Leveraging big data for strategic decision-making

Frank Mourinho

26298628

A research project submitted to the Gordon Institute of Business Science, University of Pretoria, in partial fulfilment of the requirements for the degree of Master of Business Administration.

6 November 2017

Abstract

As organisations find themselves in an environment of uncertainty and increased levels of competition, the rapid changes in technology has resulted in the availability of more data than ever before. The advent of big data has provided firms with the opportunity to take advantage of the increased volume, variety, velocity and veracity of data, allowing for increased levels of innovation, proactivity and risk-taking and a culture of evidence-based decision-making.

Although, the challenge remains that organisations are still struggling to successfully extract value from their data. This research aimed to understand if organisations can become more entrepreneurial and achieve a culture of evidence-based decision-making by leveraging big data for strategic decision-making. A quantitative study was used to measure the relationship between big data, evidence-based decision-making and entrepreneurial orientation using multivariate data analysis.

The results reported statistically significant positive correlations between big data and both evidence-based decision-making and entrepreneurial orientation. Furthermore, big data skills was reported as a significant predictor of both entrepreneurial orientation and evidence-based decision-making. These findings provide empirical evidence and guidance for both academics and business practitioners on the importance of skills and how organisations can leverage big data to become more entrepreneurial and drive an evidence-based decision-making culture.

Keywords

Big data, entrepreneurial orientation, evidence-based decision-making, strategy

Declaration

I declare that this research project is my own work. It is submitted in partial fulfilment of the requirements for the degree of Master of Business Administration at the Gordon Institute of Business Science, University of Pretoria. It has not been submitted before for any degree or examination in any other University. I further declare that I have obtained the necessary authorisation and consent to carry out this research.

Frank Mourinho
6 November 2017

Table of contents

List of figures

List of tables

Chapter 1: Introduction

1.1 Research Problem

“Big Data – no matter how comprehensive or well analyzed – needs to be complemented by Big Judgement” (Shah, Horne & Capellá, 2012, p. 6).

In an era of globalization, increased transparency and rapidly changing technologies, it is generally acknowledged that organisations find themselves in a state of risk and instability, as they search for new, creative and innovative ways to compete in a dynamic and competitive environment (Mathews, 2016; Purnama & Subroto, 2016; Reeves & Deimler, 2011; Prajogo, 2016; Sousa & Coelho, 2011; Slåtten & Mehmetoglu, 2011). Competition and the state of market leadership have changed, where the percentage of companies falling out of the top three rankings within their industry has increased from 2% to 14% from 1960 to 2008 (Reeves & Deimler, 2011). Furthermore, the probability that a market share leader is also a leader in profitability is significantly lower than before, decreasing from 30% to 7% from 1950 to 2007 (Reeves & Deimler, 2011).

The question then arises, how can organisations equip themselves to compete in an era of increasing complexity, changing technology and a state of hyper-competition that is unprecedented (Mathews, 2016), in order to answer the call of boards in search of competitive advantage whilst mitigating risk (Hill & Davis, 2017). Hagel (2016) refers to the rise of digital technology that is reshaping the business landscape globally as the Big Shift, and emphasizes the need for an entrepreneurial. Rauch, Wiklund, Lumpkin and Frese (2009, p. 8) sum up the above when they state: “In an environment of rapid change and shortened product and business model lifecycles, the future profit streams of existing operations are uncertain and businesses need to constantly seek out new opportunities”. Reeves and Deimler (2011, p. 3) suggest that organisations must be “really good at learning how to do new things” instead of “being really good at doing some particular thing”.

In response to this, big data has garnered an increasing amount of attention over the last 10 years as an enabler of competitive advantage, from both industry practitioners such as Etihad Airways, Walt Disney, Google, Facebook and Walmart (Alharthi, Krotov

& Bowman, 2017; Chen, Chiang & Storey, 2012; McAfee, Brynjolfsson, Davenport, Patil & Barton, 2012), as well as from academics (Chen et al., 2012; Frizzo-Barker, Chow-White, Mozafari & Ha, 2016). Big data is often described as an increase in the size, types, speed and accuracy of data (McAfee et al., 2012; Goes, 2014; Pigni, Piccoli & Waston, 2016; Alharthi et al., 2017) and academics are beginning to research its relationship to an organisation's ability to achieve their goals through better decision-making (Elgendy & Elragal, 2016).

The big data phenomenon has emerged from various advances in both physical and digital technology, which have contributed to an increased interconnectedness between both people and organisations. Social media, the Internet of Things (IoT), mobile and cloud technology, and other internet-based technologies continue to provide the impetus for the age of digitization (Ransbotham, Fichman, Gopal and Gupta, 2016). Of these innovations, the IoT has garnered a significant amount of attention given its potential for data generation. The IoT is defined as a pervasive presence of smart devices or "things" such as mobile phones, Radio-Frequency Identification tags (RFID) and sensors which interact with one another and generate useful data whilst guaranteeing trust, privacy and security (Atzori, Iera & Morabito, 2010; Chen et al., 2012). According to Gubbi, Buyya, Marusic and Palaniswami (2013), the number of interconnected devices overtook the number of people on the planet in 2011 at nine billion and this is expected to more than double to 24 billion devices by 2020. Moreover, IBM estimates that 80% of information is unstructured content (email, texts, images or videos) and that this will grow twice as fast as regularly structured databases (George, Haas & Pentland, 2014).

This trend of increasing "things" means that the variety, volume and velocity of data will continue to grow along with these devices (Akbay, 2015). Walmart generates approximately 2.5 petabytes (2.5 million gigabytes) of customer data on a daily basis (Frizzo-Barker et al., 2016; McAfee et al., 2012). According to Alharthi et al. (2017), there is a total of 1.8 zettabytes (1 trillion gigabytes) of digital data in the world, of which 90% was created in the last two years. This explosion of data being generated means that businesses have a wealth of information to assist in curating their value propositions and delivery, and run the risk of possibly conceding this competitive advantage to competitors if they are unable to successfully incorporate this data into their strategic decision-making.

Some businesses are already finding ways to take advantage of the opportunities that big data presents, such as Ernst & Young (2012) who have recognized that organisations are under increased pressure to improve their anti-corruption compliance programs seeing as only 14% of fraud and corruption is detected by internal audits. In response to this, Ernst & Young have constructed a bribery and corruption detection suite of tools known as ABC Analytics to aid businesses with anti-corruption compliance. This data-driven approach creates a new opportunity for detecting bribery and corruption, compared to traditional auditing methods, by leveraging not only financial data but text mining of emails as well. Customers of Xcel Energy are able to track energy usage of their homes in real-time thanks to smart grids that have been installed in their homes (Advanced Performance Institute, 2017). These smart grids will further allow the businesses that supply this energy to predict usage and plan for future demand and infrastructure requirements.

These are just two of the examples of how big data is changing how businesses and individuals interact and operate in this new era. Big data is not only being used for commercial business purposes such as targeted marketing or creating retail and manufacturing efficiencies (Bean, 2017). Research has also been done on how big data has assisted in addressing health issues such as the Ebola outbreak in West Africa (Amankwah-Amoah, 2016), and amongst other sectors such as law enforcement, city optimization and improving sports performance through analytics (Datameer, 2016). Further examples of big data usage are provided in Chen et al. (2012), Alharthi et al. (2017) and Pigni et al. (2016).

Although the value from data-driven decision-making may be recognized, the challenge remains that most organisations have not been successful in integrating the use of big data into their strategic decision-making (PricewaterhouseCoopers, 2015; Shah et al., 2012; George et al., 2014; Goes, 2014). This has been attributed by some authors to firstly the lack of skills and data infrastructure (Pigni et al., 2016), and secondly the inability of organisations to drive a data-driven culture (Bean, 2017). Industries such as the taxi industry, which was disrupted by Uber – the largest “taxi” company (Pigni et al., 2016), run the risk of becoming insignificant in the future if they do not acknowledge this big data phenomenon. Furthermore, organisations also recognize that they must respond to the shifts in their environment, such as customer demands if they wish to survive in the long-term (Slåtten & Mehmetoglu, 2011).

One way of defining an organisation's ability to learn and proactively achieve new innovation whilst taking risks is a concept known as Entrepreneurial Orientation (EO) and has been extensively studied by academic researchers (Gupta & Gupta, 2015). The concept of EO is explained as a strategy-making process that key decision makers use in order to achieve organisational goals, company vision and create a competitive advantage for themselves (Rauch et al., 2009). Academic literature converges on the idea that firms that can achieve an EO may benefit from "highlighting newness, responsiveness, and a degree of boldness" (Rauch et al., 2009, p. 8) and has been associated with better firm performance and competitive strategy (Lechner & Gudmundsson, 2014; Linton & Kask, 2017), all of which can be considered necessary outcomes for any organisation in the current global environment (Reeves & Deimler, 2011).

Given the suggestion that big data can provide improved internal efficiencies, additional revenues, better customer experience and overall improved profitability through better decision-making (Alharthi et al., 2017; Bean, 2017), this research study aimed to provide useful insights regarding the above research problem by placing emphasis on a possible relationship between big data and an organisation's ability to be entrepreneurial, measured as EO.

1.2 Academic and business motivation for research

1.2.1 Entrepreneurial Orientation

The literature on EO and the external environment has been considered "well-researched" (Linton & Kask, 2017, p. 169) with the number of studies conducted on EO increasing significantly since the 1990s (Rauch et al., 2009). It is noted that although the overall research has increased, less focus has been placed on the internal context of an organisation and its relationship to EO (Linton & Kask, 2017, p. 169). This is reiterated by Wales, Gupta and Mousa (2011, p. 368), where: "few studies have examined the moderating role of factors that reside within the organisation". Therefore, more research is required in understanding the relationship between the various internal aspects of an organisation and EO (Linton & Kask, 2017). One of these aspects, identified by this research, is an organisation's ability to successfully leverage big data when making strategic decisions.

A systematic review of the published EO literature conducted by Wales et al. (2011) explored 158 empirical articles and concluded that although significantly more research has already been conducted on EO, various gaps remained for further research. Certain outcomes from their review, relevant to this research, are identified below:

- Whilst research into the EO construct has increased over the last three decades, the antecedents of EO still remain unclear (Wales et al., 2011). Research into understanding the antecedents of EO has been conducted, although the focus has been placed on organisational, environmental and top manager characteristics. This presents an opportunity to study the availability of big data (as a means of providing timeous, accurate and new insights) as an antecedent of EO.

- Various moderator variables have been explored extensively in the EO literature (Wales et al., 2011). These included CEO tenure, human resources and networking to name a few. Furthermore, it is clear that certain aspects of culture have been researched, although the aspect of using data to make decisions (which can be worded as data-driven decision-making or evidence-based decision-making) has yet to be studied as a moderator variable. This highlighted a gap in the EO literature.

- Finally, Wales et al. (2011, p. 374) suggest that future research into the EO construct be done regarding "issues of practitioner relevance". They found that EO researchers often investigated factors that provided limited value to practitioners, as influencing certain variables might be extremely difficult or impossible. Based on this, this research is appropriate given that big data capabilities can be acquired through the hiring of key skills, implementation of required systems and the focus on driving a data-driven culture. All of which provide practical opportunities for business practitioners.

The question is no longer whether organisations should act entrepreneurially or not; instead a more pertinent question is how organisations can act more entrepreneurially (Covin, Green & Slevin, 2006). Previous research on EO has tended to focus on the linear effect of EO on firm performance (Wales et al., 2011), this has been criticized as an oversimplification given that EO has been concluded to increase variability in performances rather than solely increasing performance itself (Linton & Kask, 2017; Rauch et al., 2009). Wales et al. (2011), call for research which treats EO as the dependent variable; this research study aimed to answer that call.

Reeves and Deimler (2011) discuss the need for adaptability within this complex and hostile environment and suggest that organisations must be able to read and act on market signals, experiment with new products or services whilst receiving timeous and reliable feedback, and manage complex multi-company systems. The possession of dynamic capabilities can promote a firm's competitive advantage (Schilke, 2014) and this study proposes that big data is uniquely positioned as an internal component which will allow organisations to be more dynamic and ultimately attain higher levels of EO.

1.2.2 Big data and evidence-based decision-making

The concept of using data or evidence to inform and enhance the decision-making process has gained momentum in the field of medicine (Pfeffer & Sutton, 2006). Although the concept has not been found to be ubiquitous in the workplace (Rousseau, 2006), it has extended to business and management literature (Pigni et al., 2016; Popovic, Hackney, Coelho & Jaklic, 2012) and is recognized as a key component in extracting value from what is considered the raw material of the 21st century (Elgendy & Elragal, 2016). Research (Brynjolfsson, Hitt & Kim, 2011; LaValle, Lesser, Shockley, Hopkins & Kruschwitz, 2011; Popovic et al., 2012) has shown that the use of data to drive decision-making can have positive outcomes on both output and productivity, and research regarding big data and an organisation's decision-making ability is growing in focus (Elgendy & Elragal, 2016). Given the technology revolution (Purnama & Subroto, 2016), the amount of data generated continues to grow, placing emphasis on the importance of understanding how to harness and generate valuable insights from this data. Therefore, further research into understanding how big data can be used to create competitive advantage is necessary.

Based on interviews conducted with 1,800 senior business leaders, Price Waterhouse Coopers (PwC) state that organisations showed confidence in their ability to extract value from their data. Contrary to this self-belief however, through close scrutiny of these interviews PwC concluded that three in four businesses extracted little or no advantage from their data initiatives (PricewaterhouseCoopers, 2015). They attributed this to a lack of technical capabilities, skills, the focus of investment and culture within these organisations. A survey of Fortune 1000 executives showed that only 48.4% of organisations report achieving measurable results from big data initiatives (Bean, 2017). These results highlight a gap that remains in the understanding of how big data can be used effectively within organisations.

It is not just the view of the practitioner that highlights the need for research in this field. George et al. (2014) reiterate that although big data has become commonplace within business, there is still a lack of published management scholarship that aims to better understand how businesses should use the tools at their disposal. A systematic review by Frizzo-Barker et al. (2016, p. 1) highlighted that big data "remains a fragmented, early-stage domain of research in terms of theoretical grounding, methodological diversity and empirically oriented work". Even information systems (IS) literature, although well researched, lacks clarity on how business intelligence systems dimensions are interrelated or how they impact the use of business intelligence systems (Popovic, Hackney, Coelho & Jaklic, 2012).

A bibliometric study of the academic and industry publications done by Chen et al. (2012) found that the first publication of big data appeared in 2001 and only started to increase in 2007 to 26 publications. Furthermore, 76% of big data literature between 2009 and 2014 was published between 2013 and 2014, highlighting the increasing trend of interest in the field (Frizzo-Barker et al., 2016). Moreover, the majority of these articles did not define big data (59%) and were conceptual, as opposed to empirical, in design (72%). This is mainly because the big data phenomenon has for the most part been practitioner-led (George et al., 2014; Goes, 2014). For example, the Hadoop technology that allows businesses to deal with very large and complex datasets originated from Google (Goes, 2014). This presents an opportunity for further research that aims to bridge the gap between industry and academics.

1.3 Research aim

As discussed above, whilst organisations are continuously under pressure to achieve competitive advantage, big data has been suggested as a possible enabler of dynamic firm capabilities and better decision-making. Although a gap remains between academics and industry practitioners on how this can be achieved and whether a link exists between an organisation's ability to effectively make use of their data and their ability to be entrepreneurial.

The author of this research aimed to contribute to the literature by not only providing empirical evidence of the relationship between an organisation's big data capabilities and their ability to be entrepreneurial but also by exploring the importance of the different big data capabilities, such as skills or systems, in order to provide practical and useful insights to organisations to be more successful with their big data initiatives.

The following chapter discusses the current literature available regarding big data capabilities, evidence-based decision-making and EO. For the remainder of this study, the terms data-driven decision-making and evidence-based decision-making are treated as interchangeable.

Chapter 2: Literature Review

2.1 Introduction

The previous chapter has outlined the aim and need for this research and introduced the three key constructs to be studied. These are big data capabilities, evidence-based decision-making and EO. This chapter aimed to gain an understanding of the current literature regarding these constructs. The concepts of big data and evidence-based decision-making are briefly discussed to provide context, following which is an examination of the interplay between these two constructs. Furthermore, key antecedents for successful utilization of big data are discussed. Finally, the concept of EO is discussed and a model of the research is proposed.

2.2 Big Data

Big data has gained increased awareness with the evolution of data practices where data was previously considered "stock" (Davenport, Barth & Bean, 2012), to the continuous flows of near real-time data that is available today (Pigni et al., 2016). Big data literature does not offer a uniform definition and has been described in various ways, such as the creation of Digital Data Streams as an outcome of new technology (Pigni et al., 2016), as a result of the data generation from the introduction of mega trends such social media and the IoT (Akbay, 2015), as a tool that can enable organisations to track the impact of explorative ventures (Bøe-Lillegraven, 2014) and as a shift in thinking regarding key business decisions that inform strategy (Frizzo-Barker et al., 2016).

The majority of big data literature also focuses on the four V's approach when trying to define big data, which is volume, variety, velocity and veracity (McAfee et al., 2012; Goes, 2014; Pigni et al., 2016; Alharthi et al., 2017).

This is explained as:

- Volume – there is significantly more data. Google processes approximately 24,000 terabytes (24 million gigabytes) of data every day (Davenport et al., 2012).

- Variety – there are new types of data. For example, traditional forms of data have always been recorded in the form of text or numerals. Current systems allow for the recording of data in images, videos and locations (Alharthi et al., 2017).
- Veracity – this refers to the accuracy of the data, for example, some data may exhibit a high level of noise such as social media and must, therefore, be cleaned (Goes, 2014; Pigni et al., 2016).
- Velocity – the speed at which data updates has increased as well, to the point that the analysis of near real-time data is possible (Pigni et al., 2016).

It must be noted that although the big data phenomenon has only recently (since 2007, Chen et al., 2012) gained popularity, the understanding that business intelligence contributes to decision-making has been around for significantly longer (Chen et al., 2012) and has been studied by other authors (Popovic et al., 2012). The difference between big data and business intelligence is defined as big data being infinite, real-time and unstructured compared to business intelligence which is finite, offline and structured (Alharthi et al., 2017). This ever-increasing scope of sources from big data allows for a wider view of events, whether they be physical, online or mobile transactions; or interactions in the form of internet clicks, social media posts or sensor networks (Akbay, 2015; George et al., 2014).

2.3 Evidence-based decision-making culture

This section briefly discusses literature with regards to evidence-based decision-making (Pfeffer & Sutton, 2006; Rousseau, 2006), organisational culture (Gregory, Harris, Armenakis & Shook., 2009; House, Javidan, Hanges, Dorfman, 2002) and decision-making (Malakooti, 2012) as these concepts are all present in the evidence-based decision-making construct of this study, which is a culture of using data to inform decision-making.

2.3.1 Evidence-based decision-making

Evidence-based decision-making is defined as the explicit use of the best and most current evidence (as opposed to personal preference or unsystematic experience) to make decisions (Rousseau, 2006), and has gained momentum in the medical field over the last few decades (Pfeffer & Sutton, 2006). One would expect that the majority of decisions concerning a patient's wellbeing would be rooted in evidence, although surprisingly only 15% of physicians decisions are evidence-based (Pfeffer & Sutton, 2006). This dynamic is also experienced in the business workplace where managers prefer to rely on personal experience, business books and consultants over the use of evidence when making decisions (Rousseau, 2006).

Pfeffer and Sutton (2006) go on to reframe the concept of evidence-based decision-making in the context of business as evidence-based management. The task is notably more difficult for business decisions, when juxtaposed with patient decisions, as business models, unlike human conditions, are not homogenous. This serves as motivation that decisions based on evidence are more necessary for the business sector, seeing as it would be impossible to gain the experience to answer strategic decisions across all types of businesses when they differ so much.

2.3.2 Organisational Culture

Culture is defined as the set of values or ideas that inform or drive human behaviour in a specific manner (Hofstede, 1980). Whilst there is no standard definition of organisational culture, it is similarly defined in the literature as the set of norms, values and basic assumptions, shared across the organisation, that inform employees of how work is performed, how they will be evaluated and how they should relate to other stakeholders associated with the organisation (Cummings & Worley, 2015; Gregory et al., 2009; Morrison, Brown & Smit, 2006).

Organisational culture is important because it can affect how decisions are made now and in the future, as leaders respond to organisational culture over time and may alter their behaviours to fit the desired norms (House et al., 2002). The attitude or mindset of a business towards big data initiatives can either assist or hinder their ability to create value from data-driven initiatives (Alharthi et al., 2017; McAfee et al., 2012). Thus in order to achieve a form of evidence-based decision-making, a distinct mindset must be adopted (Pfeffer & Sutton, 2006). Organisations that wish to make use of value generated by big data effectively would need to implement an evidence-based decision-making culture to ensure that data is consulted before carrying out any important decisions. This would entail demanding facts, examining logic, avoiding conventional wisdom or gut-feelings, encouraging trial programs and rewarding evidence-based behaviour (Pfeffer & Sutton, 2006).

Organisational culture has been shown to have positive direct and indirect effects on the effectiveness of an organisation (Gregory et al., 2009). For example, in a study conducted by Gregory et al. (2009), a positive relationship between group culture and patient satisfaction was found (which is how the organisation within their study would be measured as effective). A study by Yesil and Kaya (2013) attempted to measure the effect of organisational culture on firm financial performance. Their findings showed that there is no effect between the two variables, although this study is limited by the researcher's inability to control for factors such as market trends and consumer spending sentiment. Although there is not enough empirical evidence available to prove the effect organisational culture has on desired outcomes, the majority of the literature agrees that a positive relationship exists between the two (Cummings & Worley, 2015; Gregory et al., 2009; Yesil & Kaya, 2013).

2.3.3 Decision-making

Decision-making is closely linked to strategy in that businesses will need to commit or allocate resources to certain product lines, market segments, resource requirements and other everyday decisions that can impact both performance and sustainability (Andrews, 1987, p.16). Decision-making is considered a multi-faceted process that concerns the evaluation, ranking and commitment to a range of possible actions (Malakooti, 2012). This process does not consist only of the final commitment of resources to a decision, instead, it is a set of dimensions that leaders must work through. Organisations have the opportunity of using evidence in the initial step of formulating a problem: this would consist of evidence about the environment, the timing of events, previous events or experiences, the availability of resources and the extent or impact of the problem (Malakooti, 2012). Evidence-based decision-making is the convergence of organisational culture and decision-making that makes use of data to empower decision-makers.

2.4 The link between big data and evidence-based decision-making

It is widely believed that the use of data can improve decision-making and researchers are attempting to address the dearth of empirical evidence available (Brynjolfsson et al., 2011; Cao & Duan, 2014; Elgendy & Elragal, 2016). For example, given the potential for extremely large datasets offered by current technological advances (big data, IoT etc.), businesses are able to create longitudinal interaction data (such as social media posts tracked in real-time) that allows analysis of patterns in brand and product sentiment (George et al., 2014). Businesses can use this data to inform decision-making on any positive or negative deviations regarding brand, product or services for example.

