The impact of complexity on the effectiveness of intuition in decision making

by

Danéel van Eck

14192226

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at the Gordon Institute of Business Science, University of Pretoria.

(7 July 2018)

Supervisor Name: Dr Karen Luyt

Abstract

This research investigated the impact of complexity on the effectiveness of intuition in decision making. Empirical experimental research was conducted to examine the relationship between the use of intuition and decision accuracy, and how this relationship could be moderated by individual differences such as faith in intuition and cognitive reflection. Software purchase decisions were used as the context for the study because of the inherent complexity of the decision and anecdotal evidence that rational techniques are not always used. Dual-process theory was used as the theoretical lens for the research, which was conducted using an experimental manipulation of the use of intuition. The research made an original contribution by investigating complexity as a moderator, by examining the role of individual differences as other researchers have proposed, and by utilising bias measures as a proxy for decision accuracy in order to avoid the normatively correct choice measures that have been highlighted for their limitations. The results showed that intuitive decisions were less accurate in the software selection context, and that complexity does moderate this effect on decision accuracy, with intuition showing lower accuracy in simple decisions, but similar results in complex decisions. Faith in intuition and cognitive reflection did not show interaction effects, but there was a significant direct positive relationship between cognitive reflection and decision accuracy. The practical implications of this research include prescriptions for the conditions under which intuitive decisions can be appropriate.

Declaration Regarding Plagiarism

I (full names & surname):	Danéel van Eck
Student number:	14192226
Degree	Doctor of Philosophy
Title	The impact of complexity on the effectiveness of intuition in decision making

1. I understand what plagiarism entails and am aware of the university's policy in this regard.

2. I declare that this assignment is my own, original work. Where someone else's work was used (whether from a printed source, the Internet or any other source) due acknowledgement was given and reference was made according to departmental requirements.

3. I did not copy and paste any information directly from an electronic source (e.g., a web page, electronic journal article or CD ROM) into this document.

4. I did not make use of another student's previous work and submitted it as my own.

5. I did not allow and will not allow anyone to copy my work with the intention of presenting it as his/her own work.

Danéel van Eck 7 July 2018

_____ _____

Acknowledgements

Thank you to everyone who has supported me through this process, and to everyone who always asked how it was going and offered encouragement along the way. Thank you to GIBS and the doctoral committee for continuously pushing us to succeed, while being always available to help. Thank you also for all the additional support that you provided by way of coaching, teambuilding, and fostering the development of supportive relationships with our peers. Vivienne Spooner and Lerushka Barath you were always supportive and I value our friendships that developed through this process. Thank you in particular to Prof Helena Barnard who pushed me harder than I would have pushed myself, and was always available on short notice when I needed advice.

Karen Luyt, my supervisor, thank you so much for all of your time and support. You were always encouraging even when I wasn't convinced, and you were always available, even late at night, and while you were on holiday - when I needed help. Thank you also for the countless reviews of my work and tireless suggestions for improvement.

Cindy van Eck, my wife and better half – words cannot express my gratitude. Thank you so much for all of your support and encouragement through this process. I'm sure you had no idea what I was signing you up for when we started this. I would not have been able to do this without you. You have had to deal with me missing family functions, being frustrated, and most importantly with me not being able to give 100% of my attention to our children during this last part. Thank you so much. To Søren and Jensen, our twin boys, thank you for putting up with my limited attention during the first 15 months of your lives. Your smiles in the morning helped push me to complete this. The more I watch you learn and grow each day, the more I realise that experimentation is the most natural and explainable way of learning about the world around us.

Table of Contents

List of Figures

List of Tables

Chapter 1 – Introduction

1.1 Background

"Clearly there are occasions when you should use your head instead of the formula. But which occasions they are is most emphatically not clear." (Meehl, 1973, p.4).

Business is becoming increasingly complex and every day managers are required to make complex decisions with limited time and information. Intuitive decisions can be made quickly whilst simultaneously bypassing in-depth analysis, and could therefore offer a solution (Dane & Pratt, 2007; Isenberg, 1984; Simon, 1987; Wierzbicki, 1997).

Business management literature provides many examples of managers and scientists who have trusted their "gut-feel" with considerable success (Hayashi, 2001; Isenberg, 1984; Simon, 1987; Sinclair & Ashkanasy, 2005). However, there is also a large portion of literature which demonstrates the fallibility of these same decisions. This has resulted in two strong but conflicting prescriptions – one from a group that argues in favour of intuitive decisions and praises experts and their ability to make successful intuitive decisions (e.g. Klein, 1993; 2008), and another that warns that intuition should never be trusted because of the bias and inaccuracy that will result (e.g. Kahneman, 2011; Kahneman & Frederick, 2002).

Dual-process theory offers a useful lens for thinking and research in this area. The theory describes two independent systems of thought: an intuitive system which is fast and effortless, and a rational system which is slow and effortful (Kahneman & Frederick, 2002; Stanovich, 1999; Stanovich & West, 2000). There has been significant support in the literature for various forms of this two-system view (Epstein, 1994; Evans & Stanovich, 2013; Kahneman & Frederick, 2002; Stanovich, 1999; Stanovich & West, 2000), and despite various criticisms (Evans & Stanovich, 2013), it continues to be the dominant view and continues to garner support (Usher, Russo, Weyers, Brauner, & Zakay, 2011).

Using this lens, some studies have demonstrated particular cases where intuition can outperform rational deliberation in decision making (Dijksterhuis, 2004; Dijksterhuis & Nordgren, 2006; Usher et al., 2011). Some of these studies have pointed to complexity as a potential reason, citing research that rational deliberation is limited by working memory capacity, and assuming that intuitive decision making would not be subject to these limitations (Dijksterhuis, 2004; Meszaros, 2007). However, these findings have raised controversy and have been criticised because of methodological issues and problems with replication in other contexts (Usher et al., 2011). One of the criticisms involves a lack of acknowledgement of individual differences in decision making approach (Usher et al., 2011). Epstein (1994) argued that some individuals are more predisposed toward intuitive thinking than others, and later research has shown that these predispositions to intuition were related to an increase in bias (Keller & Bless, 2009).

Software purchase decisions offer a useful context for this research because of the complexity of the decisions (Jadhav & Sonar, 2011) and anecdotal indications that standard recommended rational techniques are not being followed in practice. Personal communication (9 May 2014) with the Chief Information Officer of a large information technology company revealed that his process often involves examining all the information and then going with his “gut”, despite the availability of and prescription for formal rational models such as the weighted scoring method and analytic hierarchy process (Jadhav & Sonar, 2009). The software selection literature in particular points to intuition only as an example of an unacceptable decision method (Carney & Wallnau, 1998) or as a potential source of bias (Jamieson & Hyland, 2006), despite reports that managers often make software selection decisions without following rational processes (Bannister & Remenyi, 2000).

These tensions in the literature supported the need for further research to be conducted to determine the role of intuition in business decisions, and software selection and evaluation provided a particularly relevant context.

1.2 Problem Statement

In decision making literature there are contradictory views as to the effectiveness of intuition in business decision making. Since there is empirical research on both sides showing that either rational or intuitive decisions were superior, there must be other conditions that explain these differences. Some of the conditions eluded to in the literature included decision complexity and individual differences. Therefore these conditions need further investigation to move our understanding of the rational versus intuitive debate forward.

In a practical arena, senior decision makers are often seen making intuitive decisions when selecting software. The literature on software selection however prescribes only rational techniques for these multi-attribute decision making problems, and there is a scarcity of information that addresses intuitive approaches other than warning against them. Therefore, software selection offers a useful context for testing intuitive and rational decisions in order to further the current literature in this area, in addition to general literature regarding the effectiveness of intuitive decision making.

Solving this problem will aid in the understanding of the conditions which would promote, or even allow, the use of intuition instead of rational decision approaches.

1.3 Purpose Statement

This research aimed to explain the impact of complexity on the accuracy of intuitive decisions, by testing moderators such as complexity and individual differences in order to move the understanding of the effectiveness of intuition in business decision making forward. Within the positivist ontology, an objectivist stance was taken for the research and as such quantitative principles were used for what may seem to be a qualitative phenomenon. The research applied experimental survey instruments to experts in a sample of companies in South Africa. The results were analysed quantitatively to determine the relationship between faith in intuition, cognitive reflection, use of intuition, complexity, and decision accuracy.

1.4 Research question

The overall research question was stated as follows:

> "What is the impact of complexity, and individual differences, on the effectiveness of intuitive decisions?"

1.5 Research objectives

The objectives to support the research question were:

1. Determine if there is a relationship between use of intuition and decision accuracy, and whether intuitive decisions are more, or less accurate.
2. Determine whether decision complexity moderates the relationship between use of intuition and decision accuracy.
3. Determine whether faith in intuition moderates the relationship between use of intuition and decision accuracy.
4. Determine whether cognitive reflection moderates the relationship between use of intuition and decision accuracy.

1.6 Scope and Definitions

1.6.1 Context

Within the area of business decision making, the present research focused on software evaluation decisions. There are many potential criteria which could be used in evaluation, such as the fit to known requirements, vendor reliability, and price – and each of these could have different weightings. However, the present research focused on the decision-making used to evaluate those criteria, and not on the specific criteria themselves.

1.6.2 Scope delimitations

This section will address various theories which could be used to address different aspects of the overall problem within the selected context, but which were excluded from the scope of the present research because of the reasons discussed below.

The Judgment and Decision Making literature covers many aspects of judgment and choice. Literature on consumer choice, consumer behaviour, and consumer psychology also address theories of choice and preference construction which arise in choice decisions. These include choice effects, such as the attraction effect where having an asymmetrically dominant option can aid in choosing that option, and adding an inferior option can create a relatively dominant option by comparison and prevent delays in choice (Dhar & Ghorlin, 2013). The present research focused on the evaluation and judgment aspects of software selection and not the final choice, thereby limiting the scope to exclude theories of preference construction and choice.

Social contagion (Levy & Nail, 1993; Marsden, 1998) could explain some non-rational decisions and, because of the way these decisions are made, the effectiveness of these decisions might not be comparable to decisions made using expert judgement or other forms of intuition. In some cases, choosing the same software that other people did could be good because it worked for them, but in other cases it might not be suitable for the requirements, or the contagion may have started for other reasons. Because the present research only covered the judgment aspects of the decision which occur before choice, and the outcome was measured as decision accuracy, and because the methodology involved a fictitious case, social contagion was excluded.

Bounded rationality explains decision-making inside complex situations with large amounts of data and insufficient mental capacity or time to examine all of the options (Simon, 1979). These conditions often result in oversimplification and the use of heuristics in achieving a solution (Sadler-Smith, 2012). Simon (1979) proposes that while it is believed that perfectly rational individuals should optimise for the highest utility in each decision, limitations in access to information and conflict with other actors, leads to satisficing behaviour where minimum aspiration levels are attained. Using bounded rationality theory, Schwartz et al.

(2002) showed through empirical research, that some individuals have a tendency to optimise when making decisions (maximising), leading to increased levels of information-seeking and social comparison, and that others have a tendency to use shortcuts and settle for an option that passes a threshold of acceptability (satisficing). The present research did not focus on the search process where options are discovered before being evaluated, and instead focused on the evaluation of a particular option, so that maximising and satisficing were not relevant in this context.

The success of the outcome could be considered by some as the most important criteria for determining decision success, and success of information systems has been studied in many forms. DeLone and McLean (1992) reviewed studies over the previous decade and found that success had been elusive to define, and they proposed system quality, information quality, use, user satisfaction, individual impact, and organisational impact as outcome variables for success. Some of the literature focuses specifically on external attributes such as the match between system characteristics and user requirements (Jadhav & Sonar, 2011). The Technology Acceptance Model (TAM) was created to address the mediating role of perceived ease of use and perceived usefulness on the relationship between systems characteristics and the probability of system use as an indicator of success (Legris, Ingham, & Collerette, 2003), however this would not be applicable in the context defined here.

Outside of TAM, but often referencing it, popular dependant variables for measuring system success have included intention, self-reported use, choice, and actual use (Venkatesh, Davis, & Morris, 2007). More recently, Brown, Venkatesh, and Goyal (2014) in their comparison of competing models of expectation confirmation used behavioural intention, use, and system satisfaction as measures of system success. However, there are issues with temporal variances, where these variables and measured success can change over time. Aside from all these issues with measuring success, a larger problem is that of bias. Judging a decision by its

outcome, such as the success of an implementation, demonstrates outcome bias, since sometimes good decisions can still have bad outcomes (Baron & Hershey, 1988). Because of this, the research focused on the accuracy of the decision and did not consider the success of the selection and implementation, and the experimental case supported this.

1.6.3 Population

Chief Information Officers and other senior Information Technology decision makers often make final selection and evaluation decisions. Because of their position in the organisation, and with their extensive previous experience, it would be possible for them to have a gut feel about which option to go with. Because of this, the research was conducted in businesses where information is available to enable engagement of a senior Information Technology decision maker. Inclusion criteria included companies with senior level Information Technology staff such as a Chief Information Officer or Chief Technology Officer or an equivalent Information Technology decision-maker.

1.6.4 Definitions

1.6.4.1 Software Selection and Evaluation

Software selection and evaluation was defined as the selection and evaluation of commercially available off-the-shelf software products for purchase. These may require configuration or some customisation but do not need to be developed – analysing development would bring in other complexities. The delivery method was not specified and could include physical, downloaded, and cloud-delivered products.

1.6.4.2 Intuition

The general definition for *Intuition* was taken from Epstein (2010, p. 296) as "a sense of knowing without knowing how one knows".

1.6.4.3 Intuitive judgment

The specific definition for intuitive judgement will be introduced and explained in the literature review as a judgment that is expressed when System 2 endorses System 1's conclusion (Kahneman & Frederick, 2002). This is the definition that was used for intuitive decisions in the chosen context.

1.7 Importance and benefits of the research conducted

1.7.1 Academic contribution

This section summarises the areas of academic contribution of the current research and indicates where important studies will be furthered.

The literature regarding intuition in business decisions is split into arguments in favour of (e.g. Klein, 1993, 2008), and against (e.g. Kahneman & Frederick, 2002; Kahneman, 2011), the use of intuition and intuitive judgment. This tension between the two camps requires further research in specific contexts in order to understand the impact in those contexts, and this provided support for the present research into intuition in the software selection context. This work followed Kahneman and Frederick (2002) by using dual-process theory as the framework for the analysis. The results of the research conducted contribute to the body of knowledge on intuitive decision making in business in general and also contribute further to this debate.

Research on the role of complexity has provided mixed results with some arguing that complexity plays a role such as with unconscious thought theory research (Dijksterhuis, 2004; Dijksterhuis & Nordgren, 2006; Meszaros, 2007), while others have questioned the validity of this research (Usher et al., 2011) or failed to successfully replicate earlier findings (Acker, 2008; Calvillo & Penaloza, 2009). This work followed a similar vein to Dijksterhuis (2004) and later replications by using a similar experimental methodology, but without the main unconscious deliberation aspects, and while using a more objective measure for decision accuracy and addressing methodological concerns. This study also followed more recent replications such as those done by Usher et al. (2011) while extending this by implementing calls for further analysis into complexity and individual differences.

The use of dual-process theory (Evans & Stanovich, 2013; Kahneman & Frederick, 2002; Stanovich, 1999; Stanovich & West, 2000) as the theoretical framework, means that research in this area will also contribute knowledge around recent arguments raised in this area (De Neys, 2017; Evans & Stanovich, 2013).

The literature on the context of software selection and evaluation refers to intuition only as an example of an unacceptable decision method (Carney & Wallnau, 1998) or as a potential source of bias (Jamieson & Hyland, 2006), despite reports that managers often make software selection decisions without following rational deliberative processes (Bannister & Remenyi, 2000), and Jadhav and Sonar (2011) call for additional research in their review of software selection. These pointed to a gap in the literature on intuition in this context, as well as relevance of further research in this area.

The software selection context of the research provided potential for publication interest in the Information Systems literature. Vetter, Benlian, and Hess (2011) point to the importance of addressing bias, including overconfidence, in the context of Information Technology decision making. Goes (2013) proposes that there is a large opportunity for combining

behavioural economics principles in Information Systems research, and the present research addresses this with measures of cognitive bias in this context.

1.7.2 Methodological contribution

The research question assumed a predominantly objectivist view by nature of assumptions around the ability to measure faith in, and use of, intuition across multiple subjects (actors). This study did not consider the meaning that individuals attach to intuition and based on these considerations an experimental paradigm was chosen in order to determine causation.

Goes (2013, p. iii) proposes that "Behavioral economics methods can bring enormous potential to inform and complement IS research", and the present research implemented experimental methods from behavioural economics in the context of software selection.

Research involving decision accuracy (Dijksterhuis, 2004; Meszaros, 2007) has predominantly used prescriptive outcomes to test for decision accuracy and this has raised concerns around normatively correct answers (Usher et al., 2011). The present research provided a methodological contribution by using an additively inverted measurement of bias as an indicator of decision accuracy, to avoid previously raised concerns.

1.7.3 Practical contribution

Understanding the role that intuition plays in the context of software evaluation decisions could help managers to improve the effectiveness of these decisions by knowing whether and how to employ or avoid intuition. With the large and rapidly expanding number of software options, and reduced time for evaluation, this has become even more relevant.

The potential to generalise to other contexts, with further research, and to decision making in general, contributed to the understanding of intuition in management decision making overall.

The practical application is supported by Usher et al. (2011, p.10): "Future work that examines objective measures of decision quality in real-life problems, in which attributes are not explicit (e.g., Dijksterhuis, Bos, Van der Leij, & Van Baaren, 2009), and which takes into account individual differences, are important to better understand how to facilitate decision-making."

1.8 Summary and conclusion

This chapter has given the background for the problem, and explained the purpose and objectives for the present research. The scope was discussed, and academic, methodological, and practical justifications were presented in support of the present research to determine the impact of complexity on the effectiveness of intuition in decision making. In the next chapter, the literature review will introduce the relevant literature, and the conceptual model and hypotheses.

Chapter 2 – Literature review and hypothesis development

2.1 Introduction

This review will address the literature applicable to this study and present the research model and hypotheses. As the area of software selection and evaluation provides a rich context for research into complex decision making, this review of the literature will focus on parts that are relevant to this context instead of addressing only general business decision making. In addition to this, because of limitations demonstrated in this context, the potential for intuitive decisions will be discussed and the appropriate literature will be reviewed. The arguments separating the two main camps arguing in favour of and against the use of intuition will be presented. Thereafter, dual-process theory will be presented as the theoretical framework for the research, and arguments will be presented for inclusion of Faith in Intuition and Cognitive Reflection as individual difference measures. Hypotheses will be raised during the review in order to build the research model for this research. The review will end with a summary of the hypotheses raised and a visual representation of the research model. As a preview, and to provide context, the proposed research model is visually presented here first before the review of the literature.

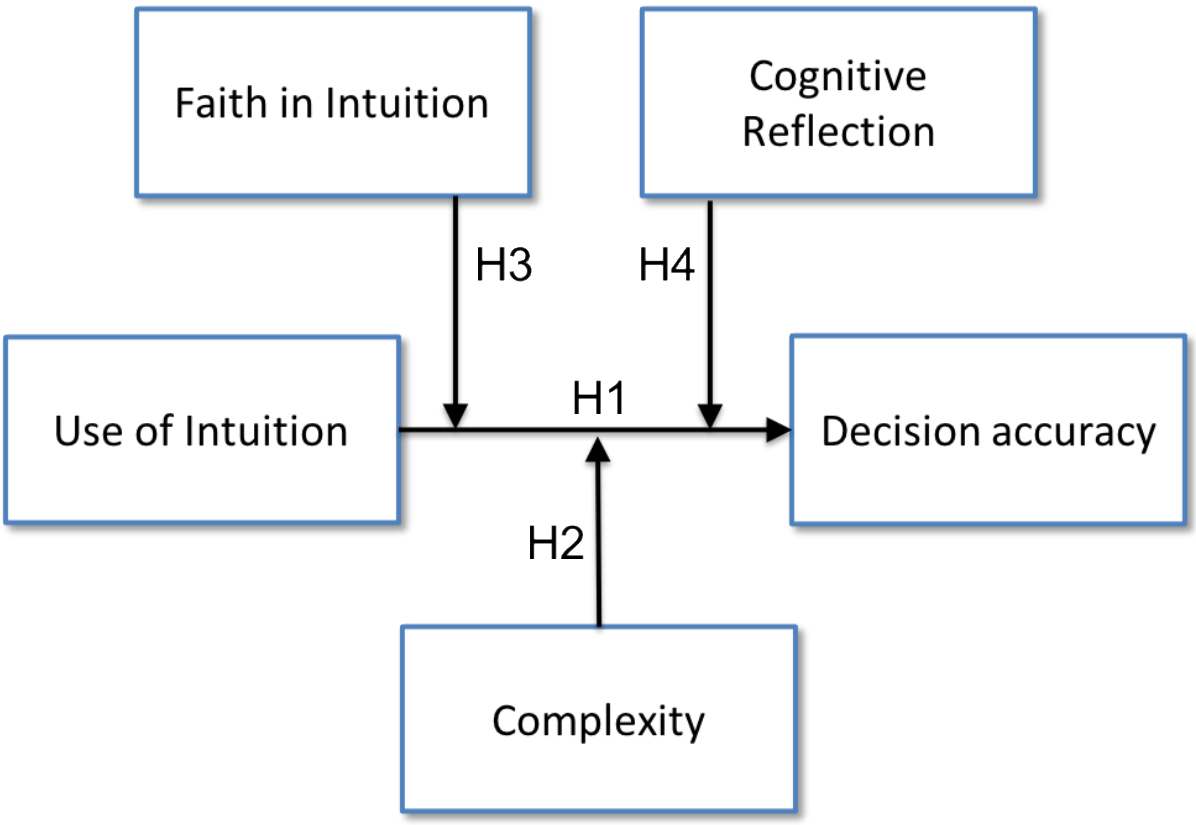


Figure 1. Research model to be developed in this review.

2.2 The context - software selection and evaluation

Software selection and evaluation is classified as a multi-criteria decision-making problem, referring to the problem of making preference decisions over available alternatives that are characterised by multiple, usually conflicting, attributes (Jadhav & Sonar, 2011). Various common decision models have been used, including: analytic hierarchy process, weighted scoring method, and fuzzy multiple criteria decision making. These will be discussed briefly here, along with some of the difficulties in applying them.

The weighted scoring method is a technique which relies on weights and rating scales being assigned to each criterion to reflect the relative importance of each (Jadhav & Sonar, 2009).

This technique produces easy to understand results based on the scoring but is difficult to accurately specify the weightings. The analytic hierarchy process is an approach which arranges the factors in a hierarchical structure, and which facilitates understanding and simplification of the selection and evaluation process (Saaty, 1990). This process handles qualitative and quantitative multi-criteria problems and is applicable to both individual and group decision making (Lai, Wong, & Cheung, 2002). However, the model becomes impractical when using a large number of options. A search for "CRM application" on SourceForge, the main open-source platform, at the time of writing, yielded over 1400 results. As an example – the analytic hierarchy process decision model requires only 54 pairwise comparisons when used to evaluate 4 alternatives with 9 specified criteria, but when the number of alternatives gets to 1000 it would require 4,495,500 pairwise comparisons and it becomes impractical. Fuzzy multiple criteria decision making enables analysis when precise performance ratings and weights are not easily determined and involves the use of fuzzy set theory to model the uncertainty of human judgements by using linguistic terms to evaluate varying options (Jadhav & Sonar, 2011). However, these models require complicated software to implement and the fuzzy appropriateness index values and ranking values are difficult to compute (Jadhav & Sonar, 2009) thereby limiting their application for decision makers.

If the models are too complicated then they are unlikely to be used, and decision makers will potentially adopt heuristics (Sadler-Smith, 2012) or base their decisions on other criteria such as the vendor's perceived stability. Adding to this problem – none of the presented decision models are able to practically handle the sheer number of choices already available in the rapidly growing marketplace for applications.

Decision support systems can provide a solution to some of the complexity and can help humans implement models that would be too computationally intensive. These tools can help

significantly with the vast number of options available. However, Wierzbicki (1997) argues that decision makers with higher levels of responsibility and professional experience are less inclined to trust tools and methods of decision analysis and support. Also, Wang and Benbasat (2009) in a study of decision aids for consumer decision making found that many users find the tools too restrictive and often will rerun the tool many times until they reach what they consider to be an acceptable decision anyway. Determining what constitutes an acceptable decision presents further difficulties and was raised in the scope delimitation section. These decision aids often still require human estimation of the attribute weighting and scoring anyway and limitations around this will be discussed in a later section.

Intuition could be used to bypass in-depth analysis and move quickly to a useful decision. This could offer a solution to the software selection decision problem where effective decisions need to be made quickly and without all the information being available. This could be especially appropriate where decisions need to be made in a short period of time to be viable, because the cost of the time spent on analysis may significantly outweigh the cost of the products being evaluated. However, most of the software selection and evaluation literature reviewed focuses on rational decision-making and does not address intuition. In some cases, the existence of intuition is acknowledged - but only as an example of an unacceptable decision method (Carney & Wallnau, 1998), or as a potential source of bias (Jamieson & Hyland, 2006). Bannister and Remenyi (2000) focus more generally on information technology investment decisions and they do argue that intuition has a role to play, and also propose that many selection decisions are made “without going through the apparent rational step by step processes which management decision-makers are expected to follow” (p.12).

The next section will address the applicable literature on intuition.

2.3 Intuition

"Gut feel" is a colloquial term used to describe intuitive judgments, and the definition used for this study is "a sense of knowing without knowing how one knows" (Epstein, 2010, p. 296). Sloman (1996) describes a process where the "mind goes off, does some work, and then comes back with the result" (Sloman, 1996, p.3), which refers to the background processing aspect of intuition. Simon (1992), one of the main researchers of intuition and decision making, describes skilled intuition by explaining that: "The situation has provided a cue: This cue has given the expert access to information stored in memory, and the information provides the answer. Intuition is nothing more and nothing less than recognition" (p.155). Sadler-Smith and Shefy (2004) define intuition as: "…a capacity for attaining direct knowledge or understanding without the apparent intrusion of rational thought or logical inference". All of these definitions describe common attributes of intuition such as background processing and arriving at an answer without being able to explain why.

However, while there are many views regarding its definition and attributes, the attitudes regarding the effectiveness of intuition are strongly divided into two main camps. Those arguing in favour of the use of intuition in complex environments (e.g. Klein, 1993, 2008) refer to research showing the brilliance of experts in their fields, while those arguing against its use (e.g. Kahneman & Frederick, 2002; Kahneman, 2011) point out the inherent biases involved in intuitive decision making. Bias is accepted as a negative aspect that reduces the accuracy of decisions. As such, an increase in bias would result in a decrease in decision accuracy.

The following section will discuss the arguments in favour of intuitive decisions.

