter 7).

CHAPTER 2: LITERATURE REVIEW & HYPOTHESIS DEVELOPMENT

2.1 Chapter Overview

In this section, we discuss the literature relevant to our research. We begin by reviewing the extant literature underlying our research, the Theories of Reasoned Action & Planned Behavior, the Technology Acceptance Models, and the Theory of Technology Dominance and how Decision Aid Reliance develops from those theories. We follow the theory discussion with the development of our research hypotheses. Finally, we visually present our hypothesis within a conceptual framework.

2.2 Theoretical Foundation of Technology Dominance

The literature review was conducted through searching Academic Source Complete and the ABI/INFORM database keyword searches to include: “decision aid(s),” “intelligent decision aid(s),” “decision aid reliance,” “decision aid complexity,” “complexity” and “TTD.” If not otherwise identified through the initial search, we also selected peer-reviewed articles available from Google Scholar, which cite the Arnold & Sutton (1998) article *“The theory of technology dominance: Understanding the impact of intelligent decision aids on decision maker’s judgments.”*. The results of these searches were filtered to include only scholarly research while specifically excluding those related to patient/disease decision aids used in clinical medical research.

While there is a general consensus in the literature on the definition of an intelligent decision aid, any computerized tool that aids in human decision making, there are two divergent approaches in how the decision aid literature views the underlying decision-making processes the aid is intended to support. Klein (2008) proposed the Naturalistic Decision-Making process, which proposes that individuals are not making decisions by evaluating alternatives but by "using prior experience to rapidly analyze situations" (Klein, 2008, p. 457). Naturalistic decision-making relies upon decision-makers having an understanding of their environment and relying on heuristics and experience to influence their decisions (Morrison, Kelly, Moore, & Hutchins, 1998). Morrison et al. (1998) developed and analyzed a decision support aid that incorporated "naturalistic" decision-making into its design. This decision-making approach and its use within decision aids are widely cited in literature relating to military or similar command and control decision-making scenarios.

The other stream of literature on decision-making processes underlying decision aid research is from the expert system and artificial intelligence domains. This view takes a position where the decision aid is intended to replicate expert decision processes through technological support. The primary distinction between the two decision-making theories when applied to a decision aid, as Morrison (1998) states is that the naturalistic decision aid "support rather than (automate) decision making and leave as much decision making with the human decision-makers as possible" (Morrison et al., 1998, p. 4).

The Theory of Technology Dominance is a mechanistic extension of the Technology Acceptance Model (TAM) attempting to further explain under which conditions users of decision aid technology choose to incorporate the tool into their decision-making process (Arnold & Sutton, 1998). In the paragraphs below, we build from the foundation of TTD, which begins with the seminal Ajzen & Fishbein (1975) Theory of Reasoned Action (TRA). We then discuss Ajzen's (1980) extension of TRA:

The Theory of Planned Behavior (TPB). This research stream is narrowed with TPB/TRA being applied to user acceptance of technology described by Davis's (1989) Technology Acceptance Model (TAM) and its subsequent extension establishing TAM 2 (Venkatesh & Davis, 2009) and Venkatesh's (2000) Unified Theory of Technology Acceptance (UTAT). We conclude this component of the literature review with Arnold & Sutton (1998) Theory of Technology Dominance. We then review the literature of decision aid reliance in general terms and how design features of decision aids specifically influence reliance.

2.2.1 The Theory of Reasoned Action

Research on the understanding of why individuals choose to undertake specific actions stems heavily from Fishbein & Ajzen's (1975) Theory of Reasoned Action which was adapted from their book "Belief, attitude, intention and behavior: An introduction to theory and research" (Fishbein & Ajzen, 1977). TRA broke from the traditional utility theory of decision-making to include psychological factors. They suggest that the most important predictor of an individual's behavior is through their intentions to undertake an action (Fishbein & Ajzen, 1977).

Behavioral Intentions, as conceptualized by Fishbein & Ajzen incorporate a wide variety of factors contributing to an individual's motivation (Ajzen, 1991). As illustrated in Figure 2, behavioral actions, such as the choice to rely upon or reject a decision aid recommendation, is influenced by the individual's attitude towards the behavior and "subjective norms" which in aggregate reflects the motivations and an individual's own disposition to influence actions.

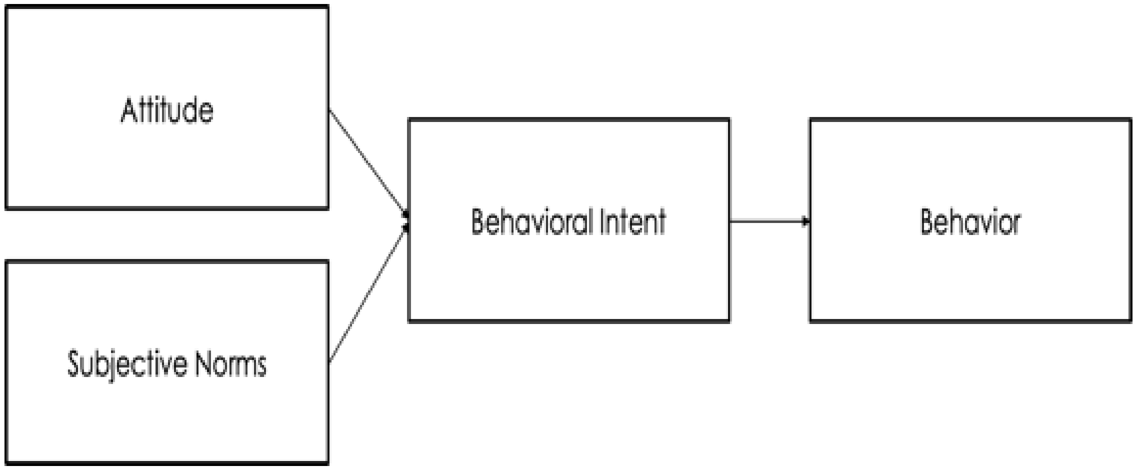

Figure 2: Theory of Reasoned Action (Madden, Ellen & Ajzen, 1992)

There are two major critiques of TRA; the first is that it does not effectively incorporate past behavior and secondly, that it incorrectly assumes individuals have complete "volitional control" or complete control over their decision-making process leading to actions (Ajzen, 1991; Rhodes, Courneya, Jones, & Exercise, 2004)

Lee & See (2004) explained the phenomenon of trust and reliance on automation using TRA by conceptualizing "trust" as an attitude as opposed to a belief or intention forming the behavioral action of reliance (Lee & See, 2004). They explain that viewing trust as a reliance behavior instead of an intention may help to ensure that it does not mask or become mixed with the effects of other factors (French, Duenser, & Heathcote, 2018)

2.2.2 Theory of Planned Behavior

The volitional control assumption of TRA asserts that individuals have complete control over their behaviors (Fishbein & Ajzen, 1977). Soon after the publication of

“Belief, attitude, intention, and behavior: An introduction to theory and research,” Ajzen (1985) addressed the limitation imposed by assuming individuals possessed full volitional control through an extension o TRA in his book *“From intentions to actions: a theory of planned behavior” (Ajzen, 1985)*.

In proposing the Theory of Planned Behavior (TPB), Ajzen included the construct of “perceived control” into the model (Ajzen, 1985). Unlike TRA, which assumed full volitional control, TPB proposed that an individual’s behaviors are also impacted by the individual’s ability to control (or perception of) outcomes (Ajzen, 1991). This change had the effect of acknowledging that there were forces that did place limitations on volitional control. Higher perceived individual control over outcomes translates to higher degrees of behavioral intent (Madden, Ellen, & Ajzen, 1992). Combined, the antecedent constructs of TRA & TPB (illustrated in Figure 3) provide the basis for a significant amount of research into human decision-making.

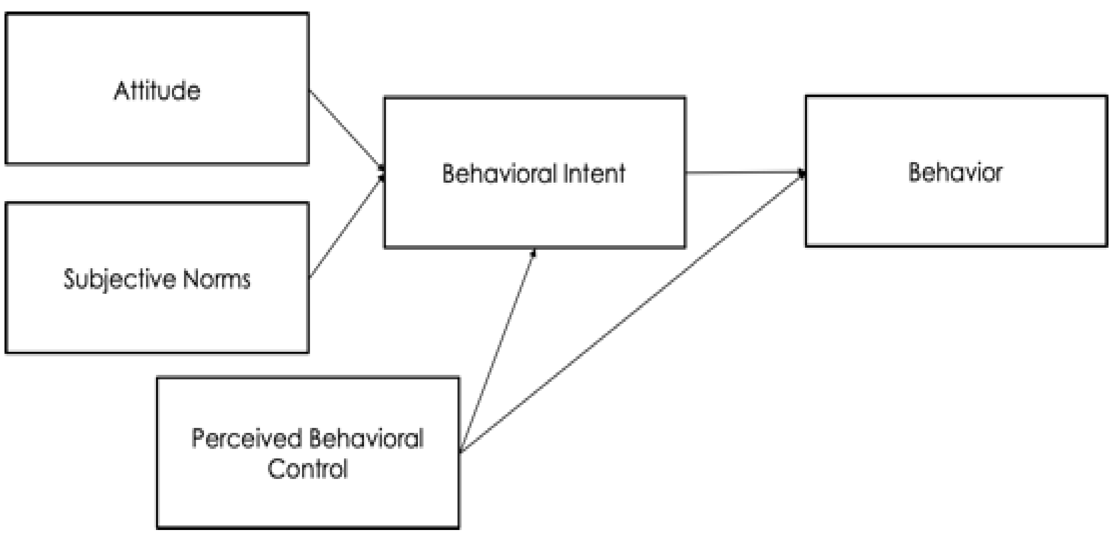

Figure 3: Theory of Planned Behavior (Madden et al., 1992)

Despite the breadth of its application, in theory, empirical support for the theory has been limited. Many studies have failed to support its propositions to the point that some researchers have questions about its ongoing usefulness in decision-making research (Sniehotta, Presseau, & Araújo-Soares, 2014).

2.2.3 Technology Acceptance Model(s)

Extending TRA and TPB into information systems research, Davis (1989) proposed the Technology Acceptance Model (TAM) (Davis, 1989). The intent of TAM was to:

"provide, and explanation of the determinants of computer acceptance that in general, is capable of explaining user behavior accords a broad range of end-user computing technologies and user populations" (Davis et al., 1989 p. 985).

Within the proposition of TAM, Davis (1989) suggests that for a user to accept technology, they must perceive it to be both easy to use and useful in the task the technology intends to address. (Figure 4). Contributing to the generalizability of TAM, numerous studies across technologies, systems, and contexts have similarly found support for the propositions of TAM (Adams, 1992; Davis, 1993; Subramanian, 1994). TAM has proven to be a more effective predictor of technology acceptance than relying upon the constructs within TPB and TRA (Davis, 1989; Hubona & Cheney, 1994; Igibaria, 1997).

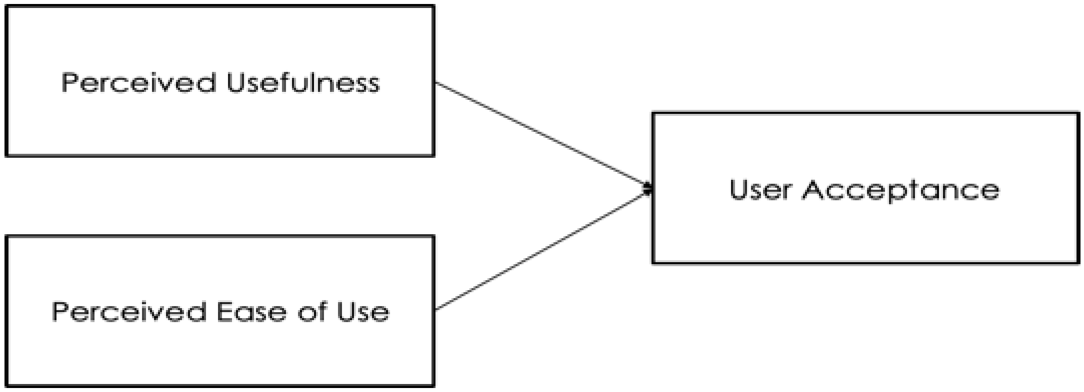

Figure 4: Technology Acceptance Model (Davis, 1989)

Venkatesh & Davis (2000) proposed an extension to TAM, known as TAM2, which considers additional Social and Cognitive factors within the decision on whether or not to accept technology (Venkatesh & Davis, 2009). TAM2 proposes that the social influencers of subjective norms, voluntariness & image, and the cognitive instrumental factors of job relevance, output quality, and result demonstrability, play a role in user acceptance. Subjective norms, as explained by Venkatesh & Davis (2000), impact user acceptance through the individual's perceptions of whether the acceptance is mandatory or if the individual feels that their decision would be looked at unfavorably by others they deem essential (Vekatesh & Davis, 2000).

Venkatesh (2003) later proposed the "unified theory of technology acceptance" or UTAUT. UTAUT synthesizes several technology acceptance theories into four constructs: performance expectancy, effort expectancy, social influence, and facilitating conditions (Venkatesh, 2003).

2.2.4 The Theory of Technology Dominance

In this section, we discuss the Theory of Technology Dominance specifically and augment that with additional literature that describes influencing factors of Decision Aid Reliance. Reliance is a step beyond user acceptance of the technology. TTD requires that acceptance be made consistent with TAM and that the user has chosen to include the Decision Aid within their decision-making process (Arnold et al., 2004). In theoretical propositions, Arnold & Sutton suggested that Task Experience is the primary influencing component of Reliance, which is then affected by the complementary levels of Decision Aid Familiarity, Cognitive Fit, and Task Complexity (Arnold et al., 2004; Arnold & Sutton, 1998). The proposition of TTD suggests that lower experienced users would rely upon a Decision Aid more than an experienced user unless the high level of experience was also accompanied by high levels of Decision Aid familiarity, cognitive fit, and task complexity (Arnold & Sutton, 1998; Hampton, 2005). Effectively they propose that a more junior user would rely blindly upon a decision aid, while an experienced user would rely upon it if it were more efficient to do so. Arnold & Sutton (1998) explanation within TTD for this occurrence is that when task experience is low, there would be no benefit to the user to reject the recommendation of the aid. This is in contrast with an experienced user who would benefit from reliance on the decision aid only when using the decision aid lowers the cognitive cost of integrating it into their decision-making process. They further propose that the lowering of the cognitive cost is a result of the experienced decision-maker possessing corresponding high-levels of Cognitive Fit and Familiarity with the decision aid (Arnold et al., 2004; Arnold & Sutton, 1998).

As an example of this phenomenon, a widely used decision support aid utilizing an embedded agent is the spell-checking and grammar tools integrated into popular word processing and email software programs. As a user composes a document within an email or word processing system, the embedded agent evaluates the content based on a series of rules. In this case, the agent looks for words that are not found in the dictionary and evaluates sentences against its grammar rules for the language it is considering. The agent then returns feedback to the user that the spelling of a word is inaccurate or that the use of the language violates some grammatic rule (i.e., improper syntax, subject-verb agreement, possession versus plurality, etc.). The user may then respond by accepting the aid's recommendation or ignore it. Galletta et al. (2003) specifically evaluated this decision aid and found that users experienced in language proficiency performed better than their more novice counterparts either while using the decision aid and when not using the aid, suggesting as TTD proposes that using the aid would benefit the experienced user only if they agreed with its recommendation through high cognitive fit and possibly familiarity with how it works (Galletta, Everard, Durcikova, & Jones, 2003)

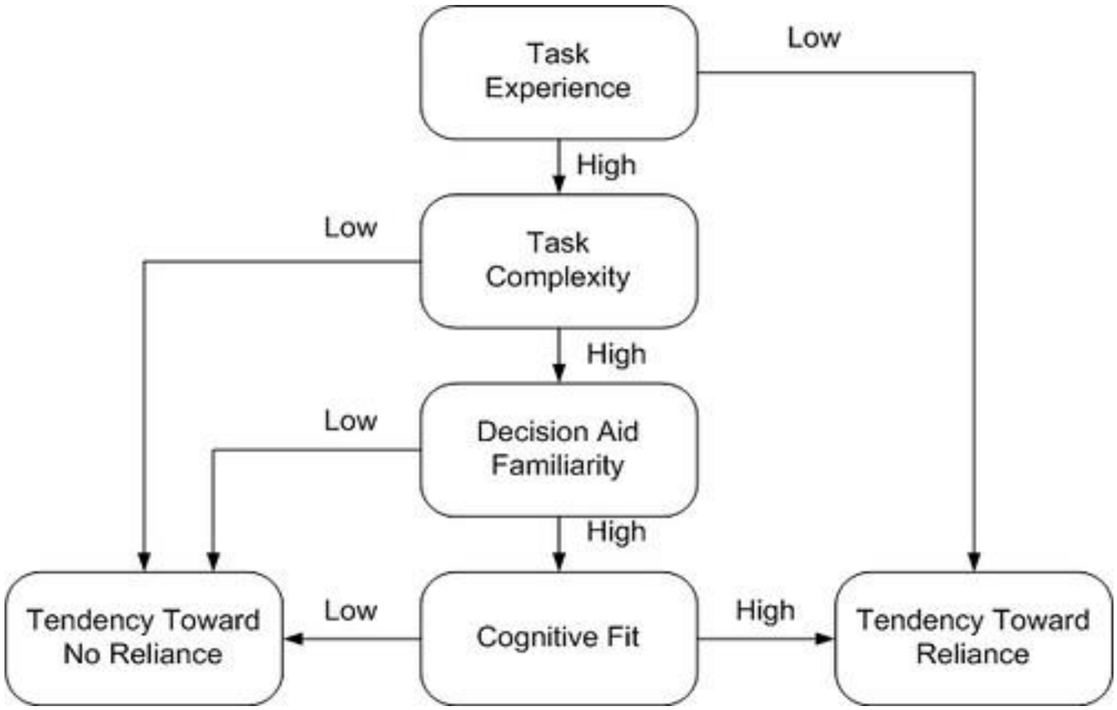

Figure 5: Theory of Technology Dominance (Arnold & Sutton, 1998)

Table 1: Select Research Foundations of Theory of Technology Dominance

Authors	Study Type	Findings
Fishbein & Ajzen, 1977	Theory	Theory of Reasoned Action (TRA) proposes that Behavioral Actions are an outcome of an individual's motivation (behavioral intent) and attitudes.
Ajzen, 1985	Theory	The theory of Planned Behavior addresses the TRA assumption of full volitional control and acknowledges control has limited, and perception of control influences behavioral actions.
Davis, 1989	Theory	Technology Acceptance Model proposes that user adopt technology based on their perceived ease of use and perceived usefulness of the technology
Klein, 1998 & Klein 2008	Theory	Naturalistic Decision-Making: individuals rely upon situational experience and heuristics to make a decision. Important to research in decision-making user stress.