Whilst the availability of big data may assist in providing an organisation with the means of accessing new and valuable insights, the actual usage of the data by decision-makers must be entrenched into the organisation's culture in order for this to translate into some form of a positive outcome (Miller & Friesen, 1982). Pigni et al. (2016) stress the need for a data-oriented mindset stating that merely having the availability of data is not enough to generate value from it, instead, an adaption to the currently established decision-making norms must be made.

Brynjolfsson et al. (2011) tested this premise, that possessing a data-driven culture may lead to better decision-making and therefore better performance, and concluded that businesses that adopt data-driven decision-making have output and productivity that is 5-6% higher than what is expected given their investment and information technology usage. Moreover, that data-driven decision-making is associated with higher profitability and market value. Brynjolfsson et al. (2011) define data-driven decision-making as the usage of data for the creation of new products or services and the usage of data for decision-making, which can also be framed as evidence-based decision-making. LaValle et al. (2011) reiterated this sentiment when they surveyed nearly 3,000 executives across 100 countries and found that top-performing businesses used analytics five times more than lower performers.

Based on the above, this research study does not aim to test or contradict the findings of Brynjolfsson et al. (2011) or LaValle et al. (2011), instead, it aims to add to the literature on data usage and its link to decision-making in order to achieve some desired outcome. For these previous studies, the desired outcomes were performance measures such as output and productivity. This research study aims to provide clarity on the link between data usage, an evidence-based decision-making culture and a firm's ability to achieve higher levels of EO, which is positioned as a desired strategic posture, as firms aim to be more innovative, proactive and risk-taking.

Although the vast potential exists for organisations to extract value from their data, many organisations are still not successfully making use of their data (Bean, 2017). A review of the literature revealed various antecedents required for organisations to start realizing value from their data. These are discussed below.

2.5 Antecedents of effective utilization of big data

At the core of data and decision-making challenges is the problem of response time latency (Pigni et al., 2016; Davenport et al., 2012). This is explained as the longer an event, or the information communicated from an event takes to result in some form of action the less value that action will have in the end. This lack of the ability to capture, analyse and inform decisions timeously is an outcome of various challenges such as outdated IT infrastructure, the lack of data science skills, privacy concerns and organisational culture (Alharthi et al., 2017). These are discussed further below.

2.5.1 Skillsets

One of the core antecedents to the effective use of big data is that of skillsets. Given the large volume, high velocity and new variety of big data, organisations must invest in new skills that are capable of managing this upsurge of data (Waller & Fawcett, 2013). This is reiterated across the literature as an essential prerequisite if organisations hope to extract value from their data (Alharthi et al., 2017; George, Osinga, Lavie & Scott, 2016; Rygielski, Wang & Yen, 2002; Pigni et al., 2016; Waller & Fawcett, 2013). Furthermore, not only will organisations that fail to employ the relevant skills fail to unlock this value, they will in addition become vulnerable to a host of challenges such as storing new forms of data; the understanding of protocols regarding security, privacy and data rights; data reliability; the inability to manage or process extremely large datasets and the lack of clarity in understanding complex patterns (George et al., 2016).

The emergent skillset from this new business necessity has been labelled as data science and data scientists are positioned to facilitate business decisions related to products, brand and services whilst maintaining a grasp of the technical requirements (Provost & Fawcett, 2013). Waller and Fawcett (2013) define data science as the use of quantitative and qualitative methods to solve business problems and predict possible outcomes. Given their high demand, considered one of the top careers in the United States (Mills, Chudoba & Olsen, 2016), this highly sought after title has even been labelled the "Sexiest job of the 21st century" (Davenport & Patil, 2012, p. 1).

Further to the required skills, the complexity involved in managing big data initiatives is immense considering the rate of how data and technology are growing, the vast amount of different sources of data and the multiple formats of data (Alharthi et al., 2017). The dearth of qualified talent within the United States is estimated to reach between 120,000 and 190,000 people by 2018 (Douglas, 2013). The proliferation of data scientist academic programs highlights that education institutions have also recognized the gap in the market for talent that specializes in data and business analytics (Goes, 2014).

Information intensive businesses such as Google, Facebook and Linkedin; as well as e-commerce platforms such as Amazon and eBay have managed to use their data as a key resource in driving business goals (Chen et al., 2012), although the majority of other businesses continue to struggle as they do not understand what to do given the fragmented environment of solutions available in the market (Goes, 2014).

Data analytics and big data initiatives rely on a broad number of skills from three broad categories, namely: information technology (IT) skills, general business management and statistical modelling skills. For organisations considering big data initiatives understanding IT concepts and infrastructure such as Hadoop; MapReduce; NoSQL; Extract, Transform and Load (ETL); machine learning; Online Analytical Processing (OLAP) and other complexities involved in dealing with increasingly more unstructured data is crucial if they plan to take advantage of big data or possibly a first mover advantage (Alharthi et al., 2017). Furthermore, skills in statistically measuring and predicting outcomes are just as important such as regression, factor analysis, clustering and discriminant analysis (Alharthi et al., 2017; Chen et al., 2012). A third skill is also emerging, demanding businesses to start finding innovative ways of collecting, organising, analysing and sharing data insights effectively across the business (Gobble, 2013).

Given the broad range of skills required by data scientists, as well as the rapid change in technology and therefore skills required (George et al., 2016), it is difficult to ascertain a set of skills which can be considered big data skills. Literature which discussed the skillsets required for big data initiatives, as well as business intelligence projects, was examined and key themes were extracted and summarised in Table 1.

Table 1: Key themes of skills highlighted from the literature

Relevant skill	**Citation**
Structured Query Language (SQL), NoSQL, Python and other programming languages	Alharthi et al. (2017), Chen et al. (2012), Mills et al. (2016)
Machine learning	Alharthi et al. (2017), George et al. (2016), Mills et al. (2016), Popovic et al. (2012)
Data mining, predictive analysis and other statistics	Alharthi et al. (2017), Chen et al. (2012), Rygielski et al. (2002), Waller & Fawcett (2013), Mills et al. (2016)
Data exploration, OLAP, reporting, visualization and the ability to share insights across business	Alharthi et al. (2017), Chen et al. (2012), George et al. (2016), Gobble (2013), Mills et al. (2016), Rygielski et al. (2002), Popovic et al. (2012)

This highlighted that although no one set of skills has been identified as the core function of data science, the literature seems to be converging on a set of fundamental requirements. This list does not propose to be exhaustive, and only includes themes which were cited by three authors or more. Other themes do exist and are not to be considered insignificant.

Finally, businesses must also decide on where big data initiatives should reside internally, as previous ways of organising analytical staff will not suffice (Davenport et al., 2012). Employing the correct skillsets that can coordinate and assemble cross-functional inputs (business, statistical and technical) into new processes, products and decision-making routines is recognized as an important step in implementing effective big data initiatives that will reduce the response time latency of decisions (Pigni et al., 2016).

2.5.2 Datasets and Toolsets

Whilst the challenges of managing business culture and human resources (employing the correct skills) are integral to effective data-driven initiatives, other factors such as toolsets and datasets, although less cited, are also identified as antecedents for effective big data initiatives. Toolsets are recognized as the unique systems required in the form of databases, enterprise resource planning (ERP) systems, data warehouses, business intelligence tools and other technology-oriented hardware and software (Pigni et al., 2016).

The complex and dynamic nature of big data means that new formats of data require new systems (or toolsets) in order to manage them (Alharthi et al., 2017). These new formats can be categorised as either structured or unstructured data, the former being data organised in traditional relational databases, whilst the latter can include images, videos, text documents and emails (Alharthi et al., 2017). Older systems may not possess the ability to process or even store these new types of data, meaning organisations must invest capital into new systems, which may pose issues related to legacy IT systems and integration with new systems (Alharthi et al., 2017). Increases in system quality and implementation of technology are suggested to lead to improved access to information for employees (Popovic et al., 2012). Furthermore, the use of business intelligence systems may lead to better levels of information quality, with faster access to information, timeous queries, consistency in the data and higher levels of interactivity (Popovic et al., 2012).

The final variable to be considered is datasets, which is defined as an organisation's ability to identify and access data that can be used for value creation within the organisation (Pigni et al., 2016). Businesses are firstly tasked with the responsibility of having the knowledge to recognize datasets that will actually be of use to their decision-making (Pigni et al., 2016). This is important as different datasets exhibit different levels of noise and must be analyzed and possibly cleansed before being integrated. This knowledge of the value of the data should provide an ambidextrous position (Bøe-Lillegraven, 2014) for businesses to exploit currently available data that is crucial to their business (i.e. actual transaction and stock data that signals performance to date) and explore the possibility of new data ventures that may add value in different ways (i.e. customer interaction data which may signal future demand).

2.6 EO as a strategy-making process

The objective in trying to understand big data and how to effectively utilise it, along with an evidence-based decision-making culture, stems from the opportunity it poses to unlock some form of competitive advantage or positive outcome (Alharti et al., 2017; Bean, 2017). Previous studies have positioned this positive outcome as firm performance, such as output and productivity, and found that the use of data in decision-making processes can lead to higher levels of performance (Brynjolfsson et al., 2011; LaValle et al., 2011). Similarly, this research aims to measure some form of positive outcome, although takes a step back before the performance of an organisation, and aims to provide clarity on the use of data and its relationship to the internal strategy-making process of the organisation, given that the strategy-making process is recognized as a source of competitive advantage (Barney, 1991; Mintzberg, 1978).

The concept of strategy has been researched for decades resulting in a host of different views (Ghemawat, 2002). These include approaches such as the resource-based view (Barney, 1991; Wernerfelt, 1984), strategy classified by an organisational mode or school (Mintzberg, 1973; Mintzberg, Ahlstrand & Lampel, 1998), strategy shaped by competitive forces within an industry (Porter, 1979), strategy in relation to an organisation's business model (Teece, 2010) and various other simplified frameworks (Ghemawat, 2002). This research study makes use of EO as a measurement of desired strategic posture for various reasons. Firstly, the use of EO allows for a measurable attribute of firm behaviour (Covin & Slevin, 1991). Secondly, the EO construct is considered a stabilized concept (Gupta & Gupta, 2015) that has been extensively researched and is widely accepted as a posture related to a firm's strategy making process (Covin, Green & Slevin, 2006; Linton & Kask, 2017; Wales et al., 2011). Finally, the concept of EO fits well within this research as it places emphasis on the possible differences in strategic posture between organisations (i.e. highly innovative or not), and instead of focusing on how organisations can reduce costs, EO serves as a gauge of the organisation's ability to meet their business objectives (Gupta & Gupta, 2015). Moreover, the use of a behavioural model allows for a simpler interpretation for practitioners in the form of the EO sub-dimensions: innovativeness, proactiveness and risk-taking.

Whilst EO itself is not a measure of the financial performance of an organisation, research has shown it to have a positive impact on firm performance (Lechner & Gudmundsson, 2014). Rauch et al. (2009) found a moderately large correlation (r = 0.242) between EO and firm performance. Furthermore, in studies of the individual sub-dimensions of EO, innovativeness was shown to be related to differentiation, with differentiation shown to be positively associated with firm performance (Lechner & Gudmundsson, 2014). Not only has EO been associated with firm performance, it has also been argued that EO is positively associated with other desirable outcomes such as employment growth and therefore economic growth (Madsen, 2007).

EO is defined as “a strategic organisational posture that captures the specific processes, practices and activities that enable firms to create value by engaging in entrepreneurial endeavors” (Wales et al., 2011, p. 357). The EO construct has been researched for decades and was initially introduced by the seminal article from Danny Miller in 1983 (Wales et al., 2011). In his research, Miller (1983), aimed to ascertain what the determinants of entrepreneurship were. He defined entrepreneurship, similarly to how strategy can be viewed, as the process by which organisations renew themselves. The focus of Miller’s (1983) work was not on the individual, instead, it was in the process of entrepreneurship at a firm-level. This was stated as: “An entrepreneurial firm is one that engages in product-market innovation, undertakes somewhat risky venture, and is first to come up with ‘proactive’ innovations, beating competitors to the punch” (Miller, 1983, p. 771). Lumpkin and Dess (1996) addressed the difference between entrepreneurship and entrepreneurial orientation by defining the former as new entry, where organisations launch new ventures and the latter as the activities that lead to this new entry. These include all processes, practices and decision-making activities that may lead to new entry (Lumpkin & Dess, 1996). Furthermore, whilst some research has discussed EO on other levels such as individual (Lyon, Lumpkin & Dess, 2000; Zahra, 1993), the majority of key research has agreed with the firm-level approach (Covin & Slevin, 1991; Lumpkin & Dess, 1996) and some have even cautioned that EO relationships remain focused at this level (Slevin & Terjesen, 2011). Therefore this research study focused on the firm-level, as per Miller (1983) and Lumpkin and Dess’ (1996) definitions.

2.7 Sub-dimensions of EO

2.7.1 Innovativeness

With organisations such as Uber and Airbnb (Ingram, 2012) finding new and innovative ways of serving customers in existing industries, organisations cannot idly sit by and expect to remain competitive. Innovativeness can be expressed as an organisation's willingness to depart from current or existing technologies, methodologies and practices to venture into new and potentially rewarding states of achieving business goals (Lumpkin & Dess, 1996). Innovation is proposed as a necessary, and disruptive, activity within hostile and dynamic markets (Miller & Friesen, 1982), which may be assumed of the current business environment. Innovativeness refers to technological innovation (Lumpkin & Dess, 1996) and product-market innovation (Miller, 1983; Lumpkin & Dess, 1996) which are both deemed vital to an organisations ability to renew itself and remain competitive.

2.7.2 Proactiveness

Whilst innovation is recognized as a key proponent to entrepreneurial behaviour, its usefulness would diminish if an organisation did not act proactively. Reactive behaviour would hardly be associated with new innovative ideas, products or marketing behaviours, therefore proactiveness is proposed as a necessary dimension of EO (Miller, 1983; Covin & Slevin, 1989; Lumpkin & Dess, 1996). Proactiveness is recognized as a key strategic approach on its own. For example, first-movers that are able to proactively address a market need can capitalize on unusually high profits and achieve a head start on competitors by establishing their brand and market dominance (Lumpkin & Dess, 1996). Organisations that are able to achieve this first-mover status may even be able to establish a resource position barrier (Wernerfelt, 1984), a barrier that prevents other organisations from replicating a certain resource, and compete strategically based purely on a resource-based view (Barney, Wright & Ketchen, 2001). Google is a good example of an organisation that has been able to establish themselves as the number one search engine in the world, due to their innovative search algorithms that were introduced very early in this age of digitization, and consequently their dominance has created a barrier for other search engines to compete on the resource of an individual's search behaviour and interests.

In this paper proactiveness is defined as an organisation's ability to be forward-thinking and anticipate possible opportunities or risks in the future (Lumpkin & Dess, 1996).

Whilst the first-mover advantage is posited as a favourable status for EO, it is not a necessary condition and organisations can be proactive and forward-thinking without being the first to market. Proactive organisations are aware of their environment and perform as market leaders given their will and ability to act on possible opportunities (Lumpkin & Dess, 1996).

2.7.3 Risk-taking

The third and final dimension of EO is identified as risk-taking (Miller, 1983; Covin & Slevin, 1989; Lumpkin & Dess, 1996). Whilst organisations must have a proactive disposition in addressing the current dynamic environment and may come up with new product-market innovations as well, it would not be possible to proactively pursue these innovations without undertaking some form of risky ventures or behaviours (Miller, 1983). Risk-taking organisations recognize the high returns and therefore opt for risky decisions in the form of committing large financial resources or even incurring high levels of debt to obtain these returns (Lumpkin & Dess, 1996).

An excellent example of an entrepreneurially oriented organisation of the 21st century to consider is Elon Musk's SpaceX (Thompson, 2011). Not only have they proactively identified possible future benefits to reusable rockets, they have taken significant risks, financial and non-financial, to develop and test new innovative technology. SpaceX has very publicly exhibited risk-taking, innovativeness and proactiveness that is to be associated with EO.

2.8 The link between EO and big data

This paper proposes that big data is uniquely positioned as a component in allowing organisations to achieve higher levels of EO when coinciding with an evidence-based decision-making culture.

Lumpkin and Dess (1996) refer to new combinations that may result from innovativeness within an organisation. An organisation's ability to scan the environment and recognize the needs of the external environment has been posited as a primary limitation of innovativeness (Miller & Friesen, 1982). However, while this categorises organisations from the 1990s, the data streams from big data allows for organisations to scan the environment in new and more efficient ways than ever before. The variety of big data (Alharthi et al., 2017) allows for new ways of receiving feedback on product-market innovations, the volume and velocity of big data (Davenport et al., 2012; Pigni et al., 2016) allow for this feedback to be near real-time ensuring timeous conclusions on the effectiveness of these innovations and the veracity of big data (Goes, 2014; Pigni et al., 2016) ensures that this feedback is accurate and objective, given that it is based on facts and data. For example, the initial measurement of customer sentiment towards a product or marketing activity would have previously only been reported sales. Today organisations can also measure the increase in foot traffic, the number of unique visits to a website, click-through rates of mailers and even sentiment of social media posts through the mining of text.

This value, generated from data, also applies to proactiveness as the view of the world that the organisation possesses becomes clearer. Non-transactional data adds a new dimension to how organisations can view a consumer and therefore allow for new insights which can possibly be turned into new opportunities. Furthermore, the strong link between big data, data science and statistics means that organisations do not only have near real-time and accurate views of historic trends but also have the ability to predict and model possible future outcomes, timeously, using regression analysis and other statistical tools. Whilst this does not remove the need for experience in decision-making, it allows for time to be spent on thinking proactively, opposed to relying on traditional methods of collecting, compiling and modelling data.

The final dimension, risk-taking, also stands to gain from big data within this model. The difficulty in risk-taking is that one must venture into the unknown and commit a significant amount of resources (Lumpkin & Dess, 1996). This is unfortunately inevitable, although big data may allow for managers to be in a state of, not absolute, but more certainty when making decisions. Thus increasing their propensity to take more, calculated, risks.

2.9 Conflicting views in measuring EO

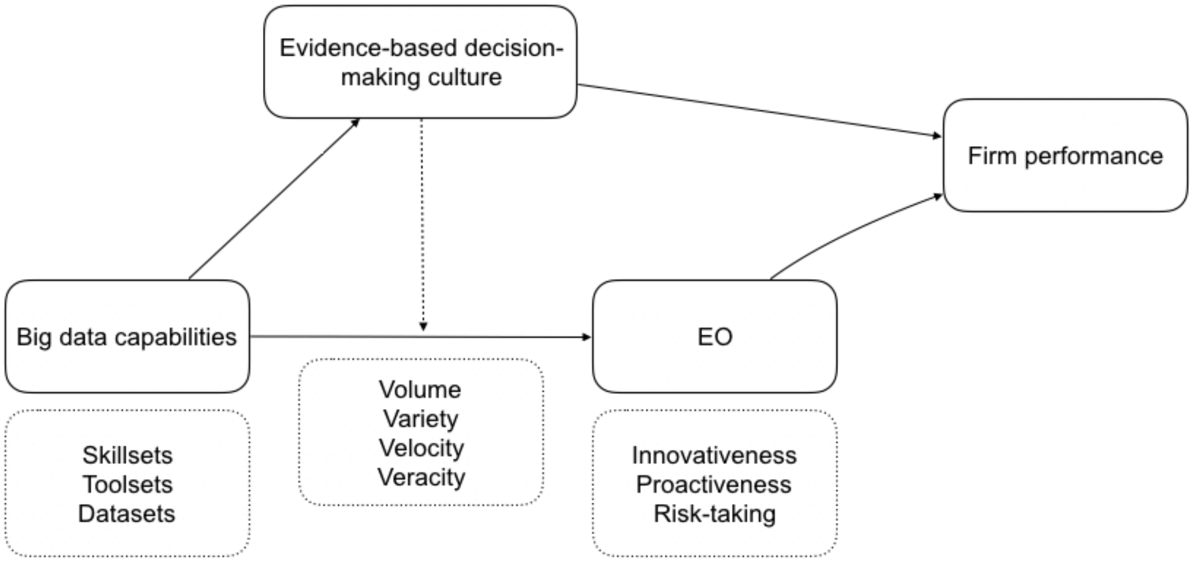
One of the most important decisions for researchers is which dimensions to adopt in measuring the EO construct (Wales et al., 2011; Gupta & Gupta, 2015), as two broad views exist and this will ultimately affect the research design and results. The initial construct introduced by Miller (1983) used the dimensions of innovativeness, proactiveness and risk-taking to test entrepreneurship. This was built on by Covin and Slevin (1989) when they measured entrepreneurial strategic posture using a scale based on the same three dimensions. Lumpkin and Dess (1996) went on to propose a further two dimensions, autonomy and competitive aggressiveness. Further to this, not only did Lumpkin and Dess (1996) suggest an extra two dimensions for the construct of EO, they also proposed a multidimensional view opposed to Covin and Slevin's (1989) unidimensional view. The latter proposed that in order for a firm to be considered entrepreneurial they would need to exhibit high scores on each of the dimensions, whilst the former implies that a firm can be entrepreneurial when any of the dimensions are evident (Gupta & Gupta, 2015).

Although the majority of previous EO research has adopted a unidimensional approach to measuring EO (Gupta & Gupta, 2015; Rauch et al., 2009; Saeed, Yousafzai & Engelen, 2014; Wales et al., 2011; Covin & Slevin, 1989), several other authors treated EO as a multidimensional construct (Lechner & Gudmundsson, 2014; Naldi, Nordqvist, Sjöberg & Wiklund, 2007). It is argued that it is important to analyze the individual constructs separately, using a multidimensional approach, as some dimensions may vary differently (Linton & Kask, 2017). A review of the literature found that 133 of 177 EO research studies operationalized the construct of EO as unidimensional, whilst only 35 studies opted to treat EO as multidimensional and the remaining nine research studies opted for a method that used both (Saeed et al., 2014). Some research has also argued that there is no single correct conceptualization of the EO construct (Wales et al., 2011) and that the choice between the use of a unidimensional and multidimensional conceptualization of EO is to be led by the research question being investigated. This "major schism" regarding the measurement of EO has only recently begun to be discussed by researchers (Gupta & Gupta, 2015, p. 59) and represents an opportunity for research to help elucidate the EO construct.

2.10 Conclusion of Literature Review

The aim of this research was to understand if a link exists between an organisation's big data capabilities and their ability to be entrepreneurial. Furthermore, the concept of evidence-based decision-making was introduced as a benefit of big data as well and a possible moderator of EO. Finally, this research also aimed to explore which big data capabilities are most important when it comes to implementing big data initiatives. A graphical representation of the research model has been provided in Figure 1.

Figure 1: Graphical representation of research model



The model is explained as big data capabilities allowing for the effective use of big data and producing more (volume), new (variety) and accurate (veracity) data at near real-time speeds (velocity). This can assist organisations to be more innovative, proactive and risk-taking with their internal strategy-making process. In order to achieve this, organisation's must address certain big data capabilities identified within the literature as the necessary skills (named skillsets in this research), such as the data scientist, in order to extract, manipulate and transform all this (big) data into insights, systems such as an ERP or business intelligence tool that can report, model and predict expected outcomes (named toolsets in this research), and gain access to the appropriate data that is clean and contains valuable and relevant information (named datasets in this research).