2.4 Arguing in favour of intuition

Isenberg (1984) proposes that managers can use intuition to bypass in-depth analysis and move quickly to a useful solution and describes intuition as an "almost instantaneous cognitive process" (p.85). In further support of this view, Hayashi (2001, p.60) describes how "…executives routinely rely on their intuitions to solve complex problems when logical methods (such as cost-benefit analysis) simply won't do". Others refer to decision makers often using intuition in ambiguous situations, such as where there are conflicting facts or inadequate information (Sinclair & Ashkanasy, 2005). Simon (1987) refers to executives using intuition for decision making in situations where there is insufficient time to follow an orderly sequential analysis of the situation, whilst Dane and Pratt (2007) refer to decision making using intuition as being very fast compared to rational decision making.

As a strong supporter of intuition, Gary Klein represents a large part of the camp arguing in favour of intuition and has spent most of his career focusing on "ways to promote reliance on expert intuition in executive decision making" (Kahneman & Klein, 2009, p. 515). Klein's naturalistic decision making framework was initiated at a 1989 conference sponsored by the Army Research Institute (Lipshitz, Klein, Orasanu, & Salas, 2001). The early work in naturalistic decision making tried to describe and analyse the decision making of firefighting commanders, a context providing uncertainty and time pressure, and which makes generation and evaluation of option sets impractical (Klein, Calderwood, & Clinton-Cirocco, 1986). The authors found that many firefighters usually generated only one option based on matching a large number of patterns that had been accumulated over many years of experience and then considered this option first. They would evaluate the option using mental simulation, and if there were shortcomings then they would modify the option or replace it with the next most

likely option. This strategy took advantage of firefighters' tacit knowledge and is explained by the recognition-primed decision model (Klein et al., 1986).

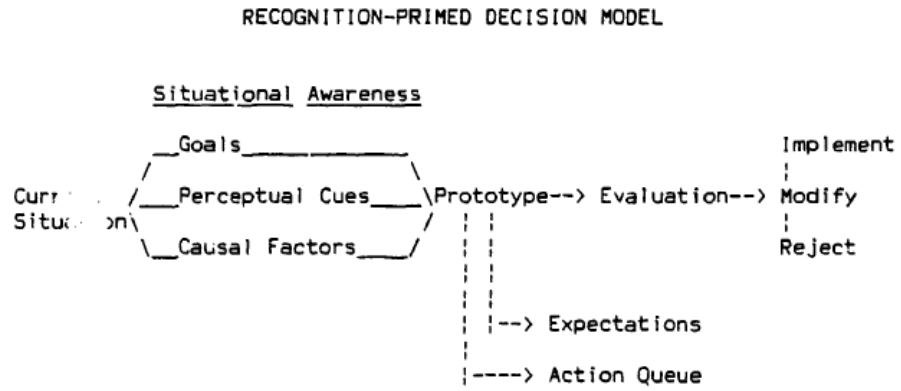


Figure 2. The original Recognition-Primed Decision Model (Klein et al., 1986, p.24)

Klein (1998) refers to standard advice for making better decisions given as being to "identify all the relevant options, define all the important evaluation criteria, weight the importance of each evaluation criterion, evaluate each option on each criterion, tabulate the results, and select the winner" (p.103). He notes that this paradigm is included in most training programs, with the message that "careful analysis is good" (p.103). However, in his experience this advice is mostly ignored, because of the fact that this approach cannot be used in most natural settings. He does mention however that the rigorous approach can be useful where decision makers lack the experience to make effective intuitive decisions. One of the main goals of naturalistic decision making is to "demystify intuition by identifying the cues that experts use to make their judgments" (Kahneman & Klein, 2009, p.516), even where these involve tacit knowledge and are difficult to articulate. Naturalistic decision making researchers use semi-structured interview techniques including cognitive task analysis methods to elicit the cues

and context influencing the decisions in complex conditions that would be difficult to recreate in laboratory experiments (Kahneman & Klein, 2009).

Simon (1955) argues that "we cannot, of course, rule out the possibility that the unconscious is a better decision maker than the conscious" (p.104), and later he explains expert judgement as refined intuition which works through virtually instantaneous visual processing methods (Simon, 1987). Dijksterhuis (2004) showed that intuition could outperform rational thought, with his concept called unconscious thought theory which showed how a period of incubation without rational deliberation resulted in higher decision accuracy.

Wilson and Schooler (1991) showed in a series of experiments how reasoning about choices can reduce the quality of the decisions and the satisfaction with them. Wilson et al. (1993) also found that introspection reduced satisfaction with decisions, and the results were replicated in later studies by Dijksterhuis and van Olden (2006).

Schooler, Ohlsson, and Brooks (1993) proposed that intuition could also be better at determining the relative importance of different attributes since conscious thought can give disproportionate weight to attributes that are easy to access and verbalise.

The next section presents the opposing views of the other camp that argues against the use of intuition.

2.5 Arguing against intuition

In rebuttal to the article from Hayashi (2001) titled “When to trust your gut”, Bonabeau (2003) published an article titled “Don’t trust your gut” where he presented a view of intuition as a “fickle and undependable guide” which is “as likely to lead to disaster as to

success" (p.117). This is the overall view of the counter camp which argues against intuitive decision making.

One of the strongest proponents of this view is represented in a program of research known as the heuristics and biases approach. This approach favours a sceptical attitude toward intuition and offers a sharp contrast to the naturalistic decision making approach. Unlike the naturalistic decision making research which is mostly executed by practitioners in organisations, most heuristics and biases research is performed by academics in well-controlled laboratory experiments (Kahneman & Klein, 2009). In support of this attitude, Rosenzweig (2009) proposes that despite the popularity in business literature of celebrating major business decisions made intuitively, many of the successes could also be attributed to luck.

Kahneman (2009) traces the origins of the heuristics and biases view back to a monograph by Paul Meehl, published in 1954. According to Kahneman, Meehl reviewed about 20 studies comparing forecasts made by humans (mostly clinical psychologists) with forecasts made by simple statistical models, and found that the algorithms were more accurate than the human predictions in most of the cases, despite being based on a subset of the information available to the human subjects. Meehl later explained his refusal to attend clinical conferences, citing uncritical use of intuition and commitment of common fallacies as reasons: "Clearly there are occasions when you should use your head instead of the formula. But which occasions they are is most emphatically not clear. The best evidence is that these occasions are much rarer than most clinicians suppose." (Meehl, 1973, p. 4). The heuristics and biases approach started with a survey of 84 participants at meetings of the Mathematical Psychology Society and the American Psychological Association (Kahneman & Frederick, 2002). Despite their experience with statistics and research, the participants consistently placed too much

confidence in the results of small samples and showed significant bias in their answers - a demonstration of the representativeness heuristic (Kahneman & Frederick, 2002).

Tversky and Kahneman (1971) noted that the respondents in their surveys seemed to have access to two separate approaches to answering the questions – one that is "spontaneous, intuitive, effortless, and fast", and another that is "deliberate, rule-governed, effortful, and slow" (Kahneman & Frederick, 2002, p.49). This led Tversky and Kahneman (1971) to conclude that intuitions should be regarded with "proper suspicion" and that researchers should "replace impression formation by computation wherever possible" (p.31). This two-system view used in the heuristics and biases literature originates from labels assigned by Stanovich (1999) and adopted by Kahneman and Frederick (2002) to show System 1 as the intuitive system and System 2 as the reasoning system, and this is presented graphically in the following figure.

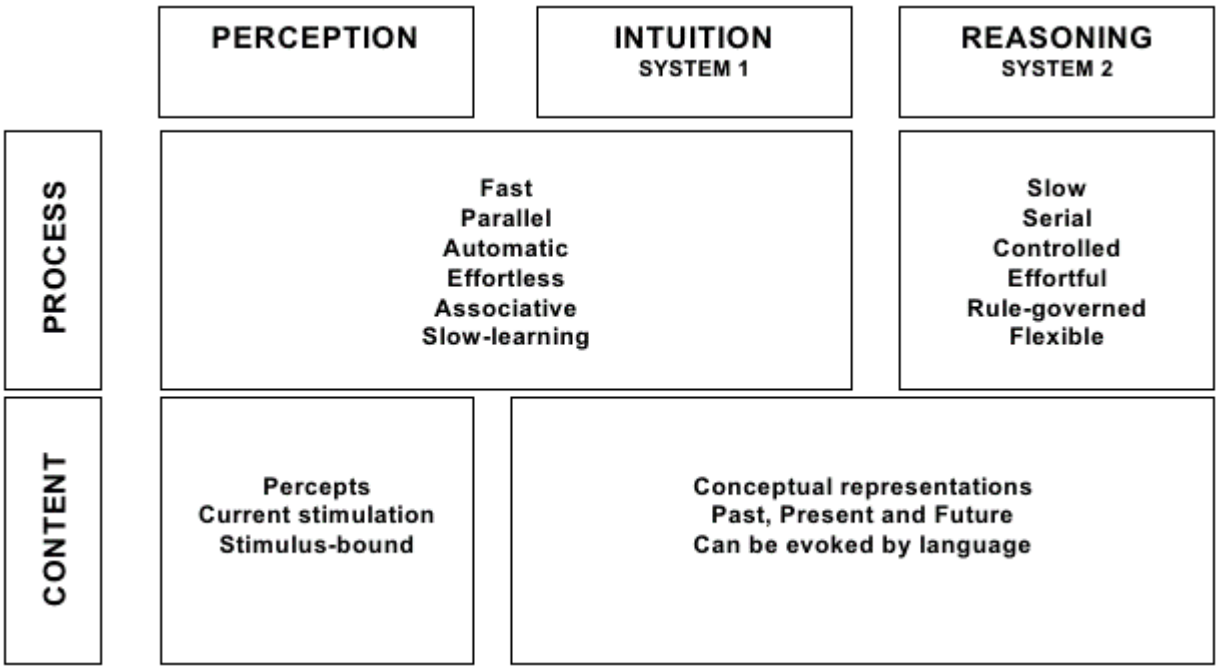


Figure 3. System 1 and System 2 (Kahneman, 2002)

Stanovich and West (2000) propose that System 1 is responsible for automatic processing that places a very low demand on computational capacity, whereas System 2 places a large demand on this capacity and is therefore limited by it. This dual-process theory was used as the theoretical framework for the present research.

Based on this framework, Kahneman and Frederick (2002) propose that intuitive judgments are expressed when System 2 endorses System 1's conclusion – and this was used as the definition of an intuitive judgment for the present research. Based on this, in the heuristics and biases literature, errors in intuitive judgment are considered to be caused by a failure of System 2 to monitor and correct an erroneous conclusion generated by System 1.

Recognising a person that walks up to you is an example of an effortless System 1 process, whereas finding the answer to $\sqrt{19163}$ requires effortful System 2 processes and there is no intuitive solution (Frederick, 2005).

Kahneman and Klein (2009) explain anchoring as one of the potential biases, and they use an example of pricing for German cars. Some participants are asked if the average price for German cars is below or above $100,000, and others are asked a similar question but with $30,000 as the anchor. The anchor can be expected to affect the estimates of the two groups such that they differ by up to half of the difference - $35,000 (Jacowitz & Kahneman, 1995).

In sharp contrast with the view held by Schooler et al. (1993) that intuition could be better at weighting attributes, the heuristics and biases literature shows how biases such as attribute substitution bias could mean that intuition (System 1) could be far worse than rational thought (System 2) in this context. Interestingly, some of the rational models for solving multi-criteria decision making problems, such as the weighted scoring method presented earlier, still require human estimation of the weightings and human judgment of the appropriate scores for each attribute, and even decision support systems will still require

these estimates. This means that the some of the supposedly rational models would also be susceptible to some of the same biases that purely intuitive decisions would be, and this could present an interesting avenue for further research.

Despite the many arguments in favour of intuition, the literature arguing against intuition is compelling, especially in terms of demonstrated increased bias. Therefore, Hypothesis 1 is presented as:

Hypothesis 1:

> There will be a negative relationship between use of intuition and decision accuracy, such that rational decisions will be more accurate than intuitive decisions.

As discussed earlier, complexity could be a potential moderator which could explain how both arguments, in favour of, and against, intuition, could be correct depending on the context. This next section will discuss the support for complexity as a moderator.

2.6 Complexity as a moderator

Epstein (1994) asserts that the intuitive system is capable of processing information of significant complexity. Smith and DeCoster (2000), in their integration of dual-process theories in cognitive and social psychology, concluded that the associative/intuitive system is highly adaptive and capable of "thoughtful processing". Dijksterhuis (2004) and later replications found through experimentation that unconscious thought resulted in consistently better decisions in problems with multiple attributes. Strick et al. (2011) also suggested

complexity as important to consider. Meszaros (2007) found through experimentation, that when participants were given large amounts of information in multi-attribute judgment tasks, the judgments were better when the participants were distracted from deliberation. However, this research has been called into question because of failed replications and criticisms of memory affects (Thorsteinson & Withrow, 2009), and because of failures to replicate the effects in larger samples (Abbott, 2015). Nieuwenstein et al. (2015) conducted a large scale replication and found no evidence of the unconscious thought affect. Their meta-analysis also showed that previous reports were limited to underpowered studies with small sample sizes, however research continues in this direction and especially with attempts to find moderators (Whillock, 2017).

Ignoring the methodological criticisms mentioned above, the concept of limited capacity of working memory required for rational deliberation could still mean that intuition is more suited to complex processing. Complexity could therefore provide an answer explaining the separation between the two camps arguing for and against the use of intuition.

“When making a decision of minor importance, I have always found it advantageous to consider all the pros and cons. In vital matters however . . . the decision should come from the unconscious, from somewhere within ourselves” (Sigmund Freud in Dijksterhuis, 2004).

Therefore, in line with these arguments that higher complexity can require more working memory than is available, especially in terms of the number of attributes to be considered, and that this will affect the accuracy of decisions, the following hypothesis is presented:

Hypothesis 2:

> Decision complexity will moderate the relationship between use of intuition and decision accuracy, such that the effect of using intuition will be greater for simple decisions than for complex decisions.

This section has reviewed some of the arguments for complexity as a moderator. The next section will go into more detail on dual-process theory as the selected theoretical framework for this research.

2.7 Dual-process theory

The two-system view or dual-process theory (Kahneman & Frederick, 2002; Stanovich & West, 2000) provides a useful theoretical framework within which to conduct the research. In terms of this framework, the definition of intuition judgement was taken from Kahneman and Frederick (2002), who propose that intuitive judgments are expressed when System 2 endorses System 1's conclusion. Similarly, errors in judgment are considered to be caused by a failure of System 2 to monitor and correct the erroneous conclusion generated by System 1. Epstein (1994, 2010), another eminent scholar on the subject of intuition also proposes a similar framework, called Cognitive-Experiential Self-Theory (CEST) which assumes two parallel, but interacting, modes of processing, lending further support to the selected

framework. Sloman (1996) also provides empirical support for two systems of reasoning and proposes a dual-process view, and Smith and DeCoster (2000) suggested a new dual-process model involving two memory systems and provided a review of the existing dual-process models.

Some criticisms have been levelled against dual-process models, where scholars have questioned various aspects such as the distinction between systems, processes, modes, and types, which some models fail to address (Evans & Stanovich, 2013; De Neys, 2017). One of these criticisms is that System 1 could be a misnomer because it could represent multiple systems. Stanovich (2005) addressed this by explaining that System 1 contains multiple autonomously operating systems – the Autonomous Set of Systems (TASS) – which respond automatically to relevant stimuli and are not dependent on input from System 2. By countering many of these criticisms, Evans and Stanovich (2013) have strengthened the theory. More recent research has proposed further potential complications such as acquiescence, where a person might realise their judgment is incorrect, but maintain it anyway (Walco & Rison, 2017), and other research proposing that there could be three stages involved in determining which systems will be engaged (Pennycook, Fugelsang, & Koehler, 2015).

The next section will present literature regarding individual difference measures for preferences toward particular reasoning styles, and intuition in particular.

2.8 Individual Differences – Faith in Intuition

Research on individual differences has shown that individuals have predisposed tendencies to approach decision situations in particular ways, known as decision styles (Appelt, Milch, Handgraaf, & Weber, 2011). Decision style has been defined by Arroba (1978, p. 219) as "the way in which a decision is approached and made". Scott and Bruce (1995) showed through their research that individuals have tendencies toward five decision styles: rational, avoidant, dependant, intuitive, and spontaneous. Research on individual differences has also shown that individuals have different ways of thinking, referred to as cognitive styles (Appelt et al, 2011). Allinson and Hayes (1996) showed through their research that individuals have different tendencies toward analytic versus intuitive cognitive processes. Similarly, Pacini and Epstein (1999) created the rational-experiential inventory to measure these predispositions on axis' of rationality and intuitiveness, and this has been widely used (Appelt et al., 2011). The rational-experiential inventory consists of sub-scales of need for cognition for rational thinking, and faith in intuition for experiential thinking, with faith in intuition showing trust in and a tendency toward use of intuition. Research has found a relationship between faith in intuition and particular bias measures (Kahneman, 2011; Keller & Bless, 2009; Shiloh, Salton, & Sharabi, 2002), but there has also been some recent criticism of the face validity of some of the items in the faith in intuition scale which do not seem to specifically reference intuition (Hamilton, Shih, & Mohammed, 2016). Keller and Bless (2009) found that individual differences in faith in intuition moderated the effect of manipulated ease of retrieval, with the result that increased faith in intuition resulted in reduced decision accuracy. Kahneman and Klein (2009) argues that overconfidence in intuitions can increase bias and this could be related to faith in intuition since high faith in intuition would likely be correlated with overconfident intuitive decisions. Because of these

arguments, faith in intuition was tested in the present study, as a moderator to the relationship between use of intuition, and decision accuracy, as well as for any direct effect.

Based on these arguments, the following hypothesis is presented:

Hypothesis 3:

Faith in intuition will moderate the relationship between use of intuition and decision accuracy, such that higher faith in intuition will result in lower accuracy for intuitive decisions.

2.9 Individual differences - cognitive reflection

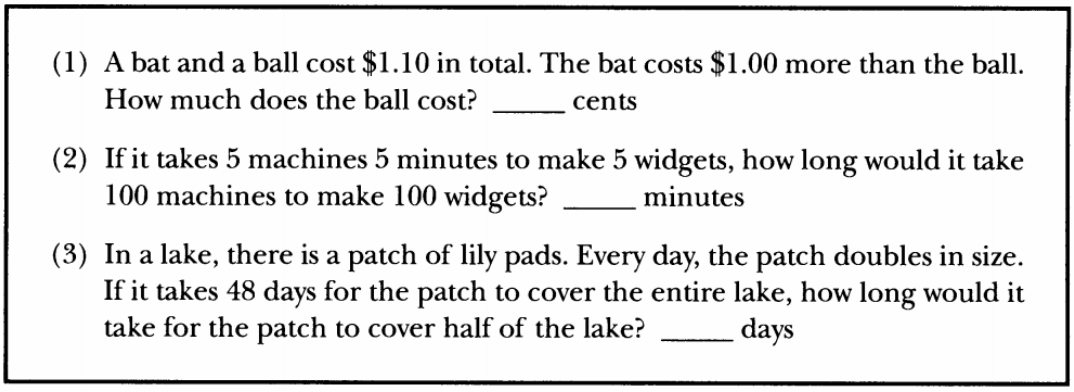
To demonstrate intuitive judgment errors – Frederick (2005, p.26) provides an example from the cognitive reflection test:

A bat and a ball cost $1.10. The bat costs $1.00 more than the ball.

How much does the ball cost? _____cents

Frederick (2005) notes that the intuitively simple answer given by most respondents in a study conducted at Princeton is 10 cents. However, with some reflection it can be seen that $1.00 – 10 cents = 90 cents – not $1.00, and Frederick points out that “catching that error” (p.27) results in the correct answer of 5 cents. This is an example of an erroneous intuition as explained by the heuristics and biases literature – where System 2 to fails to monitor and

correct the erroneous conclusion generated by System 1. The full cognitive reflection test is presented in the following figure.



(1) A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost? _____ cents

(2) If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? _____ minutes

(3) In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? _____ days

Figure 4. The cognitive reflection test (Frederick, 2005, p. 27)

The cognitive reflection test is said to measure the tendency to inhibit automatic, but incorrect, responses in reasoning tasks (Thoma, White, Panigrahi, Strowger, & Anderson, 2015). Furthermore, Pennycook, Cheyne, Koehler and Fugelsang (2016) found that cognitive reflection is a valid measure of reflective thinking and not of intuitive thinking, although they did also find correlations with faith in intuition.

Previous research has shown that cognitive reflection is directly related to bias and decision accuracy even when cognitive ability and executive functioning were controlled for (Toplak, West, & Stanovich, 2011). In this research it was proposed that cognitive reflection should moderate the relationship between use of intuition and decision accuracy. If cognitive reflection can inhibit the rapid heuristic thinking, then this measure should affect the level of bias and decision accuracy differently when intuition or rational thinking are forced and therefore it was proposed as a moderator in this study. Therefore the following hypothesis is proposed:

Hypothesis 4:

> Cognitive reflection will moderate the relationship between use of intuition and decision accuracy, such that higher cognitive reflection will result in higher decision accuracy for intuitive decisions.

This section has reviewed some arguments relating to individual differences measures and raised hypotheses for testing. The next section will present the overall research model and a summary of the hypotheses.

2.10 Summarised hypotheses and research model

The following hypotheses have been presented during the review of the literature:

Hypothesis 1:

There will be a negative relationship between use of intuition and decision accuracy, such that rational decisions will be more accurate than intuitive decisions.

Hypothesis 2:

Decision complexity will moderate the relationship between use of intuition and decision accuracy, such that the effect of using intuition will be greater for simple decisions than for complex decisions.

Hypothesis 3:

Faith in intuition will moderate the relationship between use of intuition and decision accuracy, such that higher faith in intuition will result in lower accuracy for intuitive decisions.

Hypothesis 4:

Cognitive reflection will moderate the relationship between use of intuition and decision accuracy, such that higher cognitive reflection will result in higher decision accuracy for intuitive decisions. The following figure shows a summary of the elements in the present research model and the relationships to be tested.

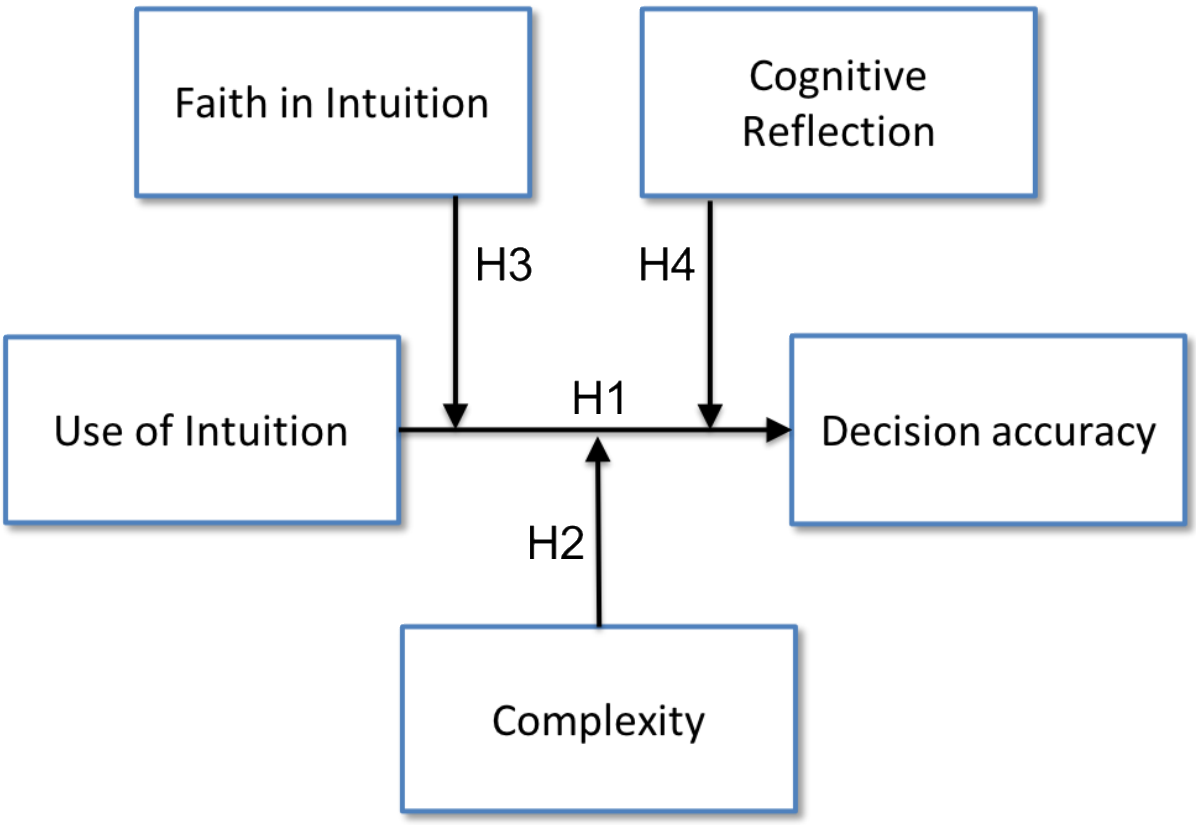


Figure 5. Research model

This chapter has presented the review of the literature and has developed the hypotheses that were tested in this research. The next chapter will present the methodology and explain how the research was designed and how it was executed.

Chapter 3 – Methodology

3.1 Introduction

The present research was conducted as an experiment, using a software tool (an online internet application) which was developed specifically for this purpose. The research utilised a true experimental, posttest-only control group, between-subjects, factorial design. An initial pilot study with a convenience sample was conducted, in order to test the survey experiment tool, and the time taken, and through that process certain weaknesses were discovered and eliminated. A second, larger pilot was conducted using a crowdsourcing platform to recruit international information technology decision makers to complete the survey. This second pilot study resulted in 56 fully completed and valid surveys which were analysed quantitatively using the methods described, however, the results of this pilot were excluded from the main sample of 90 complete and valid surveys, and the results of the pilot studies are not presented here.

This chapter explains some of these choices in more detail.

3.2 Overview

For the present research, a positivist ontology was used. Gill and Johnson (2010) refer to positivist research collecting data about observable reality and searching for causal relationships and regularities in the collected data in order to produce generalisations. Some aspects from pragmatism are acknowledged in that the researcher recognises that even the choice of the research questions and specific components of the instrument created can reflect the values of the researcher, however in-line with the positivist philosophy, the researcher tried to minimise these as far as possible.

For this particular research, a quantitative-only study was used, following an experimental survey research strategy which aligns with the mainly deductive focus of the present research.

The researcher considers social entities to exist outside of the social actors and also recognises the existence of multiple views, but chooses to work within the view which best suited the research question. The research question regarding examination of the impact of complexity on the effectiveness of intuitive decisions assumed a predominantly objectivist view by nature of assumptions around the ability to measure faith in intuition, and use of intuition across multiple subjects (actors). As such the main portion of the research did not consider or go into detail on the meaning that individuals attach to intuition. This was not meant to discount these individual meanings or assume that the subjective lens would not be appropriate or interesting for other research questions regarding attitudes toward intuition and its use – but simply delineated these as being outside the main context of this study. These ontological views provided support for selection of the positivist philosophy. The research followed a deductive approach where theory is developed from academic literature and this is then tested with data collected for this purpose. The study involved applied, cross-sectional, experimental, empirical research which collected primary data using experimental surveys applied online. The data collected included numerical data, textual data, timing, and coded responses.