2.3 Literature on Decision Aid Design Influencing Reliance Behavior

Because the substantial portion of research on Decision Aid reliance has occurred within the Accounting and Auditing domains, research on the impact of design in general and specifically those including AI or ML components is limited. Several researchers have expressed this lack of research evaluating these more advanced decision support mechanisms in this area is a limitation that needs to be addressed (Baldwin, Brown, & Trinkle, 2006; O'Leary, 2003). Our research addresses this gap by evaluating decision support aids which rely upon a supervised machine learning algorithm to perform the decision-making tasks embedded inside of a decision aid. This specifically addresses the prior call to evaluate decision aids relying upon these more advanced technologies. Supervised machine learning algorithms utilize prior data to learn as a mechanism to predict or classify potential outcomes. In our case, we use an embedded agent that contains a prediction model used to determine the likelihood of an event occurring in the future.

What is known is that decision aid features do influence user behavior by affecting the users own decision-making strategies and that enhancing the agreement between the user and decision aid improve reliance behaviors (Ashton, 1990; Boatsman, Moeckel, Pei, & Processes, 1997; Eining, Jones, Loebbecke, & Theory, 1997). Whitecotton et al. (1998) observed that when users of decision aids were involved in the decision-making process, they were more likely to rely upon the decision aid than those who were not involved (Whitecotton, Butler, & Houston, 1998). To that end, much research has occurred on the evaluation of various design aspects of decision aids that enable users to be part of the process by either the way in which they interact with the decision aid mechanically or through how the decision aid provides feedback to the user. Decision Aid users generally place more confidence in their own capabilities and tend to underestimate the capabilities of the decision support tools (Eining et al., 1997). Ashton (1990) found that decision aids which possess high "face-validity" were relied upon by users more than those with lower "face validity (Ashton, 1990). As a result, users who are not adequately made aware of the tool's capabilities tend away from reliance. Face validity is a subjective understanding by the user that the Decision Aid will support the decision-making process as the user expects. Decision Aids designed to help support establishing this understanding of capability may increase their "face validity," or the presumption that the Decision Aid will perfume as it is intended and improve reliance.

In a controlled experiment using a Decision Aid specifically designed for testing different "structural restrictions," Seow (2011) found that limiting the choices available to a user is positively associated with user reliance (Seow, 2011). By prepopulating or limiting decision choices, users make determinations as to how the decision aid is interpreting inputs to shape its recommendation. Mălăescu & Sutton (2015) examined how structural restrictiveness influences the impact on the cognitive load of a user and their intention to reuse a Decision Aid (Mălăescu & Sutton, 2015). They found that while highly restrictive decision aids reduce cognitive load for all users to some extent, they reduce reuse intentions for experienced users while increasing reuse intentions for lesser experienced users (Mălăescu & Sutton, 2015). Though users appear to be positively disposed to relying on decision aids when they are engaged in the decision-making process, this does not extend to having prior knowledge of the performance of the decision aid (Kaplan, Reneau, & Whitecotton, 2001). Kamis, Koufaris & Stern (2008) evaluated the interface design of Decision Support Systems in relation to the propositions within the Technology Acceptance Model. Their experiment found that attribute-based decision aid interfaces, those which enabled the user to see all of an outcomes attributes as opposed to requiring a user to make their own inferences, positively moderated a user's perceived usefulness of the technology (Kamis, Koufaris, & Stern, 2008).

Design features that build trust and that are persuasive have an influence on reliance behavior by establishing an agreement between the user and the decision aid (D. Brown & D. R. Jones, 1998; Muir, 1987; Parkes, 2009). It has been found that users of decision aids will select design features that they feel most competent in using even to the exclusion of those which would actually be more beneficial (Wheeler & Jones, 2003). To that end, research has found that features that provide guidance or include feedback mechanisms to the user positively influence reliance by being persuasive and building trust. When given the ability to have insight into the decision-making process by providing an explanation of the recommendation, even experienced users have been found to increase their levels of reliance on decision aids used in exceptionally complex decisions (Jensen, Lowry, Burgoon, & Nunamaker, 2010).

In a study of user interactions with TurboTax, a widely used commercial tax return preparation software, researchers found that providing guidance on audit flags and in-time refund or additional tax calculation notices positively influences user reliance (Masselli, Ricketts, Arnold, & Sutton, 2002)Additionally, Mackay, Barr & Kletke's (1992) evaluation of decision-making processes utilizing a Decision Aid suggests that the design and implementation of the Decision Aid could have interaction effects with other influencing constructs such as task complexity (Mackay, Barr, & Kletke, 1992).

What all of these features have in common is that the user must be engaged in the decision-making process for them to have an influence on reliance. What appears to be underexplored in this stream of literature is in reference to the design elements that do not use guidance or feedback to influence reliance. This gap in our understanding forms the basis of how we intend to develop our research hypothesis

Table 2: Select Literature on Decision Aid Design & Reliance Behavior

Authors	Study Type	Findings
Ashton (1990)	Quantitative (n=60)	Decision Aid with high "face validity" was relied upon more than DA with lower levels of "face validity."
Brown & Jones (1998)	Theory	Decision Aid Features, Decision-Maker Characteristics, Task Factors, and strategy selection factors all influence decision aid reliance behaviors
Kaplan & Reneau (1995)	Quantitative (n=135)	Prior knowledge of a decision aid's predictive ability decreased reliance
Kaplan, Reneau, & Whitecotton (2001)	Quantitative (n=91)	1) Decision-makers were more likely to rely on a decision aid when its predictive validity was not disclosed. 2) Decision-makers with an external "locus of control" were more likely to rely upon a decision aid recommendation. locus of control
Mălăescu & Sutton (2015)	Quantitative (n=109)	Novice users found a highly restrictive system substantially reduces cognitive load, increases the usefulness of the decision aid, and increases reuse intentions.
Masselli, Ricketts, Arnold, & Sutton (2002)	Quantitative (n=120)	Guidance and feedback mechanisms within decision aid embedded agents have the potential to change the reporting behavior of users.
Muir (1987)	Theory	In addressing how to design decision aids to increase user trust, the paper develops propositions that extend models of trust between humans to apply to trust between humans and machines.
Parkes (2009)	Quantitative (n=70)	Both suggestive and informative guidance is positively associated with reliance through building persuasiveness. Decision-makers provided with suggestive guidance reported higher levels of reliance than those with informative guidance.
Seow (2011)	Quantitative (n=94)	More structurally restrictive decision aids increase the user's decision-making bias towards prompted items.
Wheeler & Jones (2003)	Quantitative (n=136)	Users choose between decision aid features that they are most comfortable with base don their perceived capabilities.
Whitecotton, Butler, & Houston (1998)	Quantitative (n=112)	Explores the relationship between information choice and decision aid reliance. Users who were allowed to choose information would rely more on the resulting decision aid than participants who were provided with information, even if the information provided was better than what the user chose.

2.3 Hypothesis Development

2.3.1 Hypothesis 1 - Decision Aid Complexity

Figure 5 illustrates the conceptual framework that we examine within our research presented here. In exploring the literature gap with regards to decision aid elements that do not rely upon user engagement, we evaluate differences in a decision aid characteristic that is invisible to the end-user, embedded agents within intelligent decision aids. Our proposed theoretical model suggests that there is a relationship between the complexity of embedded agents within the design of a Decision Aid and the tendency for end-users to either rely upon or reject the recommendation provided by the Decision Aid. Furthermore, we expect the complexity of the Decision Aid to moderate previously validated relationships between the end user's levels of Task Experience, Familiarity, and Cognitive Fit. Though the construct of IDA Familiarity has not been empirically supported in previous research, we will examine its relationship within our experiment's context and any moderating effect Decision Aid Complexity may have upon its relationship with Decision Aid Reliance.

TTD and others have suggested that complexity does play a role in Decision Aid user reliance (Arnold & Sutton, 1998; Parkes, 2017). How the complexity of intelligent Decision Aids influences reliance behavior is a smaller subset of the overall reliance literature. Within these studies, complexity, often specifically referenced in the literature, including TTD, referred to as "task complexity," has been found to influence reliance behaviors (Arnold & Sutton, 1998; Parkes, 2017). In each case, the complexity construct is evaluated as a property of the d ecision being made, not how the decision is made.

As previously described, the Theory of Reasoned Action and Theory of Planned Behavior suggests that there must be motivation or behavioral intent for a user to choose to utilize a Decision Aid in their decision-making process. Assuming the outcomes were expected to be similar, there would be little motivation to use the Decision Aid for very low complexity decisions (Hampton, 2005). Complexity's role is primarily derived as a choice of whether or not the tool can perform the task more efficiently or at a "lower cognitive cost," than if the user were to make the decision unaided (Todd & Benbasat, 1992). As researchers agree that both task complexity and technology design elements influence Decision Aid reliance, we infer that the complexity of the Decision Aid can also influence reliance in a manner consistent with how TTD would expect Task Complexity to influence reliance. We suggest, however, that the premise of TTD, where the relationship between task complexities is proposed to be positive primarily due to the user ability to evaluate cognitive cost (through experience levels) that condition does not manifest itself within a context of a high-complexity Decision Aid use because the complexity of the decision and the complexity of the task is rendered irrelevant. With cognitive costs being equal, we would expect that the Decision Aid is performing those activities out of view of the end-user. Therefore, we hypothesize the opposite outcome:

H_1: Given constant Task Complexity, the probability that a user relies upon a decision aid is lower when the complexity of the decision aid is high.

This hypothesis suggests that there is both a relationship between Decision Aid Complexity and User Reliance, whereby the more complex the Decision Aid is, the more likely that the user will reject the recommendation. A simple example of this is completing a simple 1040 return. An individual needing to complete this task could use a decision aid. However, if the decision aid were designed to perform preparations for tax filings up to and including C-Corporations, our hypothesis suggests that the complexity of the decision aid would cause the user to reject utilizing the decision. Aid in their decision-making process and seek other options.

2.3.2 Hypothesis 2 - Task Experience

Task Experience, or expertise, has been identified as the primary individual characteristic influencing reliance on a Decision Aid (Parkes, 2016; Arnold,2000). TTD explained the concept of task experience as "level of experience a decision-maker has with respect to the completion of a given decision task and the degree to which the decision-maker has formed strategies for completing or solving the task" (Arnold & Sutton, 1998). Jensen explored a complex IDA designed to aid a user in assessing credibility, a task that has significant variation and frequently lacks complete input information (Jensen et al., 2010). In an experiment using students as novices and police officers as experts, the groups used the experimental Decision Aid, which did provide guidance mechanisms. Consistent with TTD, more novice users chose to rely upon the system more frequently than experts (Jensen et al., 2010). The propositions of TTD, and subsequently empirically validated by Hampton (2005), suggest that experience in the decision-making process is important because less experienced decision-makers are more likely to rely blindly on a Decision Aid. As described earlier, IDA designs include feedback and guidance mechanisms to close the gap between novice and expert users (Parkes, 2009). Our research assumes a context in which the decision-making agent is housed within the IDA and is invisible and autonomous to the user. The user provides the inputs and evaluates the output but is an unnecessary component of the decision-making process. The proposition of TDD, which has empirical support, is that Task Experience is negatively associated with Decision Aid Reliance except when it is also accompanied by high levels of DA Familiarity, Cognitive fit, and Task Complexity (Hampton, 2005). Because the experience of the user is unnecessary when the Decision Aid is performing the decision task autonomously, we could expect the complexity of the embedded agent to negatively moderate the TTD proposition of a negative relationship between Task Experience and Decision Aid Reliance reducing the reliance gap seen in prior research between high and low experience levels.

H_2: Given constant Task Complexity, the probability is lower that a user with high levels of task experience will rely on a decision aid recommendation.

Experienced users are more likely to reject a Decision Aid recommendation. We examine the hypothesis that when utilizing a complex Decision Aid, the likelihood that the user will reject is lessened or will tend towards reliance.

2.3.3 Hypothesis 3 - Decision Aid Familiarity

Research into the factors influencing reliance suggests that it is contingent upon the interaction between task complexity and familiarity with the Decision Aid (Mackay et al., 1992). Though the sample size was small (18 individuals), Mackay et al. (1992) performed an experiment with two groups; the first with high familiarity with the tool, Lotus 1-2-3[tm], and little task experience while the second group possessed high task experience and low tool familiarity. Though not a focus of their analysis, participants in Mackay et al. (1992) experiment provided feedback suggesting a link between the complexity of the task and use of the Decision Aid, where insufficient familiarity with the tool restricted the application of expertise (Kletke, Mackay, Barr, & Jones, 2001). Hampton (2005), however, while empirically testing the propositions of TTD, found no significant support for familiarity with the Decision Aid influencing user reliance (Hampton, 2005). The premise behind familiarity with the Decision Aid and its influence on user reliance is based upon the logic that if a user is2 familiar with the tool, they would be free to spend more time evaluating the recommendation. Unlike MacKay et al. (1992), Hampton (2005) utilized a scaled questionnaire which effectively evaluated the "comfortability" of the Decision Aid interface, similar to the "perceived ease of use" concept within the TAM (Davis, 1989; Hampton, 2005).

Hampton (2005) cautions that while the mean differences between the groups relating to IDA Familiarity were not significant, there were observed differences, and they assert that the experiment design could have imposed limitations on effectively evaluating the construct (Hampton, 2005). In the presence of a complex Decision Aid, utilizing technology which operates independently and autonomously of the user, the anecdotal evidence provided by Mackay & Kletke and the propositions of TTD suggest that a user more familiar with a Decision Aid would spend more time evaluating whether to accept or reject the recommendation of the Decision Aid, thus influencing reliance (Mackay et al., 1992). We submit that this effect is increased when the Decision Aid is more complex because the user has fewer interaction points with the Decision Aid. Assuming as prior experiments have, that the benefit of fewer required interactions associated with more complex Decision Aids is more user time freed up to spend on evaluating recommendations. Therefore:

H_3: Given similar Task Complexity, the probability that a user will rely upon a decision aid when the user is more familiar with the IDA is higher when utilizing a more complex decision aid.

This hypothesis is suggesting that though users familiar with Decision Aids already tend to rely on the results, the more complex the Decision Aid is, the more likely it is that the results are relied upon. Practically this is intuitive as users who have been utilizing systems over time become more comfortable with them; reliance becomes routine. Complexity would add to the cost-benefit as complexity drives a higher cognitive load and therefore cost of non-reliance.

2.3.4 Hypothesis 4 - Cognitive Fit

Cognitive Fit Theory ("CFT") has been explained as a special-case of cost-benefit analysis as it suggests that users will alter their decision-making strategy based upon the cognitive effort required to analyze information to arrive at a decision (Vessey, Zhang, & Galletta, 2006). The original propositions of CFT were based upon work that Vessey & Weber (1986) had previously conducted, finding differences in outcomes of similar decision-making tasks when the decision-maker was presented with information in either tables or graphs (Vessey, 1991). These earlier observations established the empirical support for CFT. Consistent with other literature in Decision Aid Reliance, providing a tool or displaying information to a user, which makes their ability to organize and evaluate information, reduces the overall cognitive cost of making the decision (Todd & Benbasat, 1992).

Figure 6, adapted from Vessey (1991), illustrates Cognitive Fit Theory. CFT suggests that problems are evaluated by individuals as both a component of how the problem is presented, referred to as the “Problem Representation,” and the actual method of solving the problem or “Problem Solving Task.” Vessey (1991) explains this phenomenon as a result of humans being “limited information processors” (Vessey, 1991). To that end, mechanisms that support problem-solving methods will reduce the complexity of the task and result in more effective and efficient problem solving (Vessey, 1991). More simply, users will utilize decision-making mechanisms in which enable them to balance (or minimize any imbalance) between the problem at hand with the work needed to make the decision (Vessey, 1991). This match allows users to more effectively and efficiently make decisions due to the match generating a lower cognitive load of decision-making (Bacic & Henry, 2018).