It is further noted that although big data capabilities can allow for the appropriate and relevant data to be collected, analyzed and distributed, if an organisational culture of using data to make decisions is not present amongst decision-makers then this data will not serve the proposed purpose. Therefore, evidence-based decision-making culture is included in the model as a moderator of the relationship between the big data capabilities and EO, although also enabled by the big data capabilities.

The construct of firm performance was included in the model in order to create a broader view of what other literature has shown or measured in relation to EO and a data-driven decision-making culture, although it was not measured or tested within this research.

Further to this, the author of this research opted to investigate the formation of the EO construct in order to provide further empirical research on the dimensionality of the construct. The literature review revealed that two schools of thought (unidimensional vs. multidimensional) existed and therefore an opportunity exists to further investigate the EO construct providing further evidence of its sub-dimensions. Frese, Bausch, Schmidt, Rauch and Kabst (2012) call for several studies in order to provide better evidence when measuring empirical relationships and this research aimed to contribute to the field of EO research as well.

This chapter identified the key literature related to the constructs of big data capabilities, evidence-based decision-making and EO, and outlined a proposed relational model in Figure 1. The following chapter frames the relevant research questions which were used to understand the relationship between these constructs.

Chapter 3: Research Questions

3.1 Introduction

The previous chapters outlined the aim of this research: to better understand the relationship between big data (specifically the big data capabilities: skillsets, toolsets and datasets), EO and evidence-based decision-making.

A review of the literature highlighted that the use of data is believed to improve decision-making (Chen et al., 2012; Popovic et al., 2012) and that organisations that employ a data-driven or evidence-based decision-making culture can achieve higher firm performance (Brynjolfsson et al., 2011; LaValle et al., 2011). This provides a possible avenue for business practitioners to answer the call for innovation given the dynamic and competitive environment (Mathews, 2016; Purnama & Subroto, 2016; Reeves & Deimler, 2011; Prajogo, 2016; Sousa & Coelho, 2011; Slåtten & Mehmetoglu, 2011).

Big data was posited as an enabler of both better data-driven decision-making and EO, given the volume, variety, veracity and velocity of data streams available to organisations (Alharthi et al., 2017; Davenport et al., 2012; Goes, 2014; Pigni et al., 2016) and three antecedents were identified for the effective implementation of big data initiatives, namely: skillsets, toolsets and datasets (Alharthi et al., 2017; George et al., 2016; Rygielski et al., 2002; Pigni et al., 2016; Popovic et al., 2012; Waller & Fawcett, 2013). This provided an opportunity to measure these various antecedents of big data (referred to as big data capabilities in this study) and their relationship to other desirable organisational postures, such as evidence-based decision-making and EO.

Further to understanding the relationship between big data and evidence-based decision-making, the concept of EO was also introduced in order to measure the relationship between the three constructs. The construct is useful as it measures firm-level behaviours such as innovativeness, proactiveness and the propensity to take risks, which have been linked to both firm performance (Lechner & Gudmundsson, 2014; Rauch et al., 2009) and employment growth (Madsen, 2007). Furthermore, it places emphasis on behaviours (Gupta & Gupta, 2015) that assist organisations in renewing themselves (Miller, 1983), as opposed to just achieving profit through cost-cutting, and become more competitive by unlocking new combinations within the organisation (Lumpkin & Dess, 1996).

Although the main aim of this research was to assess the relationship between EO and the big data capabilities, the researcher opted to make use of the opportunity to further contribute to the current literature (Gupta & Gupta, 2015; Linton & Kask, 2017; Saeed et al., 2014; Wales et al., 2011) regarding the dimensionality of the EO construct by providing empirical evidence regarding the dimensionality of EO. This is considered an important area of EO research (Linton & Kask, 2017) and therefore a question regarding the dimensionality of EO was included in this research study.

Based on the review of the literature various research questions were established and are articulated below. This is then followed by a discussion of the research methodology employed in this study.

3.2 Research Question 1

The first research question aimed to measure the relationship between the various sub-dimensions of EO, namely: innovativeness, proactiveness and risk-taking. A review of the literature highlighted that the EO construct has not only been studied extensively over the last few decades (Gupta & Gupta, 2015), but various views exist regarding the measurement of the construct (Gupta & Gupta, 2015; Saeed et al., 2014; Wales et al., 2011). The initial view of unidimensionality was introduced by Miller (1983) and adapted by Covin and Slevin (1989), whilst the opposing view of multidimensionality was suggested by Lumpkin and Dess (1996). Whilst the majority of EO literature has been conducted using the unidimensional view (Saeed et al., 2014; Wales et al., 2011), conflicting views still exist between researchers (Linton & Kask, 2017) and therefore the researcher aimed to contribute to the literature by assessing the sub-dimensions of the EO construct. The first research question was thus articulated as:

Research Question 1: Do the sub-dimensions of EO covary?

3.3 Research Question 2

The second research question focused on the relationship between the big data capabilities, namely: skillsets, toolsets and datasets and an organisation's level of evidence-based decision-making culture. This question aimed to provide empirical evidence on the notion that the presence of the correct skills, available systems and relevant data can assist organisations in using data when making decisions (Elgendy & Elragal, 2016; Popovic et al., 2012; Pigni et al., 2016). This question was articulated as:

Research Question 2: What is the relationship between big data capabilities and evidence-based decision-making?

3.4 Research Question 3

Similar to research question 2, the third research question aimed to assess the relationship between the big data capabilities and EO. This research argued that big data is uniquely positioned as an enabler of innovativeness, proactiveness and risk-taking (measured by the EO construct) within organisations, given that it allows for increased and new forms of feedback regarding products or services, that is near real-time, accurate and useful (Alharthi et al., 2017; Davenport et al., 2012; Goes, 2014; Pigni et al., 2016). Therefore the third research question was articulated as:

Research Question 3: Are the big data capabilities antecedents for EO?

3.5 Research Question 4

The final research question brought the various elements of the research together and aimed to determine if an interaction effect existed on the relationship between big data and EO. Previous studies found a host of benefits from the use of data in decision-making (Brynjolfsson et al., 2011; LaValle et al., 2011) and this research aimed to contribute to the literature by assessing the relationship between an organisation's big data capabilities and their EO, moderated by an evidence-based decision-making culture. This question was articulated as:

Research Question 4: Is evidence-based decision-making culture a moderator for big data capabilities on EO?

Chapter 4: Research Methodology

4.1 Introduction

The previous chapter has outlined the various research questions proposed by this research to address the research aim. In order to achieve this, this study made use of multivariate data analysis as it is considered a powerful tool that can assist in revealing relationships between variables that would not have otherwise been identified (Hair, Black, Babin & Anderson, 2009). This chapter discusses the methodology employed to complete this research and includes:

- The research approach
- The population
- Unit of analysis
- Sampling technique and size
- Research instrument
- Data analysis
- Limitations

4.2 Research approach

A layered approach to the design of this research was employed, where the research philosophy and approach of this were discussed and used to inform the research type and strategy. This is discussed below.

In order to facilitate replication, a structured methodology was employed and the philosophy of this research can be defined as positivism (Saunders & Lewis, 2014). This philosophy was deemed appropriate as the researcher aimed to obtain generalizable results that could be replicated in future.

Given the proliferation of research regarding EO and more recently big data and evidence-based decision-making, this study aimed to explore research questions based on current established literature and can, therefore, be described as deductive research (Saunders & Lewis, 2014). Moreover, this study aimed to provide an accurate representation of the organisations measured and can, therefore, be defined as a descriptive study. In order to accurately describe the sample, an online survey was identified as the research strategy, as this provided a structured approach to collecting data from respondents accurately (Saunders & Lewis, 2014).

A snapshot of the data from respondents was collected over a period of time and therefore this study can be described as cross-sectional (Saunders & Lewis, 2014; Zikmund, Babin, Carr & Griffin, 2009). Furthermore, this can be classified as a quantitative study as it made use of research questions, which aimed to provide further insights on the relationship between the constructs, which were addressed by the measurement and testing of quantitative data in a structured approach (Zikmund et al., 2009).

4.3 Population

Zikmund et al. (2009, p. 387) defined a population as “any complete group” that share a common set of characteristics. This group is not limited to individuals only and can be organisations or places as well (Saunders & Lewis, 2014), and must be defined at the outset of the sampling process in order to properly identify the relevant sources from which data will be collected (Zikmund et al., 2009).

The population for this study was identified as all professionals employed as managers. The definition of managers was chosen on account of these individuals being considered to be decision-makers within the workplace, who can use data, information and intelligence in order to make vital business decisions such as product or service offerings, markets to be served and resource allocations (Andrews, 1987).

4.4 Unit of analysis

The unit of analysis is defined as the “what or who” that will provide the data (Zikmund et al., 2009, p. 119). For this study the unit of analysis was defined as professionals employed as managers. This included all levels from junior to executive management. In order to ensure that the respondents used fell into this category, a question regarding the respondent’s position within their organisation was included in the questionnaire and only respondents who indicated themselves as managers were included in this study.

4.5 Sampling technique

Measuring the entire population was considered impractical given that measuring every element of the population would be expensive, inconvenient and very time consuming (Zikmund et al., 2009). Therefore a subset of the larger population was used, known as a sample (Saunders & Lewis, 2014; Zikmund et al., 2009).

The population for this study was identified as all professionals employed as managers. Therefore the population size was deemed to be extremely large, making it impossible to collect data from the entire population. Furthermore, because a full list of the population was impossible to obtain, this study made use of non-probability sampling (Saunders & Lewis, 2014). Non-probability sampling refers to sampling where the probability of a particular member of the total population being chosen is not known (Zikmund et al., 2009), which was the case within this study.

Various non-probability sampling techniques exist and for this study purposive or judgement sampling, as well as snowball sampling, was utilised (Saunders & Lewis, 2014). Judgement sampling is described as a technique where an experienced individual will use their personal judgement to select the sample (Zikmund et al., 2009). This technique was used to ensure that responses collected were congruous with the population of this study. The second non-probability sampling technique employed was snowball sampling, which is described as a technique where additional respondents are identified by information received from initial sample respondents (Zikmund et al., 2009). This was necessary as the researcher of this study relied on respondents to identify future possible respondents given the limitation of his network.

In order to obtain the necessary sample, the researcher of this study made use of personal networks, which were contacted via email to complete and then forward on the questionnaire to other managers within their network. Furthermore, respondents were contacted via business professional networking applications (such as LinkedIn) and asked to complete the questionnaire. The use of this sampling technique resulted in the inability to accurately calculate a response rate for the questionnaire.

4.6 Sample size

When conducting research, a larger sample size is associated with more accurate research (Zikmund et al., 2009) and the power of a test is very dependent on the size of the sample that was collected and used (Pallant, 2007). Furthermore, the possibility of a non-significant result may be caused by this lack of power. When conducting factor analysis, small samples may also result in less reliable correlation coefficients among variables (Pallant, 2007).

The sample size is also important when conducting multiple regression tests. In order for research results to add scientific value, they must be generalisable to other samples, although generalisability may be an issue with small samples when conducting multiple regression (Pallant, 2007. Tabachnick and Fidell in Pallant (2007) suggest a sample size of $N > 50 + 8m$ (where m = the number of independent variables used). A summary of the required responses based on this suggested formula is summarised in table 2. This study reported a final sample size of 82 responses, which was enough based on research questions 1 – 3, although fell short for research question 4.

Table 2: Minimum sample size calculation

Research Question	Number of Independent variables	Responses required
Research Question 1	3	(50 + 24) = 74
Research Question 2	3	(50 + 24) = 74
Research Question 3	3	(50 + 24) = 74
Research Question 4	8	(50 + 64) = 114

4.7 Research Instrument

4.7.1 Data collection

A self-administered online questionnaire was used to collect data from respondents. The use of Internet-based surveys allowed for a variety of advantages such as a high speed of data collection, low costs of administering the questionnaire, large geographic flexibility and the ability to provide anonymity and confidentiality (Zikmund et al., 2009).

Providing anonymity for respondents, via an online questionnaire, means that respondents were more likely to provide sensitive or embarrassing information (Zikmund et al., 2009) and avoided the possibility of social desirability bias from respondents (Thomas & Kilmann). This approach, therefore, resulted not only in a cost-effective and efficient survey process but also allowed for honest and possibly more accurate feedback.

4.7.2 Questionnaire design

This research study made use of various questions from previous studies in order to measure the constructs accordingly. Appendix 1 provides a detailed breakdown of each question used. Multiple-choice questions were used in order to standardise and control responses.

The first section labelled D in Appendix 1 included various questions regarding respondent demographics, as well as information regarding their current employment. This allowed the researcher to provide descriptive information regarding the sample, determine respondent relevance (manager or not) and establish a level of diversity. Section E (see Appendix 1) included questions that aimed to measure the intensity of the organisation's evidence-based decision-making culture. The researcher made use of established measures taken from Brynjolfsson et al. (2011) and (Center for Evidence-Based Management, 2013) and therefore opted to emulate their use of a 7-point Likert scale.

Sections S, T and DS (see Appendix 1) aimed to measure the various big data capabilities within an organisation, identified as skillsets, toolsets and datasets. These questions were taken from Popovic et al. (2012) and once again a 7-point Likert scale was used as per their study. It is further noted that although certain skills are used in broader contexts than big data, such as question S4: "Analytical applications, including trend analysis, 'what-if' scenarios", the online questionnaire was clearly labelled as a big data questionnaire and respondents were informed, via a consent form, that the research concerned big data in order to ensure that questions were answered in the context of big data.

The final sections, IV, PA and RT (see Appendix 1) were used to measure the organisation's level of innovativeness, proactiveness and risk-taking in order to measure the EO construct. These questions were taken from Barringer and Bluedorn (1999), which like the majority of other EO research studies, made use of Covin and Slevin's (1989) 9-item scale to measure the EO construct. Furthermore, a 7-point Likert scale was used, similar to Barringer and Bluedorn (1999).

4.7.3 Pre-testing of questionnaire

Before administering a questionnaire, it is strongly recommended that a pre-testing of the questionnaire be performed (Pallant, 2007; Saunders & Lewis, 2014; Zikmund et al., 2009). This will help to ensure that the questionnaire is easily understood, that scale items are clear and that respondents are able to complete the questionnaire with ease. The use of a pre-testing will also assist with any questions or items that may come across as offensive to certain respondents (Pallant, 2007).

This study made use of a pre-test in order to achieve the outcomes outlined above. The researcher of this study used judgement sampling (Saunders & Lewis, 2014; Zikmund et al., 2009) to identify potential managers as respondents. A total of 17 responses were collected and analysed during the pre-testing of the questionnaire. Moreover, the researcher's supervisor was also consulted for feedback on the questionnaire. All pre-test responses were excluded from the actual sample used for this study.

Feedback from the pre-test highlighted that the questions were understood and respondents were able to answer all the questions. Furthermore, the average duration of the questionnaire was 20 minutes and was therefore deemed acceptable. Recommendations regarding the flow and aesthetics of the questionnaire were acknowledged and implemented where necessary.

Finally, some respondent feedback highlighted the use of reverse Likert scales as a concern regarding the accuracy of responses. This feedback was acknowledged although, given the reliability tested by previous researchers, the decision was ultimately taken by the researcher of this study to maintain the question design as per previous research performed (Center for Evidence-Based Management, 2013).

4.8 Data Analysis

4.8.1 Data preparation and completion rate

An online survey tool called SurveyMonkey was used to administer the questionnaire for this study. A total of 110 responses were received through the online questionnaire. The total sample was extracted from SurveyMonkey into Microsoft Excel. Responses were then checked for completion and 27 responses were eliminated from the sample given that they were incomplete. This resulted in a final sample size of 82 responses, which equates to a 75.5% completion rate.

Responses were then coded from interval data to continuous data. Finally, the coded data was imported into IBM SPSS in order to evaluate the data.

4.8.2 Principal Components Analysis

Principal Components Analysis (PCA), often used interchangeably with Factor Analysis (Pallant, 2007), is a method that attempts to reduce larger sets of variables into a smaller group of linear combinations of the original variables, such that these new smaller groups still capture most of the variability in the data (Pallant, 2007). PCA is useful when trying to remove unnecessary variables, reducing redundancy in the data set and removing possible multicollinearity. This study aimed to make use of multiple regression in order to better understand the amount of variability in the dependent variables explained by the independent variables. Therefore, PCA was deemed useful to reduce the number of explanatory variables required and assist in avoiding possible multicollinearity.

Research question 2 treated evidence-based decision-making as an independent variable and therefore PCA was used to reduce the 18 items used to measure this scale. Before conducting the PCA, the correlation matrix was inspected in order to ascertain whether every variable evidenced at least one correlation greater than 0.3 (Pallant, 2007). These correlations greater than 0.3 are bolded in Appendix 2 and show that a positive manifold was observed.

Further to the above, two more tests were used to ascertain whether PCA is appropriate. These were Bartlett's test of sphericity and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy (Pallant, 2007). The KMO measure was calculated to be 0.790 and was therefore deemed middling, between mediocre and meritorious. Furthermore, the KMO measures for the individual measures were calculated and analysed using the Anti-image correlation matrix. A KMO value of 0.6 is required as a minimum to ensure a good factor analysis (Pallant, 2007). The KMO measures from the Anti-image correlation matrix are summarised in Appendix 3. Moreover, the questionnaire administered in this research (see Appendix 1) provides the details of each item (E1-18) used in the anti-image correlation matrix. All the variables showed KMO measures greater than 0.7, except for variable E6. Therefore, variable E6 was removed from the dataset when using the evidence-based decision-making items for the multiple regression.

Finally, Bartlett's test of sphericity was conducted as a test for the suitability of PCA. The test was deemed statistically significant ($p < .0005$). Therefore PCA was deemed suitable based on the above tests.

Given the researcher's lack of experience when conducting PCA, both the eigenvalue rule (where only factors with an eigenvalue of one or more are retained) and a scree test were used to perform the PCA (Pallant, 2007). Appendix 4 shows the results from the PCA, with an eigenvalue of 1.003 for component six. Furthermore, 72.41% of the total variance was explained by the first six components, all components explained at least 5%, and adding an additional component would only explain a further 4.05% of the total variance and was therefore deemed unnecessary.

When analysing the scree plot, one must "find a point at which the shape of the curve changes direction and becomes horizontal" (Pallant, 2007, p. 182). The scree plot in Appendix 5 showed several possible inflection points, although component number seven seemed to highlight the most obvious inflection point where the remainder of the curve becomes horizontal. Therefore, based on both the eigenvalue rule and the scree test, the reduction to six components was deemed appropriate.

The final and most important step to PCA is the rotation and interpretation of the factors (Pallant, 2007). This merely presents the loading patterns in a manner that is easier to assess. Two approaches exist when rotating factors, orthogonal

(uncorrelated) and oblique (correlated), although often result in similar results (Pallant, 2007). An orthogonal approach was utilised considering it is easier to interpret. Appendix 6 shows the Rotated Component Matrix and only correlations greater than 0.3 were presented.
In summary, a PCA was run on the 18 questions that measured evidence-based decision-making of respondents. Before the test was run, suitability of PCA was established through inspection of the correlation matrix, which showed that a correlation coefficient of at least 0.3 was observed for all variables. The overall KMO value was reported as 0.790 and all individual KMO values were greater than 0.7, except for E6, which was removed from the construct before conducting multiple regression for research question 2. Finally, Bartlett's test of sphericity was statistically significant ($p < .0005$), which indicated that the data was likely factorisable.

Therefore, the PCA concluded six components (with eigenvalues greater than 1), which explained 34.5%, 10.4%, 8.5%, 7.6%, 5.9% and 5.6% of the total variance. This conclusion was supported by the visual inspection of the scree plot, which also indicated that six components be retained. This solution explained a total of 72.4% of the total variance. Finally, a Varimax orthogonal rotation was used to assist with interpreting the rotated factors. Variables for evidence-based decision-making were then grouped and used as six separate explanatory variables for research question 2.

Based on this analysis the individual items used to measure the evidence-based decision-making construct were grouped into six components. This is summarised in table 3. Whilst all the questions aimed to measure a culture of using evidence for decision-making, components were assessed for themes amongst the items.

EBDM1 included items that asked if organisations used external sources such as academic research and the Internet and if managers were capable of appraising this research before using it. Furthermore, an item regarding internal power struggles and their impact on decision-making, as well as learning from mistakes were also included in this component. The EBDM2 component included three items, which asked respondents questions regarding the organisation's tendency to systematically evaluate internal data, exploring a variety of solutions and the potential effectiveness of decisions. EBDM3 included two items, which asked respondents if the use of data was prevalent when creating new products, services and general decision-making. EBDM4 included two items and referred to the respondent's perception of their manager's view of using data for decision-making. EBDM5 included three items that addressed an organisations tendency to look externally and learn or benchmark progress against other organisations, as well as the belief that adopting new practices was important. The final component, EBDM6, included three items and included questions that addressed the access managers had when accessing information systems and extracting accurate data from them. Furthermore, it included an item referring to the use of consultants when making decisions.

Table 3: Summary of components and factor loadings

Component	Items included
Evidence-based decision-making 1 (EBDM1)	E10, E11, E12, E15, E18
Evidence-based decision-making 2 (EBDM2)	E7, E16, E17
Evidence-based decision-making 3 (EBDM3)	E1, E2
Evidence-based decision-making 4 (EBDM4)	E13, E14
Evidence-based decision-making 5 (EBDM5)	E3, E4, E5
Evidence-based decision-making 6 (EBDM6)	E6, E8, E9

4.8.2.1 Reliability of the measurement model

Reliability indicates how free a scale is from random error and is most commonly measured using Cronbach's Alpha (Pallant, 2007). Cronbach's coefficient measures by how much the different items used in a scale all measure the same dimension by providing an average correlation between all of the items, with values closer to one indicating greater reliability (Pallant, 2007). This was necessary for this research given that this study's research design made use of Likert-type scales and it is imperative to check any scales or subscales for internal consistency reliability when using Likert-type scales (Gliem & Gliem, 2003).

A questionnaire was used to collect data and measure the various items from respondents. Several scales were then constructed from these items for each research question. For the evidence-based decision-making scale, all items were averaged to form one scale and tested for reliability for research question 1, where it was used as a dependent variable. Given that evidence-based decision-making was used as an explanatory variable for research question 4, a PCA was conducted and six components were created and all tested for reliability. The big data capabilities were used as explanatory variables for all research questions, and therefore all items were averaged to form the relevant scales. Table 4 summarises these scales:

Table 4: Cronbach Alpha values for research question 2

Scale	Cronbach Alpha
Evidence-based decision-making	.738
Skillset	.876
Toolset	.872
Dataset	.777

When dealing with the concept of EO, multiple approaches have been observed from the literature. The first proponents of EO proposed a unidimensional view of EO, where the individual constructs (innovativeness, proactiveness and risk-taking) would all need to be present in order for a firm to be entrepreneurially orientated (Covin & Slevin, 1989; Gupta & Gupta, 2015). It was later adapted by some to be considered a multidimensional construct where a firm can be considered entrepreneurially orientated when at least one of the individual constructs is evident (Lumpkin & Dess, 1996; Gupta & Gupta, 2015). In line with the majority of previous EO research, this study adopted the unidimensional approach (Saeed et al., 2014; Wales et al; 2011) and therefore individual items for innovativeness, proactiveness and risk-taking were averaged to form the individual scale: EO.