The research followed a quantitative approach with an experimental research design and survey strategy. Because this research involved not only variable observation but also manipulation, where use of intuition and complexity were varied purposefully, this research is classified as an *experimental design* and allowed for identification of *cause-and-effect relationships* (Leedy & Ormrod, 2014).

In terms of standard experimental designs identified by Campbell and Stanley (1963) in five general categories: pre-experimental, true experimental, quasi-experimental, ex post facto,

and factorial designs. Since the subjects were *randomly assigned* to groups, this study followed a *true experimental design*. Because of the variables being measured, it was not possible to include a pre-test, and therefore a *posttest-only control group design* was used for the study. The table below indicates the paradigm for the posttest-only control group approach, showing Tx for treatment related to use of intuition and Obs for observation.

Table 1 – General paradigm for chosen posttest-only true experimental design

	Group	Time ->	
Random assignment	*Group 1 (control)*	-	Obs
	Group 2 (rational)	Tx(Deliberate)	Obs
	Group 3 (intuitiveA)	Tx(Fast)	Obs
	Group 4 (intuitiveB)	Tx(Distract)	Obs

Because the research model included complexity as a moderator, the design either needed to vary complexity within each experiment as a within-subjects variable, or vary complexity as another between-subjects variable. Because of the importance of complexity as a moderator in the research model in influencing the effectiveness of intuition, this research varied complexity as a between-subjects variable while use of intuition was also varied as a between-subjects variable. The manipulation of multiple independent variables led to the classification of the study as a factorial design (Leedy & Ormrod, 2014). The table below shows the updated paradigm as a factorial design.

Table 2 - Paradigm showing full factorial design

Random assignment	Group	Subgroup	Time ->		
			use of intuition	decision complexity	
	Group 1 (control)	1S	-	Tx(simple)	Obs
		1C		Tx(complex)	Obs
	Group 2 (rational)	2S	Tx(Deliberate)	Tx(simple)	Obs
		2C		Tx(complex)	Obs
	Group 3 (intuitiveTimed)	3S	Tx(Fast)	Tx(simple)	Obs
		3C		Tx(complex)	Obs
	Group 4 (intuitiveDistracted)	4S	Tx(Distract)	Tx(simple)	Obs
		4C		Tx(complex)	Obs

3.3 Population and sampling

The research was applied to senior Information Technology decision makers at listed and private enterprises in South Africa. The population includes all companies that fit the inclusion criteria. The initial list of companies provided by the Johannesburg Stock Exchange was used for this purpose. At the time, there were 393 companies listed on the Johannesburg Stock Exchange. Because of the small population size, the entire population of listed entities was approached, and the sample was then expanded by including non-listed organisations.

Because of the generally limited response rates of postal and email surveys, the experimental surveys were applied in person, as far as practical, by a combination of the researcher and

various research assistants which were hired for this purpose. The geographic spread of companies required some surveys to be completed online but these were still set a specific appointment and were monitored remotely during completion by either the researcher or one of the research assistants.

The ITWeb/Brainstorm Chief Information Officer Directory (2014-2017) was cross-referenced with the list of companies to find additional contact information where it was available. Those not in the Chief Information Officer Directory were contacted by obtaining their details on LinkedIn or on the company website, or by calling the organisation directly and asking for the Chief Information Officer or Chief Technology Officer or any other senior Information Technology decision maker. Although various techniques were employed to ensure an acceptable response rate, it became necessary to expand the sample beyond listed companies, and questions were included in the experimental questionnaire to cater for this by asking for details of the company. An email was sent out to each person in the list to explain the benefits of the study to them and to inform them that the researcher or a research assistant would contact them to arrange an appointment. This email also communicated each of the usual consent criteria, such as noting that they may withdraw from the study at any time. Any responses to this email indicating a wish to withdraw from the study were collated so that those people were not contacted to arrange the in-person meetings. The updated list was continuously synchronised between the researcher and research assistants, using an online spreadsheet, and a few days after the email was sent they would contact the participants and arrange a suitable meeting time. In some cases where the potential respondents were not available, more than 20 attempts were made to contact them. The research assistants were trained by the researcher and provided a protocol document detailing the engagement. The research assistants were not made aware of the hypotheses so that they could not

communicate these to the subjects before the engagement, but they were given some background to the research so that they could communicate this to the subjects if required.

The final response was completed 19 months after the start (from 19 April 2016 to 22 November 2017). In total 139 unique surveys were generated for respondents that agreed to participate and this resulted in 90 fully complete and valid surveys. More details on the resulting sample are described in the results section.

Subjects were randomly assigned to different groups for the between-subjects variable use of intuition and within this to the between-subjects manipulation of decision complexity. This random assignment was achieved using block random assignment in order to try to ensure that there was an equivalent response rate across the groups which would help to reduce the effects of significantly different group sizes.

3.4 Unit of analysis

The main unit of analysis for the study was at the individual-level. The focus on the contribution of individual differences in terms of faith in intuition and cognitive reflection in the model grounded the analysis in the individual. As such, the analysis revolved around the individual senior information technology decision makers.

3.5 Variable mapping

The figure below shows each of the main variables in the research model.

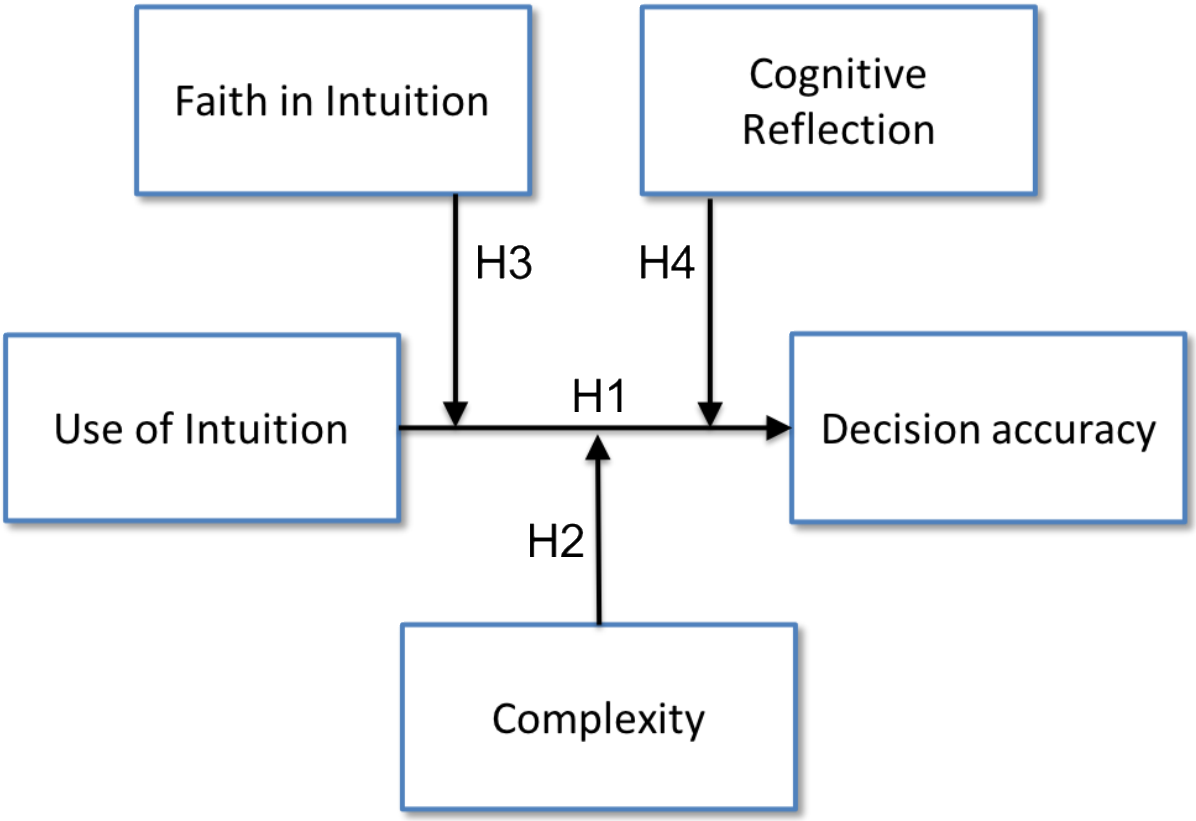


Figure 6. Main variables

The following sections will discuss the measurement for each of these variables.

3.5.1 Faith in intuition

The faith in intuition variable was generated from the application of a scale taken from the "Rational-Experiential Inventory (REI)" (Epstein, Pacini, Denes-Raj, & Heier, 1996). The scale is called the "Faith in Intuition" scale and has been validated separately from the other constructs in the REI (Epstein, et al., 1996). The scale reflects the level of experiential thinking style and is developed as a self-report measure to assess individual differences in this style of thinking – specifically the degree of confidence and engagement in intuitive abilities (Pacini & Epstein, 1999). Participants rate all items on 5-point Likert-type items

ranging from completely false to completely true. The faith in intuition scale is presented in the questionnaire in the appendices. There are multiple versions of the faith in intuition scale, and the 5-item version from Epstein et al. (1996) was chosen since the face validity of the questions matched the construct well, and testing in this research showed a higher reliability and fewer factor loadings than the 12-item scale.

The following five items make up the faith in intuition scale (each answered as 5 point likert-type items ranging from “strongly disagree” to “strongly agree”):

1. My initial impressions of people are almost always right.
2. I trust my initial feelings about people.
3. When it comes to trusting people, I can usually rely on my “gut feelings”
4. I believe in trusting my hunches.
5. I can usually feel when a person is right or wrong, even if I can’t explain how I know.

3.5.2 Use of intuition

Using dual-process theory as the lens for the research provided methods for elucidation of intuitive judgments by suppressing the rational System 2 in various ways. Two of these ways which have been used in previous experimental research on intuitive judgments (Dijksterhuis, 2004; Meszaros, 2007) included shortening the time so that deliberation (System 2) was not possible since System 2 is slow, and preventing System 2 deliberation of the required task by distracting the subject with another task that used the full capacity of System 2 - by asking the subject or respondent to solve anagrams for a period of time. There are also techniques to ensure processing by System 2 by warning them of influences on working memory or potential bias, by bringing their attention to a bad font, suggesting comparison with another option set, or instilling a feeling of high risk (Dhar & Gorlin, 2013). For the present research, similar techniques such as time given for deliberation, along with a warning of potential

trick-questions, and suggestions to be wary of first impressions were to be used to elicit a System 2 response, and both limited period and distraction interventions were used to elicit a System 1 response.

3.5.3 Decision complexity

Decision complexity was experimentally controlled as a between-subjects variable, by creating two separate cases – one with seven attributes for simple decision complexity, and one with fourteen attributes for complex decision complexity. The cases are presented in the section discussing the questionnaire.

3.5.4 Decision accuracy

Attribute substitution bias, anchoring bias, and other measurable biases grouped under general bias were summed to give a total bias as the construct for decision bias. This score was then scaled to between 0 and 1 and additively inverted to determine a decision accuracy measure. Therefore, decision accuracy = 1- decision bias, where decision bias is the sum of attribute substitution bias, anchoring bias, and general bias.

3.5.4.1 Attribute substitution bias

Specific attributes were presented in the case and questions were presented after the case to test for attribute substitution bias (Kahneman & Frederick, 2002). Attributes were tested separately in a way that gave a level of bias instead of purely a biased/not-biased result as other bias measures have been used before. There were three attribute substitution bias question-sets:

- Cultural fit attribute substitution bias (scored on estimated cultural fit of package X, estimated cultural fit of package Y, and estimated cultural fit of package Z items in the questionnaire – on 100 point items from “Very bad” to “Very good”)

  - "Please rate the expected cultural fit of: Package X?"
  - "Please rate the expected cultural fit of: Package Y?"
  - "Please rate the expected cultural fit of: Package Z?"

- Scalability attribute substitution bias (scored on estimated scalability of package X, estimated scalability of package Y, and estimated scalability of package Z items in the questionnaire – on 100 point items from "Not very scalable" to "Very scalable")
  - Please rate the scalability of: Package X?
  - Please rate the scalability of: Package Y?
  - Please rate the scalability of: Package Z?
- Remote support attribute substitution bias (scored on estimated remote support of package X, estimated remote support of package Y, and estimated remote support of package Z items in the questionnaire – on 100 point items from "Not very good" to "Very good")
  - Please rate support for remote workers with: Package X?
  - Please rate support for remote workers with: Package Y?
  - Please rate support for remote workers with: Package Z?

3.5.4.2 Anchoring Bias

Anchors were presented in various parts of the case in order to elicit an anchoring bias reaction (Kahneman & Frederick, 2002). The influence of the anchor on other answers was then tested to determine the anchoring bias. There were five groups of questions which attempted to elicit a biased reaction, and these were based on the following variables:

- estimated software value (vs an experimenter-generated R5m anchor)
  - Anchor: "Do you expect that the required software would cost more or less than R5m? Less ( ) More ( )"

    - Item: “Please give an estimate of the amount that you would expect to pay for the license cost (excluding first year maintenance) for software that would suit the requirements of the company (in millions of Rand, rounded to one decimal place - e.g. R#.# million). R [ ]”

- expected satisfaction with purchase (group 1), expected satisfaction with purchase (group 2) (same question placed differently – group 1 has the question before the group of questions regarding the expected satisfaction with package X,Y, Z – on a 100 point item from “Very unsatisfied” to “Very satisfied”)
    - Anchor: “Think back to the last actual software purchase you were involved in (not part of this case). How satisfied are you right now with that product?”
    - Items:
        - “How satisfied would you expect to be with: Package X?
        - How satisfied would you expect to be with: Package Y?
        - How satisfied would you expect to be with: Package Z?”
- expected satisfaction with package X, expected satisfaction with package Y, expected satisfaction with package Z (vs self-generated and experimentally controlled anchor in expected satisfaction with purchase (group 1))
    - Anchor: “Think back to the last actual software purchase you were involved in (not part of this case). How satisfied are you right now with that product?”
    - Items:
        - “How satisfied would you expect to be with: Package X?
        - How satisfied would you expect to be with: Package Y?
        - How satisfied would you expect to be with: Package Z?”
- previous purchase cost (group 2) and expected cost of package X, expected cost of package Y, expected cost of package Z, previous purchase cost (group 1)

  - Anchor: “Think back to the last software purchase decision you were involved in for software costing at least R500k (preferably similar to this case). Try to remember the approximate cost of the software and please fill it in here. If you cannot remember then fill in an approximate amount that you would guess it could have cost. [R]”
  - Items: “Please estimate the approximate amount that you would expect to pay (including implementation) for:
    - Package X: [R]
    - Package Y: [R]
    - Package Z: [R]”
- number of previous purchases (group 2), number of previous purchases (group 1) vs each other and vs estimate of number of months to implement the software
  - Anchor: “Approximately how many software purchase decisions have you been involved in (over the past three years)? [ ]”
  - Item: “What would your rough estimate be of the number of months it would take for implementation of any of the three options? (if different then the longest one) [months]”

3.5.4.3 *General bias*

There were three separate questions which attempted to elicit a biased reaction, using modified versions of previously used bias measures. These were all modified to take the story in the case study into account so that the subjects would consider them part of the case.

They were based on the following variables:

- Program favour general bias (based on favouring program 1, and favouring program 2).

- Operation general bias (based on evaluation of successful operation, and evaluation of failed operation)
- Estimated analysis weeks general bias (based on estimation of number of weeks of analysis for package A).

The program favour general bias variable was based on the answers to two questions which comprise a modified version of the disease framing problem (Tversky & Kahneman, 1981; West, Toplak, & Stanovich, 2008). The results should be the same despite the different wording of the two questions.

Different answers indicated a descriptive invariance bias (Tversky & Kahneman, 1981):

> favouring program 1 - "Consider that the company is currently undergoing severe financial difficulties because of severe issues with the current ERP system. Because of these difficulties it is expected that 60 people will need to be retrenched. Two alternative programs to improve profitability have been proposed. An analysis has shown the following exact consequences of each of the programs:
>
> If program A is adopted, 20 of the people will avoid retrenchment. If program B is adopted, there is a 1/3 probability that all 60 people will avoid retrenchment, and 2/3 probability that none will avoid it.
>
> Which of the two programs would you favour?"
>
> favouring program 2 - "Consider again that the company is currently undergoing severe financial difficulties because of severe issues with the current ERP system. Because of these difficulties it is expected that 60 people will need to be retrenched. Two additional alternative programs to improve profitability have been proposed. An analysis has shown the following exact consequences of each of the programs:

> If program C is adopted, 40 people will be retrenched. If program D is adopted, there is a 1/3 probability that nobody will be retrenched, and 2/3 probability that all 60 people will be retrenched.
>
> Which of the two programs would you favour?"

The operation general bias variable was based on the answers to two questions which comprise a modified outcome bias test (Baron & Hershey, 1988; Toplak & Stanovich, 2011). The choices should be the same despite the outcome or the participant will have demonstrated an outcome bias.

> evaluation of successful operation – "A 55-year old man in another department had a heart condition. He had to stop working because of chest pain. He enjoyed his work and did not want to stop. His pain also interfered with other things, such as travel and recreation. A type of bypass operation would relieve his pain and increase his life expectancy from age 65 to age 70. However, 8% of the people who have this operation die from the operation itself. His physician decided to go ahead with the operation. The operation succeeded. Evaluate the physician's decision to go ahead with the operation."
>
> evaluation of failed operation – "Another person in the company had a liver condition. He also had to stop working because of pain. He enjoyed his work and did not want to stop. His pain also interfered with other things, such as travel and recreation. A particular operation would relieve his pain and increase his life expectancy. In this case, only 2% of the people who have this operation die from the operation itself. His physician decided to go ahead with the operation. In this case the operation failed and the man died. Evaluate the physician's decision to go ahead with the operation."

The estimated analysis weeks general bias variable was based on a modified version of the bat and ball question from the cognitive reflection test (Frederick, 2005), subjects were tempted to answer 9 which seemed correct, when actually the correct answer was 4.5 weeks.

> "For package X, the analysis and implementation together are estimated to take 29 weeks. Implementation takes up most of the time and will take 20 weeks more than the analysis. How long will the analysis take? "

3.5.5 Cognitive reflection

The cognitive reflection test is a 3-item measure that was introduced by Frederick (2005). The cognitive reflection test provided a measure of general susceptibility to cognitive bias and an indication of System 1 errors and System 2 processing. The cognitive reflection test was not controlled to elicit a particular response as with the experimental case, and instead was used to provide a comparative measure of predisposition. The items in the cognitive reflection test are:

1. A bat and a ball cost R1.10 in total. The bat costs R1.00 more than the ball. How much does the ball cost? _____ cents
2. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? _____ minutes
3. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

3.5.6 Demographic variables

Additional data was collected related to personal information such as gender, language, age, experience, as well as organisation size and employment.

3.6 Data Collection

3.6.1 First pilot study

The questionnaire was pre-tested with a small pilot study (15 subjects). This allowed the researcher to test the survey experiment instrument and check the time taken on average for completion. Various improvements were made to the tool in order to ensure reliability and to ensure that all of the data required were collected. This process also allowed the researcher to discover any obvious weaknesses in the design. Although some preliminary analysis was performed using data from this pilot – the source data, individuals involved, and results of the preliminary analysis were excluded from analysis and were only used for the pre-test.

3.6.2 Second pilot study

The second larger pilot, that was conducted with 56 subjects, informed further changes to the questionnaire after analysing the results.

Some of these changes included:

- For the case study rating questions, some of the Likert-type items were not showing enough variation and so various options including increasing the number of Likert-type items were considered, however there are additional issues that are raised when increasing the items – especially linearity of labelled items (Carifio & Perla, 2007). Therefore, the items were changed to a Visual Analog Scale (deVellis, 2012) which results in a number from one to one hundred. This allowed more resolution and sensitivity in the answers instead of the very limited Likert-type items, some of which weren't

showing enough variation in the pilot study analysis. Each of these items only had an indication on each side such as very bad to very good and the rest of the range was not labeled. This was also done in order to obtain a more linear response which would be easier to analyse, compared to labeled Likert-type items where different people may interpret the intermediate label "somewhat good" differently.

- The distraction mode was relaxed somewhat after reports of subjects becoming irritated with the anagrams (one of the questions asked for comments about the case study).
    - "There were too many anagram checks."
    - "It was quite annoying that as you are taking the survey, the constant need of solving anagrams is just so aggravating. "
- Questions asking for an amount of money were changed to indicate the currency, after the international pilot sample showed confusion and were answered with multiple currencies – which were difficult to compare and analyse. The amount questions were also changed to the same slider used to replace Likert-type items except that instead of a scale from one feeling to another, they allowed a selection of an amount between R0 and R20m in multiples of R100,000 to simplify the results, and this also limited the outliers. Similarly, other questions which could be ambiguous and which resulted in multiple denominations or units in the answer were made specific or the answers were limited to specific units (e.g. number of weeks).
- Insufficient anchoring effect displayed by the first experimenter-generated anchor caused the addition of an extra question asking half of the respondents (treatment group) whether they thought the software would be more or less than R5m (Tversky & Kahneman, 1974) instead of the original technique where the anchor was presented as a range of R1m to R5m inside the case text – although both techniques are valid (Kahneman & Knetsch, 1993; Wilson, Houston, Etling, & Brekke, 1996). Using these standard techniques with a

high anchor and a low anchor was not possible, and so a decision was made to use a single anchor and a control group due to the expected sample size and having to separate into three groups instead of two (in addition to the current experimental groups). Additional, less formal, anchor questions were also added. All of the anchor questions were changed to include their own control groups, and to add additional questions to balance the groups so that the members of the control group for one anchor would be the treatment group for another anchor, to ensure sufficient data.

3.6.3 Questionnaire

The questionnaire consists of multiple sections and only some of these were experimentally manipulated.

The first section of the questionnaire has the 3 questions from the cognitive reflection test scale (Frederick, 2005), with the only adaptation being the change to "R" to represent the local currency instead of dollars in the first question. Each answer was timed so that this data could be used in the analysis. The timing was done by the software which applied the instrument, and the data resulting from the timing was stored as the number of milliseconds taken to answer from the time of presentation of a particular question.

The second section contained the experimentally manipulated case. The case was presented as the requirements of a company that needed to purchase software. By specifying the importance of various attributes in the case description, it was possible to largely exclude previous criticisms (e.g. Usher et al., 2011), of the initial studies by Dijksterhuis (2004), which stated that the multi-criteria decision making problems relied on subjective weighting of attributes by the respondents. After presentation of the case, three shortlisted options were presented along with their attributes. To avoid complications from the effects of choice, the case was structured around asking questions about the options and asking for overall

evaluations of the options instead of asking subjects to choose. Either a simple, or a complex version of the case was presented, depending on the random assignment to the decision complexity group in the experiment.

These were the two versions of the case:

3.6.3.1 Complex version

The company has less than 500 employees and most of their sales come from distribution of less than 30 sports products with multiple configurations, which they import from European countries and distribute across South Africa. The company is situated in Johannesburg and not all of the employees speak English, other languages include Sotho, Zulu, and Afrikaans. A significant component of the current administration cost in the company comes from having to handle and escalate approvals of varying nature including purchase approvals, discount approvals, and leave approvals. The employees are mostly millennial's and respond well to modern, social network type interfaces and they expect to be able to communicate with co-workers and key customers using the ERP application. The vendor for the current system is planning to sunset the product and has frozen all development except for bug-fixing for the last three years. The overall culture of the company supports long hours and many employees spend additional time on weekends working on sales proposals to channel clients and require the ability to work remotely. The company is currently experiencing severe financial difficulties as a direct result of inconsistencies in the data and processing of the current ERP. The current system has proven inflexible and cannot cater for new business requirements. Because of the single location and significant internal infrastructure there are no current requirements for a cloud-based application. CRM is not currently being requested and the current system has no CRM capability but the sales director has expressed interest in CRM.

The following three packages have been shortlisted:

1. Package X:

a. Multi-language support

b. Cloud-ready

c. Successful track record of implementation in Europe and some track record in distribution.

d. Standard interface with corporate design - HTML5

e. No social media integration

f. Built in CRM

g. Vendor provided roadmap with yearly releases

h. Ability to access remotely using HTML5 web client

i. Priced in euro – expensive

j. Cannot be customised but can be extended using graphical tool and API's for various languages

k. Adequate system performance, reasonable reporting speed

l. Built in integration for Windows Workflow Foundation

m. Although upgrades are included in the annual maintenance, upgrade implementation is for the cost of the client and has been known to be expensive

n. Web-oriented-architecture

2. Package Y:

a. Single-language support (English/Portuguese/French)

b. No cloud support but can be hosted

c. Successful track record of implementation in Africa and some track record in retail

d. Traditional Forms interface - Windows 7, 8, 10

e. Social media integration with Facebook, Twitter, and Yammer

f. Integrates fully with Microsoft Dynamics CRM

g. Reports from reference sites that the vendor responds quickly to changing customer requirements and very often releases minor updates with no charge to the requesting organisation as long as the request can benefit a reasonable proportion of the client base

h. Can work over RDP or with locally installed client

i. Priced in Rand – expensive

j. Customisable using standard .net code

k. Excellent performance, fast reporting

l. Custom workflow module with graphical process editing

m. Version upgrades are normally simple because of the incremental releases, however upgrades can break any customised .net code

n. Object-oriented client-server architecture

3. Package Z:

a. Single-language support (English)

b. No cloud support

c. Successful track record of implementation in distribution of sporting goods in North and South America

d. Multiple interfaces with modern design - Forms (Windows 7), Metro (Windows 8, 10), and Web (HTML 5)

e. Configurable integration - no support for Facebook, Twitter etc. yet but can be developed

f. Support for SalesForce.com CRM (integrated)

g. No roadmap for future development but a track record of releasing stable new versions every two years

h. Web client can be accessed remotely but does not support all enterprise functionality (only employee self-service HR, and executive dashboards on HTML5, with other functionality incompatible with RDP and requires local connection)

i. Priced in Dollars - cost effective

j. Customisable using a proprietary language which has been designed specifically for this package

k. Sluggish system performance, slow reporting

l. Sharepoint integration with built-in workflow editor

m. Built in upgrade tool which extracts customisations, performs the upgrade, and then re-implements customisations over a weekend, with no additional labour required other than training on new features

n. Service-oriented-architecture

3.6.3.2 Simple version

The company has less than 500 employees and most of their sales come from distribution of less than 30 sports products with multiple configurations, which they import from European countries and distribute across South Africa. The company is situated in Johannesburg and not all of the employees speak English, other languages include Sotho, Zulu, and Afrikaans. The employees are mostly millennial's and respond well to modern, social network type interfaces and they expect to be able to communicate with co-workers and key customers using the ERP application. The overall culture of the company supports long hours and many employees spend additional time on weekends working on sales proposals to channel clients and require the ability to work remotely. The company is currently experiencing severe financial difficulties as a direct result of inconsistencies in the data and processing of the current ERP.