Failure to match the task to the representation of the task leaves users without many cues to establish a mental representation, which enables them to create the balance necessary for them to address the problem (Vessey, 1994). The result is that a user must spend effort generating knowledge to develop their own representation of the task or problem-related concept applied to MIS research is Task-Technology Fit, which is also explained by “the degree to which system characteristics match user tasks” (Goodhue & Thompson, 1995).

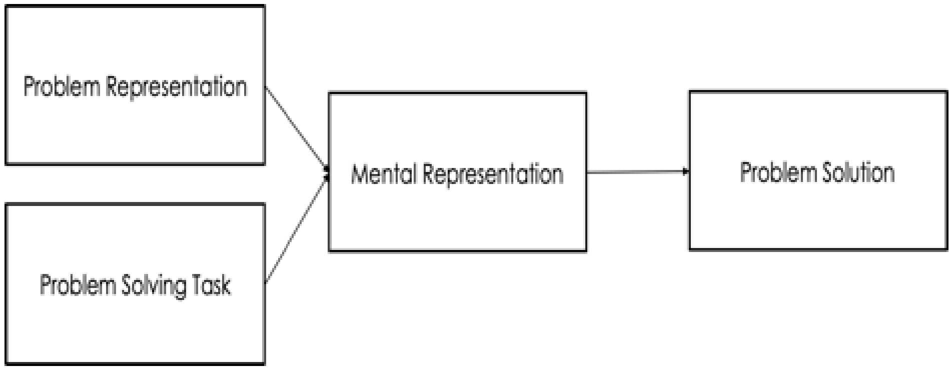

Figure 6: Cognitive Fit Theory (Vessey, 1991)

Cognitive Fit Theory has been applied in many areas beyond the charts & tables research upon which it was first established. Major applications of the theory have been found in geographical information systems, computer programming, accounting, and word processing (Dennis & Carte, 1998; Dunn & Grabski, 2001; Shaft & Vessey, 2006). In each use, CFT has been found that the way in which information is presented to users is related to the performance of their decision-making (Hong, Thong, & Tam, 2004; Vessey, 1991). Hong, Thong & Tam (2004) conducted an experiment using consumer e-commerce websites in which they found that balancing the task of browsing versus searching with the presentation of information in lists and matrices respectively, that the users were more effective in performing their task (Hong et al., 2004). Sinha & Vessey (1992) further extended Cognitive Fit Theory to encompass how technology influences a user of technology, programmers in this instance, in how they learn to balance tasks and problems given both iterative and recursive tasks (Sinha & Vessey, 1992).

What is consistent across the extant literature is that a user's understanding of the problem and the tools used in solving the problem are important to ensuring that cognitive fit is established. However, when applied to technology as we are addressing in our research, there is no consensus definition of fit. Davern (2007) provides propositions towards a unified theory of fit, which includes aspects of Cognitive Fit, as explained by Vessey (1991) and Task-Technology Fit, as explained by Goodhue (1995). As Davern (2007) explains, “fit” in either context both acknowledge that fit is a construct that contains the components of the task, the technology, and individual abilities (Davern, 2007).

Strict adherence to the TTD proposition of Cognitive Fit would lead us to expect that a complex Decision Aid, which performs the decision-making steps on behalf of the user, would negatively moderate the influence of Cognitive Fit on Reliance because the linkage between the Decision Aid and the decision-making process of the user would be severed due to the design of the aid. However, in the absence of totally autonomous decision-making, the user remains a part of the decision-making process and participates in evaluating the inputs and/or the output which may lead a user to infer the decision-making process and make their own determination regarding the mental representation of the problem irrespective of the amount of information the user possesses. Therefore, we expect that Decision Aid Complexity will not alter how the information is presented to the user or the nature of the problem and that it will positively moderate the proposition within TTD of the positive relationship between Decision Aid Reliance and Cognitive Fit.

H_4: Given consistent Task Complexity, the probability that a user will rely upon a decision aid when there is Cognitive Fit is increased when the complexity of the decision aid is high

As described earlier, Cognitive Fit is explained as the balance between the problem being addressed and how that informs the mental representation created by the user. Our hypothesis asserts that complexity is increased as a Decision Aid is designed to address its problem or task more effectively and improve the balance between the task and mental representation. Using our tax-preparation software example again, if a user is expected to use a Decision Aid to perform a complex return such as that for a C-Corporation, the balance between the user's mental representation of the tax preparation task would not establish a cognitive fit between the user and Decision Aid if the Decision Aid represented itself as if it were intended to present guidance as if the problem were an individual return. Therefore, when tasks are complex, we expect the user will have an expectation of a complex Decision Aid.

2.4 Conceptual Model

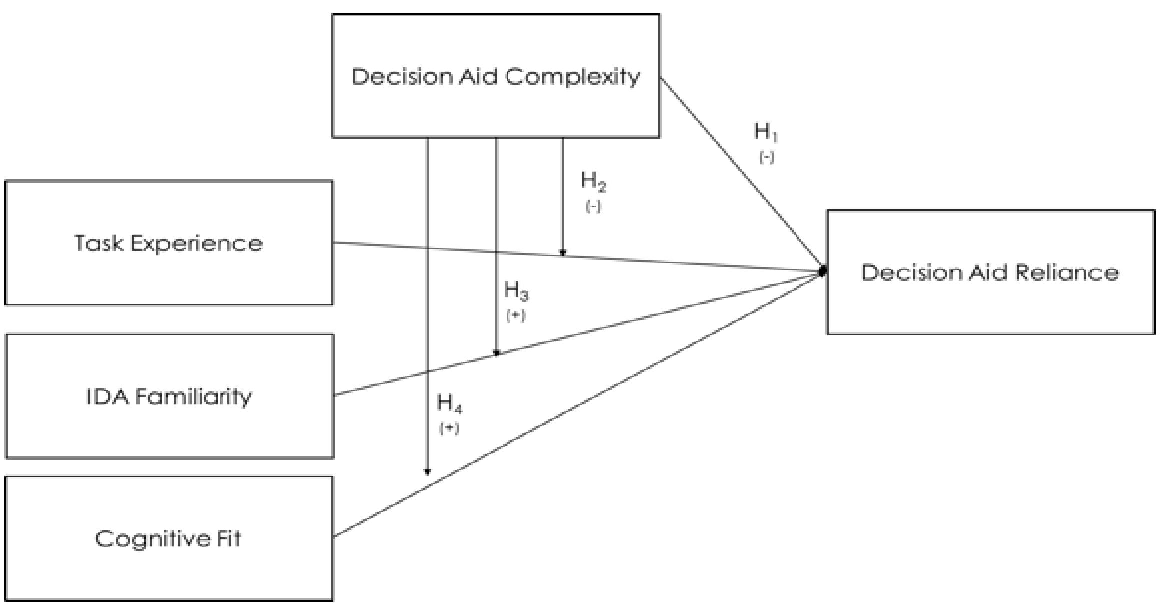

Figure 7: Conceptual Model

To test the hypotheses that we developed above, we present a theoretical model that positions Task Experience, IDA Familiarity, and Cognitive Fit as independent variables and Decision Aid Reliance as the dependent variable (Figure 7: Conceptual Model). The model explores the moderating effect that Decision Aid Complexity has on the direct effect between each independent variable and the dependent variable. The model also explores a direct effect between Decision Aid Complexity and Decision Aid reliance. A summary of these hypotheses is provided below (Table 3).

Table 3: Research Hypothesis Summary

Direct Effect of Decision Aid Complexity on Reliance	
H1	Given constant Task Complexity, the probability that a user relies upon a decision aid is lower when the complexity of the decision aid is high.
Moderating Effect of Complexity on the relationship between Task Experience, IDA Familiarity, Cognitive Fit and Reliance	
H2	Given constant Task Complexity, the probability is lower than a user with high task experience will rely on a decision aid with high complexity.
H3	Given similar Task Complexity, the probability that a user will rely upon a decision aid when the user is highly familiar with the IDA is also high when utilizing a complex decision aid.
H4	Given consistent Task Complexity, the probability that a user will rely upon a decision aid when there is Cognitive Fit is increased when the complexity of the decision aid is high.

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Sample

To test our hypothesis, we collected data from a large, multinational financial services firm ("the organization"). Headquartered in the US, the organization is one of the oldest companies of its type tracing its lineage back over 200 years. The organization has millions of customers and provides financial services to individuals and companies covering assets in excess of 1-trillion USD. As part of their normal business processes, the organization utilizes intelligent decision aids for such functions as fraud monitoring, financial management, and resource planning. We obtained data from a decision aid used by the organization to inform credit risk decisions. As our unit of analysis is the user's acceptance or rejection of a decision aid recommendation at each interaction, analyzing the organization's decision aid utilization within one of their core competencies provides a solid example of Decision Aid Reliance in practice. Researchers utilizing an organization's own data objectively is supported as an appropriate Management Information Systems research method. This is specifically supported within the Knowledge Management domain, where decision support research is positioned (Huang, Kankanhalli, Kyriakou, & Sabherwal, 2018; Kim, Mukhopadhyay, & Kraut, 2016).

During the client selection process for new business or as part of maintaining agreements on existing client's credit exposure, the organization performs evaluations of credit quality. Financial statement details provided to the organization from a current or potential client are "spread" into a database that systemically houses the data. In the making of credit-risk determinations, the spread data are input into the decision aid upon user request. By considering attributes of the transaction (ex. size, industry, structure) and business rules, the intelligent decision aid selects one of many embedded agents (models) within it. These embedded agents calculate a predicted 12-month probability of default (PD) value for the client or potential client. The value is then converted onto a standard 27-point scale meant to ensure consistency in output across the embedded agents. The 27-point scale, from -1 through 9+, is the “Obligor Risk Rating” (ORR). The ORR is returned by the decision aid to the end-user who may choose to accept the recommendation or initiate the process to override (reject).

Data from the decision aid of all actions where a customer ORR were created between 6 June 2019 and 30 August 2019 were made available for our research. With the permission and oversight of the organization, all personally identifiable information of either the client, potential transaction, or of the organization was removed. The total number of ORR scoring actions by the decision aid within the observation period was 129,614 of which 100,619 were determined to be system generated, where a user would not have been in a position to make a reliance choice, or an authorized workaround mechanism was used to manually score the client judgmentally without the support of the aid. With over 1300 unique users in the data provided by the organization, the manual nature of the data collection made collecting details on all users impractical. After confirming the required sample size through power analysis (Table 17), we randomly selected 250 users to collect the measures for Task Experience, IDA Familiarity, and Cognitive Fit. The resulting sample consisted of 4940 observations. This sample had 512 or 10.4% of its observations result in a rejection of the decision aid’s recommendation.

3.2 Measures

In this section, we discuss the dependent, independent, and moderating variables illustrated in our conceptual model. Furthermore, we present how the variables will be measured for conducting our analysis.

3.2.1 Decision Aid Reliance

The dependent variable, decision aid reliance, was calculated as a dichotomous variable by evaluating whether or not the ORR recommended by the decision aid was overridden (value of 0) by the user or remained as presented (value of 1). Prior research on Decision Aid reliance has primarily focused on decision accuracy or other measures of performance (Rose, 2002). Arnold & Sutton (1998 defined reliance not as the accuracy or performance of the decision but on whether or not the decision provided by the intelligent Decision Aid was accepted or rejected by the user (Arnold & Sutton, 1998). Our research goal is to understand factors that influence a user to include the use of a Decision Aid into their decision-making process. While not irrelevant as a potential factor in DA Familiarity, accuracy is not the specific action of interest in our study. Because we are testing propositions within TTD, we will upon the Arnold & Sutton (1998) definition and evaluate reliance as the acceptance or rejection of the recommendation.

3.2.2 Decision Aid Complexity

Complexity has as many definitions in literature across domains as varied as Management, Biological Sciences, Psychology, Operations Management, and

Management Information Systems. Broadly, complexity has been defined as a collection of components that interact in a non-simple, non-linear way (Simon, 1957). Bruce Edmunds, while being critical of complexity as a construct, defined complexity as “That property… which makes it difficult to formulate its overall behavior even when given almost complete information about its atomic components and their inter-relations” (Edmonds, 1995, p. 1). More simply, that even if we understand something down to its most basic components, the more complex it is, the more difficult it is to explain. Flood & Carson defined complexity as “anything we don’t understand” and explained it as a phenomenon of how people interact with any object of interest (Flood & Carson, 2013).

What is generally agreed to is that complexity is a function of the number of the object of interest’s components and how those components interact. Casti (1979) explored complexity in systemic terms through what he defined as “Static Complexity” or the complexity generated inherently by a system that includes interconnected components in a way that makes them difficult to understand (Casti, 1979). The static structure of any system can provide insight into its inherent complexity (Deshmukh, Talavage, & Barash, 1998) But Casti (1979) expands this single explanation singularly insufficient to adequately explain Complexity and decomposes it into four component pieces which together are formative of “Static Complexity”; Hierarchical structure, Connective pattern, Variety of Components and Strength of interactions. The levels of the structure, the difficulty or number of paths available in which they are intertwined, the variation in component types, and their relationships all contribute to static complexity (Casti, 1979). Makui & Aryanezhad (2003) expand upon the understanding of static complexity by mathematically supporting the premise that while static complexity is an inherent risk measure, that the “states’ or conditions that each system component may take, and the probability distribution of what that state is, contribute to inherent risk as well (Makui & Aryanezhad, 2003).

Bozarth et al. (2009) also incorporate the concept of uncertainty in the component interactions which they adopted from Senge (1990) as the number of variables within the component interactions (Bozarth, Warsing, Flynn, & Flynn, 2009; Senge, 1990). MIS research on knowledge-based systems continues this two-dimensional view of complexity by classifying knowledge-based systems, which include intelligent Decision Aids, based upon components which contribute to knowledge complexity and components which contribute to technical complexity (Meyer & Curley, 1991).

Others have defined complexity based upon a system possessing specific attributes. Yates (1975) described complexity as possessing any of five specific attributes where the evidence of anyone could explain complexity. In addition to component property interactions and nonlinearity previously explained, Yates includes the additional properties of “asymmetry and non-holonomic constraints” which constrain particular paths available to the system (Yates, 1978). Though Meyer & Curley (2001) did follow a definition similar to the component-interaction definition, they also incorporated the aspect of measuring the scale of specific properties such as “domain depth,” “input ambiguity,” and “programming sophistication” (Meyer & Curley, 1991). The summation of these attribute “scores” serves as the mechanism to evaluate complexity.

Specifically, with regards to model complexity as with our embedded agents under analysis, model complexity is generally only explained in terms of differences within models. However, similarly to the static definition of complexity within systems, model complexity has been defined by some by measuring the model's constituent pieces. Brooks explains complexity in this context as the "amount of detail" or how many aspects of the phenomenon being modeled are included in the model itself (Brooks & Tobias, 1996). We follow Brooks & Tobias' (1996) method for measuring model complexity as a combination of its size, connectedness, and computational complexity.

Consistent with these prior definitions, we will define the complexity of the Decision Aid in terms of the number of constituent components and interactions. Specifically, we will look to the components of the embedded agent within the IDA and leverage the number of those structural and interactive components to measure complexity. Within each embedded agent, we will aggregate the number of variables used, the number of factors used, and the number of upstream and downstream interactions using exact counts of each from the organization's internal documentation. Consistent with the approach used by Meyer & Curley (2001), we will tabulate the values of the variables, factors, and interactions and establish a "score" of complexity for each IDA embedded agent. Therefore: $Complexity = EA_{Varriables} + EA_{Factors} + EA_{Interactions}$.

Moderating variables influence the direction or strength of the independent variables on the dependent variable (Baron & Kenny, 1986). As explained in our hypothesis development, the goal of our research is to determine if 1) there is a direct-effect between Decision Aid Complexity and Decision Aid Reliance and 2) determine if the additional construct of Decision Aid Complexity moderates the propositions identified in the Theory of Technology Dominance.

3.2.3 Task Experience

Wood & Bandura defined Self-Efficacy as the collection of belief's in one's own intelligence, capabilities and actions to meet given demands (Wood & Bandura, 1989) Further refined, Self-Efficacy is explained as an individual's own perception of their competencies in completing or performing an action (Gist & Mitchell, 1992). Task Experience as it relates to performing actions influence Self-Efficacy as it positively moderates "enactive mastery" through task accomplishment (Appelbaum & Hare, 1996). Therefore, Task Experience is a mechanism by which individuals can increase their perception of competency in completing a task, which provides a possible explanation for why those with differing levels of Task Experience rely upon external decision aid support to differing degrees.

Defining Task Experience has been generally measured as a function of amounts (Quińones, Ford, & Teachout, 1995). Ford identified three components of work experience as it related to Air Force Trainees. These were the number of tasks performed (unique tasks), the number each task was performed (repetition), and the difficulty of the task (Ford, Quiñones, Sego, & Sorra, 1992). Consistent with Ford, Lance, et al. (1992) explained experience as the number of times a task has been repeated (Lance, Hedge, & Alley, 1989). Many studies have concluded that time on the job is an appropriate measure of experience. Though measures of experience through task frequency has been found to have a stronger relationship than time measures (tenure), time measures are still found to be significant and a positive indicator of performance (Quińones et al., 1995).