In order to achieve a higher level of internal consistency and therefore ensuring the reliability of the scales Question E15 was removed from the EBDM1 scale to increase the Cronbach Alpha value from 0.521 to 0.839.

Table 5: Cronbach alpha values for research question 4

Scale	Cronbach Alpha
EBDM1	0.839
EBDM2	0.809
EBDM3	0.701
EBDM4	0.637
EBDM5	0.657
EBDM6	0.764
EO	0.753

Note: EBDM refers to the individual evidence-based decision-making components identified from the principle component analysis.

A minimum Cronbach Alpha value of 0.7 is recommended (Pallant, 2007), although scales with a smaller number of items can result in smaller Cronbach Alpha values (Pallant, 2007). Cronbach Alpha values less than 0.7 were observed for two of the scales in Table 5, namely: EBDM4 and EBDM5, and it was not possible to increase the Cronbach Alpha value any further. Therefore, the mean inter-item correlations were evaluated and resulted in mean scores of 0.471 and 0.391 respectively. Optimal mean inter-item correlation values are considered to range between 0.2 and 0.4 (Pallant, 2007). Therefore EBDM4 was removed from the analysis, as it did not meet any of the above requirements, whilst the scale: EBDM5 was retained for analysis.

4.8.3 Statistical tests conducted

Descriptive statistics such as the mean, median and standard deviation, as well as demographic and company/industry related data, were evaluated in order to describe the characteristics of the sample (Pallant, 2007). Correlational techniques are useful when trying to explore associations between variables and predict the outcome of a dependent variable based on one or many independent variables (Pallant, 2007). This research followed a similar approach conducted by Miller & Friesen (1982) and Linton and Kask (2016), which involved tests of association between variables (correlations) and the use of multiple regression.

This study aimed to understand how the various constructs, such as big data capabilities, evidence-based decision-making and EO were related and therefore the use of correlation analysis was appropriate as it allowed the researcher not only to measure the strength of these relationships but also whether the correlations were significant or not. This would address the research questions which aimed to provide clarity on the relationships identified in the model in Figure 1.

The strength of the relationships was evaluated based on the interpretation suggested by Cohen in Pallant (2007), which is summarised in Table 6.

Table 6: Correlation coefficient measures

Correlation coefficient value	Classification
0.10 - 0.29	Small
0.30 – 0.49	Medium
0.50 – 1.00	Large

In addition to calculating and evaluating the correlation coefficients, multiple regression was used to measure how much variability observed in the dependent variables could be explained by the explanatory variables. While various types of regression are available (Pallant, 2007), this study made use of standard multiple linear regression for research questions 2 and 3, as well as moderated multiple regression for research question 4. This test was considered appropriate as it provides further insight into the relationship between the various constructs of this study. Whilst the correlation analysis provided a view on the strength, direction and significance of the relationships, the use of multiple regression provides further clarity by measuring whether big data capabilities can be considered predictors of EO and evidence-based decision-making. Furthermore, it would also assist in understanding the different impacts between the big data capabilities,which would assist in understanding how big data can be leveraged by providing guidance on the importance of the big data capabilities.

Finally, the use of a moderated multiple regression provided clarity on whether an interaction effect existed between evidence-based decision-making and the effect big data capabilities had on EO. This would assist in answering research question 3 and understanding the importance of having an evidence-based decision-making culture.

Before conducting the relevant tests of association and prediction, various assumptions needed to be tested (Pallant, 2007). None of the assumptions were violated and therefore it was deemed appropriate to conduct correlation techniques and multiple regression. These assumptions are elaborated upon below.

4.8.4 Testing of assumptions

This section addresses the need to test the assumptions associated with using correlation tests and multiple regression. The structure follows the testing of each construct, or sub-dimension of a construct, as per assumption. Table 7 is provided as a guide of each construct or sub-dimensions and how it was used, per research question. The results of each test are then discussed.

Table 7: Summary of research questions, constructs and variable types

Application	Construct or sub-dimension tested	Type of variable
Research Question 1	Innovativeness	Independent variable
	Proactiveness	Independent variable
	Risk-taking	Independent variable
Research Question 2	Skillsets, toolsets and datasets	Independent variables
	Evidence-based decision-making	Dependent variable
Research Question 3	Skillsets, toolsets and datasets	Independent variables
	EO	Dependent variable
Research Question 4	Skillsets, toolsets and datasets	Independent variables
	EBDM1, EBDM2, EBDM3, EBDM5, EBDM6	Moderator variables
	EO	Dependent variable

Note: EBDM refers to the individual evidence-based decision-making components identified from the principle component analysis.

4.8.5 Tests for linearity and outliers

Before proceeding with the necessary tests for association and prediction, the preliminary analysis was required in order to test if various assumptions had been violated (Pallant, 2007). The first of these tests was that of linearity. This assumption states that in order for tests of association to be performed a straight-line or linear relationship must exist between the independent variable and the dependent variable. Further to the assumption of linearity, the sample was also evaluated for any outliers, which can have a "dramatic effect" on the correlation results by over or under-estimating the true relationship (Pallant, 2007, p. 121).

In order to assess the sample for linearity and outliers, scatterplots were generated and evaluated between each of the independent variables. Where applicable, independent variables were placed on the x-axis and the dependent variable on the y-axis. From Appendix 7, 8 and 9 the below conclusions were formed:

- The first research question dealt with the three sub-dimensions of EO. The scatterplots seemed to reveal a positive linear relationship between innovativeness and proactiveness, as well as between innovativeness and risk-taking although there was no clear linear relationship between risk-taking and proactiveness evident when assessing the scatterplot. The observations reflected a relatively dispersed spread and therefore outliers were deemed not present within the data.

- The second research question dealt with the relationship between the evidence-based decision-making construct and the sub-dimensions of big data capabilities. A linear relationship seemed to exist between each big data capability and evidence-based decision-making. The scatterplots were also evaluated for any outliers and none of the observations were deemed to be outliers.

- The third research question sought to understand the relationship between the big data capabilities (skillsets, toolsets and datasets) and EO. A visual inspection of the scatterplots revealed seemingly positive relationships between all the big data capabilities and EO. Furthermore, there did not seem to be any outliers evident in the data.

Based on the above analysis, linearity was confirmed for all of the constructs and sub-dimensions to be used in the required analysis. Furthermore, the data was deemed free of any outliers.

4.8.6 Tests for normality

Parametric and non-parametric tests exist when conducting tests for association and therefore the assumption of normality must be tested before proceeding to calculate the correlation matrix. A normal distribution is described as a “bell-shaped” curve, where the majority of observations fall in the middle of the distribution (Pallant, 2007, p. 57). Normality can be assessed either by calculating a test statistic or visually inspecting the histograms and Q-Q plots (Pallant, 2007). The researcher of this study opted to check all three tests, in search of consistency in the results.

The first test for normality conducted was the evaluation of the Shapiro-Wilk statistic. Table 8 highlights the results of this test, showing that EO, evidence-based decision-making and toolsets were all normally distributed, as assessed by Shapiro-Wilk's test ($p > 0.5$). Although, the remainder of the sub-dimensions, skillsets and datasets, were not normally distributed, as assessed by Shapiro-Wilk's test ($p < 0.5$).

Table 8: Shapiro-Wilk test for normality

Scale	Shapiro-Wilk p-value	Result
Skillset	0.001	Not normal
Toolset	0.203	Normal
Dataset	0.036	Not normal
EO	0.114	Normal
Evidence-based decision-making	0.082	Normal

Further to assessing Shapiro-Wilk's test for normality, both the histograms and Q-Q plots were visually inspected for each scale. Figure 10 show the histograms for each scale and confirm similar results to the initial tests for normality, showing roughly bell-shaped curves for evidence-based decision-making, EO and toolsets, whilst skillsets and datasets seemingly showed skewness in their distributions.

The Q-Q Plots in Appendix 11 highlights that the plotted points formed a "reasonably straight" line (Pallant, 2007, p. 62) for evidence-based decision-making, EO and toolsets, whilst the remainder of the dimensions did not seem to follow this trend.

Therefore, based on the preliminary analysis, evidence-based decision-making, EO and toolsets were all deduced to being approximately normally distributed, whilst the remaining dimensions, skillsets and datasets, were not. As a consequence of not all variables being normally distributed, this study made use of Spearman's Rank Order Correlation as a non-parametric alternative to the Pearson's product-moment correlation (Pallant, 2007).

4.8.7 Tests for normality of residuals

In order to run inferential statistics, one must check to see if the residuals are normally distributed. This must be done for each dependent variable by visually inspecting the histograms and normal P-P plots of the regression standardised residual. This study made use of two dependent variables, EO (research question 3 and 4) and evidence-based decision-making (research question 2). Appendix 12 and 13 highlight these results. Based on this analysis the distributions were deemed to be approximately normal, and therefore multiple regression would be appropriate for research questions 2, 3 and 4.

4.8.8 Tests for multicollinearity

When conducting multiple regression it is important to check for multicollinearity, which refers to the relationship or extent to which independent variables are correlated with one another (Pallant, 2007; Zikmund et al., 2009). High multicollinearity will negatively impact the regression model as it can make interpreting parameter estimates difficult (Zikmund et al., 2009). In order to test if the assumption of multicollinearity had been violated, the correlations between the independent variables were evaluated, followed by other collinearity statistics. Correlation coefficients higher than 0.7 would be signals of possible multicollinearity, whilst variance inflation factors (VIF) greater than 10 would further indicate possible multicollinearity (Pallant, 2007). Zikmund et al. (2009) suggest looking for VIF values not greater than 0.5. In this study, two sets of independent variables were used.

For research questions 2 and 3, the independent variables consisted of skillset, toolset and dataset. For research question 4, the independent variables also included the various evidence-based decision-making components as moderator variables. When analysing the correlation coefficients values ranged from 0.210 to 0.730, with only one correlation coefficient value higher than 0.7. This strong correlation was observed between toolsets and datasets. Further to this, the VIF are summarised in Tables 9 and 10 for each research question. From these tables, it is clear that none of the VIF exceeded 0.5. Even though one of the correlation coefficients slightly exceeded 0.7 - toolsets and datasets - an evaluation of the VIF suggested that multicollinearity did not exist between the independent variables. Therefore it was deemed that the assumption of multicollinearity was not violated and multiple regression analysis would be appropriate.

Table 9: Variance inflation factors for research question 3

Independent variable for research question 3	Variance inflation factor
Skillset	1.201
Toolset	2.232
Dataset	2.111

Table 10: Variance inflation factors for research question 4

Independent variable for research question 4	Variance inflation factor
Skillset	1.861
Toolset	2.800
Dataset	2.365
EBDM1	1.924
EBDM2	2.237
EBDM3	1.475
EBDM5	1.506
EBDM6	1.849

Note: EBDM refers to the individual evidence-based decision-making components identified from the principle component analysis.

4.8.9 Tests for homoscedasticity

The final assumption checked, in order to ascertain the appropriateness of conducting multiple regression, in this study was that of homoscedasticity. This assumption checked the equality of the residuals for all values of the dependent variables and was therefore tested for EO and evidence-based decision-making. This was done by visually inspecting the relationship between the standardised residuals and unstandardised predicted values, where increasing or decreasing funnel type shapes would indicate heteroscedasticity. This relationship is shown in Appendix 14. Based on a visual inspection of the scatterplots the assumption of homoscedasticity was deemed not to be violated, supporting the use of multiple regression.

4.9 Limitations

Various possible limitations were considered when conducting this research and are outlined below:

- The possibility of random sampling error (Zikmund, 2009) was acknowledged given the low sample size observed. This was reported as N = 82, which may have resulted in a lack of power in the tests (Pallant, 2007) or skewed results.

- The use of snowball sampling was identified as a possible limitation as this may have resulted in respondents with similar characteristics and, therefore, low variance in the sample (Biernacki & Waldorf, 1981).

- The cross-sectional nature of this study could have introduced some bias, as it was a snapshot of the data at a point in time. There, it was not possible to understand the impact of big data on evidence-based decision-making and EO over time.

Chapter 5: Results

5.1 Introduction

The previous section outlined how this research was conducted, what method was used to collect the data and a discussion regarding the statistical tests used in this research study. This section will present a description of the characteristics of the sample collected, in order to provide context of the study, and is followed by the results of the tests completed in order to address the research questions identified in chapter 3.

5.2 Characteristics of sample

5.2.1 Respondent demographics

Figure 2 highlights the split between male and female respondents. There were significantly more male respondents (71%) compared to female respondents (29%).

Figure 2: Respondent gender

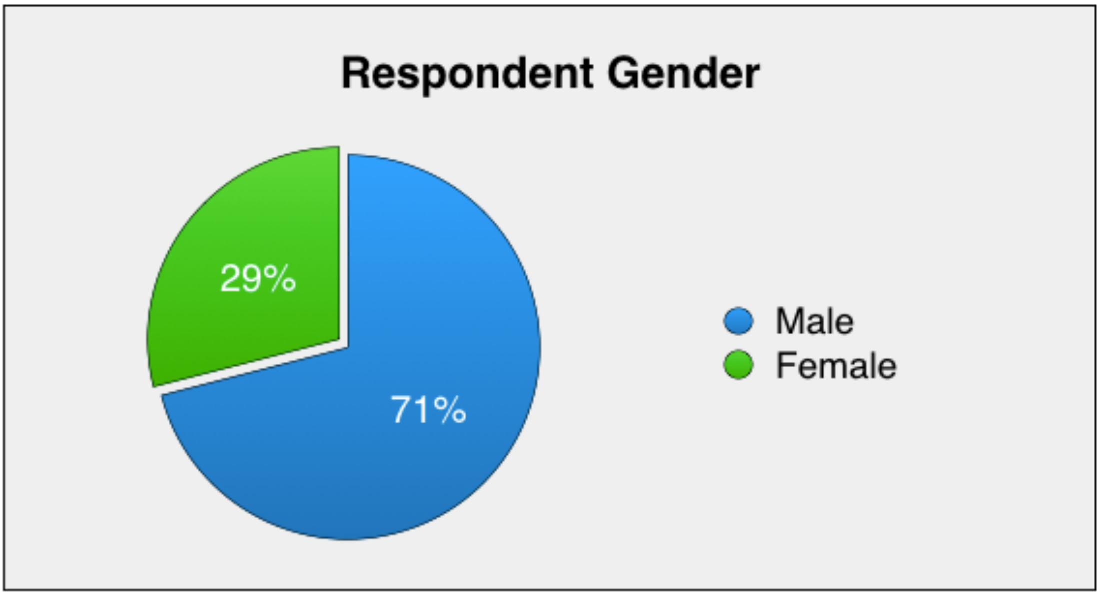
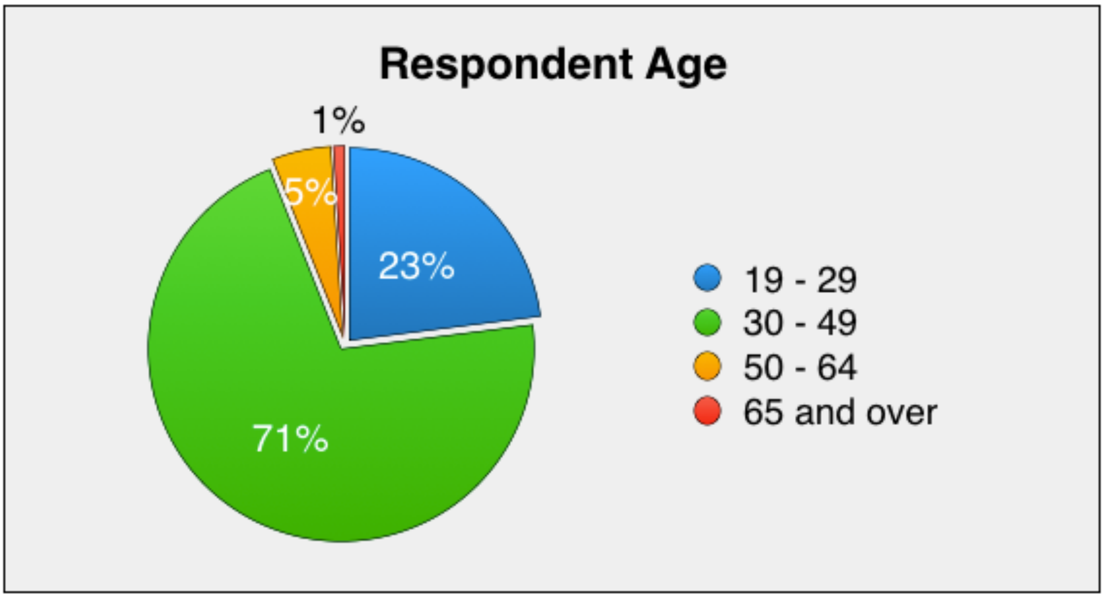


Figure 3 highlights the split of respondents by age group. The majority of respondents fell between 30 – 49 years of age (71%). The second largest group was respondents aged between 19 – 29 years of age (23%), the remainder of respondents came from the the remaining categories older than 50 years of age.

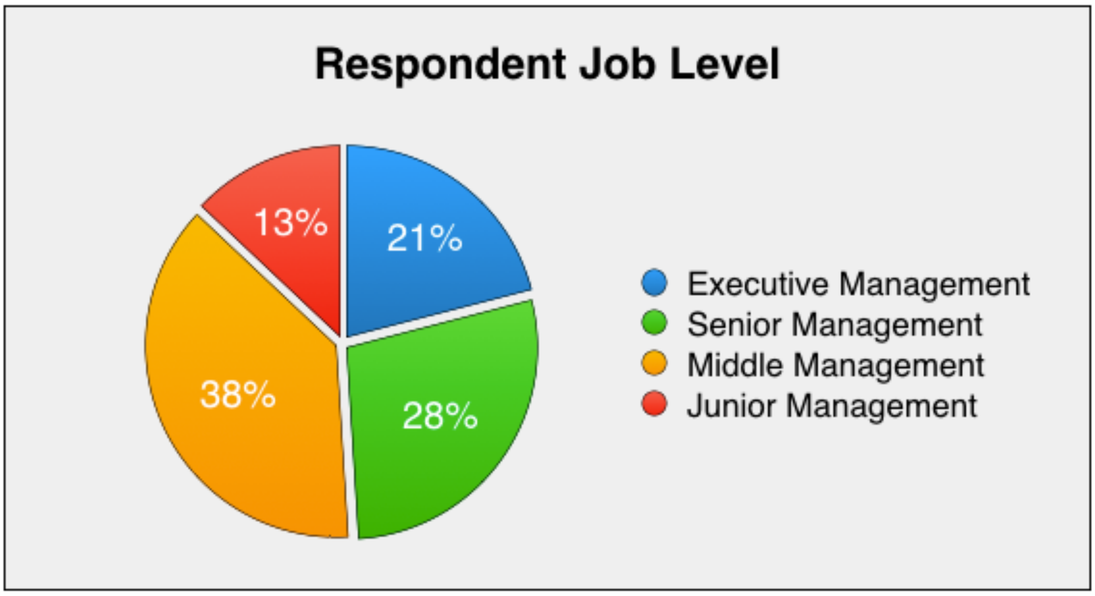
Figure 3: Respondent age



5.2.2 Respondent job level and type

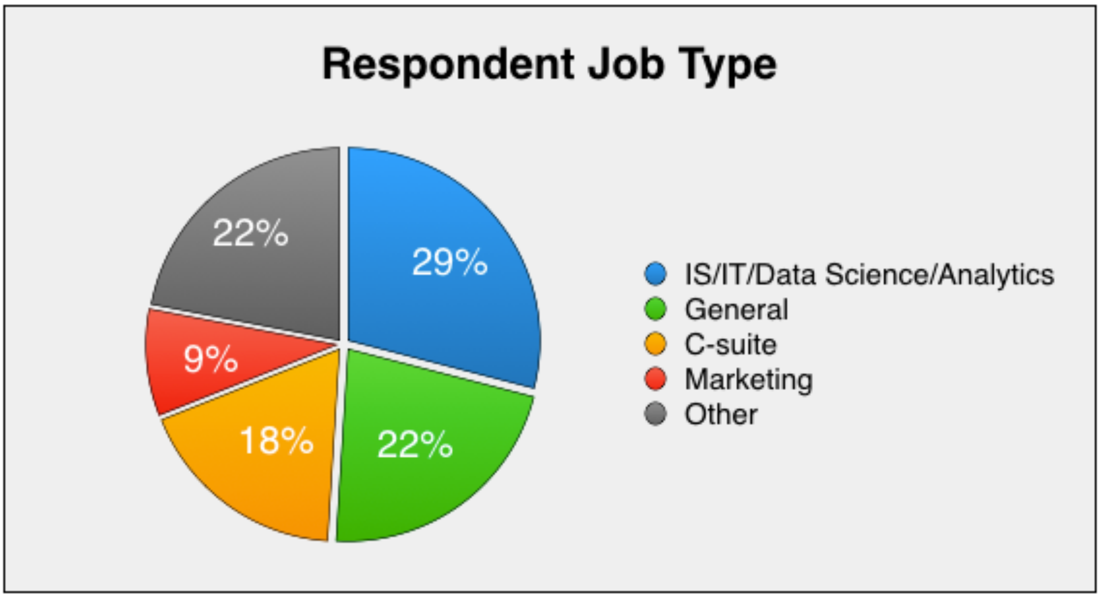
This study aimed to collect data from managers, regardless of industry or title. Therefore, all responses that did not indicate themselves within a managerial position were removed from the sample. Figure 4 shows the mix of respondents as per their level within the organisation. As shown, there was a fairly even split between the different job levels, with the majority of respondents falling into middle management (38%). It is also noted that significant portions of total respondents were from executive (21%) and senior management (28%), with a total of 87% of respondents that were not considered junior management.

Figure 4: Respondent job level



Further to respondent job level, data was also collected regarding the respondent's job type. Figure 5 shows the mix between job type. Whilst this study did not aim to restrict respondents based on job type, this data was collected to provide further context regarding the sample. The majority of respondents were from information systems (IS), information technology (IT), data science and data/business analysis functions (29%). Respondents from general management accounted for 22% of the total sample, followed by c-level respondents (CEO, COO, CFO etc.) at 18% and marketing managers at 9%. The remaining 22% of respondents were from a variety of job titles and were categorised as Other. Similar to job level, there was a level of variability from respondents regarding the job type.