The following three packages have been shortlisted:

1. **Package X:**

a. Multi-language support

b. Successful track record of implementation in Europe and some track record in distribution.

c. Standard interface with corporate design - HTML5

d. No social media integration

e. Ability to access remotely using HTML5 web client

f. Priced in euro – expensive

g. Adequate system performance, reasonable reporting speed

2. Package Y:

a. Single-language support (English/Portuguese/French)

b. Successful track record of implementation in Africa and some track record in retail

c. Traditional Forms interface - Windows 7, 8, 10

d. Social media integration with Facebook, Twitter, and Yammer

e. Can work over RDP or with locally installed client

f. Priced in Rand – expensive

g. Excellent performance, fast reporting

3. Package Z:

a. Single-language support (English)

b. Successful track record of implementation in distribution of sporting goods in North and South America

c. Multiple interfaces with modern design - Forms (Windows 7), Metro (Windows 8, 10), and Web (HTML 5)

d. Configurable integration - no support for Facebook, Twitter etc. yet but can be developed

e. Web client can be accessed remotely but does not support all enterprise functionality (only employee self-service HR, and executive dashboards on HTML5, with other functionality incompatible with RDP and requires local connection)

f. Priced in Dollars - cost effective

g. Sluggish system performance, slow reporting

After the case and the options were presented, the experimental manipulation for use of intuition was implemented. The simple and complex groups (versions of the case) were further split into four groups. Group 1, the control group, had no intervention for use of intuition. Group 2, the deliberation group, was given an intervention to evoke a system 2 response, by giving them an instruction to consider the details carefully and warning them that they may be asked to explain their answers. Group 3, the timed group, was given instructions to consider the options holistically and they were asked to give an overall impression of each of the options, in order to inhibit a system 2 response. Group 4, the distracted group, were given similar instructions to Group 3 except that they were given anagrams to solve whenever they took too long to answer a question, also to inhibit a system 2 response. This was different to many of the other similar intuition and decision-making studies where the anagrams were applied during an initial period to prevent deliberation, and then questions were answered. Also, this study displayed the case information the whole way through during the questions, whereas most other studies presented the case in the beginning, prompting criticisms about memory effects because subjects might not remember all the details of the case when answering the questions (Usher et al., 2011).

The third section applies the Faith in Intuition scale (Epstein et al., 1996; Pacini & Epstein, 1999) to provide the individual differences measure for analysis. No adaptations have been made to this measure.

The fourth and final section asked for demographic information such as gender, home language, and age, and also asks for experience to establish domain expertise. Although Reips (2002) recommends that personal information questions are placed at the start of a survey to ensure that only people who are serious about answering the survey will continue, known as the hurdle technique, because the instrument was applied in person it was decided

that the demographic questions should be moved to the end to minimise any cognitive drain effects.

The average time taken to complete the survey was 32 minutes with differences per experimental condition, and this will be explained in more detail in the results chapter in the description of the sample.

3.6.4 Access to data

For the present research, the software to administer the experimental survey was custom developed. This ensured that access to the data was controlled by the experimenter. The custom software also enabled further features required by the experiment such as the block random assignment of subjects to groups, timing of performance on the cognitive reflection test and on the case study tasks as described in the question design. Most importantly, the custom software allowed complete control of the specific functionality required for the anagrams which would distract the subjects when required.

3.7 Data analysis methods

This section presents the data analysis methods that were used. The experimental survey responses were stored in a SQL database and loaded into R and SPSS for analysis. Before the data could be analysed it needed to be cleaned, and this one done using stored procedures in SQL. Some inputs were not formatted, such as the answer to the bat and ball problem where some people would answer R0.10 and some would answer 10 and some 10c or 10 cents and all of these were converted to 10. During this phase, the tests such as the cognitive reflection test were also scored.

3.7.1 Tests for bias

A limitation of this study is that Armstrong and Overton's (1977) test for nonresponse bias was not employed in order to compare early and late respondent's scores on included constructs, however there should not be any difference between early and late responders due to the way that the respondents were contacted. Self-selection bias could also apply where only a certain type of individuals or companies would agree to take part in the study. Common method variance (CMV) is variance that is attributable to the method of measurement rather than the constructs that the method was intended to measure (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003), and this was not analysed in this study.

Group differences

For the sub-variables of decision accuracy, different tests were used to determine significant differences between groups, depending on whether the data were normally distributed or not.

Non-parametric versions of ANOVA tests (Mann-Whitney U test) were used to check for significant differences between groups where the data were not normally distributed, and studentised t-tests were used where the data did follow a normal distribution. In all cases all assumptions for each test were validated before the tests were performed.

3.7.2 Hypothesis testing

The hypotheses were tested using regression. All assumptions were tested before the regressions were performed, and these are reported in the results chapter. First, to assess linearity, a scatterplot of the outcome variable (decision accuracy) against the independent variables and interaction variables was plotted and visual inspection of the plot indicated a linear relationship. Next, homoscedasticity was assessed by visual inspection of a plot of standardised residuals versus standardized predicted values, and residuals were checked to be normally distributed by visual inspection of a normal probability plot. For each test, outliers

were checked by visual inspection of boxplots and also Shapiro-Wilk tests of normality. Outliers were dealt with in the sub-variables and no outliers were found during the hypothesis testing. Based on the study design there was no need to verify the independence of residuals with the Durbin-Watson statistic, although these were checked and recorded anyway. Once all assumptions were checked, the regressions were performed and the results reported.

3.8 Quality

This section deals with aspects of research quality.

3.8.1 Measurement validity

Measurement validity refers to the operationalisation of abstract concepts and the degree to which the measures of the variables faithfully represent the constructs they are intended to measure (Gill & Johnson, 2010).

After the research was performed, specific tests were used to ensure that the instruments did measure the constructs they were designed to measure. An exploratory factor analysis was performed to confirm construct validity on the faith in intuition variable from the faith in intuition scale. The faith in intuition five item scale loaded strongly onto a single factor – showing internal consistency reliability where all the items in a single instrument yield similar results (Leedy & Ormrod, 2014).

3.8.2 Reliability

According to Gill and Johnson (2010), reliability refers to the consistency of results obtained from the research and the extent to which replication of the research would yield the same results. The use of existing validated measures aided in ensuring reliability. Leedy and Ormrod (2014) refer to test-retest reliability as the extent to which a single instrument will

yield the same results for the same people on different occasions. The cognitive reflection test measure (Frederick, 2005) however, relies on the subject not having completed the test previously, since recalling the solution could result in a different answer.

3.8.3 Internal Validity

Gill and Johnson (2010) describe the importance of internal validity and how this can affect the ability to draw legitimate conclusions about cause-and-effect relationships. This section will discuss internal validity in the context of the present study. Internal validity is impacted by the extent to which the design and data yield this validity by eliminating the possibility of confounding variables providing other possible explanations for the measured effects (Gill & Johnson, 2010). Even the methodology employed which manipulates use each mode of thinking under the chosen dual-process theory framework, makes assumptions based on this framework. As such the internal validity is limited by the methodology and its ability to control for other explanations. Within the framework, other steps can be taken, and Leedy and Ormrod (2014) offer suggestions for countering threats to internal validity within experimental designs:

1. "*Keeping some things constant*" (p. 236): By restricting the sample to senior Information Technology decision makers of listed and private companies, the present study limited the possibility that other variables such as socioeconomic status or domain expertise could explain the measured effects. However, this restriction did also lower the *external validity* or ability to generalise the findings to other contexts (Leedy & Ormrod, 2014).
2. "*Include a control group*" (p. 236): The study included a *control group* where the use of intuition was not forced through manipulation of timing or distraction and rational deliberation was not encouraged. This allowed comparison of the results to those of

the *experimental groups* which received the interventions. However, the experimental cases were designed so that the control group was unaware of their lack of intervention, to ensure that the *reactivity* from knowing they are part of an experiment was equivalent across the groups and performance on the tasks was not influenced differently by the *Hawthorne effect* (Leedy & Ormrod, 2014).

3. "*Randomly assign people to groups*" (p. 237): In addition to keeping some variables constant by using the senior Information Technology decision maker sample mentioned in point 1, by randomly assigning people to the experimental and control groups, the researcher could make the assumption that the variation in the groups should be similar and that differences should be due only to chance (Leedy & Ormrod, 2014). This minimised the effect of other individual differences between the subjects which could affect the outcome and serve as confounding variables, by attempting to ensure that each group should have a similarly random mix of these and that the groups were therefore comparable. Leedy and Ormrod (2014) point out that random assignment is critical in the post-test-only design because without the random assignment it becomes the same as the static group comparison pre-experimental design which cannot be used to infer cause and effect relationships.
4. "*Assess equivalence before the treatment with one or more pretests*" (p.238): The study measured faith in intuition because it formed part of the research model for testing. Although this variable was an experimental pre-test because it was not affected by the treatment, it did offer the ability to comment on the equivalence of the groups in terms of this measure and on the effect on internal validity. The study also used the cognitive reflection test as a measure of a general tendency toward susceptibility to bias, not only as a control in the correlational analysis, but again also

as a method of assessing the group equivalence and commenting on the threat to internal validity.

5. "*Expose participants to all experimental conditions*" (p. 239): This involves each participant being exposed to the control treatment and experimental treatments, known as a within-subjects design (Leedy & Ormrod, 2014). Although this would be very useful in ruling out other individual differences between subjects as potential confounding variables, the design would have been overly complicated to avoid the effects of exposing the same individual to multiple similar tests in a single experiment, and the researcher felt that the random group assignment was more suited to the requirements of the study for the use of intuition. Therefore, the study implemented a ***between-subjects design*** for use of intuition. Although complexity could be affected by other unmeasured individual differences such as intelligence and working memory capacity, the benefits of random assignment and preventing issues with repeated tests (cases) for one subject were more important and therefore complexity was experimentally varied as a **between-subjects variable**.
6. "*Statistically control for confounding variables*" (p. 239): As discussed under point 4 – the requirements of the research model and the measurement of faith in intuition and cognitive reflection test also allowed for statistical control of at least two main potential confounding variables, and other standard demographic variables such as age, level of education, gender, and industry were measured for this purpose.

3.8.4 External validity

External validity refers to the extent to which the results of the study can be generalised to other contexts (Leedy & Ormrod, 2014). This section will discuss external validity in terms of the present study.

Although the controlled nature of the experimental method allowed improvement of internal validity by controlling confounding variables, the artificial nature of experiments, in general, limits the generalisability of the results to real-world contexts (Leedy & Ormrod, 2014). Gill and Johnson (2010) refer to this as ecological validity. The researcher acknowledged and accepted this limitation because of the trade-off required in terms of the nature of the phenomenon being studied where the threat to internal validity in field studies would be significant.

Leedy and Ormrod (2014) note that research which is restricted to people with a particular set of characteristics limits the generalisability to people with different characteristics, referred to by Gill and Johnson (2010) as population validity. The selection of senior Information Technology decision makers of listed and private companies in South Africa, although highly applicable in the context of the study, also limited the ability to generalise the results to other populations in that senior Information Technology decision makers may exhibit different characteristics. In the survey-experiment the demographic data was captured at the end of each survey, and so it was not possible to analyse the drop-offs to ensure that the resulting sample was not biased in relation to particular characteristics - this is a limitation of this study as it could threaten the external validity.

3.9 Ethics

The research was conducted in accordance with the approved research protocol. Informed consent was obtained from participants before they took part in the research, and it was made clear that participation is voluntary and fully confidential. Participants were also advised of the purpose, expected duration, and procedures included in the research, and their right to withdraw from the research at any time. Participants were offered debriefing following the study, and the opportunity to obtain the results of the research on completion of the study was be provided. All participants were made aware of the experimental manipulation after the study. The solution of any distraction tasks such as anagrams were provided after application of the experiment.

All data collected were encrypted and stored and archived in a manner which would protect from accidental disclosure. The study was non-intrusive, and did not endanger the participants physically or emotionally. The individual consent form for the survey experiment is included in the questionnaire in the appendix.

The researcher was directly involved with compiling data from questionnaires, and could have been exposed to sensitive information. Respondents could include customers, suppliers, or competitors of the researcher's organisation and this could lead to a conflict of interest – however the researcher is bound by ethical standards, including non-disclosure of all information obtained, and agreement to only use the information for academic purposes.

3.10 Summary and conclusion

This section has presented the design and methodology of the present research, and has also discussed potential limitations. Quality and ethics were also discussed, and motivations for each decision were presented.

The next section will present the results of the research that was conducted.

Chapter 4 – Results

4.1 Introduction

The main focus of this research was to determine the impact of complexity on the effectiveness on intuitive decision making. The research attempted to discover the conditions under which intuition would be more, or less, effective than rational decision making, in a software selection context.

The results obtained are described below, beginning with description of the sample, and then showing the results of testing the hypotheses. A more detailed account is presented in the appendices, with full details for each variable.

4.2 Sample description

The data were collected through a combination of in-person and remote applications of an online survey-experiment. All the companies listed on the Johannesburg Stock Exchange, and over 100 additional companies were invited to participate in the research. In total 139 unique surveys were generated for respondents that agreed to participate and the data were collected over a 19-month period from 19 April 2016 to 22 November 2017. Of these, nine of the surveys were generated after respondents agreed to participate, but were never started. Of those that were successfully started, 5.8% of respondents stopped during the first three questions - the cognitive reflection test. A further 18.7% of respondents completed the cognitive reflection test but stopped before completing the entire case study section. All the respondents that got as far as completing the case study also completed the faith in intuition scale and the demographic section at the end. This resulted in 92 fully completed survey-experiments. One of these was disqualified because the subject did not complete the

experiment in one sitting, and one was removed because of clearly invalid answers. This left 90 fully completed and valid survey-experiments, and this forms the analysis dataset.

The simple and complex versions of the case had similar average completion times for the survey. There was more of a difference in the average times taken depending on the use of intuition variable, with the distracted group having the longest average completion times.

Of the companies surveyed, 56% were listed on the Johannesburg Stock Exchange. Most of the companies (46%) had more than 1000 employees, and 16% had between 500 and 1000 employees. In terms of industry, 38% of the companies were in the Information Technology sector, 21% in Financial Services, 11% were in Manufacturing, and the remainder were split between various other sectors.

In terms of gender, only 10% of the respondents were female, so it was not possible to analyse gender differences. For highest academic qualifications, 4% listed doctorates, 22% had master's degrees, 27% had bachelor's degrees, and a further 27% had diplomas. In terms of age, 24% of respondents were over 50, while most (46%) of the respondents were in the 41-50 age group category. Of the survey respondents, 29% were c-level executives, 42% were in senior management, 19% were middle management, and the rest were split across other job levels. Of these, 91% had more than 10 years of work experience, and 84% had more than 10 years of Information Technology experience. Most of the respondents listed English as their primary language (58%), followed by Afrikaans (27%), with the remaining 15% split across Northern Sotho, Swazi, Tswana, Venda, Xhosa, and Zulu.

4.3 Faith in intuition scale

The 5-item faith in intuition scale was included in the survey to determine a predisposition toward trusting in, and use of, intuition. This scale is part of the shorter ten-item version of the Rational Experiential Inventory scale (Epstein et al., 1996).

Prior to performing a factor analysis, the correlation matrix was examined and most correlation coefficients were above 0.3. The Kaiser-Meyer-Olkin Measure of Sampling Adequacy was 0.765 which exceeds the recommended value of 0.6 (Kaiser, 1970, 1974), and Bartlett's Test of Sphericity (Bartlett, 1954) reached significance ($p<0.001$), providing support for the factorability of the correlation matrix. An exploratory factor analysis with principle axis factoring and Varimax rotation was performed and showed one main factor, using Kaiser's criterion, with an eigenvalue exceeding 1 and explaining 54% of the total variance. A strong bend in the Screeplot of Catell's scree test at the second factor also confirmed the single factor loading (Catell, 1966). The results of this analysis support the use of this scale as measuring the single construct of faith in intuition.

The scale had an acceptable level of internal consistency, as determined by a Cronbach's alpha of 0.78, with DeVillis (2003) recommending values of 0.7 and higher, and previous research with this scale showing levels of 0.72 for this subscale (Epstein, et al., 1996). A Shapiro-Wilk test of normality showed a normal distribution ($p=0.06$).

4.4 Cognitive reflection

The 3-item version of the cognitive reflection test was included to determine a baseline tendency to bias. Although in total there were 118 completed cognitive reflection test entries, the incomplete demographic information meant that only the 90 complete and valid records

could be used. The Shapiro-Wilk test of normality showed that the results do not follow a normal distribution ($p<0.01$).

4.5 Use of intuition

Use of intuition was an experimentally controlled factor. The experiment involved manipulations to force rational thinking and two others to force intuition.

The intuitive conditions were split between the timed-intuitive, and distracted-intuitive interventions with a smaller number of respondents in the distracted group. The distracted group had a higher drop-off rate, some of which was due to the survey-experiment taking much longer if they delayed often and had to answer many anagrams. Since the average completion time in the pilot was under 25 minutes, the appointments were usually booked for 30 minutes and some subjects in the distracted group ended up running out of time because of repeated application of the distraction anagrams.

Further analysis was then done to see if the groups could be collapsed. There were 25 subjects in the timed group and 9 in the distracted group. Since the data were normally distributed, an independent-samples t-test was run to determine if there were differences in decision accuracy between the timed and distracted groups. There were no outliers in the data, as assessed by inspection of a boxplots. Decision accuracy scores for each group were normally distributed, as assessed by Shapiro-Wilk's test for the timed group ($p = .198$) and for the distracted group ($p = .454$), and there was homogeneity of variances, as assessed by Levene's test for equality of variances ($p = .783$). Analysis of the distracted group decision accuracy score ($M = .503$, $SD = 0.211$) and timed group decision accuracy score ($M = .370$, $SD = 0.213$), resulted in a non-significant difference of 0.134 (95% CI [-0.302 to 0.034]), $t(32) = -1.622$, $p = .115$, $d = 2.970$.

Since there was no significant difference in the decision accuracy between these interventions, these were collapsed into a single intuitive group in order to increase the groups sizes and the power of the statistical analysis.

The rational and control groups also showed very little difference, which shows that the rationality intervention did not have any significant impact compared to the control group. The rationality intervention only consisted of a brief sentence just before the case study: “Please consider the case below carefully because you may be asked to explain the answers”. This might not have been strong enough and would probably lose effect more and more as the questions progressed since it was only presented at the start of the case.

Analysis was performed to determine if the rational group was different to the control group. There were 25 subjects in the control group and 31 in the rational group. Since the data were normally distributed, an independent-samples t-test was run to determine if there were differences in decision accuracy between the rational and control groups. There were no outliers in the data, as assessed by inspection of a boxplots. Decision accuracy scores for each group were normally distributed, as assessed by Shapiro-Wilk's test for the rational group ($p = .907$) and for the control group ($p = .804$), and there was homogeneity of variances, as assessed by Levene's test for equality of variances ($p = .166$). Analysis of the rational group decision accuracy score ($M = .523$, $SD = 0.211$) and control group decision accuracy score ($M = .514$, $SD = 0.269$), resulted in a non-significant difference of 0.009 (95% CI [-0.119 to 0.138]), $t(54) = 0.145$, $p = .886$, $d = 0.16$.

Many of the respondents commented that the experiment felt like a test and that they were nervous that they may not have given the correct answers. This points strongly to the context encouraging a rational approach where intuition was not forced, and would explain why the

rational intervention was not significantly different from the control, with the context encouraging rationality much more than the simple statement at the start of the case.

Because of this, and because there was no significant difference between the rational and control group decision accuracy, the rational and control groups were collapsed into one rational group, in order to increase the groups sizes and the power of the statistical analysis.

4.6 Decision complexity

The decision complexity variable was experimentally controlled. There were two experimental groups - a simple group and a complex group, and the final results were well balanced, with exactly 50% of cases in the complex condition.

4.7 Decision accuracy

The decision accuracy variable was calculated as the additive inverse of the sum of the attribute substitution, anchoring bias, and general bias measures of bias: *decision accuracy = 1 - (attribute substitution + anchoring bias + general bias)* (see Appendix B for detailed calculations).

The decision accuracy variable did not contain outliers and a test of normality showed that the variable did follow a normal distribution (p=0.49).

4.8 Hypotheses 1

> There will be a negative relationship between use of intuition and decision accuracy, such that rational decisions will be more accurate than intuitive decisions.

A linear regression was performed to understand the effect of use of intuition on decision accuracy. To assess linearity, a scatterplot of decision accuracy against use of intuition was plotted and visual inspection of the plot indicated a linear relationship. There was homoscedasticity, as assessed by visual inspection of a plot of standardised residuals versus standardized predicted values, and residuals were normally distributed as assessed by visual inspection of a normal probability plot. There were no outliers, and based on the study design there was no need to verify the independence of residuals with the Durbin-Watson statistic of 1.79. No outliers were identified. Use of intuition statistically significantly predicted decision accuracy, F(1, 88) = 5.23, p = .025, accounting for 5.6% of the variation in decision accuracy with adjusted R^2 = 4.5%, a small size effect according to Cohen (1988). The resulting equation was: decision accuracy = 0.519 – 0.114 x use of intuition.

Therefore, the decision accuracy for the Intuitive group is significantly lower than the Rational group and the null hypothesis can be rejected, and the alternative hypothesis for hypothesis 1 can be accepted.

Therefore, in the software selection context, intuitive decisions are significantly less accurate than rational decisions.

4.9 Hypothesis 2

Decision complexity will moderate the relationship between use of intuition and decision accuracy, such that the effect of using intuition will be greater for simple decisions than for complex decisions.

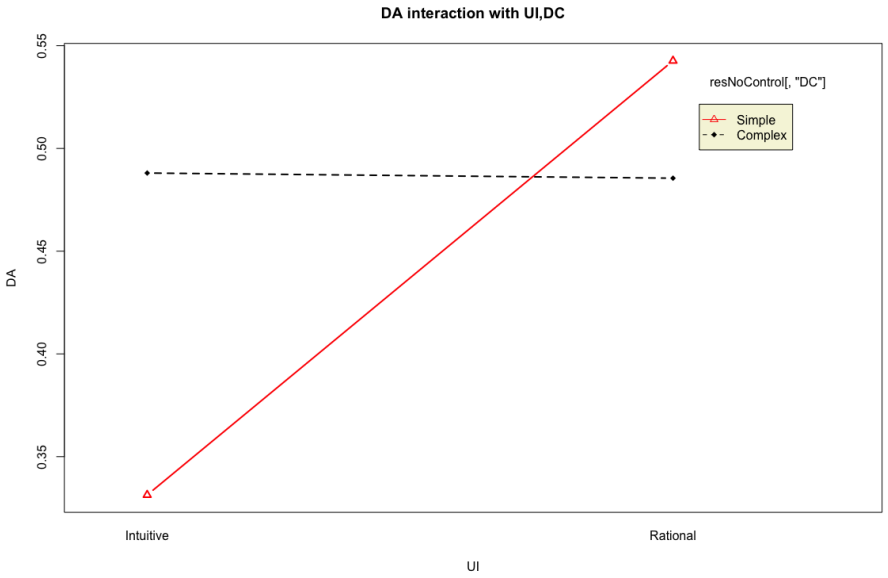


Figure 7. Interaction plot of decision accuracy by use of intuition and decision complexity.

Since the difference between the use of intuition groups is only significant in the Simple decision complexity group, it seems to imply interaction. From the interaction plot, visually it appears that there is a moderating effect (non-parallel, crossing lines).

A hierarchical multiple regression was run to assess the statistical significance of the interaction term between use of intuition and decision complexity. Linearity was established by visual inspection of a scatterplot and there was no evidence of multicollinearity since all tolerance were greater than 0.401. No unusual points were identified. There was homoscedasticity, as assessed by visual inspection of the studentised residuals plotted against the predicted values for simple and complex decisions. The studentised residuals were normally distributed, as assessed by Shapiro-Wilk's test ($p = .538$, $p = .736$). The coefficient

of the interaction term (b = 0.227, SE = 0.098) was statistically significant (p < .023) indicating that decision complexity moderated the relationship between decision accuracy and use of intuition. In addition, the hierarchical regression showed that the addition of the interaction term explained an additional 5.5% of the total variance (F(3, 86) = 3.633, p = .016). Simple slopes analysis revealed that there was a statistically significant negative linear relationship (b = -0.224, SE = 0.069) between use of intuition and decision complexity in simple decisions, p < .002, but not in complex decisions (b = 0.003, SE = 0.07), p = .970.

Therefore, the null hypothesis must be rejected, and the alternative hypothesis for Hypothesis 2 is supported – decision complexity does moderate the relationship between use of intuition and decision accuracy. While intuitive decisions were shown in hypothesis 1 to have a lower accuracy overall than rational decisions, this moderating relationship shows that while intuition was less accurate for simple decisions, for complex decisions there was no significant difference between intuitive and rational decisions. So both rational and intuitive decisions would yield acceptable software selection results.

4.10 Hypothesis 3

> Faith in intuition will moderate the relationship between use of intuition and decision accuracy, such that higher faith in intuition will result in lower accuracy for intuitive decisions.

A hierarchical multiple regression was run to assess the statistical significance of the interaction term between use of intuition and faith in intuition on decision accuracy. There was no evidence of multicollinearity, as evidenced by no tolerance values less than 0.546 on

the mean centred variables. No unusual points were identified. There was homoscedasticity, as assessed by visual inspection of the studentized residuals plotted against the predicted values. The studentised residuals were normally distributed, as assessed by Shapiro-Wilk's test ($p > .05$). The coefficient of the interaction term ($b = -0.104$, $SE = 0.086$) was not statistically significant ($p = .229$) indicating that faith in intuition did not moderate the relationship between decision accuracy and use of intuition.

Therefore, the null hypothesis must be accepted, and the alternative hypothesis for Hypothesis 3 is not supported – faith in intuition does not moderate the relationship between use of intuition and decision accuracy, and so, in this sample, the individual difference measure of faith in intuition did not show any significant impact on the relationship between use of intuition and decision accuracy. In other words, believing that "gut feel" works and trusting in your intuition, did not affect the accuracy of software decisions.

4.11 Hypothesis 4

> Cognitive reflection will moderate the relationship between use of intuition and decision accuracy, such that higher cognitive reflection will result in higher decision accuracy for intuitive decisions.

A hierarchical multiple regression was run to assess the statistical significance of the interaction term between use of intuition and cognitive reflection on decision accuracy. There was no evidence of multicollinearity, as evidenced by no tolerance values less than 0.441. No unusual points were identified. There was homoscedasticity, as assessed by visual inspection of the studentized residuals plotted against the predicted values. The coefficient of cognitive

reflection (b = 0.074, SE = 0.021) was significant (p = .001), however, the coefficient of the interaction term (b = -0. 084, SE = 0.051) was not statistically significant (p = .100) indicating that cognitive reflection did not moderate the relationship between decision accuracy and use of intuition.

Therefore, the null hypothesis must be accepted, and the alternative hypothesis for Hypothesis 4 is not supported – cognitive reflection does not moderate the relationship between use of intuition and decision accuracy.

4.12 Conclusion

This section has presented the results of the research that was conducted. Hypothesis 1 and hypothesis 2 were fully supported, but hypothesis 3 and 4 were not supported.

The following figure summarises these results, showing the significance testing with respect to the original proposed model. Those relationships that were not significant are shown in grey.