The premise of TTD is that the decision aid is an extension of the user and augments the decision-making process inherent within the expert as expressed through the reliance or rejection of the Decision Aid's recommendation (Arnold & Sutton, 1998). As a construct in this research domain, experience has been measured in empirical studies in several different ways. Researchers often rely on students as participants in empirical studies. In experiments on Decision Aids where experience or expertise was a variable of interest, researchers have differentiated undergraduate students from graduate students and students overall with professionals to segment levels of expertise (Jensen et al., 2010; Mălăescu & Sutton, 2015). Hampton (2005) operationalized this concept through a pre and post-test as the number of evaluations the participant had been exposed to (Hampton, 2005). Similarly, Mascha & Smedley (2007) measured task experience with an evaluation of the participant's command of prerequisite reading material (Mascha & Smedley, 2007). Other studies have used the number of years of experience of the participant and the level of seniority provided by their title where hierarchies were known (Arnold et al., 2004; O'Leary, 2003).

Consistent with prior research, we will use the level of expertise presumed by the title hierarchy of the organization (with some variation based upon the line-of-business the hierarchy from novice to expert would follow; Analyst, Sr. Analyst, Manager, Sr. Manager & Executive). These titles will be anonymized as a categorization variable on an ordinal (1-5) scale where the novice will be lower and the expert higher on the scale.

3.2.4 IDA Familiarity

Unfamiliarity with a Decision Aid can negatively impact the ability of a user to apply their expertise (Kletke et al., 2001). Familiarity has been examined primarily as an interaction effect sometimes defined as the specific functional capabilities a user possesses in utilizing a specific technology (Mackay et al., 1992). These have been operationalized through questionnaires seeking to assess a user's comfort with a technology (Hampton, 2005; Mackay et al., 1992). We operationalized the IDA Familiarity construct by counting the number of total interactions the user has had with the IDA within the observation timeframe of 6 June 2019 and 30 August 2019, irrespective of which embedded agent the IDA used in evaluating the decision.

3.2.5 Cognitive Fit

As we have discussed earlier, Cognitive Fit is inseparable from the technology as the IDA is consistent with Vessey's (1991) description of a mechanism that balances the complexity of the problem with the means to provide a solution (Vessey, 1991). Measuring Cognitive Fit in research has been done almost exclusively through measuring the accuracy and time needed for a user to act on a presented problem (Tan, Teo, & Benbasat, 2010; Vessey, 1991; Vessey et al., 2006). This is a result of Vessey (1991) defining performance as an outcome of Cognitive Fit as being the measure of decision quality (as measured through accuracy) and time-to-decision speed. (Vessey & Galletta, 1991). Because we have accepted Arnold & Sutton's (1991) definition of Reliance for our dependent variable, the notion of accuracy is not an acceptable measurement of Cognitive Fit for our experiment. As described earlier, CFT within technology contexts is consistent with Goodhue's (1995) theory of Task-Technology Fit (TTF). TTF asserts that for fit to occur, there must be utilization of the technology (Goodhue & Thompson, 1995). Combing both the utilization element of TTF, the time elements in the CFT, and ignoring decision accuracy, we operationalize the variable representing Cognitive Fit as the total number of utilization days (time) the IDA results were interpreted by a user performing the decision-making task relying upon the same embedded agent.

Table 4: Variable Measures

Variable	Type	Measurement
Reliance	Binary	Acceptance or Rejection of the IDA recommendation (Accept = 1, Reject = 0)
Complexity	Continuous	Sum of IDA embedded agent ("model") variables, factors and interactions
Task Experience	Ordinal (Categorical)	Rank order Novice through Executive on 5-point scale (No Title = 1, Officer =2, AVP = 3, VP = 4, DIR/SVP = 5)
Familiarity	Continuous	Sum of total interactions by the user in the dataset. Irrespective of embedded agent the IDA selects
Cognitive Fit	Continuous	Count (in Days) the user has engaged the IDA where the same embedded agent was used.

3.3 Analysis Approach

Given the ratio of Rely to Reject outcomes, some researchers could determine that the class imbalance seen between the number of Reject versus Rely outcomes in our dataset (reject observation is ~10%) presents a rare-events problem into the analysis. There is no consensus as to the definition of "rare event" other than being of a low probability of occurrence. The most widely-cited literature on rare events describes occurrences well below our observed population of rejection at 10%. King & Zeng (2001), who's seminal research on rare events in global conflicts describe events with the occurrence probability well below 1% (Breslow, 1996; King & Zeng, 2001; Lampel, Shamsie, & Shapira, 2009). For the purposes of this study, we have deemed that the rely-to-reject outcomes that we investigated do not meet the criteria of being a rare event. Putting aside this additional complexity of analyzing rare events, we decided that multi-level logistic regression is an appropriate research method to employ in this study.

Logistic regression is the appropriate method when the outcome of interest, the dependent variable, is dichotomous with either categorical and/or continuous predictor variables. One of the requirements of logistic regression modeling is that the observations within the sample are able to be considered independent. Research suggests that nesting of data occurs for two reasons: a hierarchical structure to the data such as (patients-doctors-hospitals or students-teacher-school or other less official structures) or in studies where the data contains observations of the same objects of interest which are measured repeatedly (Aguinis, Gottfredson, & Culpepper, 2013; Peugh, 2010). These repeated measures may be a result of a longitudinal study or through any data structure, which includes multiple observations of the same target. Our analysis contains repeated measures of individual/model interactions with a decision aid resulting in a decision to rely or reject its recommendation and is considered nested. Though our data was collected from one IDA, the individual observations are nested within the use of 27 embedded agents and 250 unique users, suggesting the potential need to perform a multi-level analysis of the regression. It is important to understand relationships among observations as clustering or groups inherent in how the observations occur influence their behavior. The illustration below in Figure 8 explains the nested structure of our data. Our observed relationships are not a clear as many examples of multilevel modeling present, such as students nested in classes nested within schools. However, as Kelly & Judd (1996) observed, dependence was more than merely the formal structure, but that the relationship shares “common feature, come from some common source…or arranged spatially or sequentially in time” (Kenny & Judd, 1996)

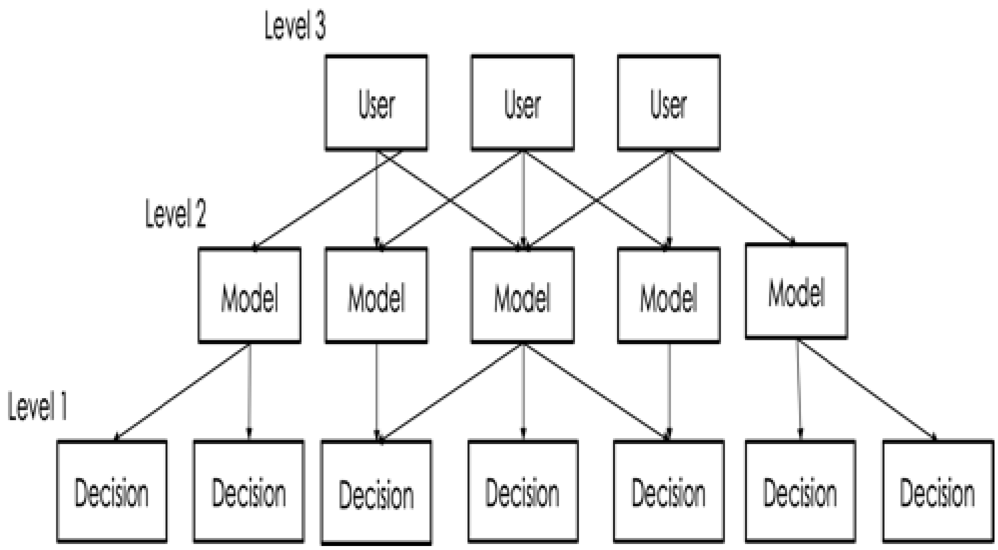

Figure 8: Multilevel Structure

As described above, our data is nested due to the repeated measures of both the User (USERID) and Embedded Agent (MODELID) with regards to the decision to rely upon or reject the recommendation of the Decision Aid (RELIANCE). To determine how our specific data were nested, we followed the explanation of nested data structures as provided by Peugh (2010). In a traditional data structure, each observation would be situated in a single row of data with the repeated measure included as a column. Nested data is different as it is characterized by its ability to be “stacked,” where you have a hierarchy of one to many relationships among variables (Peugh, 2010). Data at a lower level is stacked or clustered within a single level of a higher order. In our case, we have many RELIANCE decision observations which have relationships with both Users and Embedded Agents. With Decision being the repeated measure of analysis (Level 1), we then must evaluate at which level of nesting the User and Embedded Agent nesting occur. By stacking our data by User and Decision, we see that there is one too many "stacked" relationship of RELIANCE ~ USERID. Each Decision occurs within only one User. Therefore, we know that Decisions are nested within Users. By stacking our data by Embedded Agent and Decision, we see that there is a one to many "stacked" relationship between RELIANCE~MODELID. Each Decision aligns with one and only one embedded agent. Therefore, we know that Decisions are similarly nested within Embedded Agents. However, we do not see this one to many nesting relationships between Embedded Agent and User categorical variables (MODELID & USERID). The structure of the data does not change with the nesting order of USERID or MODELID. We know through the calculation of the interclass correlation coefficient (ICC) illustrated in Table 5 that there are fair to substantial impacts of both the USERID and MODELID categories on RELIANCE (Landis & Koch, 1977). Ignoring the contribution to the variation seen within RELIANCE demonstrated by the ICC would not be appropriate as failure to account for nested impacts increases the likelihood that our model would reject a predicted outcome of RELIANCE even when it was the accurate outcome (type 1 error) (Aguinis et al., 2013). Tasca et al. (2009) suggest that when designing experiments, "one should consider the highest level of nesting as defining the unit of analysis" (Tasca, Illing, Joyce, & Ogrodniczuk, 2009, p. 455). As our unit of analysis considers decision aid reliance behavior by individuals, then our highest-level order should tend towards

Users (as categorized by USERID) as opposed to tending towards Embedded Agents (MODELID) or Decisions (RELIANCE). This approach supports placing USERID as the highest level of the nesting structure.

Additionally, research on multilevel modeling recommends that it is appropriate to have a larger number of groupings at the highest level as “sample size at the highest level is the main limiting characteristic of the (models) design” (Snijders, 2005, p. 2). The data consists of 250 unique Users and 27 unique Embedded Agents. By selecting USERID as the highest level of nesting, followed by MODELID, we improve the model design by eliminating more of the risk of Type 1 error than would be with the smaller cluster size of the Embedded Agents at the highest level. Considering both the unit of analysis and sample size impacts of the nesting order, we suggest that the appropriate nesting structure of our data is that Decisions are nested within Embedded Agents which are nested among Users.

The levels at which the predictor variables are aligned is important to MLM modeling. Hecht & Thomas (2015) explain that misaligning covariates in a multilevel structure may lead to the “aggregation” of variable variation above the level at which it naturally occurs or “disaggregation,” which unwinds variable variation and confounds results of the testing (Heck & Thomas, 2015). Following Hecht & Thomas (2015), we rely on the conceptualization of our nested structure (Figure 8) to align the covariates and define the levels of analysis. Variables unique to the decision and which vary by decision (IDA Familiarity and Cognitive Fit) are aligned to Level 1, which is the micro-level of our analysis. Covariates which are unique to the embedded agent/model (Complexity), and which do not vary from observation to observation, are positioned at Level 2.

Finally, predictor variables unique to the user, which do not vary by observation or model are aligned to Level 3. Level 3 was selected as the User as users may interact with several models (Level 2) to support decisions (Level 3). Classification variables for Level 2 (MODELID) and Level 3 (USERID) were included to define these levels within the structure.

We relied upon the method provided by Tasca et al. (2009) to determine multilevel modeling structure appropriateness. As the data is considered nested, that there is evidence of dependence and that it consists of repeated measures, the recommended model for our analysis is a three-level MLM (Tasca et al., 2009, p. 456). Application of the model followed Aguinis et al. (2013) using the steps provided to evaluate cross-level interactions in multilevel models; Step 1) fit the null model to determine the interclass coefficient, Step 2) Random Intercept (Fixed Slope) and Step 3) Random Intercepts & Slopes (Aguinis et al., 2013). Research in multilevel modeling suggests evaluating the Interclass Correlation Coefficient or ICC to ascertain if the hierarchical nature of our variable relationships is contributing to the variation in the model. This value, between 0 and 1, estimates the amount of variation driven from the nested structure (Landis & Koch, 1977).

Consistent with previous research, we fit a “Null” model (results in Table 11) or intercept-only model with our dependent variable, Reliance, the Level 3 (USERID) and Level 2 (MODELID) indicator variables included. Researchers refer to this as the “null” model as it has no predictor variables included (Heck & Thomas, 2015). The results of calculating the Interclass Correlation Coefficient is presented in Table 5. The ICC of the Level 3 grouping contributes 0.277 to the variation in Reliance, while the Level 2 grouping contributes 0.638. ICC values above .277 are considered to have a fair impact, and value of .638 is considered a substantial impact (Landis & Koch, 1977).

Table 5: Interclass Coefficient

Level	ICC	Std. Err.	95% Conf. Interval	
userid	0.277	0.071	0.161	0.434
modelid|userid	0.638	0.039	0.559	0.711

Data were analyzed through a series of multilevel, hierarchical logistic regression model using STATA 16 software. To determine the best fit of MLM to continue our analysis, we fit both a random-intercepts model and a random-intercepts and slopes model then evaluated if there were any improvement in the model's estimation capability by allowing the slopes to vary across the data resulted in a significant improvement in the model. A random intercept model includes the covariates and interactions while including the Level 2 and Level 3 classifiers to enable the intercepts to vary. The random intercepts and slopes take the additional step to tie each covariate to the level in which it occurs within the nested structure. The results of the random intercept and random intercepts & slopes models are provided in the appendix (Table 12 & Table 13). Along with the random intercepts and slopes model failing to find convergence, the Likelihood Ratio Test of the change in the squared log-likelihood estimate with the addition of allowing the slopes as well as the intercepts to float randomly across the levels did not result in a statistically significant improvement in the model (p $|chi^2| > 0.5$). Additionally, research does not support cross-level interaction analysis in three-level MLM, which requires the random-intercept and slopes model to be utilized (Aguinis et al., 2013). Therefore, to account for the nested structure of our data and test our hypothesis, we will not evaluate covariates at differing levels of analysis; however, we will allow the intercepts of the level's categorical variable to float with the random intercepts model. Therefore, the model used in our analysis is the random-intercept model noted as follows:

$$Logit(RELIANCE_{ijk}) = \beta_0 + \beta_1 EXPERIENCE_{ijk} + \beta_2 IDA_{FAMILIARITY\,ijk} + \beta_3 COGNITIVE_{FIT\,ijk} + \beta_4 COMPLEXITY_{ijk} + (\beta_{14} COMPLEXITY_{ijk} * EXPERIENCE_{IJK}) + (\beta_{24} EXPERIENCE_{IJK} * IDA_{FAMILIARITY\,ijk}) + (\beta_{34} COMPLEXITY_{ijk} * COGNITIVE_{FIT\,ijk}) + v_i + e_{ijk} + k_{ijk}$$

where :

$$The\ unexplained\ Level\ 3\ (USER)\ variation = k_{ij} \sim (0, \sigma_k^2)$$

$$The\ unexplained\ Level\ 2\ (MODEL)\ variation = ee_{ij} \sim N(0, \sigma_e^2)$$

$$The\ unexplained\ Level\ 3\ (USER)\ variation = v_i \sim (0, \sigma_v^2)$$

CHAPTER 4: RESULTS

This section provides the results of testing the four hypothesized relationships from our conceptual framework. First, the result of the multi-level logistic regression models is presented. Next, the result of the marginal analysis of the direct effect hypothesis is presented to determine the direction and significance of the relationship. The results from calculating the true interaction effects of significant interactions are provided to determine their direction and significance as well as the marginal impact of the moderating influence. This section closes with an updated conceptual model indicating our findings.

4.1 Logistic Regression Results

Table 6 presents the results of a multi-stage, hierarchical regression model first consisting of only the Theory of Technology Dominance propositions (Model 1), our hypothesized direct effect (H_1) of Decision Aid Complexity (Model 2) and finally our hypothesized interaction effects (H_2, H_3 & H_4) of Decision Aid Complexity on the relationships between the TTD propositions and Decision Aid Reliance (Model 3). As discussed, the nested structure is accounted for by using a multilevel random intercept logistic regression model, which allows the intercepts of USERID and MODELID to float among the variables. Evaluating hypothesis is done by comparing each model using the differences between their likelihood ratio Chi^2 test to determine if the inclusion of the direct-effect hypothesis and then interaction hypothesis result in significant improvements. Findings are that there is a significant improvement at the $p<0.05$ level between Model 1 and Model 2 but do not find significant improvement between Model 2 and Model 3. Model 3, however, does support a significant improvement ($p<0.05$) over Model1.

Model 1 includes the variables representing the Theory of Technology Dominance propositions that influence Decision Aid Reliance: Task Experience, IDA Familiarity, and Cognitive Fit. The model is significant at the $p<0.05$ level. While not hypothesized as control variables, as presented in our research, these covariates have been found to be influential to Decision Aid Reliance. The model coefficients for Task Complexity, IDA Familiarity, and Cognitive Fit are all positive, but only Task Experience is found to be significant (Beta = 0.369; $p<0.05$). These results are inconsistent with the propositions of TTD, which proposes a negative relationship between Task Experience and Decision Aid Reliance.