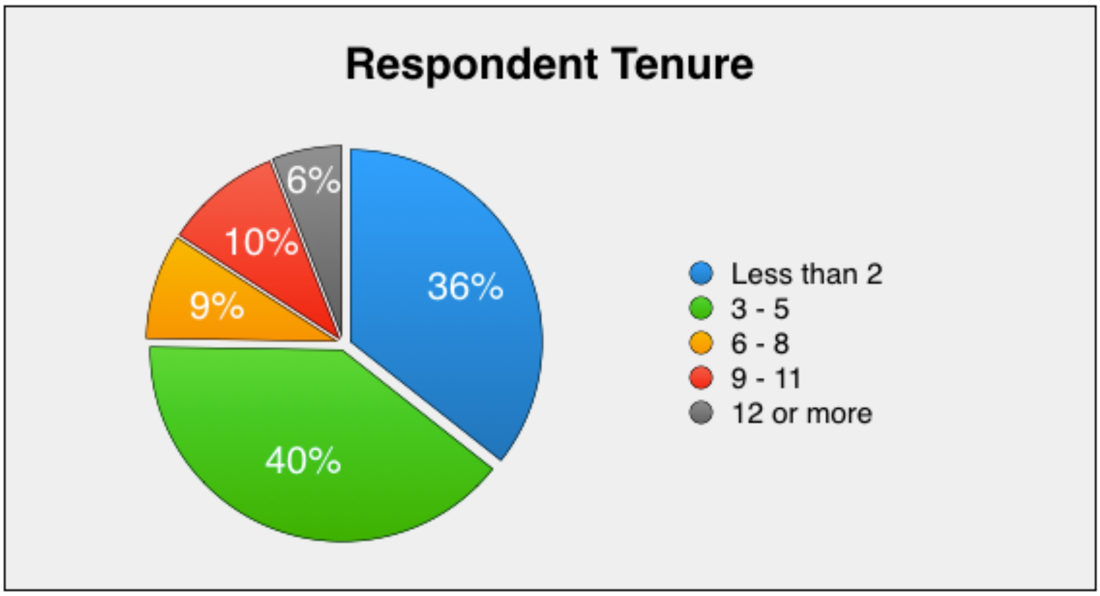
Figure 5: respondent job type



5.2.3 Respondent tenure

The questionnaire also sought to understand the tenure of respondents from their respective organisations. This is summarised in Figure 6 and highlights that majority of respondents (64%) had been at their organisation for three or more years. This means that the majority of respondents had been at the organisation long enough to develop a firm understanding of the organisation.

Figure 6: Respondent tenure



5.3 Results for Research Question 1

The first research question aimed to understand the relationship of the various sub-dimensions of EO in order to provide clarity on how EO would be measured for the remainder of the study. Therefore this question involved no more than the sub-dimensions innovativeness, proactiveness and risk-taking. The question was articulated as: Do the sub-dimensions of EO covary?

5.3.1 Descriptive analysis for research question 1

Descriptive statistics of the sample regarding the sub-dimensions of EO were evaluated and summarised into table 11. Several respondents did not complete the questions associated with the EO sub-dimensions and were therefore excluded from the analysis. Similar to previous research conducted on EO (Linton & Kask, 2016), respondents were given a scale from one (Statement A best described their situation) to seven (Statement B best described their situation), with the option of three (unsure). In each scenario, statement A reflected a situation that was less innovative, proactive or risk-taking and statement B reflected a situation that was more innovative, proactive or risk-taking. Therefore, higher values observed can be interpreted as higher levels of innovativeness, proactiveness and risk-taking. As per table 11, all three sub-dimensions of EO averaged below 4 (unsure) with innovativeness, proactiveness and risk-taking reporting mean values of 3.4, 3.5 and 3.1 respectively. This highlights that the average for each EO sub-dimension tended towards being less innovative, proactive and risk-taking.

Table 11: Descriptive statistics for research question 1

Variable	N	Mean	Median	Std. Deviation
Innovativeness	77	3.3983	3.6667	1.1522
Proactiveness	76	3.4803	3.5000	1.2448
Risk-taking	76	3.1184	3.3333	1.3129

5.3.2 Correlation analysis for research question 1

An analysis of the correlation coefficients was conducted using Spearman's rank-order correlation and is summarised in table 12. All three EO sub-dimensions showed positive correlations with one another. Innovativeness showed positive large correlations with both proactiveness and risk-taking (rs = 0.713 and rs = 0.506 respectively), whilst proactiveness and risk-taking showed a positive correlation that was deemed just between small and medium with rs = 0.299. Further to the above, the correlations were all shown to be statistically significant, $p < 0.01$.

Table 12: Correlation matrix for research question 1

Spearman's rho	Innovativeness	Proactiveness
Proactiveness	0.713*	
Risk-taking	0.506*	0.299*

Note: * = statistically significant at $p < 0.05$ level

Based on the above it was concluded that the sub-dimensions of EO covaried and therefore the EO construct is a unidimensional construct.

5.4 Results for Research Question 2

The second research question sought to understand how big data capabilities and evidence-based decision-making are related. This was articulated as: What is the relationship between big data capabilities and evidence-based decision-making? It is also noted that for research question 2, evidence-based decision-making was treated as the dependent variable and the big data capabilities as the independent variables.

5.4.1 Descriptive analysis for research question 2

The dimensions used for research question 2 and their respective descriptive statistics are summarised in Table 13. Responses for 2 of the big data capabilities (skillset and toolset) tended towards "Somewhat agree" when asked if their organisation possessed the respective big data capability. Whilst dataset tended towards "Unsure" with a mean of 4.2764. When respondents were asked about their level of evidence-based decision-making, the average respondent tended towards answering "Somewhat true".

Evidence-based decision-making also exhibited a relatively low standard deviation at 0.64, which showed that homogeneity existed within the sample responses.

Table 13: Descriptive statistics for research question 2

Variable	N	Mean	Median	Std. Deviation
Evidence-based decision-making	82	4.7173	4.7222	0.6347
Skillset	82	5.0476	5.3000	1.4510
Toolset	82	4.9220	5.0000	1.0789
Dataset	82	4.2764	4.0000	1.4069

5.4.2 Correlation analysis for research question 2

As discussed in chapter 4, certain assumptions were tested before proceeding with any tests for association. Given that normality was not confirmed for all four scales used in research question 2, the decision was made to proceed with Spearman's rank-order correlation. The results from this test are summarised in Table 14.

Table 14: Correlation matrix for research question 2

Spearman's rho	Evidence-based decision-making	Skillset	Toolset
Skillset	.604*		
Toolset	.520*	.434*	
Dataset	.399*	.433*	.730*

Note: * = statistically significant at $p < .05$ level

As evident in Table 14, the dependent variable showed a strong positive correlation to skillsets, with $r_s = .604$, a strong positive correlation to toolsets with $r_s = .520$ and a moderate positive correlation to datasets with $r_s = .399$. These correlations were all shown to be statistically significant, $p < .01$.

Furthermore, it is interesting to note that skillsets and toolsets, as well as skillsets and datasets both showed moderate positive correlations, $r_s = .434$ and $r_s = .433$ respectively, $p < .01$. Whilst toolsets and datasets showed a strong positive correlation with $r_s = .730$, $p < .01$.

5.4.3 Multiple Linear Regression results for research question 2

The results for the multiple regression model are summarised in Table 15. The overall model yielded an adjusted R-square of 0.428, suggesting a relatively good fit of the data and therefore showing that 42.8% of the variability in evidence-based decision-making can be explained by the big data capabilities. Furthermore, the p-value was reported to be less than 0.05, thus confirming that the big data capabilities were statistically significant predictors of evidence-based decision-making, $F(3,78) = 21.214$, $p < .0005$.

Table 15: Multiple linear regression results for research question 2

Summary of results	
Adjusted R-square	.428
F-test statistic	21.214
Regression degrees of freedom	3
Residual degrees of freedom	78
Probability of obtaining F-value if null hypotehsis is true	p < .0005

The regression model function was calculated as:

Evidence-based decision-making = 2.698 + 0.193(Skillset) + 0.234(Toolset) – 0.025(Dataset).

The coefficients were also analysed in order to understand the statistical significance of each independent variable, as well as their strength in predicting the dependent variable. This is summarised in Table 16 and shows that both skillsets and toolsets were considered to be statistically significant, with p-values less than 0.05. Dataset was considered not significantly with a p-value of 0.657, which is greater than 0.05.

Table 16: Beta coefficients for research question 2

Variable	B	Sig.	Outcome
Intercept	2.698	p < .0005	
Skillset	.193	p < .0005	Significant
Toolset	.234	0.002	Significant

Dataset	-.025	0.657	Not significant

5.5 Results for Research Question 3

The third research question aimed to measure the relationship between EO and the big data capabilities. This was articulated as: Are the big data capabilities antecedents for EO? This research question was addressed through the use of correlation techniques and standard multiple linear regression.

5.5.1 Descriptive analysis for research question 3

In light of the fact that EO was constructed from the sub-dimensions of innovativeness, proactiveness and risk-taking, the observed mean highlighted a similar result. This was observed to be 3.3 and highlighted that the sample tended towards less EO.

The big data capability dimensions: skillsets, toolsets and datasets, were also measured on a 7-point Likert scale and reported mean scores of 5.1, 4.9 and 4.3 respectively. This showed that respondents tended to "Somewhat Agree" to possessing the skillsets and toolsets for big data capabilities, whilst tending towards "Undecided" regarding datasets.

These descriptive statistics are highlighted in table 17.

Table 17: Descriptive statistics for research question 3

Variable	N	Mean	Median	Std. Deviation
Skillsets	82	5.0476	5.3000	1.4510
Toolsets	82	4.9220	5.000	1.0789
Datasets	82	4.2764	4.000	1.4069
EO	77	3.3297	3.3889	1.0089

5.5.2 Correlation analysis for research question 3

An analysis of the correlation coefficients was performed and is summarised in table 18. The big data capabilities all showed positive correlations with one another, with skillsets observed to have a medium positive correlation to both toolsets and datasets, reported as rs = 0.434 and rs = 0.433 respectively. A large positive correlation was observed between toolsets and datasets where rs = 0.730. The correlations between the big data capabilities were all reported to be statistically significant, $p < 0.01$. Skillsets was observed as the most correlated variable to EO.

The EO construct was also included into this analysis and yielded medium positive correlations to all three big data capabilities, with the correlation coefficients with skillsets, toolsets and datasets observed to be 0.411, 0.313 and 0.374 respectively. Moreover, the correlations between EO and the big data capabilities were all reported to be statistically significant, $p < 0.01$.

Table 18: Correlation matrix for research question 3

Spearman's rho	EO	Skillsets	Toolsets
Skillsets	0.411*		
Toolsets	0.313*	0.434*	
Datasets	0.374*	0.433*	0.730*

Note: * = statistically significant at p < .05 level

5.5.3 Multiple linear regression for research question 3

The multiple linear regression results reported an adjusted R-square of 0.170, which was considered a poor fit with only 17% of the variability in EO explained by the big data capabilities. The p-value was reported to be less than 0.05 and therefore the big data capabilities were reported to be statistically significant predictors of EO, $F(3,73) = 6.175$, $p = 0.001$. These results are summarised in table 19.

Table 19: Multiple linear regression results for research question 3

Summary of results	
Adjusted R-square	0.170
F-test statistic	6.175
Regression degrees of freedom	3
Residual degrees of freedom	73
Probability of obtaining F-value if null hypotehsis is true	0.001

The regression model function was calculated as:

EO = 1.523 + 0.210 (Skillsets) – 0.018 (Toolsets) + 0.190 (Datasets)

The coefficients were also analysed in order to understand the statistical significance of each independent variable, as well as their strength in predicting the dependent variable. Table 20 summarises these results and shows that only skillsets was reported to be statistically significant when predicting EO, with a p-value less than 0.05. The remaining big data capabilities, toolsets and datasets, were observed to be not significant with p-values of 0.907 and 0.084 (both greater than 0.05).

Table 20: Beta coefficients for research question 3

Variable	B	Sig.	Outcome
Intercept	1.523	0.008	
Skillset	0.210	0.009	Significant
Toolset	-0.018	0.907	Not significant
Dataset	-.190	0.084	Not significant

5.6 Results for Research Question 4

The final research question followed on from the third question and aimed to understand if there was an interaction effect between evidence-based decision-making and the relationship between big data capabilities and EO. This question was articulated as: Is evidence-based decision-making culture a moderator for big data capabilities on EO?

For research question 4 the evidence-based decision-making construct was used as an explanatory variable (as opposed to research question 2) and the PCA conducted in chapter 4 identified various components which were used as moderators in the moderated regression model. Correlation coefficients are repeated for EO and the big data capabilities in order to provide insight into their relationship with the various evidence-based decision-making components.

5.6.1 Descriptive analysis for research question 4

Table 21 highlights the observed descriptive statistics for research question 3. Components EBDM1, EBDM2, EBDM5 and EBDM6 all tended towards "Sometimes True" with mean scores of 4.6, 4.9, 5.2 and 5.4 respectively. All components addressed the level of evidence-based decision-making within an organisation, although individual themes amongst the components were highlighted in chapter 4. Whilst majority of components tended towards the same outcome, EBDM3 exhibited a mean score of 4.1 and therefore tended towards "Neutral".

Table 21: Descriptive statistics for research question 4

Variable	N	Mean	Median	Std. Deviation
EBDM1	82	4.5640	4.7500	1.3216
EBDM2	82	4.9817	5.3333	1.3042
EBDM3	82	4.1341	4.5000	0.9265
EBDM5	82	5.1951	5.3333	1.0383
EBDM6	82	5.4085	6.0000	1.2961

5.6.2 Correlation analysis for research question 4

Table 22 summarises the correlation coefficients for research question 4 and highlights a positive manifold amongst the various constructs and components. The majority of the correlation coefficients between the big data capabilities and the evidence-based decision-making components yielded medium to large positive correlations, with the exception of datasets and EBDM1, and toolsets and EBDM5. Furthermore, EO also exhibited medium positive correlations with all but one evidence-based decision-making component (EBDM6). The various components also showed positive correlations with one another ranging from 0.297 to 0.570.

The majority of the evidence-based decision-making components were statistically significant and correlated with EO and the big data capabilities (with $p < 0.01$), with the exception of toolsets and EBDM5 which was not significant.

Table 22: Correlation matrix for research question 4

Spearman's rho	EBDM1	EBDM2	EBDM3	EBDM5	EBDM6

EO	0.316*	0.402*	0.302*	0.340*	0.252*
Skillsets	0.446*	0.593*	0.445*	0.384*	0.496*
Toolsets	0.503*	0.530*	0.313*	0.210	0.519*
Datasets	0.287*	0.376*	0.311*	0.301*	0.554*
EBDM2	0.568*				
EBDM3	0.297*	0.495*			
EBDM5	0.320*	0.490*	0.570*		
EBDM6	0.447*	0.465*	0.348*	0.263*	

Note: * = statistically significant at p < .05 level.

5.6.3 Moderated multiple linear regression for research question 4

The results from the moderated multiple linear regression model, reported in table 23, yielded a higher adjusted R-square (to that of research question 3) at 0.209. This suggested a better fit with the inclusion of the moderator variables: evidence-based decision-making components, with 20.9% of the variability explained by the big data capabilities and moderator variables. Moreover, the p-value was reported to be less than 0.05, thereby confirming that the model was statistically significant, $F(8,68) = 3.510$, p-value = 0.002.

Table 23: Moderated multiple linear regression results for research question 4

Summary of results	
Adjusted R-square	0.209
F-test statistic	3.510
Regression degrees of freedom	8
Residual degrees of freedom	68
Probability of obtaining F-value if null hypotehsis is true	0.002

The regression model function was calculated as:

EO = 1.050 + 0.158 (Skillsets) – 0.133 (Toolsets) + 0.228 (Datasets) + 0.170 (EBDM1) + 0.064 (EBDM2) – 0.058 (EBDM3) + 0.177 (EBDM5) – 0.118 (EBDM6)

Table 24 highlights the analysis of the beta coefficients and showed that the majority of the explanatory variables were deemed insignificant, with p-values greater than 0.05.

Only one explanatory variable – dataset - reported a p-value = 0.046, which was less than 0.05 and was therefore deemed significant

Table 24: Beta coefficients for research question 4

Variable	B	Sig.	Outcome
Intercept	1.050	0.144	
Skillset	0.158	0.114	Not significant
Toolset	-0.133	0.429	Not significant
Dataset	0.228	0.046	Significant
EBDM1	0.170	0.121	Not significant
EBDM2	0.064	0.594	Not significant
EBDM3	-0.058	0.674	Not significant
EBDM5	0.177	0.161	Not significant
EBDM6	0.118	0.280	Not significant

Chapter 6: Discussion of Results

6.1 Introduction

The objective of this study, as outlined in Figure 1, was to provide further clarity into the relationships between the big data capabilities, evidence-based decision-making culture and EO. This was achieved by addressing the four research questions identified in chapter 3, which are elaborated on in this chapter. The previous chapter reported and summarised the descriptive analysis and statistical tests conducted in order to answer these research questions and achieve the research aim. This chapter discusses the results reported in chapter 5 and is also structured as per research question.

A review of the literature highlighted that big data was uniquely positioned to contribute to an organisation's ability to achieve its goals through better decision-making that is based on evidence. A research model was proposed in Figure 1 and highlights the proposed relationships between the various constructs, showing that big data is posited as an enabler of a firm's ability to be more entrepreneurial and a culture of evidence-based decision-making. Research questions 2 and 3 sought to measure these proposed relationships between an organisation's big data capabilities and their ability to be more entrepreneurial and base decisions on evidence. Research question 4 was then aimed to measure whether an organisation's evidence-based decision-making culture acted as a moderator on the relationship between big data capabilities and EO.

Furthermore, the literature revealed that conflicting views were present in the current EO literature regarding the measurement of EO. The researcher of this study opted to contribute to the literature by providing empirical evidence regarding the EO sub-dimensions. This was addressed by the first research question.

6.2 Discussion on Research Question 1

The first research question focused on the sub-dimensions of the EO construct: innovativeness, proactiveness and risk-taking, and was articulated as:

Do the sub-dimensions of EO covary?

A review of the EO literature highlighted a “major schism” (Gupta & Gupta, 2015, p. 59) amongst researchers, that although research regarding the construct had increased over the last three decades, conflicting views regarding how the construct should be measured are present (Gupta & Gupta, 2015; Linton & Kask, 2017; Saeed et al., 2014; Wales et al., 2011). Majority of research conducted on EO has utilised the unidimensional view (Saeed et al., 2014; Wales et al., 2011) established by Covin and Slevin (1989) and suggests that the sub-dimensions of EO covary, implying that organisations that are considered entrepreneurial will exhibit high levels of all three sub-dimensions. Other authors (Lechner & Gudmundsson, 2014; Naldi et al., 2007) treated EO as a multidimensional construct, as established by Lumpkin and Dess (1996), arguing that the sub-dimensions of EO can covary and, therefore, the sub-dimensions should be measured independently (Linton & Kask, 2017).

The main aim of this research was not to address the question of how EO should be measured, although, the design of this research provided the opportunity to measure the relationship between the sub-dimensions. Furthermore, when measuring relationships, the proliferation of empirical evidence in research assists and is necessary in providing clarity regarding the construct (Frese et al., 2012). In order to address this research question descriptive analysis was conducted, accompanied by tests for association.

An analysis of the descriptive statistics showed that on average respondents tended towards being less innovative, less proactive and less risk-taking, with mean scores reported of 3.4, 3.5 and 3.1 respectively (based on a 7-point Likert sale). These results were similar to that of Linton and Kask (2017) who also employed the Covin and Slevin (1989) 9-item scale, measured on a 7-point Likert scale. Similar to this study, they observed mean values of 3.3, 3.7 and 3.1 for innovativeness, proactiveness and risk-taking respectively. Furthermore, Lechner and Gudmundsson (2014) reported mean scores of 2.3, 2.7 and 3 for innovativeness, proactiveness and risk-taking respectively (based on a 5-point Likert scale), which was congruent with the findings of both this study and that of Linton and Kask (2017).

Interestingly, this suggests that the perception of respondents tended towards a view that their organisations were not necessarily innovative, proactive or risk-taking. Given that this research study relied on the perception of respondents when measuring the sub-dimensions of EO, it was not possible to ascertain whether these perceptions reflected the reality of the organization.

A possible explanation of these findings is that given the rapid change in technology (Purnama & Subroto, 2016), market instability (Reeves & Deimler, 2011) and heightened focus on creativity and innovation given the increased levels of competition (Mathews, 2016; Prajogo, 2016), employees perceptions of their own entrepreneurial activity may be negatively influenced by the perceived change and pressure from the external environment.

An analysis of the correlation coefficients highlighted statistically significant correlations between all three sub-dimensions, with medium to large positive correlation coefficients reported as 0.713, 0.506 and 0.299 between innovativeness and proactiveness, innovativeness and risk-taking and proactiveness and risk-taking respectively. This suggested that the sub-dimensions of EO do covary and, therefore, supported the view that EO is a unidimensional construct, which was aligned to the majority of current EO research (Rauch et al., 2009; Saeed et al., 2014; Wales et al., 2011). This implies that organisations that aim to be more innovative, in order to achieve some form of competitive advantage, must also proactively seek out opportunities and be willing to take risks.

Lechner and Gudmundsson (2014) did not report correlation results, although Linton and Kask (2017) reported findings similar to this research with all three coefficients exhibiting positive and statistically significant correlations. This was found to be contradictory to the possibility of a multidimensional view, proposed by Covin and Slevin (1989), where the sub-dimensions can covary. Whilst this outcome does not conclusively show that EO cannot covary, it does contribute to the view of EO as a unidimensional construct.

Whilst results from prior studies reported relatively higher standard deviations (Linton & Kask, 2017; Lechner & Gudmundsson, 2014), table 14 highlighted the low standard deviation values reported in this study and highlighted that little variance was evident in the sample, constraining the generalisabililty of the results obtained. Furthermore, these results may have been skewed given the low sample size obtained in this study, which was reported in total as 82, although reduced to 76 for research question 1 as several respondents did not complete the relevant section on EO.

6.3 Discussion on Research Question 2

The second research question sought to understand the relationship between the big data capabilities and the firm's evidence-based decision-making culture. This was articulated as:

What is the relationshp between big data capabilities and evidence-based decision-making?

The importance of maintaining an evidence-based or data-driven decision-making culture has been emphasized as a necessity in the workplace (Rousseau, 2006) as organisations are vastly more complex and not homogenous, and therefore relying on experience to answer strategic decisions can be limiting, as an individual may be bound by cognitive frames and cannot accumulate enough experience to understand all businesses and situations (Pfeffer & Sutton, 2006).

This premise has been tested and researchers found that organisations that exhibited a data-driven culture with regards to decision-making achieved higher firm performance (Brynjolfsson et al., 2011; LaValle et al., 2011) and the use of an analytical decision-making culture can result in better information usage (Popovic et al., 2012). This places emphasis on understanding what internal organisational aspects can support a culture of evidence-based decision-making. This research study aimed to contribute to the literature by examining various big data capabilities, such as skillsets, toolsets and datasets and their relationship to an organisation's evidence-based decision-making culture.