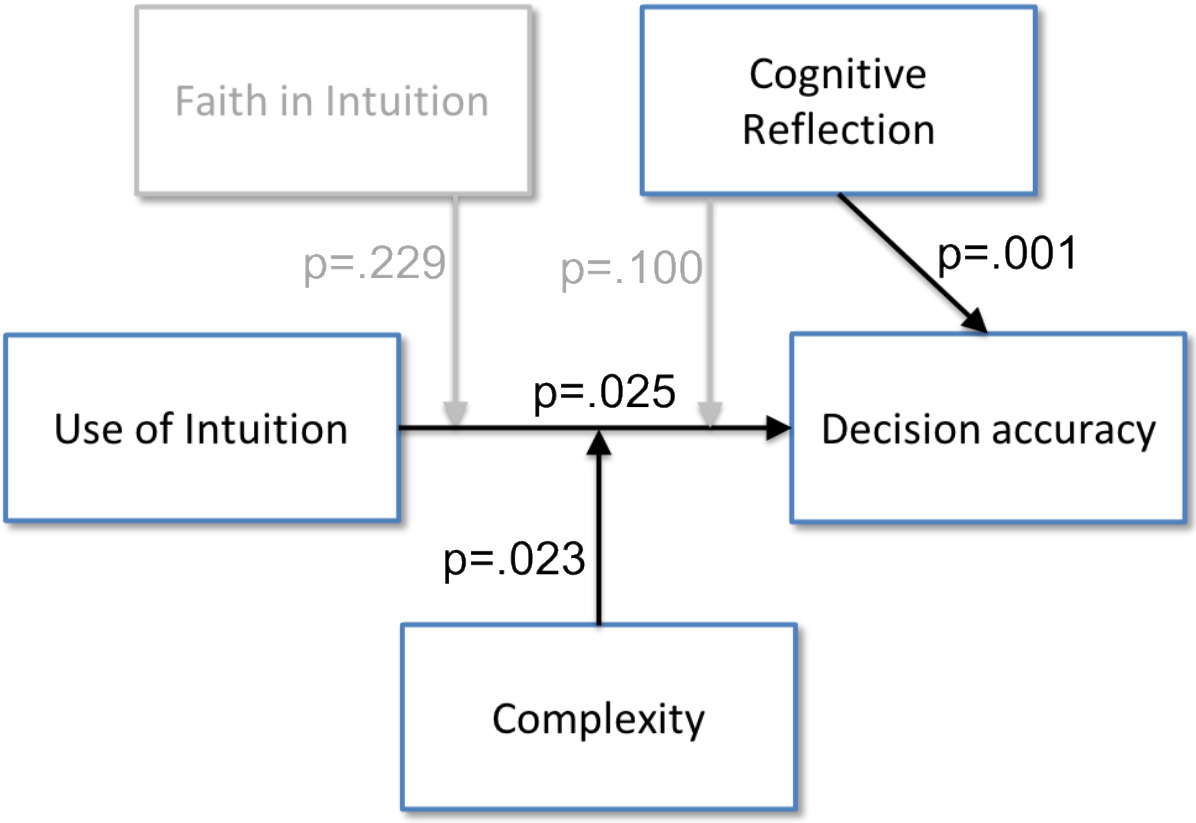


Figure 8. Summary of results of analysis.

The next section will discuss these results with respect to the literature.

Chapter 5 - Discussion

5.1 Introduction

Since the literature on the use of intuition offers contradictory views as to how effective it can be for business decision making, and barely addresses the software context, this research was motivated in terms of further determining the conditions under which intuition can provide accurate decisions, especially in a software decision context. Previous literature had pointed to complexity and individual differences as potential areas to examine, and so this research focused on these in relation to intuitive decision making.

The overall research question was: “What is the impact of complexity, and individual differences, on the effectiveness of intuitive decisions?”, and the research objectives are listed below:

1. Determine if there is a relationship between use of intuition and decision accuracy, and whether intuitive decisions are more or less accurate.
2. Determine whether decision complexity moderates the relationship between use of intuition and decision accuracy.
3. Determine whether faith in intuition moderates the relationship between use of intuition and decision accuracy.
4. Determine whether cognitive reflection moderates the relationship between use of intuition and decision accuracy.

This discussion chapter is structured around discussing the support for the hypotheses that address each of the objectives.

5.2 Research Objective 1 – The relationship between use of intuition and decision accuracy

The literature is divided on whether intuition can provide accurate decisions, and since the decision problem was designed to test bias as a reverse proxy for accuracy, and since there is literature that points to the bias inherent in intuitive decision making, the corresponding hypothesis for this research objective was stated as:

Hypothesis 1:

> There will be a negative relationship between use of intuition and decision accuracy, such that rational decisions will be more accurate than intuitive decisions.

The results showed that there was indeed a significant difference between the intuitive and rational groups for the decision accuracy variable, and decision accuracy was significantly lower for the intuitive group than for the rational group. This provided support for Hypothesis 1 – that rational decisions would be significantly more accurate than intuitive decisions.

Since this research used bias measures as a proxy for decision accuracy, this result corresponds with the heuristics and biases literature that suggested that intuitive judgement is more subject to bias (Kahneman, 2011).

This result was also in line with assumptions presented in the software selection literature that proposed that intuition would be an unacceptable decision method (Carney & Wallnau, 1998), and a source of decision bias (Jamieson & Hyland, 2006).

However, the result is in opposition with studies that demonstrated cases where intuition outperforms rational deliberation (Dijksterhuis, 2004; Usher et al., 2011). This could be because these studies used very simple decision problems with pre-decided normatively correct outcomes, whereas the present research purposefully avoided these, as discussed in the methodology section, by using bias as a reverse proxy. When the experimenter pre-decides the correct outcome, it will be based on their own weighting of the various attributes and these are mostly not made explicit. An example is where different cars are presented and the experimenter decides ahead of time that Car A should be chosen because it has the lowest fuel consumption and the longest service period, whereas the subject might value a different attribute, such as visual appeal, much higher.

The result is also in opposition to the literature that suggests that many other types of decisions, such as chess moves, are best made intuitively (Klein, 2008). However, these decisions are usually much more complex than software purchase decisions and are usually made under a lot more pressure than a simulated laboratory case-study type experiment where the stakes are significantly lower. Also, although this research did introduce complexity as a factor, even the simple case used in this study was more complex than the simple car choice type problems in previous studies (e.g. Dijksterhuis, 2004), and most likely less complex than chess moves. Future research could examine more levels of complexity to determine where the boundaries lie.

This significant result does also add to the body of literature that has used experimental techniques such as the ones employed in this study in order to elicit intuitive responses, or to encourage rational ones, further strengthening the support for these. However, based on the experience with this sample, the distraction technique should be used with care and with shorter experiments in more controlled settings – since this mode resulted in the most partially completed experiments because of respondents becoming frustrated and terminating

the experiment, corresponding to similar results found in other studies (e.g. Meszaros, 2007).

5.3 Research Objective 2 – decision complexity as a moderator

While certain studies, most notably those relating to unconscious thought theory, demonstrated cases where intuition can outperform rational deliberation in certain situations (Dijksterhuis, 2004; Dijksterhuis & Nordgren, 2006; Usher et al., 2011), there were also calls for investigating complexity as a reason for contradictory results (Dijksterhuis, 2004; Meszaros, 2007). These studies cited psychological research showing that rational deliberation is limited by the capacity of working memory and they proposed that intuitive decisions are not subject to these limitations. While addressing various methodological concerns with this type of research, this study sought to determine the relevance of complexity in this context, and so the corresponding hypothesis for this research objective was stated as:

Hypothesis 2:

> Decision complexity will moderate the relationship between use of intuition and decision accuracy, such that the effect of using intuition will be greater for simple decisions than for complex decisions.

The results showed that complexity did indeed moderate the relationship between use of intuition and decision accuracy and Hypothesis 2 was supported, with the interaction term significant ($p = 0.02$) and also with the addition of the moderator significantly improving the model fit in the regression ($p = 0.02$).

The overall result of use of intuition interacting with decision complexity on decision accuracy, is in line with the propositions made by those who argued that complexity plays a role in the effectiveness of intuitive decisions (Dijksterhuis, 2004; Dijksterhuis & Nordgren, 2006; Meszaros, 2007). Meszaros (2007) also found similarities in performance in their complex group and large differences in their simple groups.

Some researchers have raised methodological issues with the work around unconscious thought theory (Usher et al., 2011), and others have failed to replicate the results (Abbot, 2015; Acker, 2008; Calvillo & Penaloza, 2009; Nieuwenstein et al., 2015). However, there does seem to be validity in the calls for investigation into complexity, while addressing these methodological concerns as has been done in this study.

As far as the researcher is aware, this is the first research to have reproduced this effect without using a normative test-type experiment. This lends further support for the use of bias as a reverse-proxy, and this methodology should be tested further in other contexts.

This research has added significantly to research on intuition by showing that complexity does indeed play an important role when considering the effectiveness of intuitive decisions.

Counter-intuitively, there is a lot more difference in accuracy when the decisions are simpler, with less attributes. In simple decisions, with only 7 attributes, this research showed that rational approaches are far less biased and therefore far more accurate. Although the complex decisions, with 14 attributes, did not show intuition performing better than rational thought as some other studies have shown, they did show that intuitive decision accuracy was indistinguishable from rational decision accuracy in complex conditions. This finding therefore supports the use of either type of decision strategy for complex problems.

5.4 Research Objective 3 – faith in intuition as a moderator to use of intuition and decision accuracy

Some research has shown that increased faith in intuition is correlated with an increase in bias (Keller & Bless, 2009; Shiloh et al., 2002), and reduced decision accuracy, and others have shown that overconfidence in intuition leads to an increase in bias (Kahneman, 2011). Keller and Bless (2009) argued that faith in intuition moderates ease of retrieval, affecting an anchoring bias. The hypothesis for this part of the research was stated as:

Hypothesis 3:

> Faith in intuition will moderate the relationship between use of intuition and decision accuracy, such that higher faith in intuition will result in lower accuracy for intuitive decisions.

The results for this sample showed no significant relationship between faith in intuition and decision accuracy overall, or in any of the experimental groups relating to intuition, and the results showed that faith in intuition did not moderate the relationship between use of intuition and decision accuracy. Therefore, hypothesis 3 was not supported.

De Neys and Bonnefon (2013) argued that the individual differences in reasoning may only present themselves at a late stage in the reasoning process, and this would mean that the interventions designed to force intuition in this study, including time limits, could prevent getting to the stage where those differences would present themselves. However, if that were the case then there could still have been a relationship in the groups where intuition was not forced, and that was not the case in this study.

A more likely explanation could be that faith in intuition measures not only how much you trust your intuition, but also how likely you are to use it, then this would explain why no relationship was found in this study – since the use of intuition was experimentally controlled – and the subject's intention for use would be mostly irrelevant. It would also mean that the trust itself is not related to how accurate the resulting decisions are, but that rather that the trust could cause the use of intuition and then that would result in a particular level of accuracy or bias. If this is the case then it could be argued that use of intuition mediates the relationship between faith in intuition and accuracy, and this could be tested in further research.

This result contradicts existing theory, but an explanation has been provided that could explain why. Experimentally controlling faith in intuition could have prevented a relationship between faith in intuition and decision accuracy. From a practical standpoint, the results of this study would indicate that faith in intuition does not affect decision accuracy. However, as discussed, use of intuition could mediate the relationship, in which case, further research should be done before practical principles are decided in terms of the effect of faith in intuition.

5.5 Research Objective 4 – cognitive reflection as a moderator to use of intuition and decision accuracy

Hypothesis 4:

> Cognitive reflection will moderate the relationship between use of intuition and decision accuracy, such that higher cognitive reflection will result in higher decision accuracy for intuitive decisions.

The results showed that cognitive reflection did not moderate the relationship between use of intuition and decision accuracy. However, there was a significant direct relationship between cognitive reflection and decision accuracy ($p = 0.001$), and this is in line with previous studies which found that cognitive reflection predicted performance on bias tasks even after cognitive ability and executive functioning were controlled for (Toplak, West, & Stanovich, 2011). Because of this strong direct relationship, the addition of the interaction term did not add significantly to the fit of the linear regression and so no moderating effect was found. Overall, cognitive reflection had a very strong impact on how accurate decisions were, despite the intuitive manipulations. Practically, the results suggest that a measure of cognitive reflection will significantly predict decision performance in real world situations. This result could mean that cognitive reflection could be more important than use of intuition when predicting decision bias and accuracy, and again provides an interesting area for more research. Potentially, high cognitive reflection individuals could be left to use intuition or rational approaches due to lower levels of bias.

5.6 Summary of discussion

This chapter discussed the results and findings of the research and the relationship to existing literature. Hypothesis 1 was fully supported and there were interesting findings which opposed existing literature that promotes intuition as being superior to rational deliberation. Hypothesis 2 was also fully supported and again in opposition to current literature reporting intuitive decisions as being superior, but there was support for criticisms of that work, and intuition proved similar in accuracy when complexity was high. Hypothesis 3 and 4 were not supported, but cognitive reflection did show a strong direct relationship to decision accuracy.

The following figure presents the updated model. This model has been adjusted, based on the original proposed model, to reflect the relationships that were found in the sample.

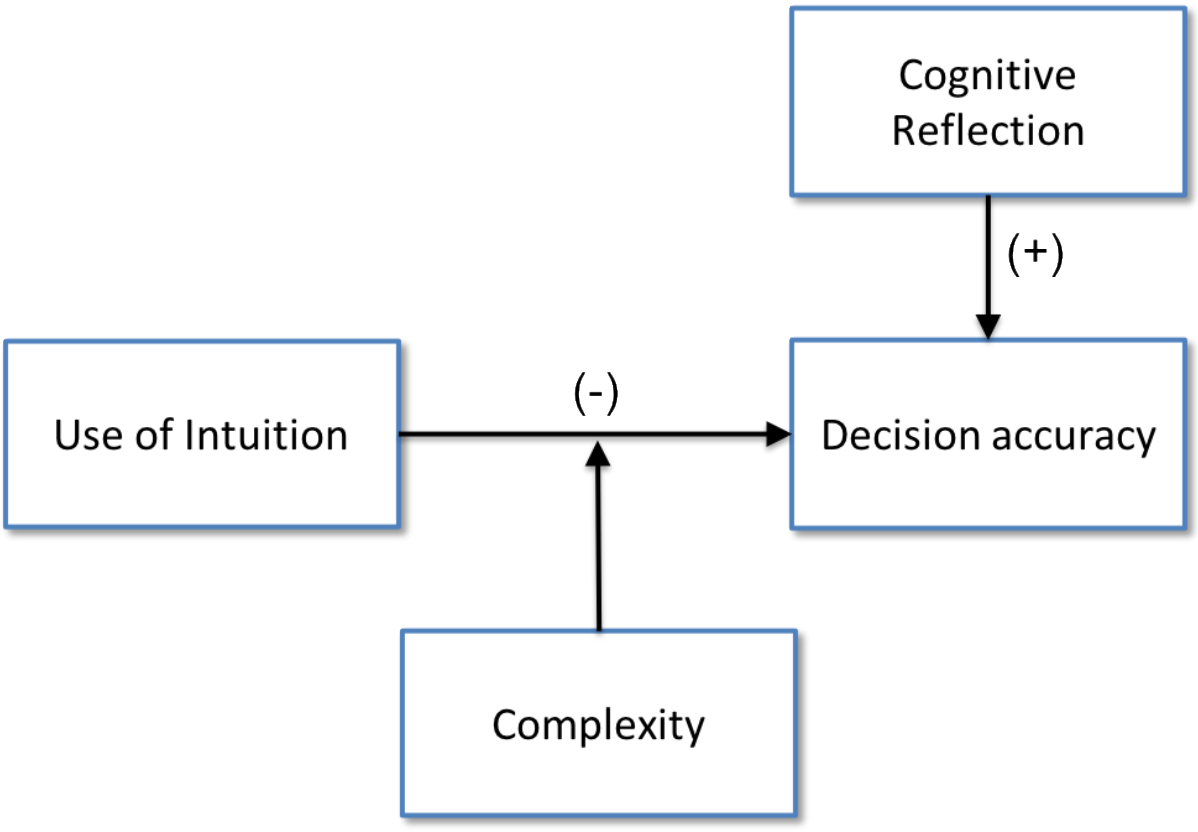


Figure 9. Model summary matching findings.

Chapter 6 - Conclusion

6.1 Introduction

The existing literature on the use of intuition offers contradictory views as to how effective it can be for decision making. This study investigated the impact of complexity on the effectiveness of intuition in decision making, in order to find out what the conditions are under which intuitive decisions could be accurate. Tensions in the literature prompted analysis of the boundary conditions, and particularly complexity and individual differences, since there are proponents arguing respectively in favour of intuitive and rational decision making, while some research supports both.

An experiment was applied in survey form using a case study, in a software selection context, with senior information technology decision makers. Using dual-process theory as the lens, and as a basis for the experimental methodology, subjects were randomly assigned to intuitive, rational, or control groups – and to simple and complex versions of the same case study.

The research made an original contribution by investigating complexity as a moderator, by examining the role of individual differences such as faith in intuition and cognitive reflection, and by utilising bias measures as a proxy for decision accuracy in order to avoid the normatively correct choice measures that have been highlighted for their limitations.

The results showed that intuitive decisions were somewhat less accurate in this context. The results also showed that complexity does moderate the relationship between use of intuition and accuracy. This could imply that information technology managers, when making decisions about software, should in general rather use a rational than a “gut” feel approach where the decisions are simple. However, the results of this research suggest that with more

complex decisions the accuracy could be similar between rational and intuitive decisions, and most practical software decisions should fall into the latter category.

The following sections will detail some of the contributions of the study, discuss the limitations, and offer suggestions for future research.

6.2 Contribution

6.2.1 Theoretical contribution

As indicated in the problem statement, there is a gap in the knowledge around the conditions under which intuition can provide accurate decisions. This research has contributed to filling that gap by considering complexity and individual differences as conditions. The results have shown that complexity does have a moderating effect on the relationship between use of intuition and decision accuracy.

This research also contributed to existing research using Dual Process Theory as a lens, by using experimental interventions for controlling intuition. The significant results achieved strengthen support for the theory and for the methodology.

In addition, this research has contributed to the debate on whether intuitive or rational decisions would be more accurate. In the software selection context where this was tested, the result was that rational decisions are more accurate overall, but this is moderated by the complexity of the decision. This moderating relationship resulted in rational decisions being significantly more accurate in simple decisions, with almost no difference when the decisions were more complex.

While unconscious thought theory has been criticised for methodological concerns and spurious results from small sample sizes (Nieuwenstein et al., 2015), this study has shown a

clear difference in performance between rational and intuitive decisions for simple tasks, and no significant difference for complex tasks. This has been achieved without using the period of deliberation that the unconscious thought theory paradigm requires, and while addressing various concerns regarding unconscious thought theory – such as memory effects (where subjects forget the attributes if they are only presented initially). Most importantly, this result has been achieved while using bias as a proxy for decision accuracy, thereby addressing all concerns around experimenter-generated normatively correct outcomes in test-like settings, which do not take into account the very likely difference in weighting of attributes between the experimenter and the subject. This last aspect forms the most significant methodological contribution of this study, which will be discussed further in the next sub-section.

6.2.2 Methodological contribution

The main methodological contribution is that, this is the first study that has used bias measurements as a reverse-proxy for decision accuracy when testing the effectiveness of intuitive decisions. As mentioned in the previous section, this methodological change addresses all of the criticism about normatively correct outcome testing for decision analysis - because the bias in the decision is being measured, instead of the experimenter deciding ahead of time which answer is correct. As discussed in the methodology chapter this prevents issues where the experimenter's pre-decided answers might not be correct for the subject, because their attribute weightings may be different. It also creates greater variability in a single case because multiple measurements of bias on continuous levels can be performed instead of a few binary variables indicating a pass or fail for each item being combined to form an indication.

This study also involved the development and testing of new attribute substitution bias measures, which showed very good variation and significant results, and which could be tested with other samples and modified for use in other contexts.

Another contribution of the study is a completely custom application to administer the experiments and the ability to create the distractions in a custom way that wasn't possible with existing tools. This can be replicated in further research and can easily be extended to handle other highly customised interactions with participants.

6.2.3 Practical contribution

This study has shown that rational decisions generally show greater accuracy, in a software selection context. With further research this could be generalised to other problems with attribute-type decisions where it would be important to reduce bias and improve accuracy. With this result, practitioners could look to encourage rationality in these types of decisions in order to improve the accuracy, especially with simpler decisions – where the differences in accuracy have been shown to be significant. The results also showed that complexity did affect the effectiveness of intuition, and so for more complicated decisions the accuracy could be very similar between rational and intuitive approaches, and this could be taken into account in practical situations when prescribing for or against the use of intuition.

One of the findings of this research is that cognitive reflection is significantly positively correlated with decision accuracy, and therefore one of the forms of the cognitive reflection test could potentially be considered an employment screening tool where the most accurate decisions are required.

Practically it seems, from this study and in this context, that that simple problems should be approached rationally, and more complex problems could be solved faster, and with similar accuracy, with intuition than with rational approaches.

This could have implications for the current demands for system, application, and "app" selection processes. According to the results of this study, simple "app" type selection processes should follow rational approaches and avoid intuition, since the number of attributes are usually lower. However, more complex system and application selection processes could be equally accurate with either rational, or intuitive, approaches.

As discussed in the introduction, even when the number of attributes is low, the sheer number of options will make many rational models impractical. However, despite the fact that intuition would certainly offer a faster route to making a decision, the results of this research show that the complexity of the problem is very important to take into account in order to determine whether intuitive decisions could be acceptably accurate. In addition, as shown in this research, the level of cognitive reflection of the decision maker could be even more important in determining tendency toward bias, and resultant decision accuracy.

This study was conducted in a software selection context, which limits the generalisability to other contexts. However, if these results do hold in other contexts, then it could be that simpler problems such as driving would be better served with rational approaches, whereas more complex unstructured problems could be served intuitively with much faster results. With cognitive reflection having a strong relationship to decision accuracy, this could be especially true with individuals showing a high level of cognitive reflection.

6.3 Limitations of the study

The study was purposefully limited to a particular population (senior information technology decision makers), but this could limit the generalisability of the results. Without another population to compare to, it is not possible to tell if characteristics of the population impacted the result. As indicated in the previous paragraphs, the context of software selection, though

allowing for interesting speculation, does somewhat limits the generalisability, and so future research should test these results in other contexts and with different populations.

As with most bias research and many decision studies, the study was limited in that it was designed as a laboratory-type experiment with a fictional case study. This reduces the external validity and the generalisability, since people may react differently in a real situation when there are real consequences.

6.4 Suggestions for further research

Future studies could expand on the number of attribute substitution questions. The three composite items (with nine questions) showed good variation on the experimental variables, but more items would strengthen this result.

Future research could consider an alternative (later) form of the cognitive reflection test since there are reports (Pennycook et al., 2016) that there is significant exposure to this test and that exposure reduces validity.

Not all of the anchoring sub-variables were effective in this study and different ones could be chosen in future studies. Some anchors could be replaced with more relevant ones, and it has been shown that relevant anchors have more effect (Wilson, Houston, Etling, & Brekke, 1996) and some of the anchors might have benefited from being set at the 15th or 85th percentile (Jacowitz & Kahneman, 1995).

The entire experiment took approximately half an hour to complete (exact times in results chapter) and utilising a shorter experiment in future research would reduce drop offs and probably improve the response rate since it was difficult to convince executives to complete a

long study. Future studies could utilise the items from this study that did show good variation and remove the ones that didn't work well, thereby shortening the experiment time.

A further consideration is the construct used for complexity. The construct used is based on the recommendation by Dijksterhuis (2004) to base complexity on the number of attributes presented, and in this regard, the complexity in this experiment was varied by changing the number of attributes in the case study. In this study, this varying of number of attributes did have a significant effect on the effectiveness of intuition and the accuracy of intuitive decisions. However, there was a marked difference in the results between the simple and complex conditions in this study, and it would be useful for further research to be conducted to the behaviour between those two points by varying the number of attributes between the simple and complex conditions. In addition to this, there are other complexity constructs in the literature and these should be investigated further as potential moderators. As an example, the difference in complexity of a simple attribute-type decision, compared to a chess move does not match up well to the number of attributes construct and could explain the difference in results found by different researchers.

In this study, no relationship was found between faith in intuition, and decision accuracy, despite previous findings that linked the two. As discussed, this could be because in this study the use of intuition was experimentally controlled, and this could point to use of intuition mediating the relationship between faith in intuition and decision accuracy. Further research should be done to test if this assertion has value.

Most importantly, the use of bias as a reverse proxy for decision accuracy should be tested further in other contexts, and with different populations. This methodology has shown significant results, while addressing various methodological concerns in similar types of research, and should be tested further.

This research attempted to find the conditions which would indicate when an intuitive decision should be made instead of a rational one. Complexity in terms of the number of attributes did change the effectiveness of intuition, and in many cases in complex decisions there were non-significant difference in bias and accuracy between intuitive and rational decisions. Overall, further research needs be done to determine what these boundary conditions are that could explain the differences found in the literature. It would seem that there must be explanations for why some literature can show that intuition is very powerful and accurate, and other literature that decries intuition as fallible, biased, and inaccurate. Both cannot be correct unless there are further conditions which explain the differences, and which would help us understand which mode to use in each circumstance.

6.5 Conclusion

"Clearly there are occasions when you should use your head instead of the formula. But which occasions they are, is most emphatically not clear." (Meehl, 1973, p.4)

This research attempted to discover the conditions under which intuition could be more, or less, effective than rational deliberation in decision-making. The results showed that complexity moderates how effective and accurate intuition decisions could be. In this study it was found that intuition was significantly less accurate in simple decisions, but that rational and intuitive approaches were similar when the decisions were more complex..

The initial research question asked:

> "What is the impact of complexity on the effectiveness of intuitive decisions?"

In this study, it was found that complexity, in terms of the number of attributes, did have an impact on the effectiveness of intuition in decision making, in that the more complex, the less differences between rational and intuitive decisions. Also that cognitive reflection measures are related to decision accuracy.

Despite answering the research question in this context, and providing some additional occasions when you should "use your head instead of the formula", various additional questions were raised. Future research should be conducted to find even more conditions which would help to explain the differences in effectiveness of intuition in decision making.

Chapter 7 - References

Abbott, A. (2015). Unconscious thought not so smart after all. *Nature*, 517, 537-538.

Acker, F. (2008). New findings on unconscious versus conscious thought in decision making: Additional empirical data and meta-analysis. *Judgment and Decision Making*, *3*(4), 292-303.

Allinson, C. W., & Hayes, J. (1996). The cognitive style index: A measure of intuition-analysis for organizational research. *Journal of Management studies*, *33*(1), 119-135.

Appelt, K. C., Milch, K. F., Handgraaf, M. J., & Weber, E. U. (2011). The Decision Making Individual Differences Inventory and guidelines for the study of individual differences in judgment and decision-making research. *Judgment and Decision Making*, *6*(3), 252-262.

Armstrong, J. U. S. T., & Overton, T. (1977). Estimating nonresponse bias in mail surveys. *Journal of marketing research*, *14*, 396-402.

Arroba, T. Y. (1978). Decision-making style as a function of occupational group, decision content and perceived importance. *Journal of Occupational Psychology*, *51*(3), 219-226.