Table 6: Model Results

Reliance	Model 1	Model 2	Model 3
Experience	0.369*	0.377*	0.343*
IDA Familiarity	0.006	0.003	0.007
Cognitive Fit	0.006	0.012	0.007
Complexity		-0.036*	-0.027
Complexity*Experience			-0.257*
Complexity*IDA Familiarity			0.193
Complexity*Cognitive Fit			0.019
Constant	2.242	2.985	2.840
userid			
var(_cons)	2.038	1.935	1.965
userid>modelid			
var(_cons)	3.255	3.306	3.351

No. of Observations = 4940

Likelihood Ratio Chi2 5.6* 5.55(11.15*)

* p<0.05
** p<0.01
*** P<0.001

Model 2 demonstrates a significant model improvement over Model 1 when including our main-effects hypothesis; that Decision Aid Complexity has an influencing effect on Decision Aid Reliance. The estimated coefficient of Decision Aid Complexity in Model 2 is negative and significant (Beta = 0.036; p<0.05), providing support for Hypothesis 1.

Model 3 includes our moderating hypothesis of the interaction of Decision Aid Complexity on the relationship between Task Experience, IDA Familiarity, and Cognitive Fit. Our analysis finds that Decision Aid Complexity has a significant interaction effect (Beta = -0.257; $p<0.05$) on the relationship between Decision Aid Reliance and Task Experience.

4.2 Hypothesis 1: Marginal Effect of Decision Aid Complexity

Though Model 2 results show support for H1 & and Model 3 shows support for H_2, Hoetker (2007) and Wiersema & Bowen (2009) caution researcher's using logistic regression or similar methods with categorical or binary dependent variables that reliance on the model coefficients or signs as directional relationship indicators may not be representative of the actual relationship or direction at all values of the predictor (Hoetker, 2007; Kistruck, Morris, Webb, & Stevens, 2015; Wiersema & Bowen, 2009). This analysis relied on the method prescribed in Wiersema & Bowen (2009) and followed the reporting & analysis recommendations, which examines marginal effects to interpret the true nature of the hypothesized relationship (Wiersema & Bowen, 2009). This includes a visual analysis of the model's marginal effect and the true interaction effect across all of the predicted probability ranges (Wiersema & Bowen, 2009).

To test the direct effect hypothesis of Decision Aid Complexity on Decision Aid Reliance, we performed the marginal effect analysis described by Wiersema and Bowen (2009). Hypothesis 1 proposed that given constant Task Complexity, the probability that a user relies upon a decision aid is lower when the complexity of the decision aid is high. As presented in our conceptual model, this hypothesis suggests that the relationship is both significant and negatively associated. To evaluate the directional hypothesis within our logistic regression, we evaluate the incremental significance and direction of each of the observations of Decision Aid Complexity. It is generally agreed that the most suitable mechanism for conducting marginal impact analysis on limited dependent variable main effects is to use illustrations (Hoetker, 2007).

Figure 9 illustrates the marginal effect of Decision Aid Complexity on the probability of Decision Aid Reliance. As the graph illustrates, all marginal effects across the range of probabilities are negative (all y-axis values are negative). This supports the directional statement of H_1 as well as confirming the direction of the regression coefficients. As for significance, we do see that there are both significant and insignificant values across the range of probabilities of Decision Aid Reliance. All values below the line where z > 1.96| are significant; however, this does not hold across all values with the majority of effects being insignificant. Continuing to follow the Wiersema & Bowen (2009) method, we calculate a summary of the marginal effect of Decision Aid Complexity on Decision Aid Reliance at the means of the model variables.

The marginal effect is found by differentiating the regression with respect to Decision Aid Complexity. Adapting the equation provided by Wiersema & Bowen (2009) from our regression model, we calculate the marginal effect as specified in the equation below (Wiersema & Bowen, 2009).

[1] $$Marginal\ Effect\ of\ Complexity = \frac{\partial \Pr\ (Rely = 1|\mathbb{V}, \beta}{\partial\ Complexity}$$

$$= \frac{\partial \Pi(\mathbb{V}\beta')}{\partial\ Complexity} = \pi(\mathbb{V}\beta')\beta_x$$

[1] Equation is adapted from Wiersema & Bowen (2009) "equation 2" (Wiersema & Bowen, 2009, p. 683).

This calculation resulted in a marginal effect of -0.0035 and a z-statistic value of -1.743, corresponding to a p-value of 0.041. The main-effect hypothesis (H_1) is supported as the relationship between Decision Aid Complexity and Decision Aid Reliance has been found to be both negative and significant.

With the relationship shown to be statistically significant, we expected to see a distribution where the level of significance across the range of predicted probabilities was more discrete instead of the varying levels we see. This may not have occurred in this analysis due to the small effect size (Beta = -0.036), causing other effects to influence the predicted probability more at the marginal effect. This could also explain, as you can see in Figure 9, that there are many observations across the range of predicted probabilities all positioned at the same significance levels.

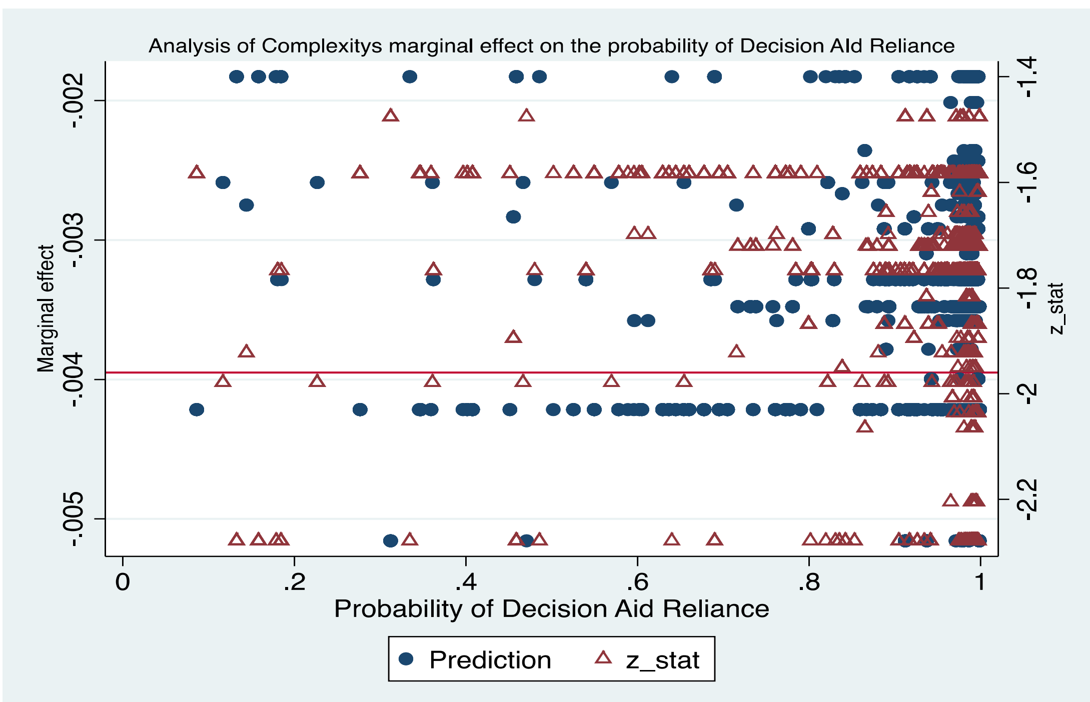

Figure 9: Marginal Effect of Complexity

4.3 Hypothesis 2: True Interaction Effect of Decision Aid Complexity

In the regression results, the interaction estimates (coefficients) of Decision Aid Complexity on the relationship between Decision Aid Reliance and Task Experience show a negative effect, which is significant at the $p<0.05$ level. In Hypothesis 2, we proposed that when Task Experience is high, the likelihood, and therefore probability, of reliance would be less given high complexity. This suggests a negative moderating effect of Decision Aid Complexity on the relationship. As with the main-effect hypothesis, we cannot rely upon the estimates of the logistic regression coefficients to make determinations regarding the direction and significance of interactions (Hoetker, 2007; Wiersema & Bowen, 2009). In evaluating the moderating effect, we again follow the method suggested by Wiersema & Bowen (2009) to test the moderating hypothesis, which requires calculating the "True Interaction Effect," which is calculated as the cross-partial derivative of the marginal effect (Wiersema & Bowen, 2009). The standard-error and significance are calculated at each value of the interaction effect. To calculate the true interaction effect, we calculate the marginal effect using our regression equation to include the moderating variables as described in evaluating H1. We then solve for the true interaction effect by differentiating the marginal effect by the interaction of Task Experience and Decision Aid Complexity:

$$^{2}True\ Interaction\ Effect = \frac{\partial(Marginal\ Effect\ of\ EXPERIENCE)}{\partial\ (COMPLEXITY)}$$

$$= \frac{\partial\Pi(\mathbb{V}\beta')}{\partial EXPERIENCE * \partial COMPLEXITY}$$

$$= \Pi(\mathbb{V}\beta')(1 - \Pi(\mathbb{V}\beta'))[\beta_{COMPLEXITY*EXPERIENCE)} + 1(-2\Pi(\mathbb{V}\beta')))$$

$$= (\beta_{EXPERIENCE} + \ \beta_{EXPERIENCE*COMPLEXITY}COMPLEXITY)(\beta_{COMPLEXITY}$$

$$+ \beta_{EXPERIENCE*COMPLEXITY}EXPERIENCE$$

Figure 10 illustrates the True Interaction Effect of Decision Aid Complexity on the relationship between Task Experience and Decision Aid Reliance. The solid Figures represent the values of the True Interaction Effect, while the triangle-shaped symbols represent the corresponding z-statistic. The interaction effect is plotted on the left y-axis, with the z-statistics plotted along the right y-axis. As illustrated, the true interaction effect of Complexity on the relationship between Task Experience and Decision Aid Reliance is both negative and significant as well as positive and significant at values across the range of the predicted probabilities of Decision Aid Reliance.

Because we find both the significance and direction of the true interaction effect changing between being both positive and significant as well as negative and significant across virtually all levels of the predicted probability of Decision Aid Reliance, relying upon the illustration alone is inconclusive to determine effect significance and direction. Had we found all values to indicate there was not a change in the direction, a summary statistic at the mean value of Decision Aid Complexity could be relied upon (Wiersema

---

[2] Equation is adapted from Wiersema & Bowen (2009) "equation 5" (Wiersema & Bowen, 2009, p. 686)

& Bowen, 2009). As we do see the direction and significance changing, we calculate the significance of the effect at the minimum, mean, and maximum values of Decision Aid Complexity. The results, shown in Table 7, demonstrate that the effect of the interaction decreases as the value of Decision Aid Complexity increases, supporting a generally negative moderating effect. However, this observation is only significant at or below the mean values of Decision Aid Complexity. At levels of Decision Aid Complexity between its mean and maximum values, the moderating effect is rendered insignificant. The graphical illustration also shows that while the maximum value of Decision Aid Complexity is not significant, there are values of the predicted probability of Decision Aid Reliance, generally when the predicted probability is above 0.80 where the direction of the true interaction effect is both positive and significant shown by the values above $z>1.96$ and above the true interaction effect >0. This suggests that at some but not all levels of Decision Aid Complexity above its mean value, its marginal effect on the relationship between Task Experience and Decision Aid Reliance increases the likelihood of user reliance consistent with H_2

Hypothesis 2 proposes that the probability of Decision Aid Reliance is reduced when accompanied by higher levels of Decision Aid Complexity. Our results have found that the direction and significance of the moderating effect support accepting $H_{2,}$ however not completely. When Decision Aid Complexity is increasing, the log-odds and, therefore, the probability of the relationship between Task Experience and Decision Aid Reliance is reduced to a point. However, as the interaction effect is not consistently significant at the maximum levels of Decision Aid Complexity, we claim that H_2 is only partially supported.

Practically this finding suggests that while complexity is impacting the relationship between Task Experience and Decision Aid Reliance, it is not consistent and indicates that other conditions or considerations not under analysis for this paper may be impacting the relationship. In discussion with subject-matter experts within the organization, the complexity of the model may belie complexity in the decision due to several factors such as historical experience with defaults (specifically thin-defaults), systemic data capture capabilities to obtain a reasonable training sample to build the model with and legacy operating effects.

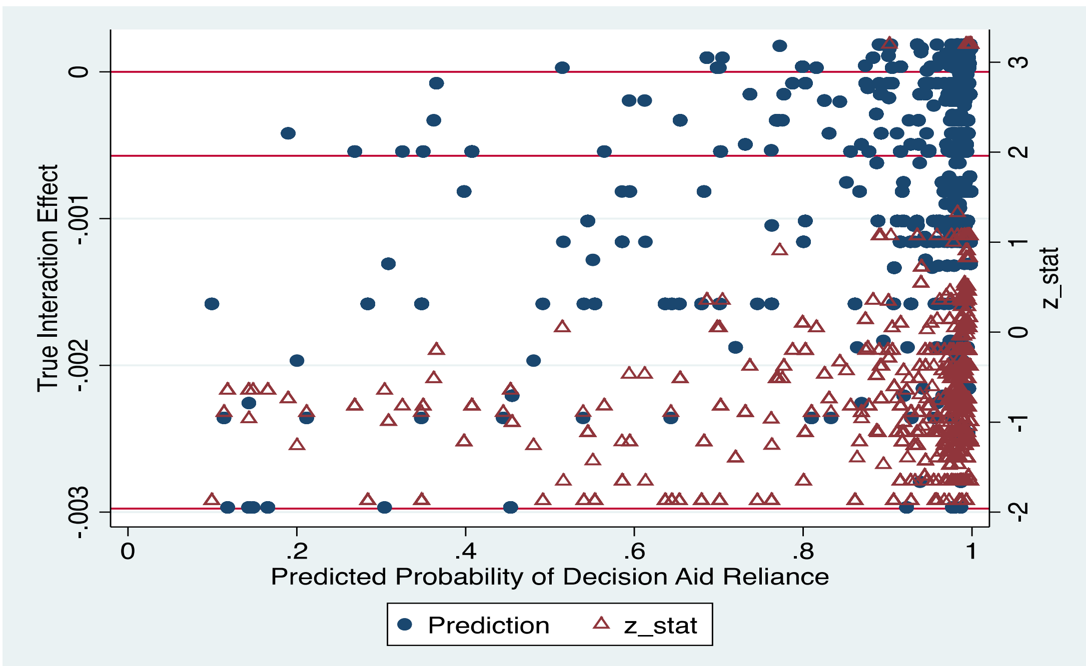

Figure 10: True Interaction Effect

Table 7: Effect of Complexity on the marginal effect of Experience on the probability Decision Aid Reliance

Value of Decision Aid Complexity	Marginal Effect of Task Experience	z-Statistic
Minimum	0.012*	2.76
Mean	0.009*	2.23
Maximum	0.004	0.61

The minimum value of Decision Aid Complexity has a marginal effect on the relationship between Task Experience and Decision Aid Reliance of 0.012 and a z-statistic of 2.76, which corresponds to a p-value of 0.0029. The mean value of Decision Aid Complexity has a marginal effect on the relationship between Task Experience and Decision Aid Reliance of 0.009 and a z-statistic of 2.23, which corresponds to a p-value of 0.0128. The maximum value of Decision Aid Complexity has a marginal effect on the relationship between Task Experience and Decision Aid Reliance of 0.004 and a z-statistic of 0.61, which corresponds to a p-value of 0.4756. A z-statistic with an absolute value of 1.96 corresponds to a p-value of 0.05, which is the point where we determine whether or not the observation is statistically significant.

4.4 Hypothesis 3 & 4

While we understand that because the original regression estimates did not find significant interactions between Complexity and the Relationships between IDA Familiarity and Decision Aid Reliance or Cognitive Fit and Decision Aid Reliance, we do present the graphical representation of the True Interaction Effects for both as confirmation. Unlike the interaction of Complexity on the Relationship between Task

Experience and Decision Aid Reliance, as H_3 & H_4, there should not be a secondary interaction effect to evaluate.

Figure 11 is the plot of the True Interaction Effect of Complexity on the relationship between IDA Familiarity and Decision Aid Reliance. As the illustration shows, all values of the predicted probability are of the same direction (negative) and insignificant at all levels. Therefore, as discussed in Wiersema & Bowen (2009), we may use the summary statistics at the mean value of Complexity to evaluate overall significance. Results are a marginal effect of -2.359 with a standard error of 2.435 and a z-statistic of -0.97 corresponding to a p-value of 0.333. These results confirm the lack of moderating effect of Decision Aid Complexity on the relationship between IDA Familiarity and Decision Aid Reliance.

Figure 12 is the plot of the True Interaction Effect of Complexity on the relationship between Cognitive Fit and Decision Aid Reliance. As the illustration shows, all values of the predicted probability are of the same direction (negative) and insignificant at all levels. Therefore, as discussed in Wiersema & Bowen (2009), we may use the summary statistics at the mean value of Complexity to evaluate overall significance. Results show a marginal effect of -2.832 with a standard error of 5.936 and a z-statistic of -0.48 corresponding to a p-value of 0.633. These results confirm the lack of a moderating effect of Decision Aid Complexity on the relationship between Cognitive Fit and Decision Aid Reliance.