This research question was addressed by analysing the descriptives, correlation coefficients and the evaluation of the big data capabilities as possible predictors of evidence-based decision-making using multiple linear regression.

It was noted that all 82 responses were retained for this research question and the mean scores were reported as 4.7, 5.1, 4.9 and 4.3 for evidence-based decision-making, skillsets, toolsets and datasets respectively. Therefore, respondents tended towards "Somewhat agree" when responding and this highlighted that on average respondents felt relatively positive regarding their organisations stance towards the possessing the relative skillsets, toolsets and evidence-based culture. Although, respondents tended to be “Unsure” when asked about their organisation’s possession of the relevant and appropriate datasets.

The evaluation of the correlation matrix in table 17 highlighted statistically significant correlations between all four constructs. The big data capabilities reported medium to strong positive correlation coefficients of 0.434, 0.433 and 0.730 between skillsets and toolsets, skillsets and datasets and toolsets and datasets respectively. This confirmed a positive relationship between the different big data capabilities but in no way confirmed a causal relationship. Furthermore, it was still unclear which capability preceded the others.

Further to the above, evidence-based decision-making reported large positive correlations of 0.604, 0.520 and 0.399 with skillsets, toolsets and datasets respectively. These findings further contribute to the big data and evidence-based decision-making literature (Elgendy & Elragal, 2016) in support of the view (Popovic et al., 2012) that data usage can result in the improved use of information.

The multiple linear regression results reported an adjusted R-square of 0.428, which is considered a relatively good-fit of the model with the big data capabilities explaining approximately 43% of the variability in evidence-based decision-making. Furthermore, the model was reported as statistically significant. These findings contribute to the literature by providing insight into how organisations can become more data-driven, something organisations are still struggling with (Bean, 2017). Furthermore, these findings provide much needed empirical evidence (Frese et al., 2012), highlighting the relationship between the usage of data (Elgendy & Elragal, 2016; Popovic et al., 2012) and an organisation’s ability to maintain an evidence-based decision-making culture.

It is futher noted that the big data capabilities: skillsets and toolsets were both reported as statistically significant with positive beta coefficients of 0.193 and 0.234 respectively. Therefore, highlighting an important notion that an organisation’s ability to accumulate the correct and relevant skillsets can result in higher levels of evidence-based decision-making culture. Moreover, the stronger an organisation’s ability to provide data and insights to employees through the relevant systems and business intelligence tools (toolsets) can also result in higher levels of evidence-based decision-making.

The majority of the big data literature seemed to focus on the importance of skillsets (Alharthi et al., 2017; Chen et al., 2012; George et al., 2016; Gobble, 2013; Waller &

Fawcett, 2013; Mills et al., 2016), which is justified by these findings as not only did big data skillsets exhibit the largest correlation coefficient, it was also reported as a significant predictor, with a positive beta coefficient, of evidence-based decision-making culture in the multiple regression results. This contributes to the literature (Bean, 2017; Elgendy & Elragal, 2016; Popovic et al., 2012) by confirming the focus on big data skillsets as both necessary and useful, given its positive relation to evidence-based decision-making culture. Furthermore, these findings suggest that further emphasis be placed on an organisations available systems (toolsets), which showed similar results to skillsets.

Therefore, organisations that hope to achieve increased firm performance (Brynjolfsson et al., 2011; LaValle et al., 2011) in a world that is rapidly becoming inundated with new technology and volumes of data must understand how they can effectively employ the correct skills that can lead data projects and the implementation of relevant systems.

6.4 Discussion on Research Question 3

The main aim of this research was to understand how big data can be leveraged for strategic decision-making, in order for organisations to achieve a competitive advantage through the effective use of their data. EO was identified as a stabilised (Gupta & Gupta, 2015) and appropriate firm-level measurement (Covin, Green & Slevin, 2006; Linton & Kask, 2016; Wales et al., 2011) of the strategic decision-making posture and used to address research question 3. This was articulated as:

Are the big data capabilities antecedents for EO?

A review of the literature highlighted EO as a well researched construct since it was first established by Miller (1983) and Covin and Slevin (1989). Since then the majority of EO research has focused on the formation of an accurate measurement tool for EO (which was discussed extensively in research question 1), the link between EO and firm performance (Lechner & Gudmundsson, 2014; Rauch et al., 2009) and employment growth (Madsen, 2007), and the role of external moderators on EO (Linton & Kask, 2016; Wales et al., 2011).

The construct was considered appropriate for this research as it places emphasis on the different strategic postures that an organisation can exhibit (Gupta & Gupta, 2015) and focuses on how organisations can become more successful through innovation, proactiveness and risk-taking, that creates competitive advantage, as opposed to a strategy that aims to cut costs. Similar to business model literature that explores how organisations deliver value to customers and how an enterprise can effectively organise itself to meet those customer needs (Teece, 2010), the EO construct measures a firms ability to achieve the goals of the firm and is, therefore, deemed relevant and useful.

The proliferation of EO research seems to be an acknowledgement of the value placed on the EO construct, with several studies attempting to understand the antecedents of EO (Wales et al., 2011). Although the majority of the research regarding the antecedents of EO has focused on the CEO, the management team, the organisation's strategic orientation and environmental considerations (Wales et al., 2011). This provided an opportunity to contribute to the literature by examining whether an organisation's ability to effectively utilise big data internally impacted their ability to be entrepreneurial.

This research study posited big data capabilities as antecedents to EO, creating dynamic capabilities for organisations (Schilke, 2014), measured by EO, through the effective use of their data that can promote competitiveness. The variety, velocity, veracity and volume of big data (Alharthi et al., 2017; Davenport et al., 2012; Goes, 2014; Pigni et al., 2016) is discussed extensively across the literature as an enabler of internal efficiencies, better customer experience and improved revenues and profitability (Alharthi et al., 2017; Bean, 2017).

The literature revealed three key areas of focus for organisations that wished to utilise their data effectively. The first was identified as the necessary skills required to address big data (Pigni et al., 2016; Waller & Fawcett, 2013). These skills were identified as IT proficiency, statistical modelling skills and the ability to deliver useful insights across the business (Alharthi et al., 2017; Chen et al., 2012; George et al., 2016; Gobble, 2013; Mills et al., 2016). Secondly, an organisaiton's ability (labelled as toolsets) to effectively implement and maintain the appropriate databases, ERP systems, data warehouses, business intelligence tools and other technology is highlighted as a necessary antecedent as well (Pigni et al., 2016). Organisations can gain faster and improved access to information, higher levels of consistency and interactivity and timeous queries with the use of a business intelligence system (Popovic et al., 2012), although it is important that the systems used are capable of managing these new forms of data (variety), whether they be structured or unstructured (Alharthi et al., 2017). Finally, an organisation's ability to identify and access the relevant data that can be used for value creation is also identified as a necessity to successful big data initiatives (Pigni et al., 2016).

This research question aimed to contribute to the literature by measuring the relationship between an organisation's big data capabilities and EO and answer the call from researchers (Wales et al., 2011) for more EO research that provided practical relevance to business practitioners. Similarly to research question 2, research question 3 was addressed through the analysis of the descriptive statistics, the correlation coefficients and the multiple linear regression model.

The descriptive statistics for research question 3 were summarised in table 20, with the addition of the EO construct. The reported mean score for EO was reported as 3.3 and therefore indicated that respondents tended towards considering their organisations as less entrepreneurial. The big data capabilities descriptives have already been discussed in the previous section. Once again, it is noted that the sample size for EO was reduced to 77 given that several respondents did not complete the questionnaire. Furthermore, relatively low standard deviations were reported and therefore indicated low variability in the sample, which can negatively impact the generalisability of the results.

Table 21 highlights the correlation coefficients between EO and the various big data capabilities. Interestingly, the results showed medium positive correlations between the EO construct, that were all statistically significant at a 95% confidence level.

Further to this, the summary of the regression results yielded an adjusted R-square of 0.170, which indicated a poor fit as only 17% of the variation in EO was explained by the big data capabilities. Furthermore, the model was reported to be statistically significant.

Finally, an analysis of the explanatory variables highlighted that only skillsets was reported as a statistically significant predictor of EO. Furthermore, a positive beta coefficient of 0.210 was observed for skillsets.

These findings confirm the notion that a positive relationship exists between an organisation's big data capabilities and their ability to be enterpreneurial, although do not imply causality in any way. This contributes to the current literature by building the argument, with empirical evidence, that the importance of possessing the correct big data capabilities, especially skillsets, is fundamental to allowing organisation's to act more entrepreneurially. Furthermore, whilst big data literature has focused on comparing the usage of data and an organisation's ability to make data-driven decisions (Elgendy & Elragal, 2016; Popovic et al., 2012), these findings confirm and contribute to the literature by providing a different view of the possible benefits from the relevant big data capabilities.

Interestingly, the concept of skillsets was reported as statistically significant (similar to research question 2) and had a positive impact on an organisation's ability to be entrepreneurial. However, toolsets was reported as insignificant (unlike research question 2) when predicting EO. This not only highlights the intense significance of skillsets specifically, but also calls for further research into understanding which skills organisations should focus on. The requirement of IT and statistical skills within the big data field is prevalent within the literature, although the requirement for data scientists to find innovative ways of collecting, organising and sharing data insights effectively across the business is described as an emerging skill (Gobble, 2013). This emerging skills may be closer to the construct of EO than the need for IT skills and, therefore, these findings call for further research into which specific big data skills have a greater impact on a organisation's ability to be more entrepreneurial.

6.5 Discussion on Research Question 4

The final research question included all the constructs from the previous research questions, namely: big data capabilities, evidence-based decision-making and EO and aimed to understand if an interaction effect existed between evidence-based decision-making and the relationship between big data capabilities and EO. This was articulated as:

Is evidence-based decision-making culture a moderator for big data capabilities on EO?

Further to what has already been discussed in this chapter, the literature revealed that various moderator variables have been explored extensively by previous researchers (Wales et al., 2011). Various research studies have been done on the moderating effect of variables concerning the CEO, culture, environment, networking, organisation ownership and structure, organisational learning, strategy, human resources, leadership and team cohesiveness (Wales et al., 2011). The research on culture included various focus areas, although it did not include evidence-based decision-making culture. This presented an opportunity to further the literature on the role of moderating variables on EO.

For this research question the evidence-based decision-making construct was treated as an explanatory variable, which consisted of eight items that were reduced into five components following the PCA and tests for reliability. These five components were then used as moderators for research question 4.

The descriptive statistics revealed that all four components tended towards "Sometimes true", with the exception of EBDM3 which tended towards "Neutral". EBDM3 concerned the usage of data for the creation of new products or services.

A review of the correlation matrix revealed that the majority of the EBDM components were statistically significant with EO, skillsets, toolsets and datasets. With the exception of EBDM5 and toolsets. Furthermore, these correlation coefficients were all reported as small, medium and large positive correlations.

The multiple regression results reported an adjusted R-square of 0.209, which was a relatively poor fit with approximately 20.9% of the variation in EO explained by the independent variables. Although, it is further noted that the adjusted R-square in this model yielded a higher value compared to the model without the moderator variables (17%). Therefore, the inclusion of the EBDM moderator variables increased the goodness of fit.

Furthermore, an inspection of the independent variables yielded only one significant variable, datasets, with the remainder of the variables reported as not significant. Datasets was also reported to have a positive impact on EO with a beta coefficient of 0.228 and referred to an organisations ability to to identify and access the correct data, that was relevant and adequate (without "noise") and not obsolete.

6.6 Conclusion on discussion

The overall findings of this study seem to support the position of big data capabilities in the research model represented in Figure 1. The aim of this study was to understand how organisations can leverage big data for strategic decision-making. The findings from research questions 2 and 3 both exhibited positive significant correlations between the big data capabilities and EO, as well as between the big data capabilities and an organisation's evidence-based decision-making culture. Furthermore, the regression results reported skillsets as a statistically significant predictor of EO and evidence-based decision-making, and toolsets as a statistically significant predictor of evidence-based decision-making.

These findings suggest a link exists between an organisation's big data skillsets and their ability to be more entrepreneurial, which serves as a contribution to the current literature regarding EO antecendents (Wales et al., 2011) and provides practical areas of focus for business practitioners. Furthermore, the importance of the appropriate toolsets has been stressed in the literature (Alharthi et al., 2017; Pigni et al., 2016) and these findings support this view as given the positive relationship reported between toolsets and evidence-based decision-making.

Finally, the sub-dimensions of EO were also analysed and showed that the EO construct is a unidimensional construct. This contributes to the literature and provides clarity on one of the main constructs of this study, EO, which was shown to be positively related to the big data capabilities. Whilst organisations must focus on employing the correct skillsets and maintaining the appropriate toolsets, the idea of a unidimensional view suggests that organisations must also recognize that they must proactively seek out opportunities and take risks if they hope to achieve innovativeness outcomes.

Chapter 7: Conclusion

This chapter aimed to consolidate the findings of this research study and place them in context of the research problem and research aim presented in chapter 1. Organisations find themselves in an environment of rapid technology change and increased levels of competition. Whilst big data has been posited as an enabler of competitive advantage in this new era, it is not clear how organisations can leverage big data to achieve this and organisations continue to struggle in implementing successful big data initiatives.

The concept of EO was used as a measure of a possible benefit of big data that can assist firms in being more innovative, proactive and risk-taking in order to create and maintain higher levels of competitive advantage. This research aimed to establish if a relationship existed between the big data capabilities and EO in order to provide new and useful insights on how organisations can possibly achieve this proposed competitive advantage.

The literature identified various big data capabilities required by organisations who aim to implement big data initiatives and were used to measure an organisation's ability to effectively use big data. These were categorised as skillsets, toolsets and datasets and not only allowed the researcher to measure the ability of the organisations studied to manage big data but also allowed for futher analysis into the intensity of the relationship between the various capabilities and EO.

Finally, the concept of evidence-based decision-making was also included as a necessary culture when making use of data for strategic decision-making. The research also aimed to understand if a moderating effect existed between an evidence-based decision-making culture and the relationship between big data and EO.

Further to the above, the researcher made use of the opportunity to study the dimensionality of the EO construct, as well as the relationship between big data and evidence-based decision-making culture in order to provide further valuable and useful insights that may contribute to the current body of literature.

The principal findings are discussed, followed by a discussion on the implications of these findings to management, suggestions for further research and the limitations observed. Finally, the research is concluded.

7.1 Principal findings

One of the key findings of this research was between that of the big data capabilities and the EO construct. This study found positive medium correlations, that were statistically significant, between the various big data capabilities and EO, highlighting that higher levels of big data skillsets, toolsets and datasets can result in higher levels of entrepreneurial activity from firms. While research that measured the relationship between EO and big data was not available, to the researchers knowledge, this finding reaffirms the notion that big data is uniquely positioned in assisting organisations to achieve their goals (Alharthi et al., 2017; Davenport et al., 2012; Pigni et al., 2016).

The multiple regression results reported big data skillsets as a statistically significant predictor of EO. These findings are important as they provide further, and much needed, insights into which parts of big data are more significant and help direct focus of academics and business practitioners. Furthermore, these findings contribute to the EO literature (Wales et al., 2011) that aims to understand the antecedents of EO by identifying internal factors, such as big data skillsets, that exhibited a positive relationship with EO.

EO research that focuses on variables that are difficult to understand or measure can be limiting to business practitioners as finding the means to control or influence them is difficult. These findings answer the call of Wales et al. (2011) for EO research that focuses on issues of relevance for business practitioners as they focused on internal variables that are easily interpretable. Therefore, implementing strategies to increase the organisation’s big data skillsets or investment in toolsets can have an impact on the firm’s entrepreneurial activity and can be considered pragmatic solutions.

The second key finding identified in this research was between the big data capabilities and an organisation’s evidence-based decision-making culture. Large positive and statistically significant correlations were reported between all three of the big data capabilities and evidence-based decision-making. This implies that organisations with higher levels of big data skillsets, toolsets and datasets can lead to higher levels of data usage in the decision-making culture.

Moreover, this study found that approximately 43% of the variation found in evidence-based decision-making culture was explained by the big data capabilities and that both skillsets and toolsets were statistically significant, with positive beta coefficients, when used to predict evidence-based decision-making. These results are useful to business practitioners as an increase in skillsets and toolsets capabilities can result in higher levels of evidence-based decision-making.

These findings are important as data-driven decision-making has been associated with various positive outcomes (Brynjolfsson et al., 2011; LaValle et al., 2011) and understanding how to achieve higher levels of this style to decision-making is valuable. Furthermore, these findings confirm the current literature (Elgendy & Elragal, 2016; Popovic et al., 2012) that a relationship exists between an organisation's usage of data and their data-driven decision-making.

As per the research model depicted in Figure 1, this study also aimed to understand if evidence-based decision-making culture had a moderating effect on the relationship between the big data capabilities and EO. The moderated multiple regression results reported that a statistically significant interaction effect did exist. Furthermore, this improved the reported adjusted R-square from 17% to 20.9%. These results contribute to the current EO literature (Linton & Kask, 2017; Wales et al., 2011) on the moderating effect of the internal context of an organisation by presenting a new focus for EO research that focuses on big data capabilities within an organisation.

Finally, the EO construct was also determined to be a unidimensional construct, based on the findings from this study. This provides further insight into one of the main constructs of this study and implies that organisations that wish to be more innovative must also proactively seek out new opportunities and markets and be willing to take risks to extract value from them.

7.2 Implications for management

These findings suggest that business practitioners must recognize the big data phenomenon as a source of possible competitive advantage, given its link to EO and evidence-based decision-making. In an environment of intense competition, a firm's ability to innovate in a proactive manner and take calculated risks are necessary in order to remain relevant. Moreover, this study approached big data and strategic posture by assessing certain big data capabilities, EO and evidence-based decision-making. This approach allows for simple interpretation and pragmatic opportunities that can allow organisations to achieve their goals through more entrepreneurial behaviours, opposed to cost-cutting strategies.

These findings assist business practitioners by providing some direction as to which aspects of big data to place focus on. Organisations that wish to embark on big data initiatives must ensure that they are aware of the data currently available to them (datasets), that they have the relevant skills to extract, analyse and deliver information and key insights across the organisation (skillsets) and that the appropriate systems are in place to facilitate this movement of data internally (toolsets).

Skillsets was reported as a significant factor to both EO and evidence-based decision-making and is, therefore, identified as a critical factor of big data and competitive advantage in general. Business practitioners should, therefore, give consideration to the role played by internal functions such as human resources in achieving successful big data initiatives. Moreover, if not done so already, organisations should also give consideration to the inclusion of a chief data officer (CDO) or similar role within their senior management team that can assist in guiding the use of data, acquiring of further big data skillsets and implementation and management of the toolsets required.

Finally, this study found that on average respondents perceived their organisations as less entrepreneurial (based on EO). This finding was similar to previous studies (Lechner & Gudmundsson, 2014; Linton & Kask, 2017) and suggests that if organisations deem themselves as less entrepreneurial, then the importance of understanding big data skillsets and toolsets is even more prominent as there is room to grow with regards to their entrepreneurial efforts.

7.3 Limitations of the research

As stated in chapter 4, various limitations were observed during this study and are outlined below:

- The total sample size reported of N = 82 was identified as a limitation of this study. Larger samples are associated with more accurate results and the power of a test is considered to be dependent on the size of the sample (Pallant, 2007; Zikmund et al., 2009). Therefore, this may have skewed the results reported by this study.

- The use of snowball sampling may have also been a limitation of this study. Biernacki and Waldorf (1981) suggest that the use of snowball sampling may result in a sample with a low variance. The standard deviations reported in this study were considered relatively low, indicating low variance in the sample, which may have skewed the results reported.

7.4 Suggestions for future research

This study has identified a relationship between the various big data capabilities and EO, which to the researcher's knowledge has not been measured before. This calls for future research that aims to replicate and enhance the findings of this study. Skillsets was identified as a significant predictor and can possibly be explored further in relation to EO.

Whilst this study reported statistically significant correlations between the big data capabilities, evidence-based decision-making and EO, it is not clear from this study which of the big data capabilities precede the others. Further research into understanding if skillsets is a requirement for toolsets can assist organisations in understanding where to begin with regards to investing in big data capabilities.

The big data field is considered an early stage domain of research (Frizzo-Barker et al., 2016) and whilst the literature review revealed a host of definitions and means of explaining what big data is, a lack of a clear and simple definition exists. The author of this study agrees with George et al. (2014) that more published management scholarship and empirically oriented work that aims to clarify the big data phenomenon is required.

The research design of this study was cross-sectional in nature, therefore, the various big data capabilities, evidence-based decision-making and EO were measured at a point in time. This provides an opportunity to evaluate these constructs using a longitudinal study in order to understand how the level of evidence-based decision-making culture or EO could change over time.

7.5 Conclusion

As organisations find themselves in uncertain environments with increased competition and pressure from boards to perform, the search for innovative new products or ways of serving markets is prevalent in the current global context. Whilst some organisations have managed to create value from big data, others continue to struggle.

The aim of this research was to understand the relationships identified in the research model in Figure 1 in order to provide insights as to how organisations can leverage big data. Therefore, this study has been successful in elucidating the various relationships between big data, evidence-based decision-making and EO and providing further insights into which aspects of big data capabilities to focus on.

Furthermore, given the proliferation of EO research and its link to firm performance, the relationship observed by this study between big data and EO suggests an interesting relationship that calls for further research that aims to explain and empirically measure this relationship further.

References

Advanced Performance Institute. (2017). How is Big Data Used in Practice? 10 Use Cases Everyone Must Read. Retrieved from: https://www.ap-institute.com/big-data-articles/how-is-big-data-used-in-practice-10-use-cases-everyone-should-read

Akbay, S. (2015). How big data applications are revolutionizing decision making. *Business Intelligence Journal, 20*(1), 25-29.

Alharthi, A., Krotov, V., & Bowman, M. (2017). Addressing barriers to big data. *Business Horizons,*

Amankwah-Amoah, J. (2016). Emerging economies, emerging challenges: Mobilising and capturing value from big data. *Technological Forecasting and Social Change, 110*, 167-174.

Andrews, K. R. (1987). *The Concept of Corporate Strategy*. New York, NY: Richard D. Irwin, Inc.

Atzori, L., Iera, A., & Morabito, G. (2010). The internet of things: A survey. *Computer Networks, 54*(15), 2787-2805.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17*(1), 99-120.

Barney, J., Wright, M., & Ketchen Jr, D. J. (2001). The resource-based view of the firm: Ten years after 1991. *Journal of Management, 27*(6), 625-641.

Barringer, B. R., & Bluedorn, A. C. (1999). The relationship between corporate entrepreneurship and strategic management. *Strategic Management Journal*, 421-444.

Bean, R. (2017). How companies say they're using big data. Retreived from https://hbr.org/2017/04/how-companies-say-theyre-using-big-data

Biernacki, P., & Waldorf, D. (1981). Snowball sampling: Problems and techniques of chain referral sampling. *Sociological Methods & Research, 10*(2), 141-163.