Bannister, F., & Remenyi, D. (2000). Acts of faith: instinct, value and IT investment decisions. *Journal of Information Technology*, 15(3), 231-241.

Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of personality and social psychology*, *54*(4), 569.

Bartlett, M. S. (1954). A note on the multiplying factors for various χ 2 approximations. *Journal of the Royal Statistical Society*. Series B (Methodological), 296-298.

Bonabeau, E. (2003). Don't trust your gut. *Harvard Business Review*, *81*(5), 116-123.

Brown, S. A., Venkatesh, V., & Goyal, S. (2014). Expectation Confirmation in Information Systems Research: A Test of Six Competing Models. *Mis Quarterly*, 38(3).

Calvillo, D. P., & Penaloza, A. (2009). Are complex decisions better left to the unconscious? Further failed replications of the deliberation-without-attention effect. *Judgment and Decision Making*, *4*(6), 509-517.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research* (pp. 171-246). Boston: Houghton Mifflin.

Carifio, J., & Perla, R. J. (2007). Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes. *Journal of Social Sciences*, *3*(3), 106-116.

Carney, D. J., & Wallnau, K. C. (1998). A basis for evaluation of commercial software. *Information and Software Technology*, 40(14), 851-860.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate behavioral research*, 1(2), 245-276.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* 2nd edn.

Dane, E., & Pratt, M. G. (2007). Exploring intuition and its role in managerial decision making. *Academy of Management Review*, 32(1), 33-54.

De Neys, W. (Ed.). (2017). *Dual Process Theory 2.0.* Routledge.

De Neys, W., & Bonnefon, J. F. (2013). The 'whys' and 'whens' of individual differences in thinking biases. *Trends in cognitive sciences*, 17(4), 172-178.

DeLone, W. H., & McLean, E. R. (1992). Information systems success: the quest for the dependent variable. *Information systems research*, *3*(1), 60-95.

DeVellis, R. F. (2012). *Scale Development: Theory and Applications* (Vol. 26). SAGE Publications.

Dhar, R., & Gorlin, M. (2013). A dual-system framework to understand preference construction processes in choice. *Journal of Consumer Psychology*, *23*(4), 528-542.

Dijksterhuis, A. (2004). Think different: the merits of unconscious thought in preference development and decision making. *Journal of personality and social psychology*, *87*(5), 586.

Dijksterhuis, A., Bos, M. W., Van der Leij, A., & Van Baaren, R. B. (2009). Predicting soccer matches after unconscious and conscious thought as a function of expertise. *Psychological Science*, 20(11), 1381-1387.

Dijksterhuis, A., & Nordgren, L. F. (2006). A theory of unconscious thought. *Perspectives on Psychological science*, *1*(2), 95-109.

Dijksterhuis, A., & van Olden, Z. (2006). On the benefits of thinking unconsciously: Unconscious thought can increase post-choice satisfaction. *Journal of Experimental Social Psychology*, *42*(5), 627-631.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American psychologist*, *49*(8), 709.

Epstein, S. (2010). Demystifying intuition: What it is, what it does, and how it does it. *Psychological Inquiry*, *21*(4), 295-312.

Epstein, S., Pacini, R., Denes-Raj, V., & Heier, H. (1996). Individual differences in intuitive–experiential and analytical–rational thinking styles. *Journal of personality and social psychology*, 71(2), 390.

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition advancing the debate. *Perspectives on Psychological Science, 8*(3), 223-241.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives*, 25-42.

Goes, P. (2013). Editor's Comments: Information Systems Research and Behavioral Economics. *Management Information Systems Quarterly, 37*(3), iii-viii.

Hamilton, K., Shih, S. I., & Mohammed, S. (2016). The development and validation of the rational and intuitive decision styles scale. *Journal of personality assessment*, 98(5), 523-535.

Hayashi, A. M. (2001). When to trust your gut. *Harvard Business Review*, 79(2), 59-65.

Isenberg, D. J. (1984). How senior managers think. (Cover story). *Harvard Business Review*, 62(6), 81-90.

Jacowitz, K. E., & Kahneman, D. (1995). Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin, 21*, 1161-1166.

Jadhav, A. S., & Sonar, R. M. (2009). Evaluating and selecting software packages: A review. *Information and Software Technology*, 51(3), 555-563.

Jadhav, A.S., & Sonar, R.M. (2011). Framework for evaluation and selection of the software packages: A hybrid knowledge based system approach. *The Journal of Systems and Software*, 84, 1394-1407.

Jamieson, K., & Hyland, P. (2006). Good intuition or fear and uncertainty: The effects of bias on information systems selection decisions. *Informing Science: International Journal of an Emerging Transdiscipline*, 9, 49-69.

Gill, J., & Johnson, P. (2010). *Research methods for managers* (4th ed.). London: Sage.

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

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and biases: The psychology of intuitive judgment*, 49-81.

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: a failure to disagree. *American Psychologist*, *64*(6), 515.

Kahneman, D., & Knetsch, J. (1993). Anchoring or shallow inferences: The effect of format. *Unpublished manuscript*.

Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika*, 35(4), 401-415.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36.

Keller, J., & Bless, H. (2009). Predicting future affective states: How ease of retrieval and faith in intuition moderate the impact of activated content. *European Journal of Social Psychology*, *39*(3), 467-476.

Klein, G. A. (1993). A recognition-primed decision (RPD) model of rapid decision making. *Decision making in action: Models and methods*, *5*(4), 138-147.

Klein, G. A. (1998). *Sources of power: How people make decisions*. MIT press.

Klein, G. (2008). Naturalistic decision making. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 456-460.

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (1986, September). Rapid decision making on the fire ground. In *Proceedings of the Human Factors and Ergonomics Society annual meeting* (Vol. 30, No. 6, pp. 576-580). SAGE Publications.

Lai, V. S., Wong, B. K., & Cheung, W. (2002). Group decision making in a multiple criteria environment: A case using the AHP in software selection. *European Journal of Operational Research*, 137(1), 134-144.

Legris, P., Ingham, J., & Collerette, P. (2003). Why do people use information technology? A critical review of the technology acceptance model. *Information & management*, *40*(3), 191-204.

Leedy, P. D., & Ormrod, J. E. (2014). *Practical Research: Planning and Design*, Tenth Edition. Pearson, Essex.

Levy, D. A., & Nail, P. R. (1993). Contagion: A theoretical and empirical review and reconceptualization. *Genetic, Social, and General Psychology Monographs*.

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

Marsden, P. (1998). Memetics and social contagion: Two sides of the same coin. *Journal of Memetics-Evolutionary Models of Information Transmission*, *2*(2), 171-185.

Meehl, P. E. (1973). Why I do not attend case conferences. *Psychodiagnosis: Selected papers*, 225-302.

Meszaros, M. (2007). *Multiattribute Judgment and Nonconscious Cognition: Further Characterization of the Deliberation-without-attention Effect*. (Doctoral dissertation, ProQuest).

Nieuwenstein, M. R., Wierenga, T., Morey, R. D., Wicherts, J. M., Blom, T. N., Wagenmakers, E. J., & van Rijn, H. (2015). On making the right choice: a meta-analysis and large-scale replication attempt of the unconscious thought advantage. *Judgment and Decision Making*.

Pacini, R., & Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of personality and social psychology*, *76*(6), 972.

Pennycook, G., Cheyne, J. A., Koehler, D. J., & Fugelsang, J. A. (2016). Is the cognitive reflection test a measure of both reflection and intuition?. *Behavior Research Methods*, 48(1), 341-348.

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology*, 80, 34-72.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of applied psychology*, *88*(5), 879.

Reips, U. D. (2002). Internet-Based Psychological Experimenting Five Dos and Five Don'ts. *Social Science Computer Review*, *20*(3), 241-249.

Rosenzweig, P. (2009). *The halo effect:... and the eight other business delusions that deceive managers*. Simon and Schuster.

Saaty, T. L. (1990). How to make a decision: the analytic hierarchy process. *European Journal of Operational Research*, 48(1), 9-26.

Sadler-Smith, E. (2012). *Inside intuition*. Routledge.

Sadler-Smith, E., & Shefy, E. (2004). The intuitive executive: Understanding and applying gut feel in decision-making. *The Academy of Management Executive*, 18(4), 76-91.

Schooler, J. W., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. *Journal of experimental psychology: General*, 122(2), 166.

Scott, S. G., & Bruce, R. A. (1995). Decision-making style: The development and assessment of a new measure. *Educational and psychological measurement*, *55*(5), 818-831

Schwartz, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., & Lehman, D. R. (2002). Maximizing versus satisficing: happiness is a matter of choice. *Journal of personality and social psychology*, *83*(5), 1178.

Shiloh, S., Salton, E., & Sharabi, D. (2002). Individual differences in rational and intuitive thinking styles as predictors of heuristic responses and framing effects. *Personality and Individual Differences*, 32(3), 415-429.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1), 99-118.

Simon, H. A. (1979). Rational decision making in business organizations. *American Economic Review*, 69: 493–513.

Simon, H. A. (1987). Making management decisions: The role of intuition and emotion. *The Academy of Management Executive* (1987-1989), 57-64.

Simon, H. A. (1992). What is an "explanation" of behavior?. *Psychological science*, 3(3), 150-161.

Sinclair, M., & Ashkanasy, N. M. (2005). Intuition: Myth or a decision-making tool? *Management Learning*, 36(3), 353-370.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological bulletin*, *119*(1), 3.

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and social psychology review*, *4*(2), 108-131.

Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Psychology Press.

Stanovich, K. E. (2005). *The robot's rebellion: Finding meaning in the age of Darwin*. University of Chicago Press.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and brain sciences*, 23(5), 645-665.

Strick, M., Dijksterhuis, A., Bos, M. W., Sjoerdsma, A., Van Baaren, R. B., & Nordgren, L. F. (2011). A meta-analysis on unconscious thought effects. *Social Cognition*, 29(6), 738-762.

Thoma, V., White, E., Panigrahi, A., Strowger, V., & Anderson, I. (2015). Good thinking or gut feeling? Cognitive reflection and intuition in traders, bankers and financial non-experts. *PloS one*, 10(4), e0123202.

Thorsteinson, T. J., & Withrow, S. (2009). Does unconscious thought outperform conscious thought on complex decisions? A further examination. *Judgment and Decision Making*, 4(3), 235.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & cognition*, 39(7), 1275.

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological bulletin*, *76*(2), 105.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124-1131.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453-458.

Usher, M., Russo, Z., Weyers, M., Brauner, R., & Zakay, D. (2011). The impact of the mode of thought in complex decisions: Intuitive decisions are better. *Frontiers in psychology*, *2*.

Venkatesh, V., Davis, F. D., & Morris, M. G. (2007). Dead or Alive? The Development, Trajectory and Future of Technology Adoption Research. *Journal of the association for information systems*, *8*(4).

Vetter, J., Benlian, A., & Hess, T. (2011). Overconfidence in IT investment decisions: why knowledge can be a boon and bane at the same time. *In Proceedings of the 32nd International Conference on Information Systems*, Shanghai, China.

Walco, D. K., & Risen, J. L. (2017). The empirical case for acquiescing to intuition. *Psychological science*, 28(12), 1807-1820.

Wang, W., & Benbasat, I. (2009). Interactive Decision Aids for Consumer Decision Making in E-Commerce: The Influence of Perceived Strategy Restrictiveness. *MIS Quarterly*, *33*(2).

West, R. F., Toplak, M. E., & Stanovich, K. E. (2008). Heuristics and biases as measures of critical thinking: associations with cognitive ability and thinking dispositions. *Journal of Educational Psychology*, 100(4), 930.

Whillock, S. R. (2017). *The search for moderation: does anxiety enhance the unconscious thought advantage?* (Doctoral dissertation, Montana State University-Bozeman, College of Letters & Science).

Wierzbicki, A. P. (1997). On the role of intuition in decision making and some ways of multicriteria aid of intuition. *Journal of Multi-Criteria Decision Analysis*, 6(2), 65-76.

Wilson, T. D., Houston, C. E., Etling, K. M., & Brekke, N. (1996). A new look at anchoring effects: basic anchoring and its antecedents. *Journal of Experimental Psychology: General*, *125*(4), 387.

Wilson, T. D., Lisle, D. J., Schooler, J. W., Hodges, S. D., Klaaren, K. J., & LaFleur, S. J. (1993). Introspecting about reasons can reduce post-choice satisfaction. *Personality and Social Psychology Bulletin*, *19*, 331-331.

Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: introspection can reduce the quality of preferences and decisions. *Journal of personality and social psychology*, *60*(2), 181.

Appendix A - Questionnaire

Informed consent for participation in an academic research study

THE IMPACT OF COMPLEXITY ON THE EFFECTIVENESS OF INTUITION IN DECISION MAKING

Research conducted by:

Mr D.J. van Eck (u14192226, gibs510204, cel: 082-579-1997)

Dear Respondent,

You are invited to participate in an academic research study conducted by Daneel van Eck, a Doctoral student from the Gordon Institute of Business Science at the University of Pretoria. The purpose of the study is to determine the predisposition toward, and effectiveness of intuitive decision-making for software selection.

Please note the following:

- This study involves an anonymous survey. Your name will not appear on the questionnaire and the answers you give will be treated as strictly confidential. You cannot be identified in person based on the answers you give.

- Your participation in this study is very important to us. You may, however, choose not to participate and you may also stop participating at any time without any negative consequences.
- Please answer the questions in the attached questionnaire as completely and honestly as possible. This should not take more than 25 minutes of your time.
- The results of the study will be used for academic purposes as well as for lay articles and conference proceedings; however, identities of individuals will be kept confidential. The results of the study will also be published in an academic journal. We will provide you with a summary of our findings on request.

Please contact my supervisor, Dr Karen Luyt (082 895 2289; karen.luyt@bcx.co.za) if you have any questions or comments regarding the study.

I confirm that I have read and understand the information provided above and I give my consent to participate in the study on a voluntary basis. (please tick the checkbox to confirm)

[ ]

The following three questions vary in difficulty, please answer as many of them as you can.

4. A bat and a ball cost R1.10 in total. The bat costs R1.00 more than the ball.
   How much does the ball cost? _____ cents
5. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? _____ minutes
6. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?
7. Have you seen these questions before? _____ (if yes then which ones _____)

Case Study

You have been asked to evaluate ERP software packages for suitability at a particular company "QR Sports".

Click next to view the details of the case study

Software purchase case (complex)

Case study details:

The company has less than 500 employees and most of their sales come from distribution of less than 30 sports products with multiple configurations, which they import from European countries and distribute across South Africa. The company is situated in Johannesburg and not all of the employees speak English, other languages include Sotho, Zulu, and Afrikaans. A significant component of the current administration cost in the company comes from having to handle and escalate approvals of varying nature including purchase approvals, discount approvals, and leave approvals. The employees are mostly millennial's and respond well to modern, social network type interfaces and they expect to be able to communicate with co-workers and key customers using the ERP application. The vendor for the current system is planning to sunset the product and has frozen all development except for bug-fixing for the last three years. The overall culture of the company supports long hours and many employees spend additional time on weekends working on sales proposals to channel clients and require the ability to work remotely. The company is currently experiencing severe financial difficulties as a direct result of inconsistencies in the data and processing of the current ERP. The current system has proven inflexible and cannot cater for new business requirements. Because of the single location and significant internal infrastructure there are no current requirements for a cloud-based application. CRM is not currently being requested and the current system has no CRM capability but the sales director has expressed interest in CRM.

The following three packages have been shortlisted:

1. Package X:

a. Multi-language support

b. Cloud-ready

c. Successful track record of implementation in Europe and some track record in distribution.

d. Standard interface with corporate design - HTML5

e. No social media integration

f. Built in CRM

g. Vendor provided roadmap with yearly releases

h. Ability to access remotely using HTML5 web client

i. Priced in euro – expensive

j. Cannot be customised but can be extended using graphical tool and API's for various languages

k. Adequate system performance, reasonable reporting speed

l. Built in integration for Windows Workflow Foundation

m. Although upgrades are included in the annual maintenance, upgrade implementation is for the cost of the client and has been known to be expensive

n. Web-oriented-architecture

2. Package Y:

a. Single-language support (English/Portuguese/French)

b. No cloud support but can be hosted

c. Successful track record of implementation in Africa and some track record in retail

d. Traditional Forms interface - Windows 7, 8, 10

e. Social media integration with Facebook, Twitter, and Yammer

f. Integrates fully with Microsoft Dynamics CRM

g. Reports from reference sites that the vendor responds quickly to changing customer requirements and very often releases minor updates with no charge to the requesting organisation as long as the request can benefit a reasonable proportion of the client base

h. Can work over RDP or with locally installed client

i. Priced in Rand – expensive

j. Customisable using standard .net code

k. Excellent performance, fast reporting

l. Custom workflow module with graphical process editing

m. Version upgrades are normally simple because of the incremental releases, however upgrades can break any customised .net code

n. Object-oriented client-server architecture

3. Package Z:

a. Single-language support (English)

b. No cloud support

c. Successful track record of implementation in distribution of sporting goods in North and South America

d. Multiple interfaces with modern design - Forms (Windows 7), Metro (Windows 8, 10), and Web (HTML 5)

e. Configurable integration - no support for Facebook, Twitter etc. yet but can be developed

f. Support for SalesForce.com CRM (integrated)

g. No roadmap for future development but a track record of releasing stable new versions every two years

h. Web client can be accessed remotely but does not support all enterprise functionality (only employee self-service HR, and executive dashboards on HTML5, with other functionality incompatible with RDP and requires local connection)

i. Priced in Dollars - cost effective

j. Customisable using a proprietary language which has been designed specifically for this package

k. Sluggish system performance, slow reporting

l. Sharepoint integration with built-in workflow editor

m. Built in upgrade tool which extracts customisations, performs the upgrade, and then re-implements customisations over a weekend, with no additional labour required other than training on new features

n. Service-oriented-architecture

Software purchase case (simple)

Case study details:

The company has less than 500 employees and most of their sales come from distribution of less than 30 sports products with multiple configurations, which they import from European countries and distribute across South Africa. The company is situated in Johannesburg and not all of the employees speak English, other languages include Sotho, Zulu, and Afrikaans. The employees are mostly millennial's and respond well to modern, social network type interfaces and they expect to be able to communicate with co-workers and key customers using the ERP application. The overall culture of the company supports long hours and many employees spend additional time on weekends working on sales proposals to channel clients and require the ability to work remotely. The company is currently experiencing severe financial difficulties as a direct result of inconsistencies in the data and processing of the current ERP.

The following three packages have been shortlisted:

1. **Package X:**

a. Multi-language support

b. Successful track record of implementation in Europe and some track record in distribution.

c. Standard interface with corporate design - HTML5

d. No social media integration

e. Ability to access remotely using HTML5 web client

f. Priced in euro – expensive

g. Adequate system performance, reasonable reporting speed

2. **Package Y:**

a. Single-language support (English/Portuguese/French)

b. Successful track record of implementation in Africa and some track record in retail

c. Traditional Forms interface - Windows 7, 8, 10

d. Social media integration with Facebook, Twitter, and Yammer

e. Can work over RDP or with locally installed client

f. Priced in Rand – expensive

g. Excellent performance, fast reporting

3. Package Z:

a. Single-language support (English)

b. Successful track record of implementation in distribution of sporting goods in North and South America

c. Multiple interfaces with modern design - Forms (Windows 7), Metro (Windows 8, 10), and Web (HTML 5)

d. Configurable integration - no support for Facebook, Twitter etc. yet but can be developed

e. Web client can be accessed remotely but does not support all enterprise functionality (only employee self-service HR, and executive dashboards on HTML5, with other functionality incompatible with RDP and requires local connection)

f. Priced in Dollars - cost effective

g. Sluggish system performance, slow reporting

Case questions

Now that you have considered the above case, please answer the following questions:

Please give a general overall rating on each of the three shortlisted products:

How do you feel overall about Package X?

Very bad ---------- (click the slider to answer) ---------- Very good

How do you feel overall about Package Y?

Very bad ---------- (click the slider to answer) ---------- Very good

How do you feel overall about Package Z?

Very bad ---------- (click the slider to answer) ---------- Very good

Do you expect that the required software would cost more or less than R5m?

Less ( ) More ( )

Please give an estimate of the amount that you would expect to pay for the license cost (excluding first year maintenance) for software that would suit the requirements of the company (in millions of Rand, rounded to one decimal place - e.g. R#.# million). R [ ]

Please rate the expected cultural fit of: Package X?

Very bad ---------- (click the slider to answer) ---------- Very good

Please rate the expected cultural fit of: Package Y?

Very bad ---------- (click the slider to answer) ---------- Very good

Please rate the expected cultural fit of: Package Z?

Very bad ---------- (click the slider to answer) ---------- Very good

Please rate the scalability of: Package X?

Not very scalable ---------- (click the slider to answer) ---------- Very scalable

Please rate the scalability of: Package Y?

Not very scalable ---------- (click the slider to answer) ---------- Very scalable

Please rate the scalability of: Package Z?

Not very scalable ---------- (click the slider to answer) ---------- Very scalable

Please rate support for remote workers with: Package X?

Not very good ---------- (click the slider to answer) ---------- Very good

Please rate support for remote workers with: Package Y?

Not very good ---------- (click the slider to answer) ---------- Very good

Please rate support for remote workers with: Package Z?

Not very good ---------- (click the slider to answer) ---------- Very good

Think back to the last actual software purchase you were involved in (not part of this case).

How satisfied are you right now with that product?

Very unsatisfied ---------- (click the slider to answer) ---------- Very satisfied

How satisfied would you expect to be with: Package X?

Very unsatisfied ---------- (click the slider to answer) ---------- Very satisfied

How satisfied would you expect to be with: Package Y?

Very unsatisfied ---------- (click the slider to answer) ---------- Very satisfied

How satisfied would you expect to be with: Package Z?

Very unsatisfied ---------- (click the slider to answer) ---------- Very satisfied

Think back to the last software purchase decision you were involved in for software costing at least R500k (preferably similar to this case). Try to remember the approximate cost of the software and please fill it in here. If you cannot remember then fill in an approximate amount that you would guess it could have cost. [R]

Please estimate the approximate amount that you would expect to pay (including implementation) for:

Package X: [R]

Package Y: [R]

Package Z: [R]

Approximately how many software purchase decisions have you been involved in (over the past three years)? [ ]

What would your rough estimate be of the number of months it would take for implementation of any of the three options? (if different then the longest one) [months]

Based on the brief initial analysis, the vendor estimates that there is a 20% chance that the implementation project for Package Y could run over time and over budget. Please give a rough estimate of the amount you think the project might run over by: [R]

Please give an estimate of the amount that you would expect to pay for the implementation services cost for software that would suit the requirements of the company (in millions of Rand, rounded to one decimal place - e.g. R#.# million). [R]

Please indicate your order of preference from the available options by placing a "1" next to your first choice, a "2" next to your second choice, and a "3" next to your last choice for which package to implement for the company.

Package X [ ]

Package Y [ ]

Package Z [ ]

How confident are you in your rating of the three packages in the previous question?

Not at all confident ---------- (click the slider to answer) ---------- Completely confident

Please explain briefly how you got to your rating?

[ ]

Consider that the company is currently undergoing severe financial difficulties because of severe issues with the current ERP system. Because of these difficulties it is expected that 60 people will need to be retrenched. Two alternative programs to improve profitability have been proposed. An analysis has shown the following exact consequences of each of the programs:

If program A is adopted, 20 of the people will avoid retrenchment

If program B is adopted, there is a 1/3 probability that all 60 people will avoid retrenchment, and 2/3 probability that none will avoid it.

Which of the two programs would you favour?

A ( ) B ( )

A 55-year old man in another department had a heart condition. He had to stop working because of chest pain. He enjoyed his work and did not want to stop. His pain also interfered with other things, such as travel and recreation. A type of bypass operation would relieve his pain and increase his life expectancy from age 65 to age 70. However 8% of the people who have this operation die from the operation itself

His physician decided to go ahead with the operation. The operation succeeded. Evaluate the physician's decision to go ahead with the operation.

-3	-2	-1	0	1	2	3
Incorrect and inexcusable	Incorrect, all things considered	Incorrect, but not unreasonable	The decision and its opposite are equally good	Correct, but the opposite would be reasonable too	Correct, all things considered	Clearly correct, and the opposite decision would be inexcusable

For package X, the analysis and implementation together are estimated to take 29 weeks. Implementation takes up most of the time and will take 20 weeks more than the analysis.

How long will the analysis take? [weeks]

Consider again that the company is currently undergoing severe financial difficulties because of severe issues with the current ERP system. Because of these difficulties it is expected that 60 people will need to be retrenched. Two additional alternative programs to improve profitability have been proposed. An analysis has shown the following exact consequences of each of the programs:

If program C is adopted, 40 people will be retrenched.

If program D is adopted, there is a 1/3 probability that nobody will be retrenched, and 2/3 probability that all 60 people will be retrenched.

Which of the two programs would you favour?

C ( ) D ( )

Another person in the company had a liver condition. He also had to stop working because of pain. He enjoyed his work and did not want to stop. His pain also interfered with other things, such as travel and recreation. A particular operation would relieve his pain and increase his life expectancy. In this case, only 2% of the people who have this operation die from the operation itself. His physician decided to go ahead with the operation. In this case the operation failed and the man died. Evaluate the physician's decision to go ahead with the operation.

-3	-2	-1	0	1	2	3
Incorrect and inexcusable	Incorrect, all things considered	Incorrect, but not unreasonable	The decision and its opposite are equally good	Correct, but the opposite would be reasonable too	Correct, all things considered	Clearly correct, and the opposite decision would be inexcusable

If you have any comments regarding the case or the questions then please describe them here? [ ]

Indicate how much you agree with each of the following by **selecting the appropriate number**.

My initial impressions of people are almost always right.

1	2	3	4	5
Strongly Disagree				Strongly Agree

I trust my initial feelings about people.

1	2	3	4	5
Strongly Disagree				Strongly Agree

When it comes to trusting people, I can usually rely on my “gut feelings”

1	2	3	4	5
Strongly Disagree				Strongly Agree

I believe in trusting my hunches.

1	2	3	4	5
Strongly Disagree				Strongly Agree

I can usually feel when a person is right or wrong, even if I can't explain how I know.

1	2	3	4	5
Strongly Disagree				Strongly Agree

What is your gender? [select option]

What is your home language? [select option]

How old are you? [select option range]

What is your current level in the organisation? [select option]

What is your current title in the organisation? [ ]

What industry is your organisation in? [select option]

How large is the organisation? [select option]

Is the organisation listed on the Johannesburg Stock Exchange? [select option]

How many years of work experience do you have? [select option range]

How many years of I.T. experience do you have? [select option range]

What is your highest qualification? [select option]

Appendix B – Additional detail on results

This appendix presents additional information from the analysis of results, that was not presented in the results chapter.

9.1 Faith in intuition scale

The results of the analysis of the faith in intuition scale are presented in the results chapter. This section presents additional information for that analysis.