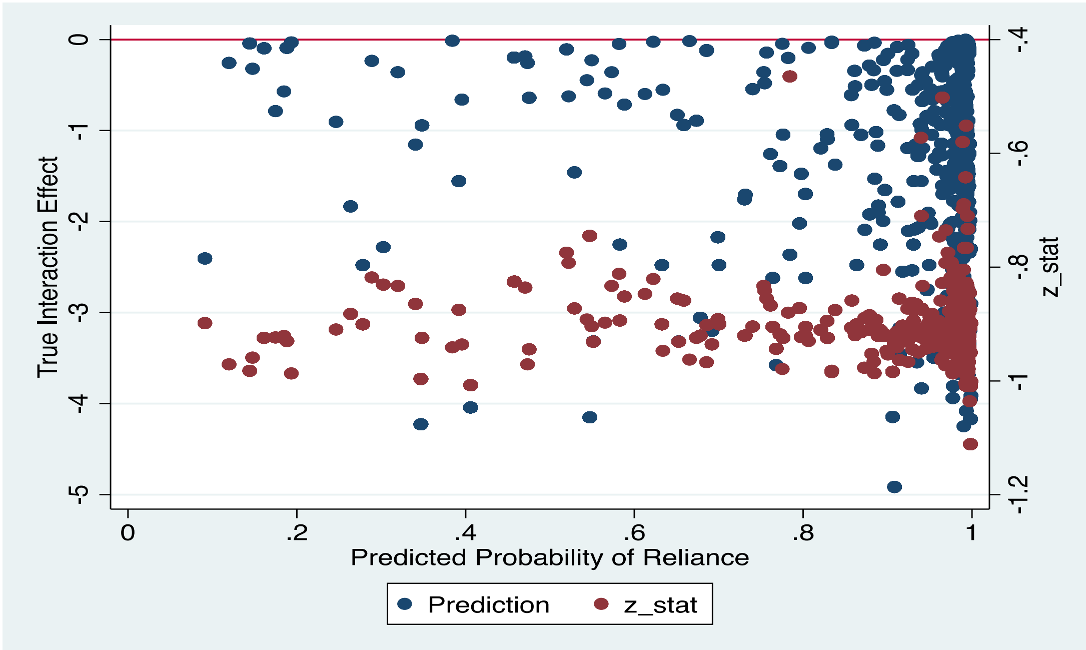

Figure 11: True Interaction Effect of Complexity on Reliance ~ IDA Familiarity

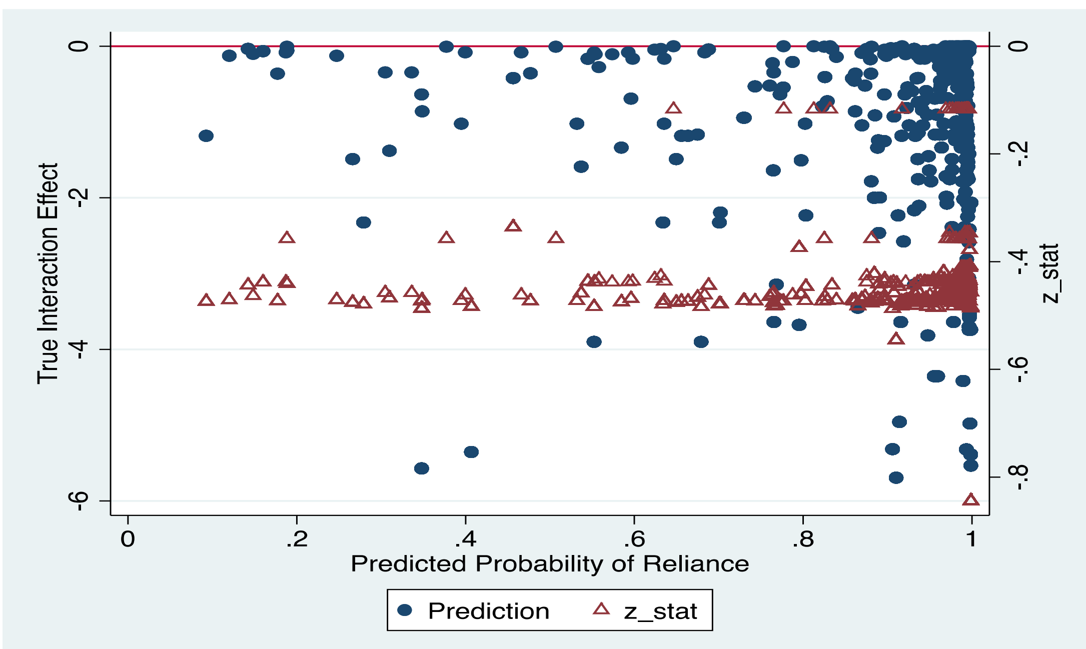

Figure 12: True Interaction Effect of Complexity on Reliance ~ Cognitive Fit

In concluding our analysis, we present in Figure 13, an updated graphic of our conceptual model presenting in Chapter 2. The bold lines represent our findings in support of H_1 & H_2; that there is a negative relationship between Decision Aid Complexity and Decision Aid Reliance and that Decision Aid Complexity negatively moderates (at some levels) the relationship between Task Experience and Decision Aid Reliance. The dashed lines, indicating the unsupported hypothesis H_3 & H_4 while the solid lines indicate the TTD proposed relationships examined in our research.

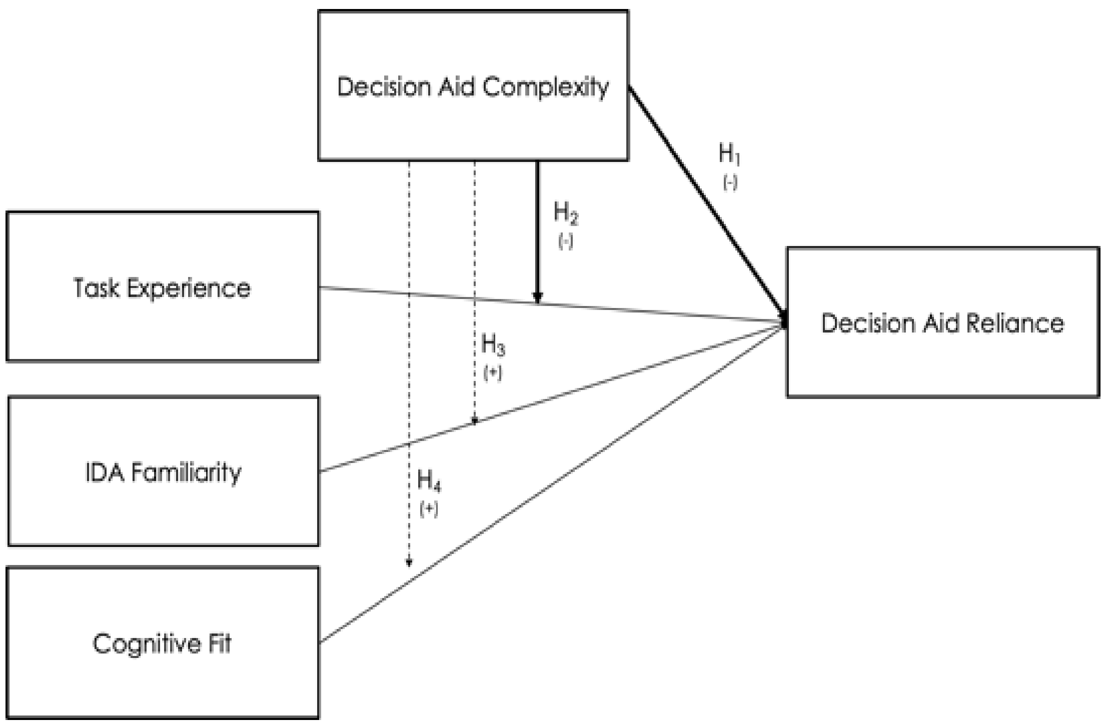

Figure 13: Conceptual Model with Results

Table 8: Hypothesis Results

Supported	
H1	Given constant Task Complexity, the probability that a user relies upon a decision aid is lower when the complexity of the decision aid is high.
Partially Supported	
H2	Given constant Task Complexity, the probability is lower than a user with high task experience will rely on a decision aid with high complexity.
Not Supported	
H3	Given similar Task Complexity, the probability that a user will rely upon a decision aid when the user is highly familiar with the IDA is also high when utilizing a complex decision aid.
H4	Given consistent Task Complexity, the probability that a user will rely upon a decision aid when there is Cognitive Fit is increased when the complexity of the decision aid is high.

4.5 Post Hoc Analysis

While we acknowledge that our experiment did not follow traditional methods with regards to multilevel modeling because our hypothesis contained a Level 2 covariate moderating Level 1 and Level 3 variables, we failed to successfully fit a model with any of three commonly used software packages (R, STATA, SPSS) which would converge a Random Intercepts and Slope Model containing the full model with the levels specified per the hierarchy illustrated in Figure 8. To ensure we adjusted for any technological or conceptual differences, instead of using STATA 16 as we did with our hypothesis testing, for this post hoc analysis, we produced results relying upon the GLMR function of the R

Studio LME4 v1.1.21 package which is used widely to fit generalized and linear mixed effect, hierarchical and multilevel models (Bates, Mächler, Bolker, & Walker, 2014).

The results of the post hoc test below show the model estimates of a random-intercepts and slope model, which manipulates the multilevel structure to force Experience as a Level 1 predictor and fixes the slope of the models by their complexity in Level 2. As with our initial attempt to fit a random-intercepts and slopes model, this model too fails to converge. What is optimistic is that these model estimates, even given its noted limitations, do not negate the findings presented previously. The significance and direction of our main effect and moderating effects remain consistent with the random intercept model.

Failures of convergence in maximum likelihood estimate models as we are utilizing here often suffer from convergence failures as a result of fitting too complex of a model to the data (Bates, Kliegl, Vasishth, & Baayen, 2015). We assume that to be the case in why we were unable to fit this model. The model structure is too complex, and with MLM effectively modeling each parameter of the model multiple times, our data just does not support the complexity of the random intercepts and slopes model. It is possible that relying on the interclass correlation coefficient to determine that MLM was required was insufficient given the lack of true nested structure that we see in much MLM analysis (e.g. "students, classes, schools" or "patients, doctor hospitals," or "city, state/province country").

Table 9: Post Hoc Analysis Results

AIC	BIC	Log-Likelihood	deviance	df residual
2487.6	2578.7	-1229.8	2459.6	4927

Scaled residuals:

Min	1Q	Median	3Q	Max
-11.7171	0.0962	0.1421	0.2815	2.1175

Random effects:

Groups	Name	Variance	Std. Dev.
userid	(Intercept)	7.449182	2.72932
experience	0.750224	0.86615	-0.86
modelid	(Intercept)	0.019176	0.13848
complexity	0.001172	0.03424	-1

Fixed effects:

	Estimate	Std. Error	z-value	Pr (>|z|)
(Intercept)	1.2548737	0.7421388	1.691	0.090858
complexity	0.0700359	0.0305527	2.292	0.021889*
ida_familiarity	-0.0008604	0.0056579	-0.152	0.879129
cognitive_fit	0.0052772	0.0065703	0.803	0.421863
experience	0.2590976	0.1633845	1.586	0.112781
INT_Experience	-0.3027973	0.0787522	-3.845	0.000121***
INT_IDAFam	-0.0869637	0.1217451	-0.714	0.475036
INT_cfit	0.704788	0.3057762	2.305	0.021171*

* p<0.05

** p<0.01

*** P<0.001

CHAPTER 5: DISCUSSION & IMPLICATIONS

5.1 Theoretical Implications

Complexity in decision making is a defining characteristic of today's management environment. It is changing how we work, interact with one another, and the technology organizations need and use. This transition can be seen in the changes from the traditional IT servicing model towards cloud-based solutions worldwide. Even in developing countries, a recent industry survey found that since 2015, cloud-based IT spending has seen an increase while traditional in-house IT has reduced (Yaokumah & Amponsah, 2019). We are not yet to the point where human decision-making has been made obsolete to AI or ML algorithms, but we are progressing along that path for many decisions made by management today. While Decision Aid research and design have focused on users interacting with or their reactions to feedback and guidance provided by the tool, these mechanisms retain the assumption that humans will continue to be the final arbiter of decisions.

Our study's aim has been to advance the scholarly understanding of the factors that contribute to individual reliance behavior using intelligent decision aids. In doing so, we extend previous theory on reliance influencers, specifically those proposed in Arnold & Sutton (1998) Theory of Technology Dominance, to include the complexity of the decision aid's embedded agent (Decision Aid Complexity). Embedded agents reside within intelligent decision aids and operate both invisibly and autonomously of the decision aid's user. We find that differences in embedded agents, specifically that the difference in their complexity, influences user reliance behavior. Taking these aspects together, developers of decision aids can make design decisions, which may result in unintended outcomes.

The finding that Decision Aid Complexity negatively moderates the relationship between Task Experience and Decision Aid Reliance is interesting. As the relationship was negative and significant at only low levels of Complexity, it leads us to consider that this has occurred, in part, due to the fact that no model of risk management can ever capture all aspects of the real-world risk at any given time. How we have measured Decision Aid Complexity suggests that the more complex the embedded agent is, the more risk factors the model is considering. Therefore, when considering a high-risk customer, the embedded agent would almost certainly be more complex. Finding that Decision Aid Complexity negatively moderates the impact of Task Experience less as it increases suggests the possibility that factors influenced by Complexity and not necessarily Complexity itself is a driver of the relationship.

A possible explanation of this is that in practice, lower-experienced users would be less likely to be evaluating risk for the most complex (high-risk) customers. Consistent with TTD, we would expect less experienced users to tend towards reliance on the decision aid. Therefore, at low levels of Complexity, we would expect the Task Experience levels to be similarly low, and as Complexity increases, Task Experience would increase similarly and somewhat mute the impact.

We see this phenomenon represented in our data. As illustrated in Table 15, the mean Complexity score of each embedded agent is higher as the Task Experience level increases. Analysis of Variance (ANOVA) results indicated that there were significant differences between mean Decision Aid Complexity measures within the Task

Experience levels. We then conducted Tukey's HSD test to isolate the significant differences in each level of Task Experience (Tukey, 1977).

The Tukey's HSD Test findings (Table 16) are that we have significant differences between the means of Decision Aid Complexity at Task Experience levels 2 & 3, levels 3 & 4, and levels 4 & 5. We do not have significant differences in mean complexity scores between Task Experience Levels 1 & 2 or 3 & 5. Though we do have a lack of homogeneity between each of our Task Experience levels with regards to the mean Decision Aid Complexity score, the data support that in general, lesser experienced users interact with less complex models. We understood that the clustering of Task Experience would influence our results, as explained by our ICC calculation (Table 5). However, the magnitude of that clustering is apparently sufficient enough to also change interaction effects as Complexity increases.

Given our results, we consider the outcome that organizations should align more complex decision aids with more experienced users in order to offset the influence of complexity as our results indicate that eventually, the interaction is eliminated and positively moderates Decision Aid Complexity.

This is consistent with the relationship between Decision Aid Reliance and Task Complexity as proposed in TTD, whereby a mismatch in the experience of the user and the complexity of the task result in diminished reliance behavior (Mascha & Smedley, 2007). Our findings suggest that even when the decision aid in use is common, and the decision-making agent is hidden, design elements leading to Decision Aid Complexity could be synonymous with Task Complexity. Research on decision aids has long held that complexity influences reliance as it adds to the overall cognitive cost of decision making. Therefore, we can reasonably expect that a user will rely upon a decision aid as long as it lessens their cognitive cost (Todd & Benbasat, 1992). Research has also supported that individuals often insert their judgment for the judgment of the tool unless design features build trust, persuasiveness or provide some other mechanism that increases users perceived validity of the decision aid (D. Brown & D. R. Jones, 1998; Kaplan et al., 2001; Mălăescu & Sutton, 2015; Mascha & Smedley, 2007). Our results find that even without differences in cognitive cost through the use of embedded agent technologies within intelligent decisions aids, users will continue to evaluate the cognitive cost of the decision in their reliance behaviors.

Our research provides several theoretical and research contributions. Research that has evaluated design impacts on decision aid reliance has focused primarily on functional aspects of how users cognitively interact with a decision aid or how guidance and feedback mechanisms within a decision aid influence reliance. Both of these features require the overt engagement of the user. In contrast, design features that are passive, which do not require overt interaction by the user, but which are inherent in the decision process, such as an embedded agent, has not been widely considered within decision aid research. In addressing this limitation in research, we illustrate that reliance behavior can be influenced by differences in a decision aid's design using a differentiator of complexity without the feature being visible to the user, i.e., covert design elements. This is important as decision aids increasingly utilize complex machine learning and artificial intelligence mechanisms, which frequently rely upon embedded agents. Our findings suggest that in addition to decision-level contributions and user-level contributions as influencing aspects decision aid reliance, components of the technology underlying how the decision aid evaluates decisions is also an important influencing aspect of reliance behavior.

Also, researchers have traditionally evaluated complexity within decision aid research as a function of the decision being made (Jensen et al., 2010; Mascha & Smedley, 2007; Parkes, 2017). This was done primarily in response to decision support research, which identified cognitive cost as a prominent determinant of decision aid reliance (Todd & Benbasat, 1992). Relying on a broader definition of complexity by incorporating research from Operations and Supply-Chain disciplines, our results support that the static complexity of a decision aid can influence reliance in process terms as well as the decision contexts. Specifically, complexity as an influencing aspect of decision aid reliance is inherent to the overall process of technology-supported decision-making when including an IDA into the user's decision-making process.