Bøe-Lillegraven, T. (2014). Untangling the ambidexterity dilemma through big data analytics. *Browser Download this Paper,*

Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011). Strength in numbers: How does data-driven decisionmaking affect firm performance? *Available at SSRN 1819486,*

Cao, G., & Duan, Y. (2014). Gaining competitive advantage from analytics through the mediation of decision-making effectiveness: An empirical study of UK manufacturing companies. *Pacis,* 377.

Center for Evidence-Based Management. (2013). Evidence-Based Management Assessment for Organizations. Retrieved from https://www.cebma.org/wp-content/uploads/EBMgt-Assessment-for-Organizations-vs-May-2013.pdf

Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly, 36*(4), 1165-1188.

Covin, J. G., Green, K. M., & Slevin, D. P. (2006). Strategic process effects on the entrepreneurial orientation–sales growth rate relationship. *Entrepreneurship theory and practice*, *30*(1), 57-81.

Covin, J. G., & Slevin, D. P. (1989). Strategic management of small firms in hostile and benign environments. *Strategic management journal*, *10*(1), 75-87.

Covin, J. G., & Slevin, D. P. (1991). A conceptual model of entrepreneurship as firm behavior. *Entrepreneurship: Critical perspectives on business and management*, *3*, 5-28.

Cummings, T. G., & Worley, C. G. (2015). *Organisation development & change* (Tenth Edition). Canada: Nelson Education Ltd.

Datameer. (2016). Top Five High-Impact Use Cases for Big Data Analytics. Retrieved from: https://www.datameer.com/pdf/eBook-Top-Five-High-Impact-UseCases-for-Big-Data-Analytics.pdf

Davenport, T. H., Barth, P., & Bean, R. (2012). How big data is different. *MIT Sloan Management Review, 54*(1), 43.

Davenport, T. H., & Patil, D. J. (2012). Data scientist. *Harvard business review*, *90*(5), 70-76.

Douglas, M. (2013). Big data raises big questions. *Government Technology, 26*, 12-16.

Elgendy, N., & Elragal, A. (2016). Big Data Analytics in Support of the Decision Making Process. *Procedia Computer Science*, *100*, 1071-1084.

Ernst & Young. (2012). Anti-bribery and corruption analytics. Retrieved from: http://www.ey.com/Publication/vwLUAssets/Anti-bribery_and_corruption_analytics/$FILE/Anti-bribery-and-corruption-brochure.pdf

Frese, M., Bausch, A., Schmidt, P., Rauch, A., & Kabst, R. (2012). Evidence-based entrepreneurship: Cumulative science, action principles, and bridging the gap between science and practice. *Foundations and Trends® in Entrepreneurship*, *8*(1), 1-62.

Frizzo-Barker, J., Chow-White, P. A., Mozafari, M., & Ha, D. (2016). An empirical study of the rise of big data in business scholarship. *International Journal of Information Management, 36*(3), 403-413.

George, G., Haas, M. R., & Pentland, A. (2014). Big data and management. *Academy of Management Journal, 57*(2), 321-326.

George, G., Osinga, E. C., Lavie, D., & Scott, B. A. (2016). Big data and data science methods for management research. *Academy of Management Journal*, *59*(5), 1493-1507.

Ghemawat, P. (2002). Competition and business strategy in historical perspective. *Business History Review, 76*(01), 37-74.

Gliem, R. R., & Gliem, J. A. (2003). Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for likert-type scales.

Gobble, M. M. (2013). Big data: The next big thing in innovation. *Research Technology Management, 56*(1), 64-66. doi:10.5437/08956308x5601005

Goes, P. (2014). Big data and IS research.[article]. *MIS Quarterly, 38*(3), iii-viii.

Gray, D. E. (2013). *Doing research in the real world* Sage.

Gregory, B. T., Harris, S. G., Armenakis, A. A., & Shook, C. L. (2009). Organisational culture and effectiveness: A study of values, attitudes, and organisational outcomes. *Journal of Business Research, 62*(7), 673-679.

Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems, 29*(7), 1645-1660.

Gupta, V., & Gupta, A. (2015). The concept of entrepreneurial orientation. *Foundations and Trends® in Entrepreneurship*, *11*(2), 55-137.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). Multivariate Data Analysis.

Hagel, J. (2016). We Need to Expand Our Definition of Entrepreneurship. Retrieved from: https://hbr.org/2016/09/we-need-to-expand-our-definition-of-entrepreneurship

Hill, L., Davis, G. (2017). The Board's New Innovation Imperative. Retrieved from:

https://hbr.org/2017/11/the-boards-new-innovation-imperative

Hofstede, G. (1980). Culture and organisations. *International Studies of Management & Organisation, 10*(4), 15-41.

House, R., Javidan, M., Hanges, P., & Dorfman, P. (2002). Understanding cultures and implicit leadership theories across the globe: An introduction to project GLOBE. *Journal of World Business, 37*(1), 3-10.

Ingram, M. (2012). Airbnb, Coursera, and Uber The Rise of the Disruption Economy. Retrieved from https://www.bloomberg.com/news/articles/2012-10-25/airbnb-coursera-and-uber-the-rise-of-the-disruption-economy

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT Sloan Management Review, 52*(2), 21.

Lechner, C., & Gudmundsson, S. V. (2014). Entrepreneurial orientation, firm strategy and small firm performance. *International Small Business Journal*, *32*(1), 36-60.

Linton, G., & Kask, J. (2017). Configurations of entrepreneurial orientation and competitive strategy for high performance. *Journal of Business Research*, *70*, 168-176.

Lumpkin, G. T., & Dess, G. G. (1996). Clarifying the entrepreneurial orientation construct and linking it to performance. *Academy of management Review*, *21*(1), 135-172.

Lyon, D. W., Lumpkin, G. T., & Dess, G. G. (2000). Enhancing entrepreneurial orientation research: Operationalizing and measuring a key strategic decision making process. *Journal of management*, *26*(5), 1055-1085.

Madsen, E. L. (2007). The significance of sustained entrepreneurial orientation on performance of firms–A longitudinal analysis. *Entrepreneurship and Regional Development*, *19*(2), 185-204.

Malakooti, B. (2012). Decision making process: Typology, intelligence, and optimization. *Journal of Intelligent Manufacturing, 23*(3), 733-746.

Mathews, J. (2016). An Information Processing View of Competition Analysis. *IUP Journal of Business Strategy*, *13*(1), 7.

Mazzei, M. J., & Noble, D. (2017). Big data dreams: A framework for corporate strategy. *Business Horizons,*

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D., & Barton, D. (2012). Big data. *The Management Revolution.Harvard Bus Rev, 90*(10), 61-67.

Miller, D. (1983). The correlates of entrepreneurship in three types of firms. *Management science*, *29*(7), 770-791.

Mills, R. J., Chudoba, K. M., & Olsen, D. H. (2016). IS Programs Responding to Industry Demands for Data Scientists: A Comparison between 2011-2016. *Journal of Information Systems Education*, *27*(2).

Miller, D., & Friesen, P. H. (1982). Innovation in conservative and entrepreneurial firms: Two models of strategic momentum. *Strategic management journal*, *3*(1), 1-25.

Mintzberg, H. (1973). Strategy-making in three modes. *California Management Review, 16*(2), 44-53.

Mintzberg, H. (1978). Patterns in strategy formation. *Management Science, 24*(9), 934-948.

Mintzberg, H., Ahlstrand, B., & Lampel, J. (1998). *Strategy Safari: A Guided Tour Through The Wilds Of Strategic Management.* New York: Free Press.

Morrison, J. M., Brown, C. J., & Smit, E. (2006). A supportive organisational culture for project management in matrix organisations: A theoretical perspective. *South African Journal of Business Management, 37*(4), 39-54.

Naldi, L., Nordqvist, M., Sjöberg, K., & Wiklund, J. (2007). Entrepreneurial orientation, risk taking, and performance in family firms. *Family business review*, *20*(1), 33-47.

Pallant, J. F. (2007). SPSS survival manual: A step-by-step guide to data analysis with SPSS.

Pfeffer, J., & Sutton, R. I. (2006). Evidence-based management. *Harvard Business Review, 84*(1), 62.

Pigni, F., Piccoli, G., & Watson, R. (2016). Digital data streams. *California Management Review, 58*(3), 5-25.

Porter, M. E., & Porter, M. E. (1979). How competitive forces shape strategy.

Popovič, A., Hackney, R., Coelho, P. S., & Jaklič, J. (2012). Towards business intelligence systems success: Effects of maturity and culture on analytical decision making. *Decision Support Systems*, *54*(1), 729-739.

Prajogo, D. I. (2016). The strategic fit between innovation strategies and business environment in delivering business performance. *International Journal of Production Economics*, *171*, 241-249.

PricewaterhouseCoopers (2015). Seizing the information advantage. Retrieved from https://www.pwc.es/es/publicaciones/tecnologia/assets/Seizing-The-Information-Advantage.pdf

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data, 1*(1), 51-59.

Purnama, C., & Subroto, W. T. (2016). Competition Intensity, Uncertainty Environmental on the Use of Information Technology and Its Impact on Business Performance Small and Medium Enterprises (SMEs). *International Review of Management and Marketing*, *6*(4).

Ransbotham, S., Fichman, R. G., Gopal, R., & Gupta, A. (2016). Special section Introduction—Ubiquitous IT and digital vulnerabilities. *Information Systems Research, 27*(4), 834-847.

Rauch, A., Wiklund, J., Lumpkin, G. T., & Frese, M. (2009). Entrepreneurial orientation and business performance: An assessment of past research and suggestions for the future. *Entrepreneurship theory and practice*, *33*(3), 761-787.

Reeves, M., Deimler,. M. (2011). Adaptability: The New Competitive Advantage. Retrieved from: https://hbr.org/2011/07/adaptability-the-new-competitive-advantage

Rousseau, D. M. (2006). Is there such a thing as "evidence-based management"? *Academy of Management Review, 31*(2), 256-269.

Rygielski, C., Wang, J. C., & Yen, D. C. (2002). Data mining techniques for customer relationship management. *Technology in society*, *24*(4), 483-502.

Saeed, S., Yousafzai, S. Y., & Engelen, A. (2014). On cultural and macroeconomic contingencies of the entrepreneurial orientation–performance relationship. *Entrepreneurship Theory and Practice*, *38*(2), 255-290.

Saunders, M. N., & Lewis, P. (2014). *Doing research in business and management: An essential guide to planning your project* Pearson Higher Ed.

Schilke, O. (2014). On the contingent value of dynamic capabilities for competitive advantage: The nonlinear moderating effect of environmental dynamism. *Strategic Management Journal*, *35*(2), 179-203.

Shah, S., Horne, A., & Capellá, J. (2012). Good data won't guarantee good decisions. *Harvard Business Review, 90*(4)

Slåtten, T., & Mehmetoglu, M. (2011). What are the drivers for innovative behavior in frontline jobs? A study of the hospitality industry in Norway. *Journal of Human Resources in Hospitality & Tourism*, *10*(3), 254-272.

Slevin, D. P., & Terjesen, S. A. (2011). Entrepreneurial orientation: Reviewing three papers and implications for further theoretical and methodological development. *Entrepreneurship Theory and Practice*, *35*(5), 973-987.

Sousa, C. M., & Coelho, F. (2011). From personal values to creativity: evidence from frontline service employees. *European Journal of Marketing*, *45*(7/8), 1029-1050.

Teece, D. J. (2010). Business models, business strategy and innovation. *Long range planning*, *43*(2), 172-194.

Thomas, K. W., & Kilmann, R. H. (1975). The social desirability variable in organisational research: An alternative explanation for reported findings. *Academy of Management Journal, 18*(4), 741-752.

Thompson, L. (2011). What NASA risks by betting on Elon Musk's SpaceX. Retrieved from https://www.forbes.com/sites/beltway/2011/05/23/what-nasa-risks-by-betting-on-elon-musks-spacex/#51d5a7d26eb2

Wales, W. J., Gupta, V. K., & Mousa, F. T. (2013). Empirical research on entrepreneurial orientation: An assessment and suggestions for future research. *International Small Business Journal*, *31*(4), 357-383.

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. *Journal of Business Logistics*, *34*(2), 77-84.

Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal, 5*(2), 171-180.

Yesil, S., & Kaya, A. (2013). The effect of organisational culture on firm financial performance: Evidence from a developing country. *Procedia-Social and Behavioral Sciences, 81*, 428-437.

Zahra, S. A. (1993). A conceptual model of entrepreneurship as firm behavior: A critique and extension. *Entrepreneurship: Theory and practice*, *17*(4), 5-22.

Zikmund, W., Babin, B., Carr, J., & Griffin, M. (2009). *Business research methods* (8th ed). Stamford, CT: Cengage Learning.

Appendices

Appendix 1: Research questionnaire

D	Demographic information	Response options
D1	Age	Open response
D2	Gender	Open response
D3	Job title	Open response
D4	Number of employees in your organisation?	Open response
D5	Position in organisation?	• Junior management • Middle management • Senior management • Executive management
D6	Tenure at current organisation?	• Less than 2 years • 3 – 5 years • 6 – 8 years • 9 – 11 years • 11 or more years
D7	Industry in which your organisation operates?	• Aerospace • Agriculture • Automotive and transportation • Chemical • Communication • Construction • Electrical equipment • Electricity • Energy • Financial services • Food products • Gas and water supply • Information technology • Machinery • Manufacturing

		• Mechanical • Media • Mining • Pharmaceuticals • Printing / paper • Steel and non-ferrous metals • Textile • Wholesale and retail • Other

E	Evidence-based decision-making	Question type	Reference
E1	We use data to create new products or services.	Likert Scale (1 – 5)	Brynjolfsson et al. (2011)
E2	We depend on data to support decision-making (work practices and environment of the entire company).	Likert Scale (1 – 5)	Brynjolfsson et al. (2011)
E3	We believe it is important to adopt new and cutting-edge practices.	Likert Scale (1 – 7)	Center for Evidence-Based Management, 2013)
E4	We make decisions by looking at what other organisations are doing, and how it's working for them.	Likert Scale (1 – 7)	Center for Evidence-Based Management, 2013)
E5	We use benchmarking to identify best practices used in other organisations to help improve our organisation.	Likert Scale (1 – 7)	Center for Evidence-Based Management, 2013)
E6	We use consultants to help us make decisions.	Likert Scale (1 – 7)	Center for Evidence-Based

			Management, 2013)
E7	Before any decision is taken we systematically evaluate internal data to better understand the nature of the problem.	Likert Scale (1 – 7)	Center for Evidence-Based Management, 2013)
E8	Our managers have access to a management information system.	Likert Scale (1 – 7)	Center for Evidence-Based Management, 2013)
E9	A data set, randomly pulled from the management information system, will be accurate and can be trusted.	Likert Scale (1 – 7)	Center for Evidence-Based Management, 2013)
E10	We use evidence from academic research to help us make decisions about how to solve our problems.	Likert Scale (1 – 7)	Center for Evidence-Based Management, 2013)
E11	Our managers know how to use the Internet to search for scientific evidence to guide their decisions.	Likert Scale (1 – 7)	Center for Evidence-Based Management, 2013)
E12	Our managers know how to critically appraise both internal data and evidence from scientific research.	Likert Scale (1 – 7)	Center for Evidence-Based Management, 2013)
E13	Managers in my organisation tend to believe that the organisation is unique and hence the outcome of scientific research is not applicable.	Likert Scale (1 – 7)	Center for Evidence-Based Management,

			2013)
E14	Managers in our organisation tend to believe that experience and knowledge gained on the job is the only important source of information when considering how to tackle a problem.	Likert Scale (1 – 7)	Center for Evidence-Based Management, 2013)
E15	Internal politics and power struggles influence the way we make decisions about policies and practices.	Likert Scale (1 – 7)	Center for Evidence-Based Management, 2013)
E16	We spend time identifying and exploring a range of possible solutions to the problems we face.	Likert Scale (1 – 7)	Center for Evidence-Based Management, 2013)
E17	We systematically evaluate the effectiveness of new policies and practices we introduce.	Likert Scale (1 – 7)	Center for Evidence-Based Management, 2013)
E18	If we make mistakes in our decision-making we try to learn from them.	Likert Scale (1 – 7)	Center for Evidence-Based Management, 2013)
S	Big data capabilities: Skillsets	Question type	Reference
S1	Data mining.	Likert Scale (1 – 7)	Popovic et al. (2012).
S2	On-line analytical processing (OLAP).	Likert Scale (1 – 7)	Popovic et al. (2012).
S3	Interactive reports (Ad-hoc).	Likert Scale (1 – 7)	Popovic et al. (2012).
S4	Analytical applications, including Trend analysis, “What-if” scenarios.	Likert Scale (1 – 7)	Popovic et al. (2012).
Q5	Dashboards, including metrics, key	Likert Scale	Popovic et al.

	performance indicators (KPI), alerts.	(1 – 7)	(2012).
T	Big data capabilities: Toolsets	Question type	Reference
T1	The information is precise and close enough to reality.	Likert Scale (1 – 7)	Popovic et al. (2012).
T2	The information is easily understandable by the target group.	Likert Scale (1 – 7)	Popovic et al. (2012).
T3	The information is to the point, void of unnecessary elements.	Likert Scale (1 – 7)	Popovic et al. (2012).
T4	The provision of information corresponds to users' needs and habits.	Likert Scale (1 – 7)	Popovic et al. (2012).
T5	The information is processed and delivered rapidly without delay.	Likert Scale (1 – 7)	Popovic et al. (2012).
DS	Big data capabilities: Datasets	Question type	Reference
DS1	Statement A: Data are scattered everywhere – on the mainframe, in databases, in spreadsheets, in flat files, in Enterprise Resource Planning (ERP) applications. Statement B: Data are completely integrated, enabling real-time reporting and analysis.	Likert Scale (1 – 7)	Popovic et al. (2012).
DS2	Statement A: Data in the sources are mutually inconsistent. Statement B: Data in the sources are mutually consistent.	Likert Scale (1 – 7)	Popovic et al. (2012).
DS3	The scope of information is adequate (neither too much nor too little).	Likert Scale (1 – 7)	Popovic et al. (2012).
DS4	The information is up-to-date and not obsolete.	Likert Scale (1 – 7)	Popovic et al. (2012).
IV	EO: Innovativeness	Question type	Reference
IV1	Generally, the top managers of my firm favour... Statement A: A strong emphasis on the marketing of tried and true products and services. Statement B: A strong emphasis on R&D, technological leadership and innovation.	Likert Scale (1 – 7)	Barringer and Bluedorn (1999).

IV2	How many new lines of products or services has your firm marketed in the past 5 years? Statement A: No new lines of products or services. Statement B: Many new lines of products or services.	Likert Scale (1 – 7)	Barringer and Bluedorn (1999).
IV3	How many new lines of products or services has your firm marketed in the past 5 years? Statement A: Changes in product or service lines have been mostly of a minor nature Statement B: Changes in product or service lines have usually been quite dramatic	Likert Scale (1 – 7)	Barringer and Bluedorn (1999).
PA	EO: Proactiveness	Question type	Reference
PA1	In dealing with it's competitors, my firm... Statement A: Typically responds to actions which competitors initiate. Statement B: Typically initiates actions to which competitors then respond.	Likert Scale (1 – 7)	Barringer and Bluedorn (1999).
PA2	In dealing with it's competitors, my firm... Statement A: Is very seldom the first firm to introduce new products/services, operating technologies, etc. Statement B: Is very often the first firm to introduce new products/services, operating technologies, etc.	Likert Scale (1 – 7)	Barringer and Bluedorn (1999).
PA3	In dealing with it's competitors, my firm... Statement A: Typically seeks to avoid competitive clashes, preferring a "live-and- let-live" posture. Statement B: Typically adopts a very competitive, "undo-the-competitor" posture.	Likert Scale (1 – 7)	Barringer and Bluedorn (1999).
RT	EO: Risk-taking	Question type	Reference

RT1	Generally, the top managers of my firm favour... Statement A: Low-risk projects with normal and certain rates of return Statement B: High-risk projects with changes of very high returns.	Likert Scale (1 – 7)	Barringer and Bluedorn (1999).
RT2	Generally, the top managers of my firm favour... Statement A: A cautious, “wait and see” posture in order to minimize the probability of making costly decisions when faced with uncertainty. Statement B: A bold, aggressive posture in order to maximize the probability of exploiting potential when faced with uncertainty.	Likert Scale (1 – 7)	Barringer and Bluedorn (1999).
RT3	Generally, the top managers of my firm believe that... Statement A: Owing to the nature of the environment, it is best to explore gradually via cautious behaviour. Statement B: Owing to the nature of the environment, bold, wide-ranging acts are necessary to achieve the firm’s objectives.	Likert Scale (1 – 7)	Barringer and Bluedorn (1999).