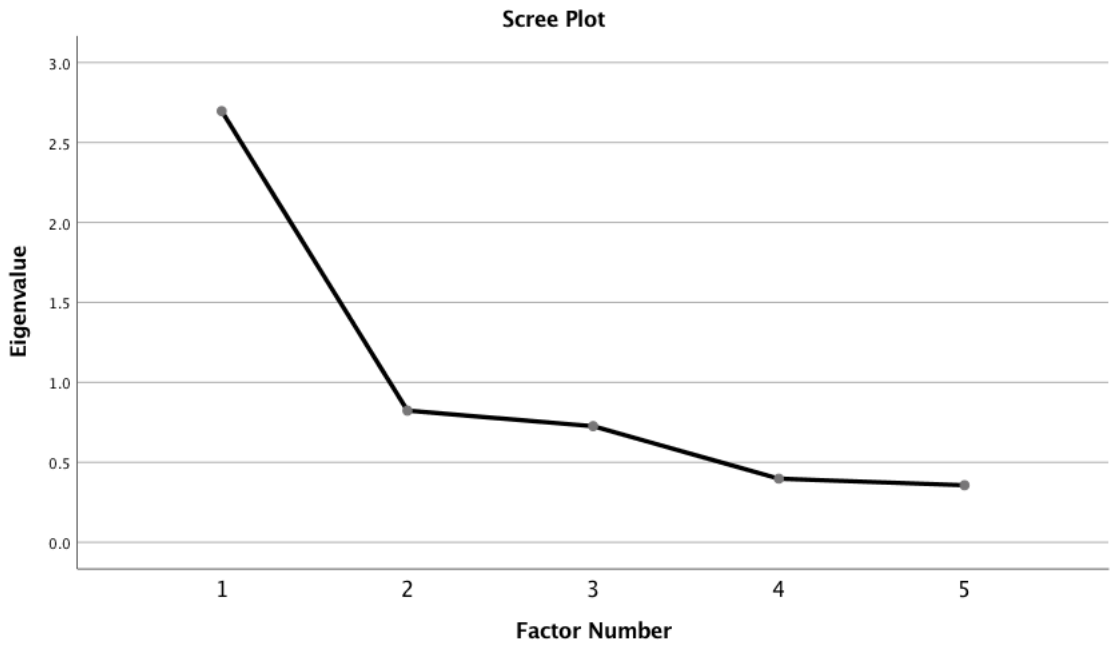


Figure 10. Screeplot of the factor analysis of the faith in intuition scale.

Table 3. Correlation matrix of the factor analysis for faith in intuition.

Correlation Matrix

	I trust my initial feelings about people	I believe in trusting my hunches	My initial impressions of people are almost always right	When it comes to trusting people, I can usually rely on my "gut feelings"	I can usually feel when a person is right or wrong even if I can't explain how I know
I trust my initial feelings about people	1.000	.524	.374	.607	.262
I believe in trusting my hunches	.524	1.000	.231	.565	.452
My initial impressions of people are almost always right	.374	.231	1.000	.438	.264
When it comes to trusting people, I can usually rely on my "gut feelings"	.607	.565	.438	1.000	.435
I can usually feel when a person is right or wrong even if I can't explain how I know	.262	.452	.264	.435	1.000

Table 4. Total variance explained for faith in intuition.

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Factor	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.696	53.926	53.926	2.195	43.905	43.905
2	.823	16.465	70.391			
3	.726	14.522	84.913			
4	.398	7.955	92.868			
5	.357	7.132	100.000			

Extraction Method: Principal Axis Factoring.

Table 5. Factor matrix for faith in intuition.

Factor Matrix[a]

	Factor 1
I trust my initial feelings about people	.702
I believe in trusting my hunches	.698
My initial impressions of people are almost always right	.474
When it comes to trusting people, I can usually rely on my "gut feelings"	.852
I can usually feel when a person is right or wrong even if I can't explain how I know	.516

Extraction Method: Principal Axis Factoring.[a]

a. 1 factors extracted. 9 iterations required.

9.2 Attribute substitution bias

There were three questions which attempted to elucidate a biased reaction, using attribute substitution bias, and these were based on the following sub-variables:

- Cultural fit attribute substitution bias (scored on estimated cultural fit of package X, estimated cultural fit of package Y, estimated cultural fit of package Z)
- Scalability attribute substitution bias (scored on estimated scalability of package X, estimated scalability of package Y, estimated scalability of package Z)
- Remote support attribute substitution bias (scored on estimated remote support of package X, estimated remote support of package Y, estimated remote support of package Z)

Each of these sub-variables will be described separately in the following sections.

9.3 Cultural fit attribute substitution bias

The cultural fit attribute substitution bias sub-variable was based on the estimated cultural fit of package X, estimated cultural fit of package Y, and estimated cultural fit of package Z items in the questionnaire. The cultural fit attribute substitution bias variable did contain a Tukey (1977) outlier (values:1), which were transformed through Winsorizing to the maximum and minimum non-outlier values. The Shapiro-Wilk test of normality for cultural fit attribute substitution bias showed that it did not follow a normal distribution ($p<0.001$).

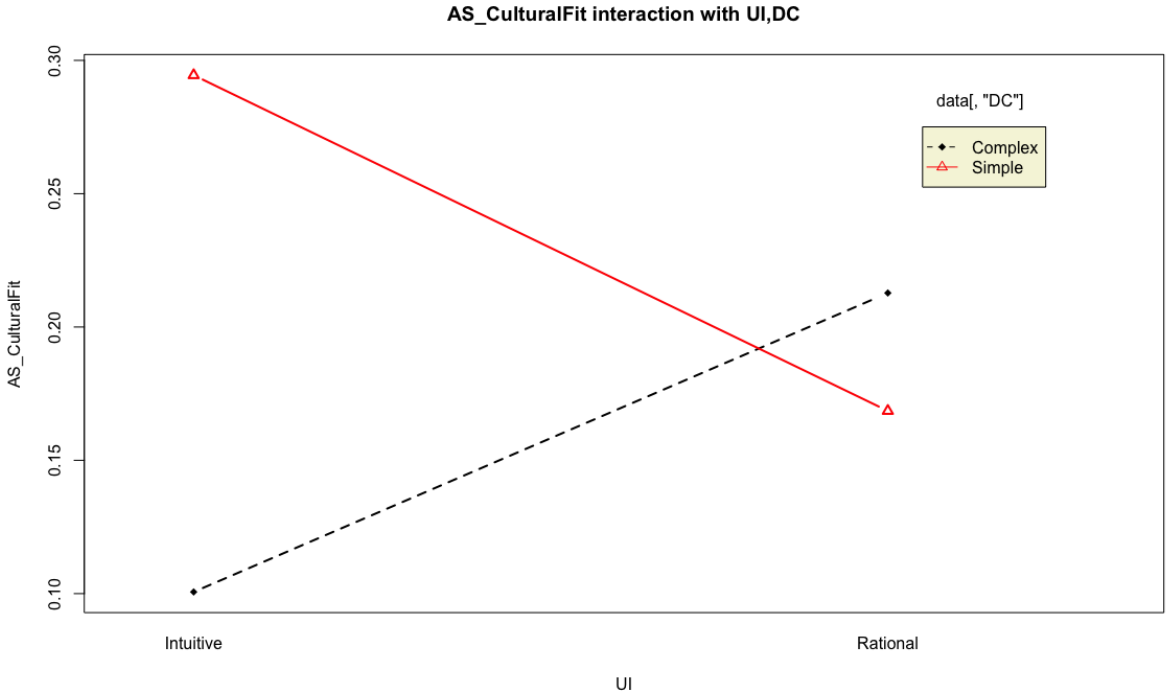


Figure 11. Interaction plot of Cultural fit attribute substitution bias against use of intuition and decision complexity.

Visually there is a difference on the interaction plot between simple and complex in the intuitive group. Since the data did not follow a normal distribution, a Mann-Whitney U test was used and it showed a significant difference for cultural fit attribute substitution bias between the intuitive-simple and intuitive-complex groups ($p=0.03$). This showed that for this cultural fit attribute substitution bias sub-variable, the intuitive group was significantly more biased for the simple case than for the complex case.

9.3.1 Scalability attribute substitution bias

The scalability attribute substitution bias variable was based on the estimated scalability of package X, estimated scalability of package Y, estimated scalability of package Z items in the questionnaire. The scalability attribute substitution bias variable did not contain any Tukey (1977) outliers. The Shapiro-Wilk test of normality for scalability attribute substitution bias showed that the results did follow a normal distribution (p= 0.12).

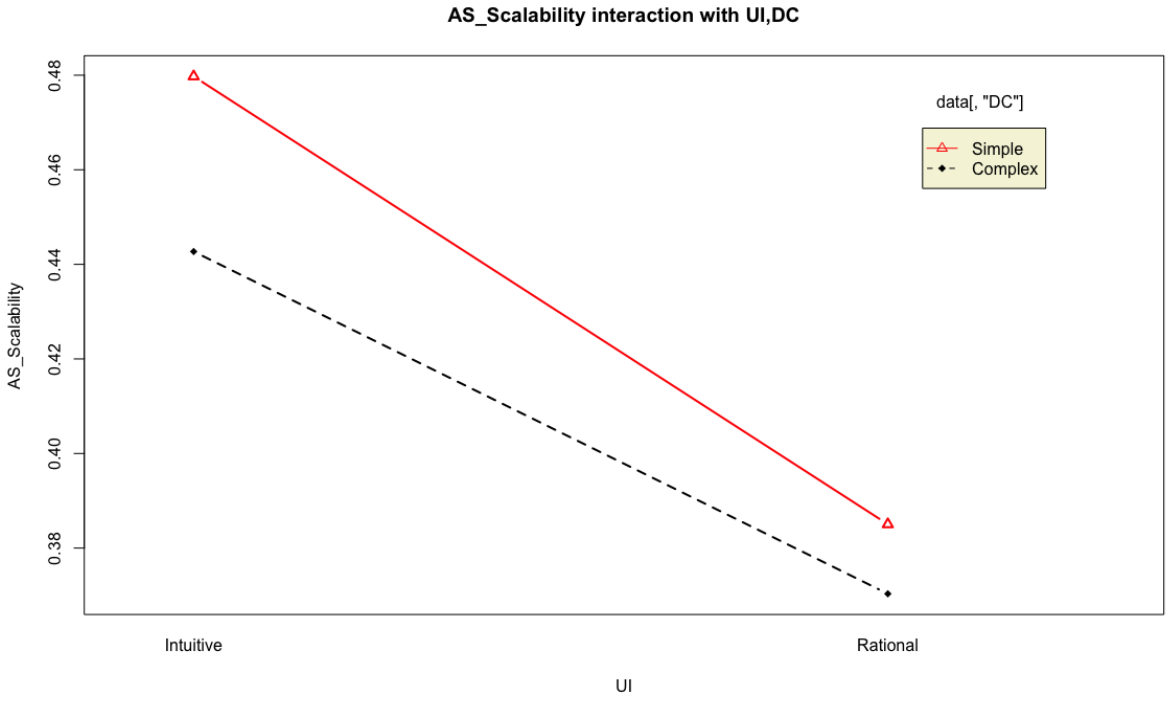


Figure 12. Interaction plot of Scalability attribute substitution bias by use of intuition and decision complexity.

9.3.2 Remote support attribute substitution bias

The remote support attribute substitution bias variable was based on the estimated remote support of package X, estimated remote support of package Y, and estimated remote support of package Z items in the questionnaire. The remote support attribute substitution bias variable did contain Tukey (1977) defined outliers which were transformed through Winsorizing to the maximum and minimum non-outlier values. The Shapiro-Wilk test of normality for Remote support attribute substitution bias showed that the results did not follow a normal distribution ($p<0.001$).

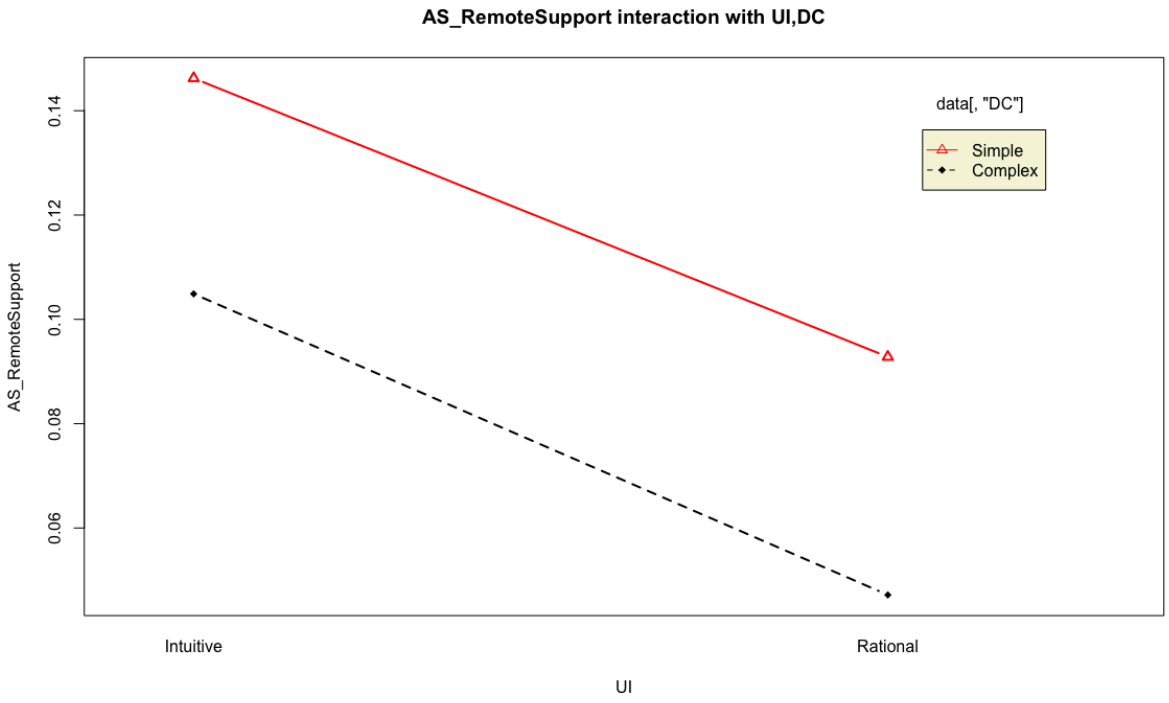


Figure 13. Interaction plot of Remote support attribute substitution bias interaction by use of intuition and decision complexity.

Since the data did not follow a normal distribution, a Mann-Whitney U test was performed and it showed a significant difference between the intuitive and rational groups ($p=0.04$), and

between the simple and complex groups ($p<0.05$), and also between both the complex-intuitive and complex-rational groups ($p=0.04$), and between the rational-simple and rational-complex groups ($p=0.03$).

9.3.3 Attribute substitution bias summary

The composite attribute substitution bias variable is based on the cultural fit attribute substitution bias, scalability attribute substitution bias, and remote support attribute substitution bias. These are summed to form the composite attribute substitution bias variable. The attribute substitution bias variable did contain a Tukey (1977) outlier (1.87). This was transformed through Winsorizing to the maximum non-outlier value. The Shapiro-Wilk test of normality for attribute substitution bias showed that the results did not follow a normal distribution ($p=0.04$).

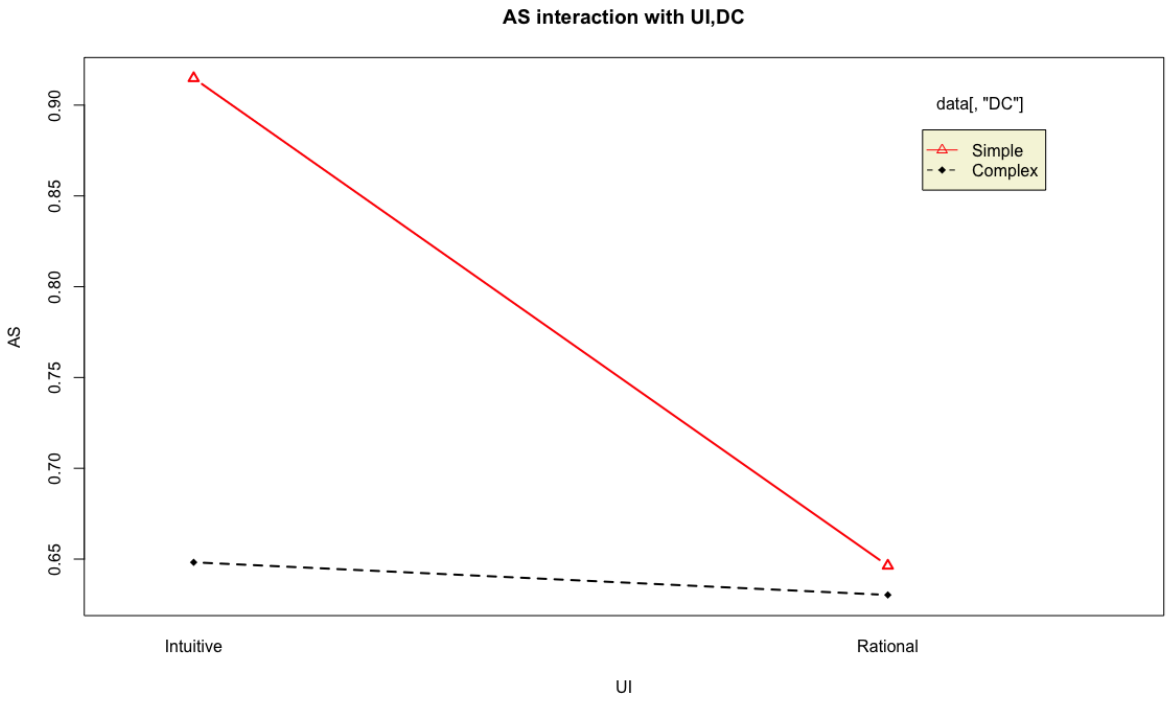


Figure 14. Interaction plot of attribute substitution bias by use of intuition and decision complexity.

On the interaction plot in the figure there is a visual difference between the complex and simple treatments in the intuitive group. Since the data did not follow a normal distribution, Mann-Whitney U tests were performed and these showed significant differences between the simple-intuitive and simple-rational ($p=0.02$) groups, and between the intuitive-simple and intuitive-complex groups ($p=0.03$).

9.4 Anchoring bias

There were five groups of questions which attempted to elucidate a biased reaction, using anchoring bias, and these were based on the following variables:

- estimated software value (vs a experimenter-generated R5m anchor)
- expected satisfaction with purchase (group 1), expected satisfaction with purchase (group 2) (same question placed differently - group 1 has the question before the expected satisfaction with package X,Y,Z questions)
- expected satisfaction with package X, expected satisfaction with package Y, expected satisfaction with package Z (vs self-generated and experimentally-controlled anchor in expected satisfaction with purchase (group 1))
- previous purchase cost (group 2) and expected cost of package X, expected cost of package Y, expected cost of package Z, previous purchase cost (group 1)
- number of previous purchases (group 2), number of previous purchases (group 1) vs each other and vs estimate of number of months to implement the software

9.4.1 Estimated software value anchoring bias

The estimated software value variable did contain Tukey (1977) outliers (values: R11,4m, R14,6m, R13m, R9.9m, R9m, R15m, R10m, R10,7m). These were transformed through Winsorizing to the maximum and minimum non-outlier values. The Shapiro-Wilk test of normality for estimated software value showed that the results did not follow a normal distribution ($p<0.001$).

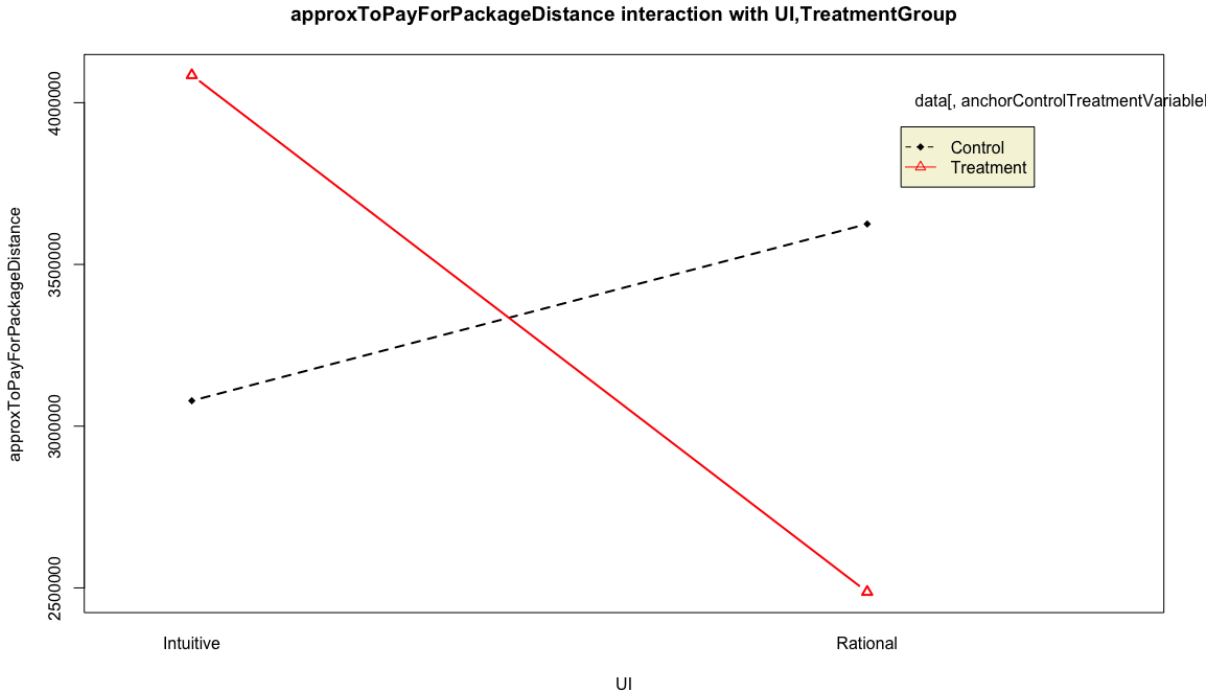


Figure 15. Interaction plot of estimated software value by use of intuition and decision complexity.

Looking at the boxplots visually, the treatment group does seem to have moved closer to a distance of 0 from the experimenter-provided anchor. Since the data did not follow a normal distribution, a Mann-Whitney U test was performed and it showed a significant difference between the control and treatment groups ($p<0.001$).

9.4.2 Purchase satisfaction anchoring bias

The first group was asked to rate their satisfaction with the last purchase they can remember (before answering about the case study packages). All groups were asked what they thought their satisfaction would be for each of the three case study packages (X, Y, Z). The second group was asked to rate their satisfaction with the last purchase they can remember after rating the case study packages. Results were tested to determine if there was any relationship between the answers given.

The variable expected satisfaction with package Y did contain Tukey (1977) defined non-parametric outliers which were transformed through Winsorizing to the maximum and minimum non-outlier values. The Shapiro-Wilk test of normality for expected satisfaction with purchase (group 2) showed that the results do not follow a normal distribution ($p<0.001$). Therefore, non-parametric tests were used for analysis.

A non-parametric Spearman correlation test showed the following significant results:

Table 6. Significant correlations for purchases satisfaction anchoring bias source items.

Group	Variable 1	Variable 2	Estimate	Sig
Overall	expected satisfaction with package X	expected satisfaction with purchase (group 2)	0.33	0.03
Simple	expected satisfaction with package X	expected satisfaction with purchase (group 1)	0.46	0.03
Complex	expected satisfaction with package X	expected satisfaction with purchase (group 2)	0.62	<0.01
Complex	expected satisfaction with package Y	expected satisfaction with purchase (group 2)	0.46	0.02

Because of the multiple variable options for this anchoring test, significance testing was used to determine the effect, and expected satisfaction with package X was chosen as the overall anchor, and expected satisfaction with purchase (group 2) as the affected variable in order to calculate a level of bias. The anchor distance (and bias value) was then calculated as the distance between the anchor and the estimate, scaled to a range of 0-1 from the lowest (0) and highest (1) values. The resulting purchase satisfaction anchoring bias variable did contain Tukey (1977) outliers (values:0.17,0). These were transformed through Winsorizing to the maximum and minimum non-outlier values. The Shapiro-Wilk test of normality for purchase satisfaction anchoring bias showed that the results do not follow a normal distribution ($p<0.001$). This final variable formed part of the composite anchoring bias variable.

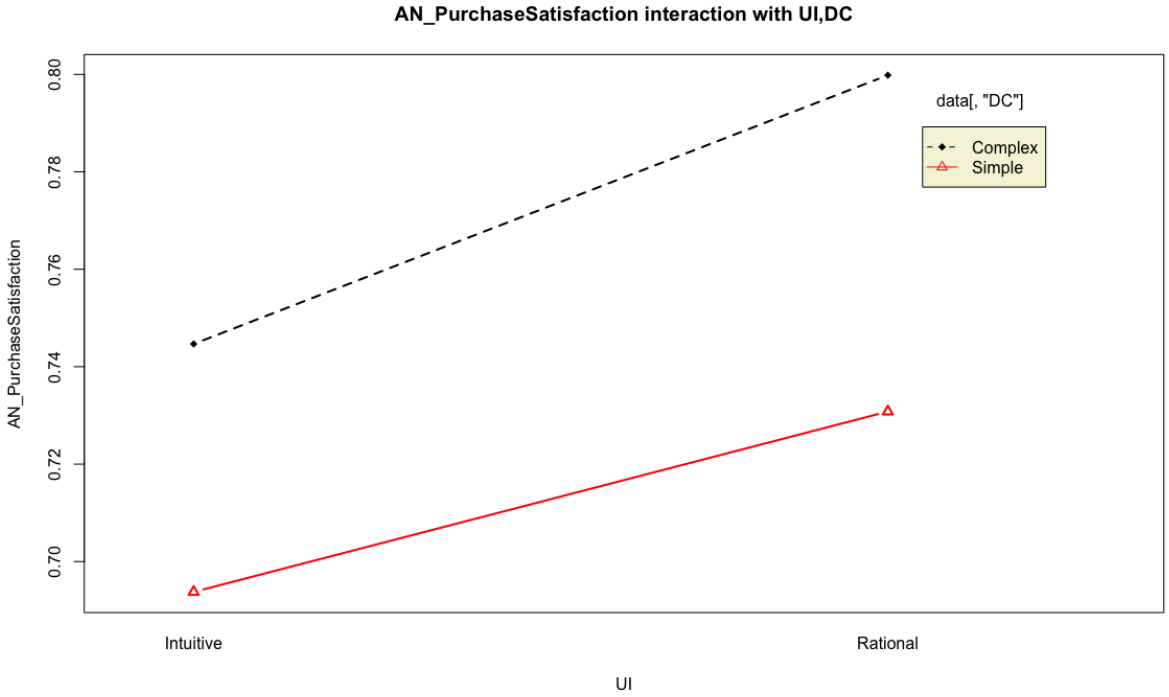


Figure 16. Interaction plot of purchase satisfaction anchoring bias by use of intuition and decision complexity.