By empirically evaluating the propositions of TTD, we intended to provide further evidence that its propositions would be supported through analysis within a practice setting. The proposition of IDA Familiarity is based upon the cognitive cost theory as described by Todd & Benbasat (1992) in that the more familiar a user is with a decision aid, the lower the cognitive cost of use, or as TAM defines as a lower "perceived ease of use" (Davis, 1989; Todd & Benbasat, 1992). As did Hampton (2014), we too failed to find a significant relationship between Decision Aid Reliance and Decision Aid Familiarity. This continued lack of significance finding may suggest that Decision Aid Familiarity may not be a sufficiently influencing factor of reliance behavior given the current technological state if intelligent decision aids. Decision Aids which heavily rely on embedded technology could render the user's familiarity with the decision aid irrelevant.

Firms that are more transactional are focused on generating sales or revenue with a customer as a series of engagements. The general belief of firms with this type of orientation is that through their actions, the firm is fulfilling demand and provide choice for consumers that, in turn, incentivize them and other firms to continually increase value to the customer (Sheth & Parvatiyar, 1995). However, there is no intent to establish a relationship with the customer other than for repeated transactional engagement and performance over competitors (Kumar, Bohling, & Ladda, 2003). Though the landscape is continuously changing, these firms would be engaged in businesses characterized by providing goods that turn-over quickly, such as commoditized products or consumer retail (Coviello & Brodie, 1998). In contrast, firms with a more relational-driven posture do strive to deepen relationships with their clients over the long-run and view the relationship as a source of competitive advantage (Buttle & Practice, 1996). These firms are characterized by longer time-horizon goods or services such as those provided to commercial & industrial customers (Coviello & Brodie, 1998).

Some firms, especially larger ones, often perform activities that are both transactional and relational (Walsh, Gilmore, & Carson, 2004). Users of the decision-support aid in our experiment, though employed by the same organization, were primarily from two separate functions within the organization under analysis. The first group consists of those charged with ensuring the overall industry or product risk management while the second is concerned with servicing and maintaining the relationship with the clients or potential clients. The "risk-management" users would resemble a more transactional posture firm or industry as the relationship is not a part of their mandate, while the "portfolio manager" would resemble a posture similar to that of a relational firm or industry type. Applying the transactional and relationship marketing postures previously discussed, if a user of a decision aid is operating within a relational market orientation, we could expect reliance behavior to take into account the relationship as a source of advantage and allow it to influence reliance behavior. Similarly, if a user of a decision aid were more transactional in their posture, the longevity of the relationship would not necessarily be a concern, but there would be motivation to execute the transaction in the firm's best interest. In either case, the relationship strategy of the firm could be a consideration by the user as an influencing aspect of Decision Aid Reliance.

As broad and large as the organization product and customer mix is, they position themselves in both transactional and relational engagement with customers frequently. Our experiment included participants who are a part of both of those customer engagement strategies. Though we did not hypothesize any effects with respect to market orientation, by including users with differing customer orientations, our results suggest that our findings would be generalizable across both transaction-focused and relationship-focused firms and industries. This would enable our findings to be theoretically applicable to firms offering products that move quickly between supplier and consumer, such as much e-commerce and those which take longer to manifest, such as a commercial real estate loan.

These results could practically apply to Actuarial functions within insurance companies, which would also possess decision aids containing embedded agents producing hazard models of risk for life, health, and automobile underwriting. Electronic

Medical Records are an emerging technology that generates vast amounts of data that could enable diagnostic decision aids for physicians. A recent change in Generally Accepted Accounting Principles has mandated the incorporation of predictive analytics into allowance provisions, which necessitate the use of decision aids in completing accounting functions for large institutions. Each of these is a unique use in which the complexity of a decision aid and its influence on reliance behavior could be of interest to both researchers and practitioners.

5.2 Practical Implications

Developers of embedded agent AI technology already address designs in algorithms that may yield unintended bias with respect to race, age, and gender (Introna & Wood, 2004). Furthermore, one of the critical findings in factors that led to the Great Recession is that financial models produced unintended consequences and that users failed to understand under what conditions they could be used and, more importantly, should not be used (Danielsson, James, Valenzuela, & Zer, 2016).

While this paper was not specifically designed as a case study, our analysis did rely upon the data and a business process used within an operating organization. The experimental design did not alter the observations or place controls on the organization to produce the data under analysis. As a result, our analysis identified an opportunity for the organization to consider changes to their policies and procedures to potentially increase the value of their monitoring of reliance behavior as a mechanism to manage model risk. Specifically, we found that a large percentage of the dataset provided created Obligor Risk Ratings within the decision aid utilizing a mechanism which bypasses the business rules that selects the embedded agent and allows for expert judgment to score the risk of the customer or potential customer. As discussed in Section 4.1, our original sample, before selecting the users to gather data on, consisted of 30,090 interactions with the decision aid. This includes all interactions which were identified as being performed by a user and not a systemically generated action. In the analysis presented in this paper, we excluded 10,904 interactions as they were the result of the expert-judgment bypass action, and therefore the construct of Decision Aid Complexity could not be measured. This mechanism does not contain the feature of an embedded agent and is not a model that can be evaluated as we did with the other models. This scoring method during the time-period our data covered was the most widely used method of scoring, with over 1900 more observations than the highest used embedded agent model.

Each use of the expert-judgment scoring mechanism could be appropriately considered to represent a rejection of the Decision Aid's recommendation as the user made the overt decision to not include the decision aid within their decision-making process consistent with Arnold & Sutton's (1998) definition of reliance. Based upon our results, had each of these methods been scored utilizing the Decision Aid, and later rejected we could have possibly seen a greater effect size of the relationship between Decision Aid Complexity and Decision Aid Reliance as the volume of the Reject class would have been increased by a factor of five.

The organization relies, in part, on the rejection rates of the embedded agents as a method to monitor and inform whether or not it is performing as intended. Bypassing the embedded agent at this high a rate may be dilutive of the override rates of the population and underrepresenting true rejections. We suggest a change in the procedure used by the organization to require each scoring transaction to occur per the standard process utilizing the embedded agents. If the user determines that the recommendation should be rejected, use the expert-judgment mechanism to inform the approval of the override, but not use in lieu of the standard process.

CHAPTER 6: LIMITATIONS AND FUTURE RESEARCH

6.1 Overview

In this section, we describe limitations found within our research and provide suggestions for future research. We identified two areas of material limitations to our analysis; data limitations and experimental design limitations, which included the inability to evaluate multiple intelligent Decision Aids and the scope of analysis relying upon a single organization's experience. We will explain these limitations in further detail, describe how our analysis may have been impacted by the limitation, and suggest how future research may be designed to address the limitations.

6.2 Data Limitations

Though we do find significance in the relationship we hypothesized between Decision Aid Complexity and Reliance, there is a possibility that the significance we found was a result of the large size of the sample we utilized. Because of the expansion of datasets paralleling the rise of Big Data availability, researchers have access to vast amounts of data. Lin et al. (2013) reviewed 98 Information Systems Research studies which relied upon p-values with large sample sizes and suggested that due to the mathematical properties of the p-value logical migration towards zero as sample size increases, relying on p-values for statistical inference could be problematic (Lin, Lucas Jr, & Shmueli, 2013). However, Lin et al. (2013) do not suggest that reliance on the p-value to infer significance is not supportable and that the "increased power of large sample-sizes means that researchers can detect smaller effects" (Lin et al., 2013, p. 1).

While we claim the latter due to the ratio of Rely to Reject outcomes in our large sample, and believe we address this sufficiently through utilizing the marginal-effect analysis, that our results have statistical significance, we acknowledge the possibility of this limitation influencing our analysis.

A second data limitation was that we did not collect user-specific data beyond our measurement of their Task Experience levels due to our inability to leverage confidential data of the organization. This resulted in our inability to evaluate common variables of interest, such as age, gender, industry, transaction scale, etc. Future research should work with organizations to obtain access to these important factors. Additionally, this limitation did not present the opportunity to evaluate user bias. We assume, based upon our results from the relationship and interaction between Task Experience, Decision Aid Complexity, and Decision Aid Reliance, that automation bias could be playing a role here that we were unable to control for or measure. Future research in this area should include a control for this type of bias.

We are generally comfortable with the measures used to operationalize our variables for Experience, IDA Familiarity, and Complexity. However, we do acknowledge a limitation in the operationalization of our measure for Cognitive Fit. Extant literature on Cognitive Fit almost universally measures this as the ability of a user to perform (accurately) as a measure of time to act, or click-speed to accurate response (Shaft & Vessey, 2006; Sinha & Vessey, 1992; Vessey, 1991, 1994; Vessey et al., 2006). Our inability to measure the user's speed-to-decision time may not be measuring the effect of this construct as intended. Had we found significant results of that measure; we would be hesitant to have claimed those results as theoretically relevant. Additionally, as discussed in explaining the findings in H_2, the variation in results indicate that our measure of complexity did not control for model development factors that informed decisions, which resulted in the complexity that we did measure. Those factors include data availability at the time the model was built, historical experience and legacy operating decisions or events

To address data limitations, we would suggest for future research that a richer dataset including more variables to control or differentiate the observations with a more balanced class. This could allow for alternative analysis controlling for factors as described above using analysis of variance methods, which could more effectively isolate and measure the true source of variation in Decision Aid Reliance.

6.3 Design Limitations

The second limitation we acknowledge is in the design of our experiment. We relied upon differences in the embedded agents within a single Decision Aid. While this design ensured that we were able to isolate the Decision Aid Complexity construct and eliminate any variation in task complexity, we do not know if these results would be repeated utilizing another Decision Aid. To add to the generalizability of our results, we suggest that additional experiments be conducted on embedded agent complexity incorporating multiple Intelligent Decision Aids. Finding similar results would provide additional support for the findings of this study.

While our analysis was a quantitative study leveraging an organization's internal MIS, case-study limitations do apply to our analysis. Anomalies particular to the organization, either through its unique cultural or managerial control environments, were not addressed in our analysis. Differences in those among organizations could impact the generalizability of these results more widely. We suggest that future research be conducted across multiple organizations to further validate the findings we have presented in this paper.

While our paper focused primarily on Decision Aid Complexity, the expansion of understanding of the influencing factors of Decision Aid Reliance given the deluge of data and processing advances that have been made in the past decade makes continued research in this area relevant and timely. While we have illustrated the influence of one type of technology's impact on Decision Aid Reliance, there are vast categories of technologies and decision solutions that need to be studied, so that decision aid users and designers have a better understanding of how and why outcomes are influenced. We used embedded agents in this study. However, evaluating classes of decision aids such as those using Random Forest, neural networks, and unsupervised algorithms should be evaluated in future research.

The Theory of Technology Dominance explicitly relies upon the premise that a user has accepted the decision aid technology consistent with Davis' (1989) Technology Acceptance Model. TTD, however, has not considered extensions of influencing aspects of technology acceptance, such as those found in the Unified Theory of Technology Acceptance (UTAUT), as presented by Venkatesh (2003). Unlike the original TAM, UTAUT considers social factors such as how individuals perceive how others will interpret their actions to use a technology or how command and control structures influence technology acceptance (Venkatesh, Morris, Davis, & Davis, 2003). In the context of our research, these factors vary across industries and firms. While the organization being evaluated in the research presented here is in a very mature, highly regulated environment, we could anticipate results to be different in early-entry industries with less regulation. There is much left to be studied in this space, and the nexus of the organization, how it enables decision-making in a Big Data world, and the technology which supports them is wide open for continuing research.

Finally, as previously discussed, our finding suggests that our model is generalizable to industries broader than only the financial services sector the organization that we analyzed. Our experiment included participants who represent both a transactional and relational engagement strategy to the firm's customers. Extant literature is limited surrounding this particular relationship and presents another opportunity to further understand the influencing aspects of Decision Aid Reliance. There is much left to be studied in this space. The nexus of the organization, how it enables decision-making in a Big Data world, and the technology which supports them remains fertile with research opportunities.

CHAPTER 7: CONCLUSION

Our research motivation set out to determine 1) whether or not there was a relationship between the complexity of an intelligent Decision Aid and users relying upon its recommendation and 2) if the propositions presented by Arnold & Sutton (1998) were moderated by the complexity of the Decision Aid. Our results were obtained by evaluating an intelligent decision aid used within a practice-setting. The context, along with the significant size of observations (n=4991) our findings add to the generalizability of the Theory of Technology Dominance and extend it by finding empirical support for Decision Aid Complexity as having both a direct-effect on Decision Aid Reliance as well as a negatively moderating impact on the relationship between Task Experience and Decision Aid Reliance. To our knowledge, we are the first to analyze the complexity of the decision aid as an influencing factor in user reliance behavior. Furthermore, to our knowledge, we are one of the very few studies to have empirically evaluated any of the propositions of the Theory of Technology Dominance within a practice setting outside of the Audit or Accounting domains.

This expands our understanding of the factors that contribute to user reliance. Specifically, we contribute an additional dimension to the Theory of Technology Dominance by finding significant support for the concept of Decision Aid Complexity as an influencing aspect of Decision Aid Reliance.

In conclusion, we find support for extending Arnold & Sutton's (1998) Theory of Technology Dominance to include Decision Aid Complexity as a potential influencing factor in Decision Aid Reliance. Specifically, we find that Decision Aid Complexity influences the previously proposed relationship between Decision Aid Reliance and Task

Experience. How decision aids are designed influences reliance behavior and should be considered by users, designers, and organizations that deploy them to fully understand the potential impact of design choices on reliance outcomes.

APPENDIX

Table 10: Descriptive Statistics

Variable	Obs.	Mean	Std. Dev	Min	Max
Rely	4,941	0.896	0.305	0.000	1.000
Task Experience		2.824	1.268	1.000	5.000
Cognitive Fit		37.193	43.716	1.000	165.000
IDA Familiarity		49.814	46.028	1.000	174.000
Complexity		23.969	8.738	3.000	39.000
Task Experience*Complexity		0.005	1.004	-4.118	3.452
Cognitive Fit*Complexity		0.469	1.033	-1.424	4.479
IDA Familiarity*Complexity		0.380	1.034	-6.475	4.642

Table 11: Covariance Matrix

Variable	1	2	3	4	5	6	7
Task Experience	1.000						
Cognitive Fit	-0.145	1.000					
IDA Familiarity	-0.154	0.918	1.000				
Complexity	0.006	0.469	0.380	1.000			
Task Experience*Complexity	-0.043	-0.014	0.012	-0.078	1.000		
Cognitive Fit*Complexity	-0.013	0.663	0.656	-0.027	-0.066	1.000	
IDA Familiarity*Complexity	0.012	0.656	0.572	0.084	-0.094	0.855	1.000

Table 12: Intercept Only (Null) Model

Intercept-Only (NULL MODEL)				
AIC	**BIC**	**logLik**	**deviance**	**df.resid**
2521.9	2541.5	-1258	2515.9	4938
Scaled residuals:				
Min	1Q	Median	3Q	Max
-9.108	0.11	0.1402	0.2859	1.8484
Random effects:				
Groups	**Name**	**Variance**	**Std.Dev.**	
USERID	(Intercept)	3.4068	1.8458	
MODELID	(Intercept)	0.5753	0.7585	
Number of Obs: 4941, Groups; USERID: 250, MODELID, 27				
Fixed effects:				
Estimate	Std. Error	z	value	Pr(>|z|)
(Intercept)	3.6878	0.2979	12.38	<2e-16

Table 13: Random Intercept Model

Random Intercept Model				
AIC	**BIC**	**logLik**	**deviance**	**df.resid**
2508.5	2554	-1247.2	2494.5	4934
Scaled residuals:				
Min	1Q	Median	3Q	Max
-8.7984	0.0976	0.1441	0.2854	1.9507
Random effects:				
Groups	**Name**	**Variance**	**Std.Dev.**	
USERID	(Intercept)	3.0044	1.7333	
MODELID	(Intercept)	0.5629	0.7503	
Number of Obs: 4941, Groups; USERID: 250, MODELID, 27				
Fixed effects				
Estimate	Std. Error	z	value	Pr(>|z|)
(Intercept)	2.4255936	0.681346	3.56	0.000371***
Complexity	0.0095263	0.0255447	0.373	0.709204
Experience	0.2740332	0.1129849	2.425	0.015292*
IDA Familiarity	0.0005495	0.0059986	0.092	0.92701
Cognitive Fit	0.0140101	0.0037724	3.714	0.000204***

Table 14: Random Slope, Random Intercepts

Random Slopes, Random Intercept				
AIC	**BIC**	**logLik**	**deviance**	**df.resid**
2511.5	2570.1	-1246.8	2493.5	4932
Scaled residuals				
Min	1Q	Median	3Q	Max
-9.5903	0.1043	0.1514	0.2812	2.072
Random effects:				
Groups	Name	Variance	Std.Dev.	Corr
USERID	(Intercept)	8.645714	2.94036	
Experience	0.647348	0.80458	-0.88	
MODELID	(Intercept)	0.236261	0.48607	
Complexity	0.001561	0.03951	-0.77	
Number of Obs:	4941,	groups:	USERID, 250;	MODELID, 27
Fixed effects:				
Estimate	Std. Error	z	value	Pr(>|z|)
(Intercept)	3.438092	0.322048	10.676	2E-16
IDA_Familiarity	-0.00233	0.005637	-0.413	0.679351
Cognitive_Fit	0.014308	0.003797	3.768	0.000165

Table 15: Complexity Score by Experience Level

Experience Level	**Mean Complexity Score (Raw)**
1	23.33
2	23.58
3	26.95
4	22.12
5	25.70

Table 16: Tukey's Honest Differences Test

Dependent Variable: Task Complexity							
	Task Experience(I)	Task Experience	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	1	2	-0.254	0.377	0.962	-1.28	0.78
		3	-3.619*	0.376	0.000	-4.65	-2.59
		4	1.206*	0.348	0.005	0.26	2.16
		5	-2.371*	0.517	0.000	-3.78	-0.96
	2	1	0.254	0.377	0.962	-0.78	1.28
		3	-3.365*	0.375	0.000	-4.39	-2.34
		4	1.460*	0.347	0.000	0.51	2.41
		5	-2.117*	0.516	0.000	-3.53	-0.71
	3	1	3.619*	0.376	0.000	2.59	4.65
		2	3.365*	0.375	0.000	2.34	4.39
		4	4.825*	0.347	0.000	3.88	5.77
		5	1.249	0.516	0.110	-0.16	2.66
	4	1	-1.206*	0.348	0.005	-2.16	-0.26
		2	-1.460*	0.347	0.000	-2.41	-0.51
		3	-4.825*	0.347	0.000	-5.77	-3.88
		5	-3.576*	0.496	0.000	-4.93	-2.22
	5	1	2.371*	0.517	0.000	0.96	3.78
		2	2.117*	0.516	0.000	0.71	3.53
		3	-1.249	0.516	0.110	-2.66	0.16
		4	3.576*	0.496	0.000	2.22	4.93

* The mean difference is significant at the 0.05 level.