Appendix 2: Correlation matrix (PCA)

E1 - 9	E1	E2	E3	E4	E5	E6	E7	E8	E9
E1	1								
E2	**.541**	1							
E3	**.428**	.264	1						
E4	.262	.273	**.370**	1					
E5	.257	**.300**	**.408**	**.397**	1				
E6	-.029	.028	-.052	**.327**	.146	1			
E7	**.352**	**.390**	**.404**	.112	**.377**	.065	1		

E8	**.334**	**.421**	**.322**	.270	.188	**.361**	**.479**	1	
E9	.190	.306	.091	.215	.078	.163	**.356**	**.626**	1
E10	.250	.165	.276	.088	.057	-.034	.265	**.438**	.219
E11	.166	.247	**.315**	.279	.179	.122	**.396**	**.353**	**.367**
E12	.292	**.357**	**.378**	**.300**	.203	.014	**.489**	**.460**	**.499**
E13	-.267	**-.315**	**-.315**	-.203	-.243	-.163	-.182	**-.465**	**-.339**
E14	-.247	-.243	-.171	-.172	-.089	.003	-.136	**-.302**	-.297
E15	-.234	-.050	-.244	.020	-.124	.204	-.288	-.120	.-.039
E16	.210	.130	**.326**	.182	**.344**	.254	**.558**	**.357**	.203
E17	**.310**	.285	**.514**	.180	**.333**	-.050	**.617**	**.430**	.225
E18	**.325**	.223	**.381**	.138	.175	-.033	**.470**	**.435**	.299
E10 - 18	E10	E11	E12	E13	E14	E15	E16	E17	E18
E10	1								
E11	**.506**	1							
E12	**.530**	**.821**	1						
E13	-.206	**-.315**	-.279	1					
E14	-.365	-.224	-.251	**.482**	1				
E15	**-.406**	**-.394**	**-.403**	.147	.276	1			
E16	.226	**.381**	.297	-.219	-.166	-.203	1		
E17	**.413**	**.513**	**.550**	-.289	-.246	**-.474**	**.590**	1	
E18	**.430**	**.592**	**.589**	-.198	-.216	**-.415**	**.468**	**.565**	1

Appendix 3: KMO measures taken from anti-image correlation matrix

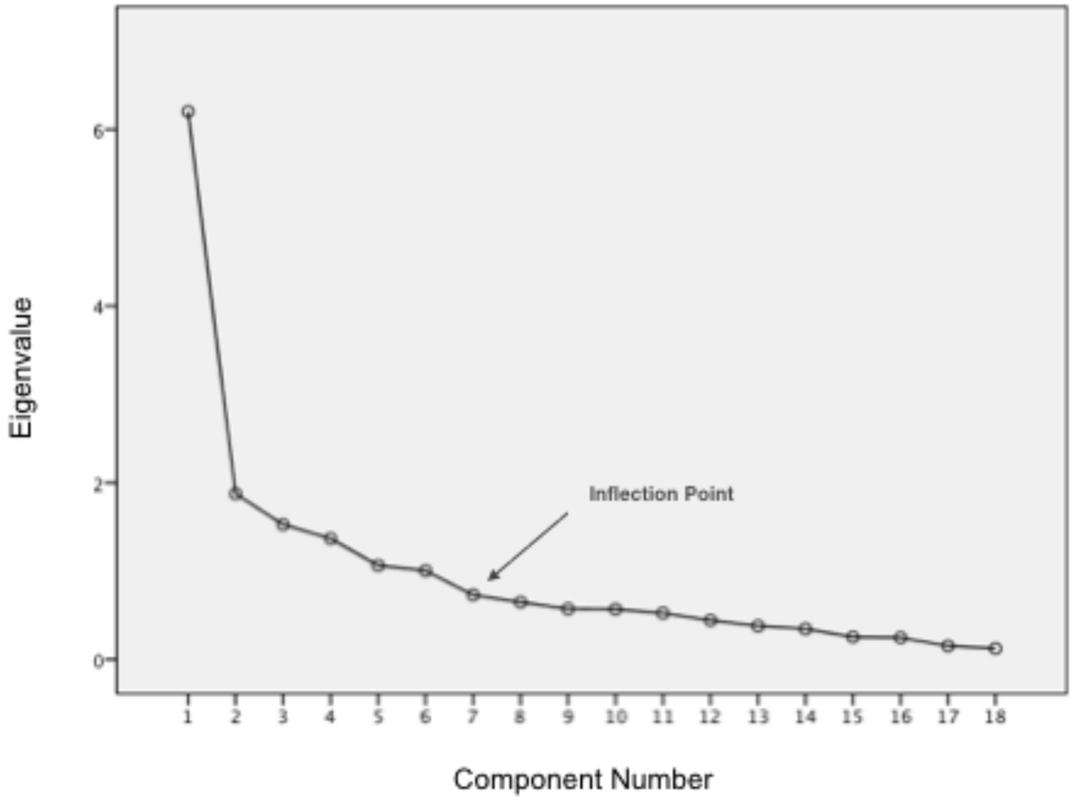
Variable	KMO Measure	Variable	KMO Measure
E1	.791	E10	.782
E2	.758	E11	.767
E3	.835	E12	.832
E4	.715	E13	.737
E5	.811	E14	.736
E6	.417	E15	.835
E7	.866	E16	.797
E8	.723	E17	.887
E9	.735	E18	.894

Note: E1 – 18 refers to the individual attributes of evidence-based decision-making construct as per Appendix 1.

Appendix 4: Total variance explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Component	Total	% of Var.	Cum. %	Total	% of Var.	Cum. %
1	6.202	34.457	34.457	6.202	34.457	34.457
2	1.874	10.409	44.866	1.874	10.409	44.866
3	1.525	8.473	53.340	1.525	8.473	53.340
4	1.368	7.600	60.939	1.368	7.600	60.939
5	1.062	5.901	66.840	1.062	5.901	66.840
6	1.003	5.574	72.414	1.003	5.574	72.414
7	.728	4.046	76.460			
8	.646	3.591	80.051			
9	.570	3.164	83.215			
10	.567	3.147	86.362			
11	.522	2.899	89.261			
12	.440	2.446	91.707			
13	.377	2.096	93.803			
14	.344	1.913	95.716			
15	.253	1.404	97.120			
16	.245	1.361	98.481			
17	.151	.841	99.322			
18	.122	.678	100.000			

Appendix 5: Scree plot



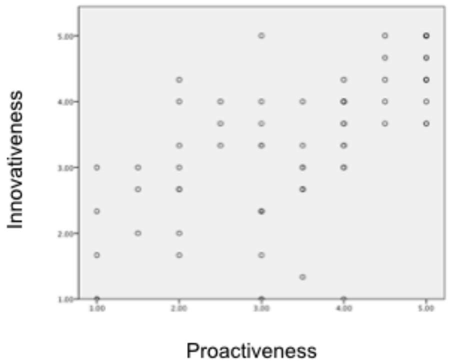
Appendix 6: Rotated structure matrix for PCA with Varimax rotation

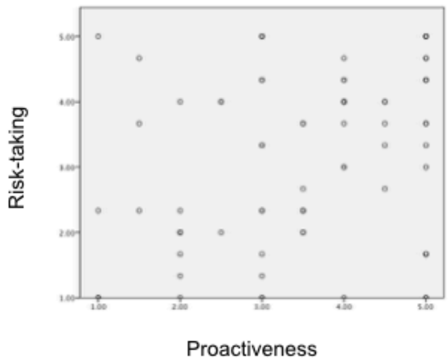
Items	Comp. 1	Comp. 2	Comp. 3	Comp. 4	Comp. 5	Comp. 6
E12	**.860**					
E11	**.852**					
E10	**.670**			.333		
E18	**.635**	.454				
E15	**-.509**					.472
E16		**.820**				
E7		**.733**	.399			
E17	.450	**.675**				
E2			**.819**			
E1			**.697**			
E14				**-.842**		
E13				**-.747**		
E4					**.780**	.332
E5		.388			**.674**	

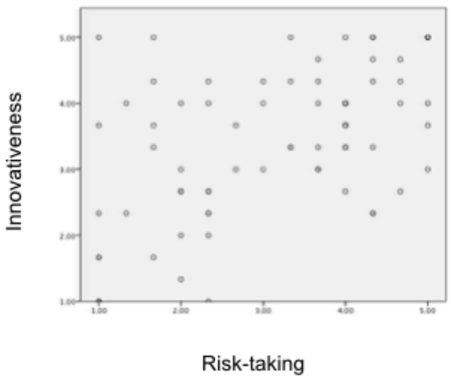
E3		.341			**.566**	
E6						**.789**
E9	.393		.433			**.536**
E8	.333	.330	.420	.339		**.506**

Note: E1 – 18 refers to individual attributes of evidence-based decision-making construct.

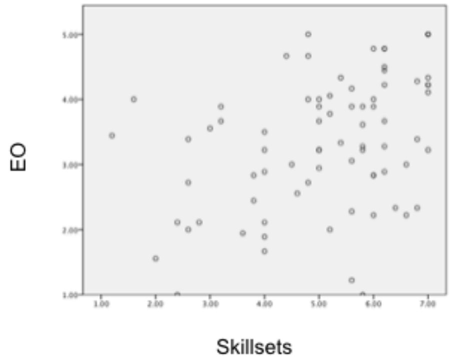
Appendix 7: Scatterplots of EO sub-dimensions

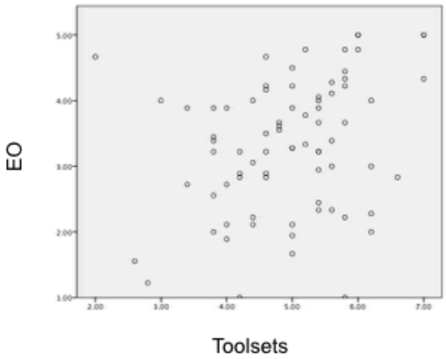


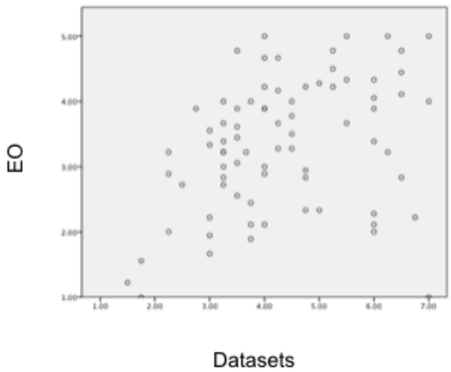




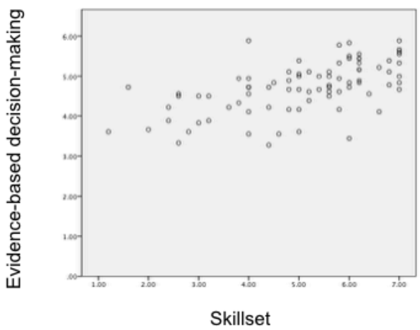
Appendix 8: Scatterplots of EO and big data capabilities

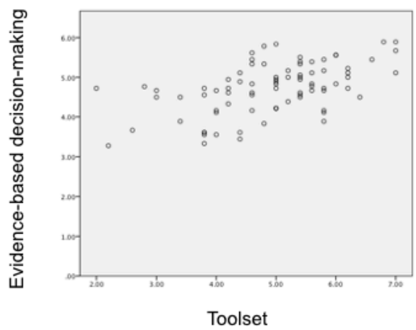


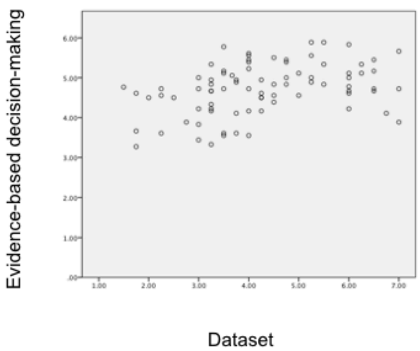




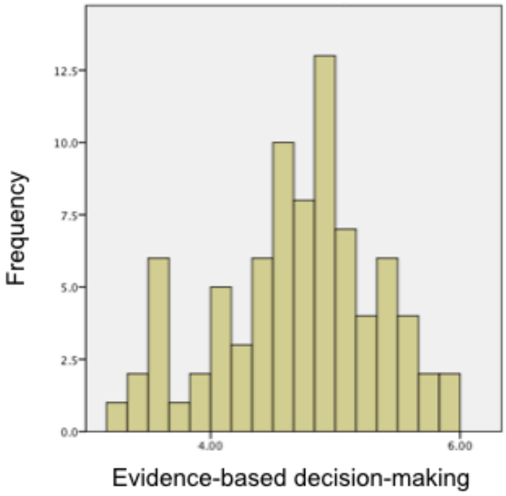
Appendix 9: Scatterplots of evidence-based decision-making and big data capabilities

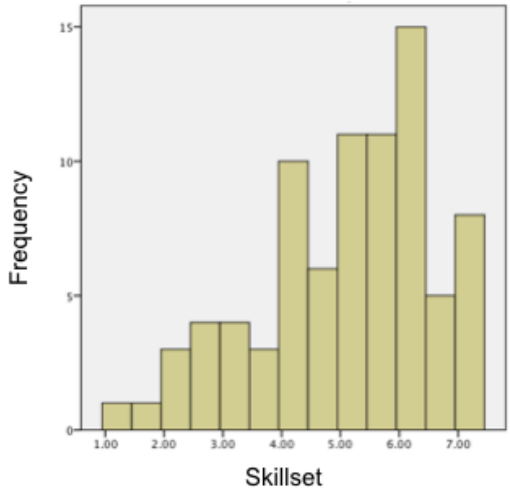


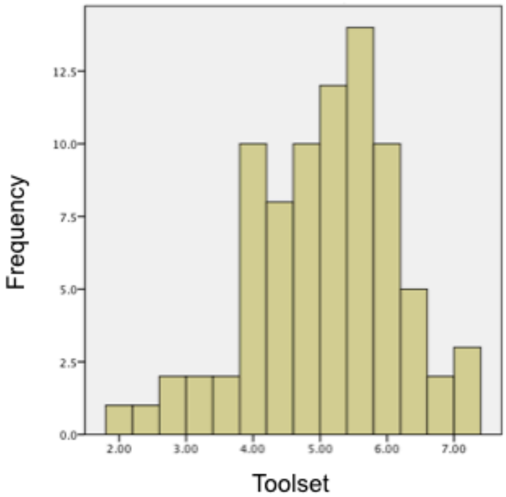


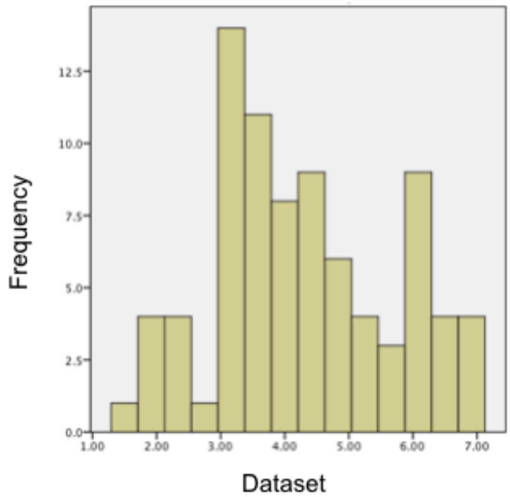


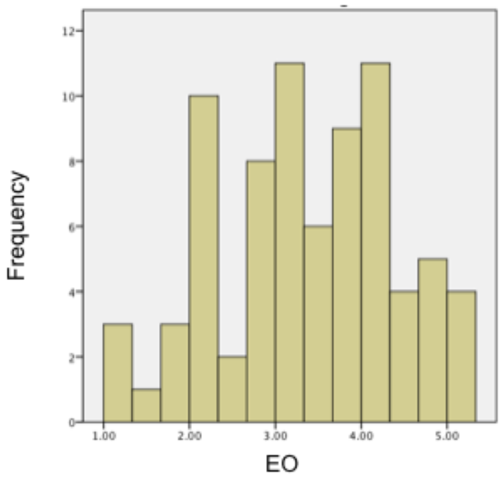
Appendix 10: Histogram - test for normality



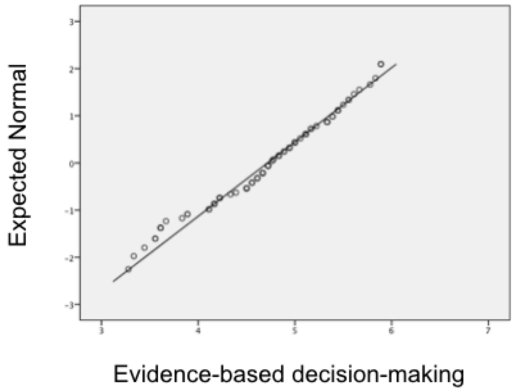


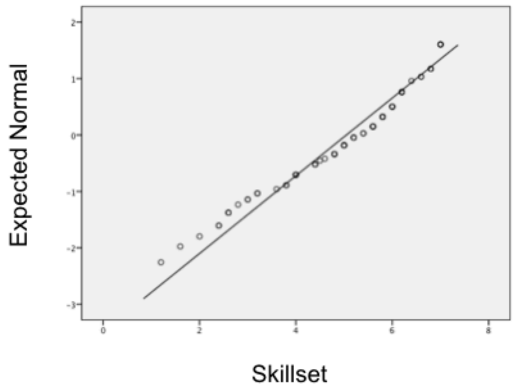


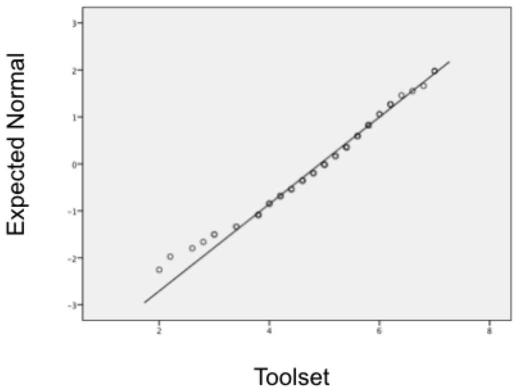


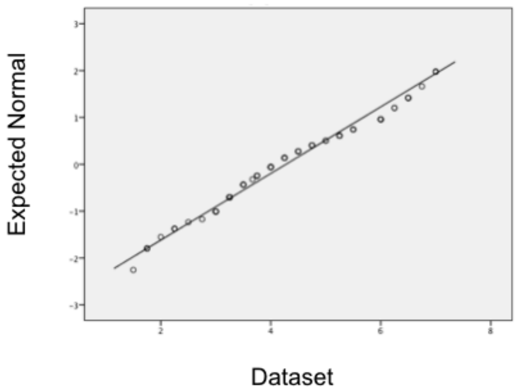


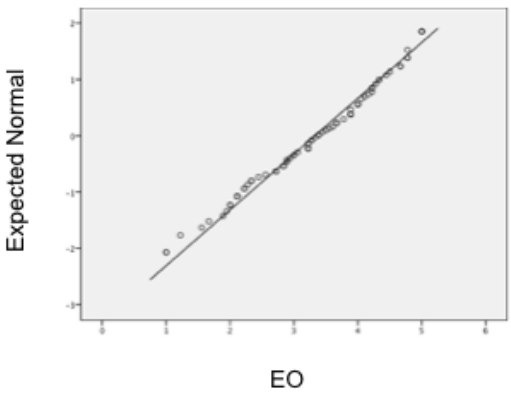
Appendix 11: Q-Q plots - test for normality



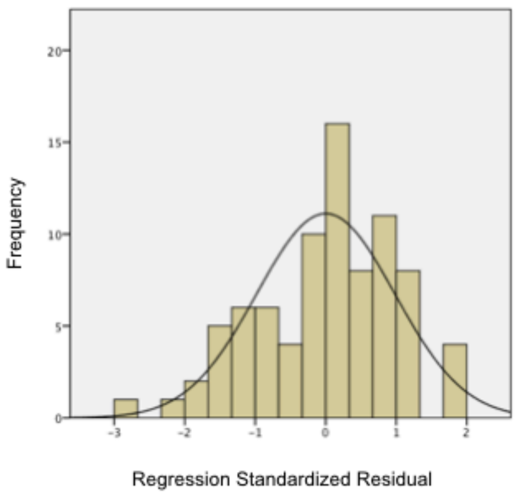


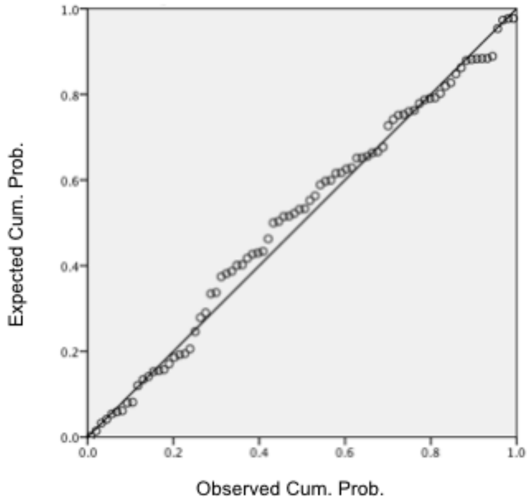




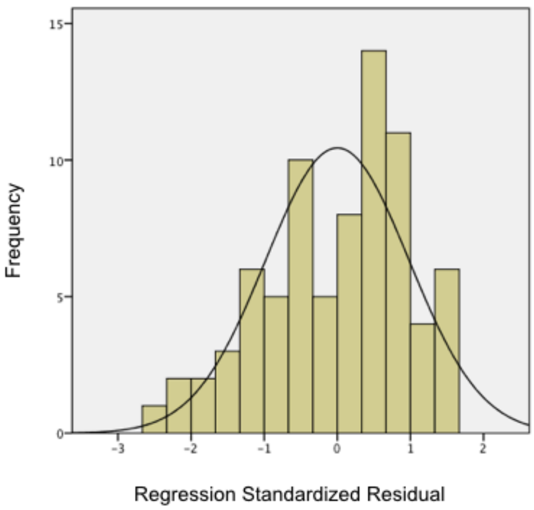


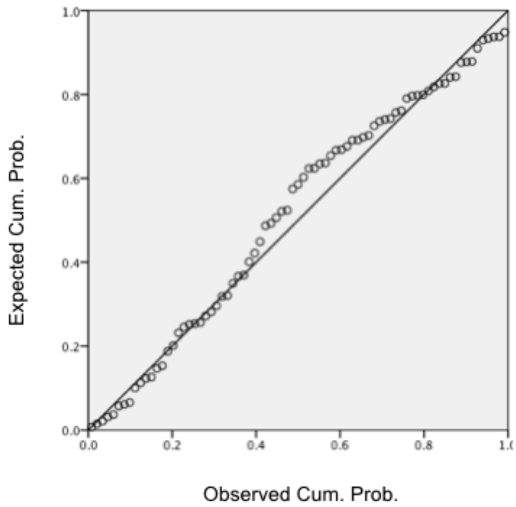
Appendix 12: Histogram and normal P-P plot of regression standardised residual or evidence-based decision-making



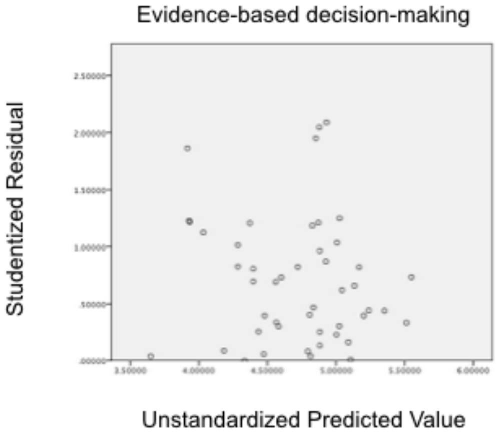


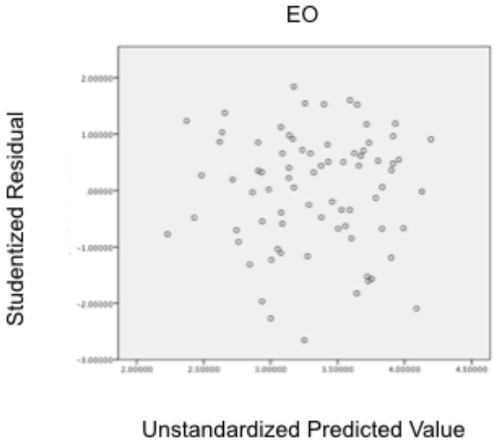
Appendix 13: Histogram and normal P-P plot of regression standardised residual for EO





Appendix 14: Plot of studentized residuals and unstandardised predicted values for evidence-based decision-making and EO





Appendix 15: Ethical clearance approval letter

Gordon Institute of Business Science
University of Pretoria

13 July 2017

Frank Mourinho

Dear Frank,

Please be advised that your application for Ethical Clearance has been approved.

You are therefore allowed to continue collecting your data.

We wish you everything of the best for the rest of the project.

Kind Regards

GIBS MBA Research Ethical Clearance Committee

Gordon Institute of Business Science
Reg. No. 99/19816/08

26 Melville Road, Illovo, Johannesburg
PO Box 787602, Sandton, 2146, South Africa

telephone **(+27) 11 771 4000**
fax **(+27) 11 771 4177**

website **gibs.co.za**
University of Pretoria