9.4.3 Previous purchase cost anchoring bias

In a similar way to the purchase satisfaction anchoring bias measure, for previous purchase cost anchoring bias, the first group was asked to estimate the cost of the last purchase they can remember (before answering about the case study packages). All groups were asked what they thought the cost would be for each of the three case study packages (X,Y,Z). The second group was asked to estimate the cost of the last purchase they can remember after rating the case study packages. Results were tested to determine if there was any relationship between the answers given. The approximate cost of previous purchase (group 1) variable did contain Tukey (1977) outliers (values: R20m, R19.9m). These were transformed through Winsorizing to the maximum and minimum non-outlier values. The Shapiro-Wilk test of normality for approximate cost of previous purchase (group 1) showed that the results did not follow a

normal distribution ($p<0.001$), and therefore, non-parametric tests were used for analysis. Spearman's correlation tests shows multiple significant correlations. Because of the multiple variable options for this anchoring test, significance testing was used to determine the effect, and expected cost of package X was chosen as the overall anchor, and approximate cost of previous purchase (group 1) as the affected variable in order to calculate a level of bias. The anchor distance (and bias value) was then calculated as the distance between the anchor and the estimate, scaled to a range of 0-1 from the lowest (0) and highest (1) values. The resulting previous purchase cost anchoring bias variable did contain Tukey (1977) defined non-parametric outliers (values:0.24, 0.37, 0.28, 0.07, 0.41, 0.06, 0.42, 0.06, 0, 0.18, 0.36). These were transformed through Winsorizing to the maximum and minimum non-outlier values. The Shapiro-Wilk test of normality for previous purchase cost anchoring bias showed that the results did not follow a normal distribution ($p<0.001$).

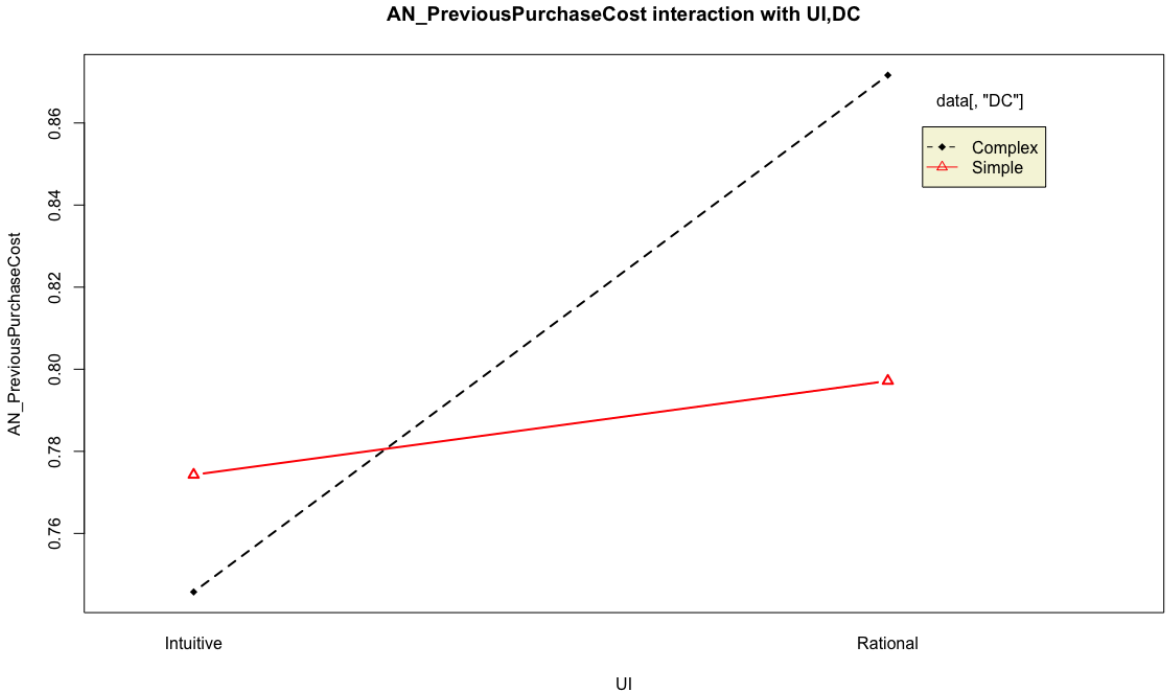


Figure 17. Interaction plot of previous purchase cost anchoring bias by use of intuition and decision complexity.

Since the data were not normally distributed, a Mann-Whitney U test was performed and it showed a significant difference between the complex-intuitive and complex-rational groups (p=0.04).

9.5 Estimate of number of months to implement the software anchoring bias

In a similar way to the previous anchoring bias measures, for the estimate of number of months to implement the software anchoring bias, the first group was asked to estimate the number of software purchases they have made (before answering about the case study). All groups were asked to estimate the number of months it would take to implement any of the packages in the case study. The second group was asked to estimate the number of software

purchases they could remember making after rating the case study packages. Results were tested to determine if there was any relationship between the answers given. The variable number of previous purchases (group 1) contained Tukey (1977) outliers (values:20, 20, 20, 25, 160, 44, 20), the number of previous purchases (group 2) variable contained outliers (values:20, 20, 25, 20), and the estimate of number of months to implement the software variable contained outliers (values:24, 36, 36, 24, 24, 36), which were transformed through Winsorizing to the maximum and minimum non-outlier values. The Shapiro-Wilk test of normality for number of previous purchases (group 1) showed that the results did not follow a normal distribution ($p<0.001$). Therefore, non-parametric tests were used for analysis. None of the variables were significantly correlated based on Spearman's correlation testing, and therefore this bias measure was set to 0.

9.5.1 Anchoring bias summary

The resulting composite anchoring bias variable did not contain any Tukey (1977) outliers, and the Shapiro-Wilk test of normality showed that the results did not follow a normal distribution ($p<0.001$).

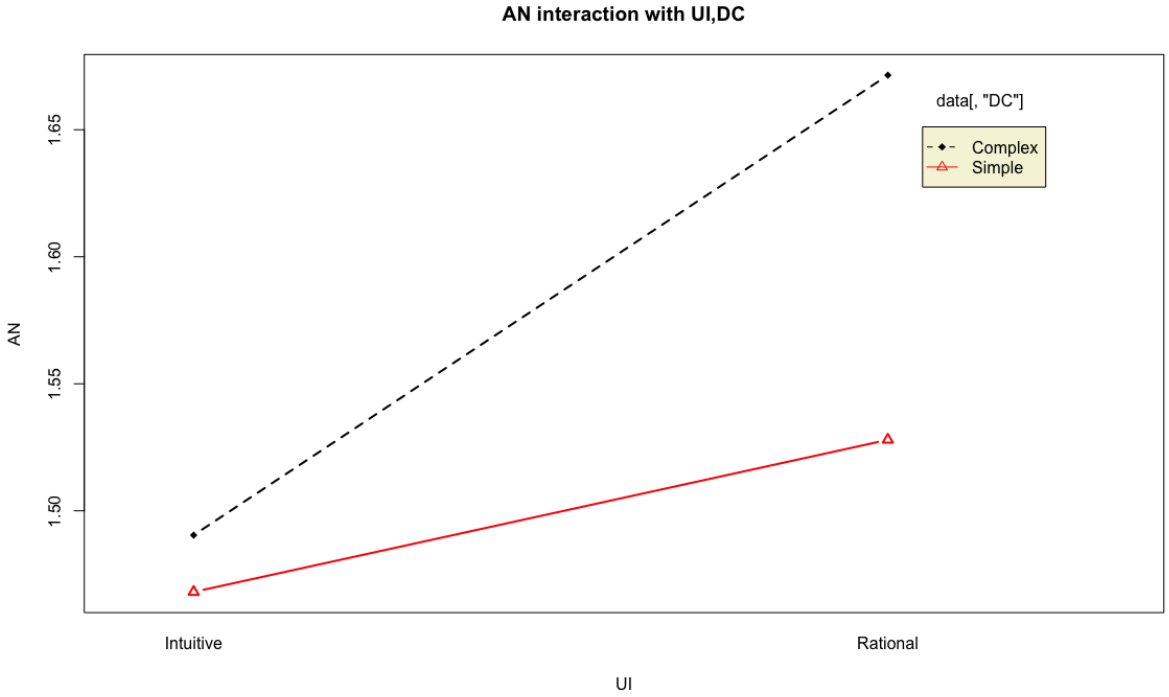


Figure 18. Interaction plot of anchoring bias by use of intuition and decision complexity.

Since the data were not normally distributed, a Mann-Whitney U test was performed and it showed a significant difference between complex-intuitive and complex-rational groups (p=0.04) and between the rational-simple and rational-complex groups (p=0.04).

9.6 General bias

There were three questions which attempted to elucidate a biased reaction, using modified versions of previously used bias measures, and these were based on the following variables:

- program favour general bias (based on favouring program 1, favouring program 2).
- operation general bias (based on evaluation of successful operation, evaluation of failed operation)

- estimated analysis weeks general bias (based on estimation of number of weeks of analysis for package A).

Each of these variables will be presented in the following sections, followed by the composite general bias variable.

9.6.1 Program favour general bias

The program favour general bias variable was based on the favouring program 1 and favouring program 2 items in the questionnaire. The problem is based on a modified version of the disease framing problem (Tversky & Kahneman, 1981; West, Toplak & Stanovich, 2008). The results should be the same despite the different wording of the two questions. Different answers indicate a descriptive invariance bias (Tversky & Kahneman, 1981). The program favour general bias variable is based on the favouring program 1 item, and favouring program 2 item. After the formula was applied, the resulting program favour general bias variable had no Tukey (1977) outliers, and the Shapiro-Wilk test of normality showed that the results did not follow a normal distribution ($p<0.001$).

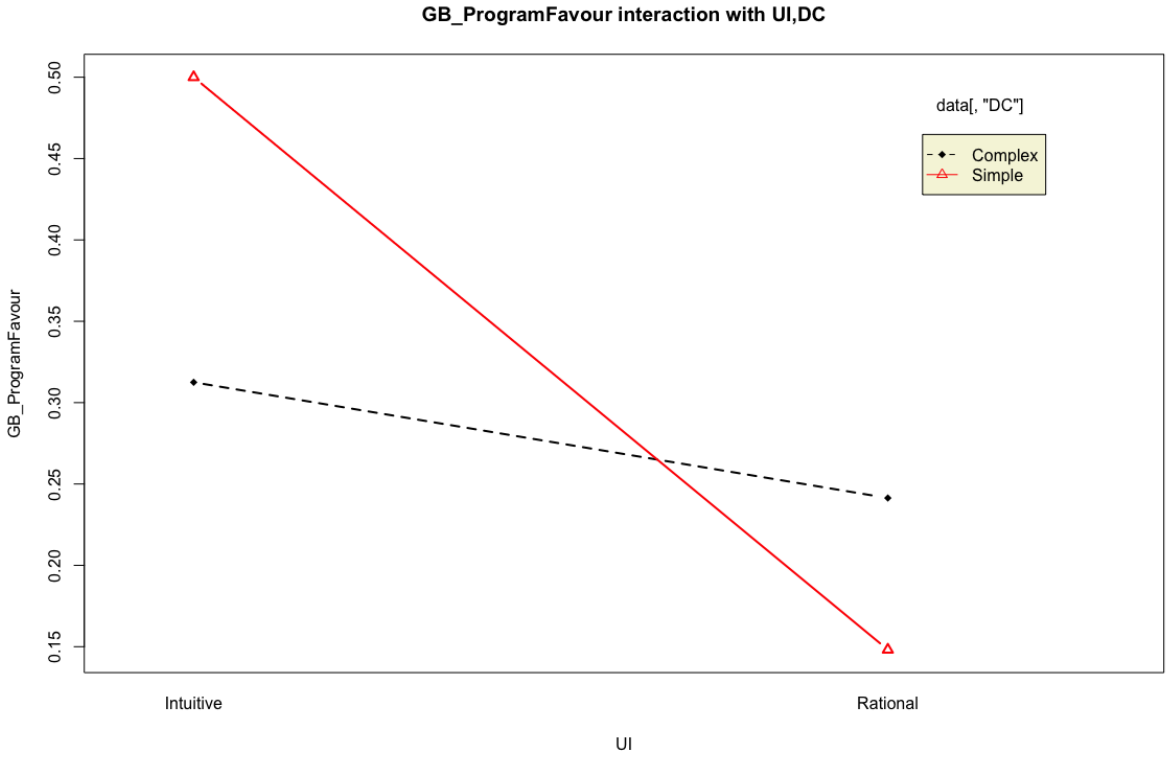


Figure 19. Interaction plot of program favour general bias by use of intuition and decision complexity.

The interaction plot visually showed some interaction between decision complexity and use of intuition, with the intuitive group showing more bias than the rational group. Since the data were not normally distributed, a Mann-Whitney U test was performed, and it showed a significant difference between the simple-intuitive and simple-rational groups (p=0.01).

9.6.2 Operation general bias

The operation general bias variable was calculated from the evaluation of successful operation and evaluation of failed operation variables. The questions were based on modified versions of outcome bias tests (Baron & Hershey, 1988; Toplak, West, & Stanovich, 2011). The choices should be the same despite the outcome or the participant will have demonstrated an outcome bias. The resulting operation general bias composite variable had too few distinct

values to check for outliers. The Shapiro-Wilk test of normality for operation general bias showed that the results did not follow a normal distribution ($p<0.001$).

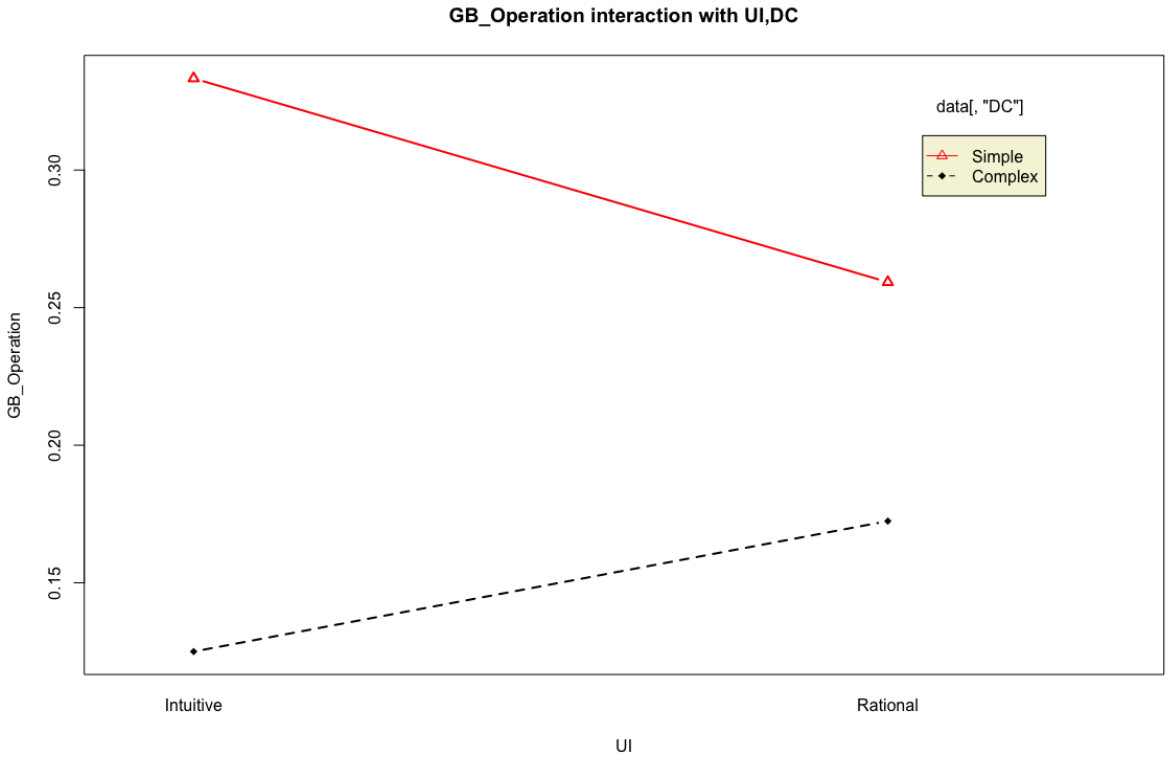


Figure 20. Interaction plot of operation general bias by use of intuition and decision complexity.

The interaction plot visually showed no interaction between use of intuition and decision complexity, but did show higher levels of bias for the intuitive group within the same complexity.

9.6.3 Estimated analysis weeks general bias

This measure was a modified version of the bat and ball question from the cognitive reflection test (Frederick, 2005). The estimated analysis weeks general bias variable was calculated from the estimation of number of weeks of analysis for package A variable which was calculated by checking if the item in the questionnaire was answered correctly. The resulting estimated analysis weeks general bias variable did not have outliers. The Shapiro-Wilk test of normality showed that the results did not follow a normal distribution ($p<0.001$).

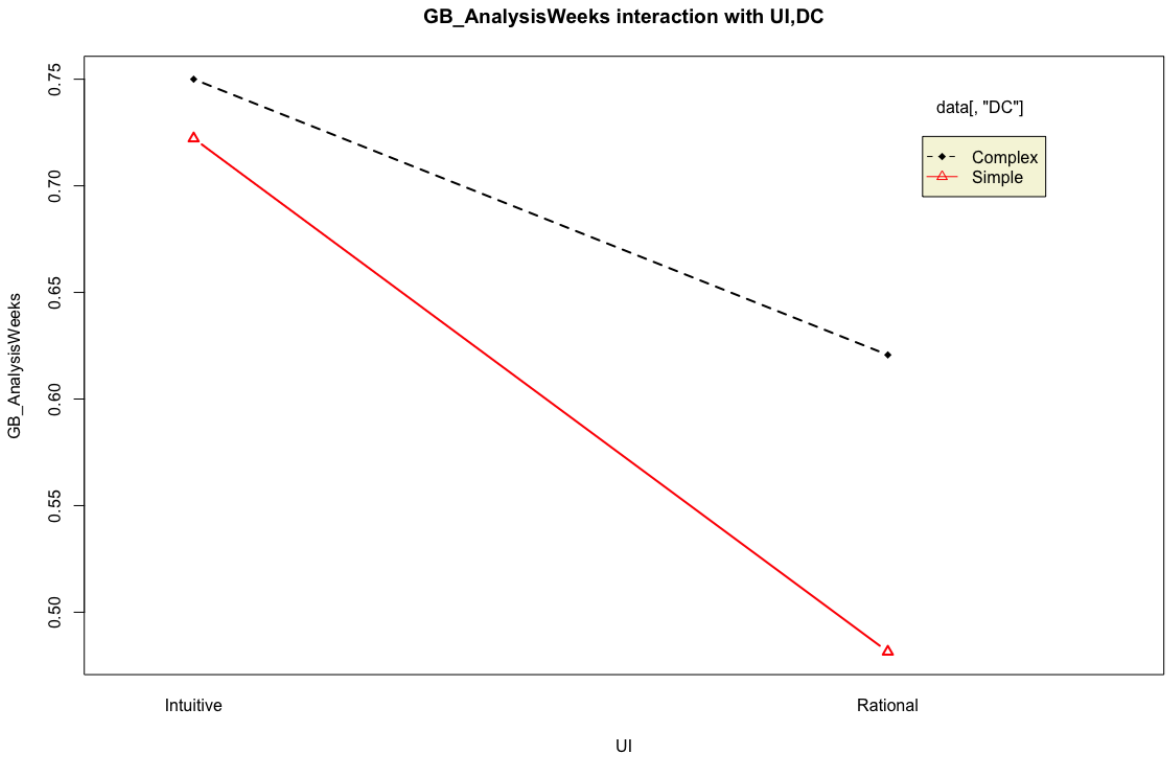


Figure 21. Interaction plot of estimated analysis weeks general bias by use of intuition and decision complexity.

The interaction plot visually showed a minor interaction between use of intuition and decision complexity and showed the intuitive group as more biased than the rational group.

9.6.4 General bias summary

The general bias variable was a composite variable calculated as the sum of the program favour general bias, operation general bias, and estimated analysis weeks general bias variables. The resulting general bias composite variable did not contain any Tukey (1977) outliers, and the Shapiro-Wilk test of normality showed that the results did not follow a normal distribution ($p<0.001$).

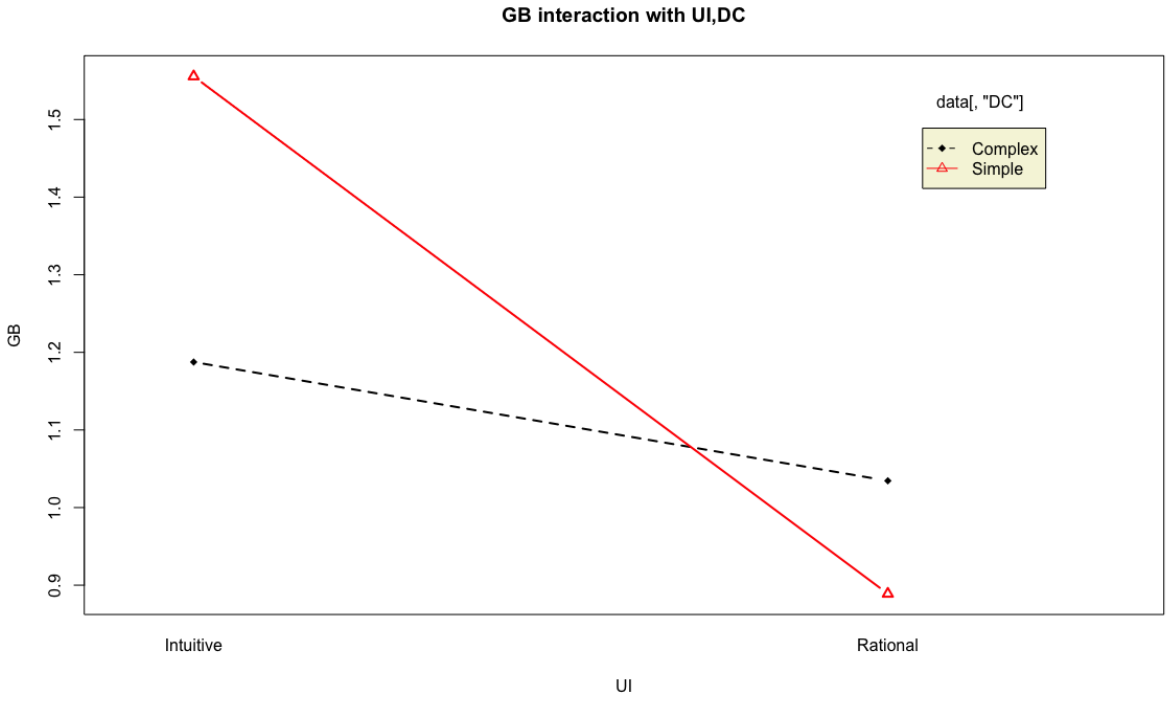


Figure 22. Interaction plot of general bias by use of intuition and decision complexity.

Visually, the interaction plot showed an interaction between use of intuition and decision complexity, with complexity seeming to affect the intuitive group more, and the intuitive group showing more bias than the rational group. Since the data were not normally distributed, a Mann-Whitney U test was performed and it showed a significant difference, for

general bias, between the intuitive and rational groups (p=0.01), and a significant difference between the simple-intuitive and simple-rational groups (p=0.01).

9.6.5 Decision bias and decision accuracy

The decision bias variable was a composite variable that was calculated as a sum of the attribute substitution bias, anchoring bias, and general bias variables. The decision accuracy variable was inverted from decision bias (1-decision bias) as a proxy for decision accuracy. The resulting decision accuracy composite variable did not contain Tukey (1977) outliers, and the Shapiro-Wilk test of normality showed that the results did follow a normal distribution (p= 0.55).

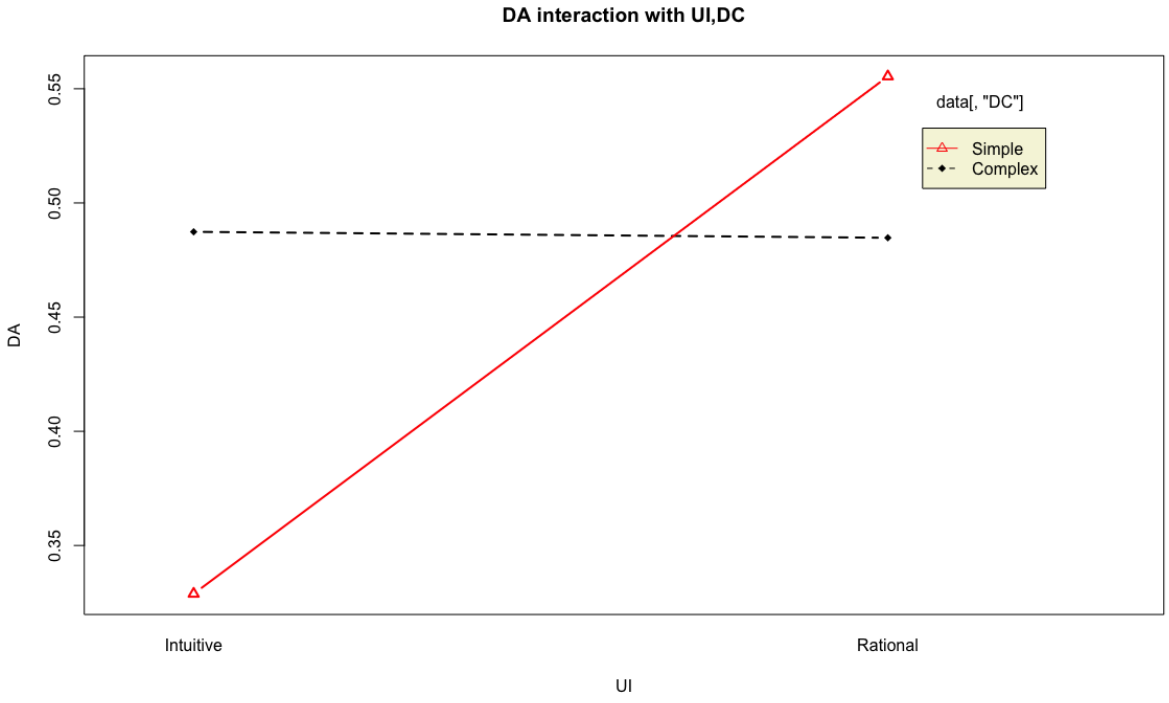


Figure 23. Interaction plot of decision accuracy by use of intuition and decision complexity. Note that this plot will appear opposite to previous bias plots since this shows accuracy as the inverse of bias.

The interaction plot visually showed an interaction between use of intuition and decision complexity, and although the results appear similar for the complex group across the intuitive and rational groups, for the simple group there appeared to be a large difference between the intuitive and rational groups, with the intuitive group showing lower accuracy. Since the data were normally distributed, a student's t-test was performed and it showed a significant difference in decision accuracy between the intuitive and rational groups ($p=0.02$). Within the simple group there was a significant difference between the intuitive and rational groups ($p<0.01$), and within the intuitive group there was a significant difference between the simple and complex groups ($p=0.04$). Therefore, for the overall decision accuracy variable, within the simple group the rational group was significantly more accurate than the intuitive group, and within the intuitive group the simple group was significantly more accurate, and overall the rational group was significantly more accurate than the simple group. The complex group did not show significant differences between intuitive and rational accuracy. Therefore, for the simpler problems, use of intuition resulted in less accurate decisions than rational deliberation.