Table 17: Power Chi-Squares Test

> pwr.chisq.test
w = 0.5
N = 250
df = 249
sig. level = 0.05
power = 0.8155161
NOTE: N is the number of observations

REFERENCES

Aguinis, H., Gottfredson, R. K., & Culpepper, S. A. (2013). Best-practice recommendations for estimating cross-level interaction effects using multilevel modeling. *Journal of Management, 39*(6), 1490-1528.

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In *Action control* (pp. 11-39): Springer.

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes, 50*(2), 179-211.

Appelbaum, S. H., & Hare, A. (1996). Self-efficacy as a mediator of goal setting and performance. *Journal of Managerial Psychology, 11*(3), 33-47.

Arnold, V., Collier, P. A., Leech, S. A., & Sutton, S. G. (2004). Impact of intelligent decision aids on expert and novice decision-makers' judgments. *Accounting & Finance, 44*(1), 1-26.

Arnold, V., & Sutton, S. G. (1998). The theory of technology dominance: Understanding the impact of intelligent decision aids on decision maker's judgments. *Advances in accounting behavioral research, 1*(3), 175-194.

Ashton, R. H. J. J. o. A. R. (1990). Pressure and performance in accounting decision settings: Paradoxical effects of incentives, feedback, and justification. *28*, 148-180.

Bacic, D., & Henry, R. (2018). Task-Representation Fit's Impact on Cognitive Effort in the Context of Decision Timeliness and Accuracy: A Cognitive Fit Perspective. *AIS Transactions on Human-Computer Interaction, 10*(3), 164-187.

Baldwin, A. A., Brown, C. E., & Trinkle, B. S. (2006). Opportunities for artificial intelligence development in the accounting domain: the case for auditing. *Intelligent Systems in Accounting, Finance & Management: International Journal, 14*(3), 77-86.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Boatsman, J. R., Moeckel, C., Pei, B. K. J. O. B., & Processes, H. D. (1997). The effects of decision consequences on auditors' reliance on decision aids in audit planning. *71*(2), 211-247.

Bozarth, C. C., Warsing, D. P., Flynn, B. B., & Flynn, E. J. (2009). The impact of supply chain complexity on manufacturing plant performance. *Journal of Operations Management, 27*(1), 78-93.

Breslow, N. E. (1996). Statistics in epidemiology: the case-control study. *Journal of the American Statistical Association, 91*(433), 14-28.

Brooks, R. J., & Tobias, A. M. (1996). Choosing the best model: Level of detail, complexity, and model performance. *Mathematical and computer modelling, 24*(4), 1-14.

Brown, D., & Jones, D. R. (1998). Factors that influence reliance on decision aids: A model and an experiment. *Journal of Information Systems, 12*(2).

Brown, D. L., & Jones, D. R. J. J. o. I. S. (1998). Factors that influence reliance on decision aids: a model and an experiment. *12*(2), 75-94.

Buttle, F. J. R. M. T., & Practice, P. C. P., London. (1996). Relationship marketing. 1-16.

Casti, J. L. (1979). *Connectivity, complexity and catastrophe in large-scale systems* (Vol. 7): John Wiley & Sons.

Chui, M., Malhotra, S. (2018). *AI adoption advances, but foundational barriers remain*. Retrieved from New York, NY: https://www.mckinsey.com/featuredinsights/artificial-intelligence/ai-adoption-advances-but-foundational-barriers-remain

Coviello, N. E., & Brodie, R. J. J. J. o. S. M. (1998). From transaction to relationship marketing: an investigation of managerial perceptions and practices. *6*(3), 171-186.

Danielsson, J., James, K. R., Valenzuela, M., & Zer, I. (2016). Model risk of risk models. *Journal of Financial Stability, 23*, 79-91.

Davern, M. J. (2007). Towards a unified theory of fit: task, technology and individual. *Information systems foundations: Theory, representation and reality*, 49-69.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340.

Dennis, A. R., & Carte, T. A. (1998). Using geographical information systems for decision making: Extending cognitive fit theory to map-based presentations. *Information systems research, 9*(2), 194-203.

Deshmukh, A. V., Talavage, J. J., & Barash, M. M. (1998). Complexity in manufacturing systems, Part 1: Analysis of static complexity. *IIE transactions, 30*(7), 645-655.

Dunn, C., & Grabski, S. (2001). An investigation of localization as an element of cognitive fit in accounting model representations. *Decision Sciences, 32*(1), 55-94.

Edmonds, B. (1995). What is Complexity?-The philosophy of complexity per se with application to some examples in evolution. In *The evolution of complexity*: Kluwer, Dordrecht.

Eining, M. M., Jones, D. R., Loebbecke, J. K. J. A. A. J. o. P., & Theory. (1997). Reliance on decision aids: An examination of auditors' assessment of management fraud. *16*(2).

Fishbein, M., & Ajzen, I. (1977). Belief, attitude, intention, and behavior: An introduction to theory and research.

Flood, R. L., & Carson, E. R. (2013). *Dealing with complexity: an introduction to the theory and application of systems science*: Springer Science & Business Media.

Fogel, A. L., & Kvedar, J. C. (2017). Benefits and risks of machine learning decision support systems. *Jama, 318*(23), 2356-2356.

Ford, J. K., Quiñones, M. A., Sego, D. J., & Sorra, J. S. (1992). Factors affecting the opportunity to perform trained tasks on the job. *Personnel psychology, 45*(3), 511-527.

French, B., Duenser, A., & Heathcote, A. (2018). Trust in Automation. In: CSIRO Report EP184082.

Galletta, D. F., Everard, A., Durcikova, A., & Jones, B. (2003). *Cognitive fit and an intelligent agent for a word processor: should users take all that advice?* Paper presented at the 36th Annual Hawaii International Conference on System Sciences, 2003.

Gist, M. E., & Mitchell, T. R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management review, 17*(2), 183-211.

Goodhue, D. L., & Thompson, R. L. (1995). Task-technology fit and individual performance. *MIS quarterly*, 213-236.

Hampton, C. (2005). Determinants of reliance: An empirical test of the theory of technology dominance. *International Journal of Accounting Information Systems, 6*(4), 217-240.

Heck, R. H., & Thomas, S. L. (2015). *An introduction to multilevel modeling techniques: MLM and SEM approaches using Mplus*: Routledge.

Hoetker, G. (2007). The use of logit and probit models in strategic management research: Critical issues. *Strategic management journal, 28*(4), 331-343.

Hong, W., Thong, J. Y., & Tam, K. Y. (2004). The effects of information format and shopping task on consumers' online shopping behavior: A cognitive fit perspective. *Journal of Management Information Systems, 21*(3), 149-184.

Huang, P., Kankanhalli, A., Kyriakou, H., & Sabherwal, R. (2018). Research Curation on Knowledge Management.

Introna, L., & Wood, D. (2004). Picturing algorithmic surveillance: The politics of facial recognition systems. *Surveillance & Society, 2*(2/3), 177-198.

Jensen, M. L., Lowry, P. B., Burgoon, J. K., & Nunamaker, J. F. (2010). Technology dominance in complex decision making: The case of aided credibility assessment. *Journal of Management Information Systems, 27*(1), 175-202.

Kamis, A., Koufaris, M., & Stern, T. (2008). Using an attribute-based decision support system for user-customized products online: an experimental investigation. *MIS quarterly*, 159-177.

Kaplan, S. E., Reneau, J. H., & Whitecotton, S. (2001). The effects of predictive ability information, locus of control, and decision maker involvement on decision aid reliance. *Journal of Behavioral Decision Making, 14*(1), 35-50.

Kenny, D. A., & Judd, C. M. (1996). A general procedure for the estimation of interdependence. *Psychological bulletin, 119*(1), 138.

Kim, S. H., Mukhopadhyay, T., & Kraut, R. E. (2016). When Does Repository KMS Use Lift Performance? The Role of Alternative Knowledge Sources and Task Environments. *MIS quarterly, 40*(1), 133-156.

King, G., & Zeng, L. (2001). Logistic regression in rare events data. *Political analysis, 9*(2), 137-163.

Kistruck, G. M., Morris, S. S., Webb, J. W., & Stevens, C. E. (2015). The importance of client heterogeneity in predicting make-or-buy decisions. *Journal of Operations Management, 33*, 97-110.

Klein, G. (2008). Naturalistic decision making. *Human Factors, 50*(3), 456-460.

Kletke, M. G., Mackay, J. M., Barr, S. H., & Jones, B. (2001). Creativity in the organization: the role of individual creative problem solving and computer support. *International Journal of Human-Computer Studies, 55*(3), 217-237.

Kumar, V., Bohling, T. R., & Ladda, R. N. J. I. M. M. (2003). Antecedents and consequences of relationship intention: Implications for transaction and relationship marketing. *32*(8), 667-676.

Lampel, J., Shamsie, J., & Shapira, Z. (2009). Experiencing the improbable: Rare events and organizational learning. *Organization Science, 20*(5), 835-845.

Lance, C. E., Hedge, J. W., & Alley, W. E. (1989). Joint relationships of task proficiency with aptitude, experience, and task difficulty: A cross-level, interactional study. *Human Performance, 2*(4), 249-272.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174.

Lee, J., & See, K. (2004). Trust in technology: Designing for appropriate reliance. *Human Factors, 46*(1), 50-80.

Lin, M., Lucas Jr, H. C., & Shmueli, G. (2013). Research commentary—too big to fail: large samples and the p-value problem. *Information systems research, 24*(4), 906-917.

Mackay, J., Barr, S., & Kletke, M. (1992). An Empirical Investigation of the Effects of Decision Aids on Problem-Solving Processes. *Decision Sciences, 23*(3), 648.

Madden, T. J., Ellen, P. S., & Ajzen, I. (1992). A comparison of the theory of planned behavior and the theory of reasoned action. *Personality and social psychology Bulletin, 18*(1), 3-9.

Makui, A., & Aryanezhad, M.-B. (2003). A new method for measuring the static complexity in manufacturing. *Journal of the Operational Research Society, 54*(5), 555-557.

Mălăescu, I., & Sutton, S. G. (2015). The effects of decision aid structural restrictiveness on cognitive load, perceived usefulness, and reuse intentions. *International Journal of Accounting Information Systems, 17,* 16-36.

Mascha, M. F., & Smedley, G. (2007). Can computerized decision aids do "damage"? A case for tailoring feedback and task complexity based on task experience. *International Journal of Accounting Information Systems, 8*(2), 73-91.

Masselli, J. J., Ricketts, R. C., Arnold, V., & Sutton, S. G. (2002). The Impact of Embedded Intelligent Agents on Tax-Reporting Decisions. *The Journal of the American Taxation Association*, *24*(2), 60-78.

Meyer, M. H., & Curley, K. F. (1991). An applied framework for classifying the complexity of knowledge-based systems. *MIS quarterly*, 455-472.

Morrison, J. G., Kelly, R. T., Moore, R. A., & Hutchins, S. G. (1998). Implications of decision-making research for decision support and displays.

Muir, B. M. (1987). Trust between humans and machines, and the design of decision aids. *International journal of man-machine studies, 27*(5-6), 527-539.

O'Leary, D. E. (2003). Auditor environmental assessments. *International Journal of Accounting Information Systems, 4*(4), 275-294.

Parkes, A. (2009). Persuasive decision support: Improving reliance on decision support systems. *PACIS 2009 Proceedings*, 11.

Parkes, A. (2017). The effect of individual and task characteristics on decision aid reliance. *Behaviour & Information Technology*, *36*(2), 165-177.

Peugh, J. L. J. J. o. s. p. (2010). A practical guide to multilevel modeling. *48*(1), 85-112.

Quińones, M. A., Ford, J. K., & Teachout, M. S. (1995). The relationship between work experience and job performance: A conceptual and meta-analytic review. *Personnel psychology*, *48*(4), 887-910.

Rhodes, R. E., Courneya, K. S., Jones, L. W. J. P. o. S., & Exercise. (2004). Personality and social cognitive influences on exercise behavior: adding the activity trait to the theory of planned behavior. *5*(3), 243-254.

Rose, J. (2002). Behavioral decision aid research: Decision aid use and effects. *Researching accounting as an information systems discipline, 5*, 111-134.

Senge, P. M. (1990). The art and practice of the learning organization. In: New York: Doubleday.

Seow, P.-S. (2011). The effects of decision aid structural restrictiveness on decision-making outcomes. *International Journal of Accounting Information Systems, 12*(1), 40-56.

Shaft, T. M., & Vessey, I. (2006). The role of cognitive fit in the relationship between software comprehension and modification. *MIS quarterly*, 29-55.

Sheth, J. N., & Parvatiyar, A. J. I. b. r. (1995). The evolution of relationship marketing. *4*(4), 397-418.

Simon, H. A. (1957). Models of man; social and rational.

Sinha, A. P., & Vessey, I. (1992). Cognitive fit: an empirical study of recursion and iteration. *IEEE Transactions on Software Engineering*(5), 368-379.

Sniehotta, F. F., Presseau, J., & Araújo-Soares, V. (2014). Time to retire the theory of planned behaviour. In: Taylor & Francis.

Snijders, T. A. J. E. o. s. i. b. s. (2005). Power and sample size in multilevel linear models.

Sutton, S. G., Holt, M., & Arnold, V. (2016). "The reports of my death are greatly exaggerated"—Artificial intelligence research in accounting. *International Journal of Accounting Information Systems, 22*, 60-73.

Tan, C. H., Teo, H. H., & Benbasat, I. (2010). Assessing Screening and Evaluation Decision Support Systems: A Resource-Matching Approach. *Information systems research, 21*(2), 305-326.

Tasca, G. A., Illing, V., Joyce, A. S., & Ogrodniczuk, J. S. (2009). Three-level multilevel growth models for nested change data: A guide for group treatment researchers. *Psychotherapy Research, 19*(4-5), 453-461.

Todd, P., & Benbasat, I. (1992). The use of information in decision making: an experimental investigation of the impact of computer-based decision aids. *MIS quarterly*, 373-393.

Tukey, J. W. (1977). *Exploratory data analysis* (Vol. 2): Reading, Mass.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478.

Vessey, I. (1991). Cognitive fit: A theory-based analysis of the graphs versus tables literature. *Decision Sciences, 22*(2), 219-240.

Vessey, I. (1994). The effect of information presentation on decision making: A cost-benefit analysis. *Information & Management, 27*(2), 103-119.

Vessey, I., Zhang, P., & Galletta, D. (2006). The theory of cognitive fit. *Human-computer interaction and management information systems: Foundations, 5*, 141-183.

Walsh, S., Gilmore, A., & Carson, D. J. I. J. o. B. M. (2004). Managing and implementing simultaneous transaction and relationship marketing.

Wang, L., Jamieson, G. A., & Hollands, J. G. J. H. f. (2009). Trust and reliance on an automated combat identification system. *51*(3), 281-291.

Wheeler, P. R., & Jones, D. R. J. J. o. I. S. (2003). The effects of exclusive user choice of decision aid features on decision making. *17*(1), 63-83.

Whitecotton, S. M., Butler, S. A., & Houston, R. W. J. B. R. i. A. (1998). Influencing Decision Aid Reliance Through Involvement in Information Choice/COMMENTARY ON Influencing Decision Aid Reliance Through Involvement in Information Choice. *10*, 182.

Wiersema, M. F., & Bowen, H. P. (2009). The use of limited dependent variable techniques in strategy research: Issues and methods. *Strategic management journal, 30*(6), 679-692.

Wood, R., & Bandura, A. (1989). Social cognitive theory of organizational management. *Academy of Management review, 14*(3), 361-384.

Yaokumah, W., & Amponsah, R. A. (2019). Examining the contributing factors for cloud computing adoption in a developing country. In *Cloud Security: Concepts, Methodologies, Tools, and Applications* (pp. 1663-1685): IGI Global.

Yates, F. E. (1978). Complexity and the limits to knowledge. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *235*(5), R201-R204.