Gordon Institute of Business Science
University of Pretoria

Big data: toward the influence of organisation culture and artificial intelligence on firm performance

Tiveshen Chetty

Student Number: 18378022

Tiveshenchetty@gmail.com

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

8th November 2019

Abstract

Big Data has become pervasive in the business environment, as the datafication of the world continues through the deployment of the internet of things. In the race to achieve data supremacy, organisations are making large investments into Big Data technologies. These projects, however, are proving difficult to implement, with substantive value in the form of return on investment, evading a large proportion of organisations. To alleviate the disconnect between Big Data implementation and improved FPer, research has highlighted the involvement of OC as a key enabler of improved FPer. Furthermore, as Big Data has evolved to the status of a factor of production, the next frontier of technology has arrived in the form of AI, allowing these vast repositories of data to be analysed for deeper insights than was previously possible. Thus, there exists a new entanglement of relationships between Big Data Analytics Capabilities, the OC allowing for these technologies to be leveraged correctly, and the fringe data science technology of AI. This study seeks to delve deeper into these relationships to understand their effect on FPer, and ultimately how organisations can best utilise them to create a sustained competitive advantage

Key Words

Big Data, Big Data Analytics Capability, AI, Organisational Culture, FPer

Key Words

Abbreviation	Meaning
BDA	Big Data Analytics
RBV	Resource Based View
OC	Organisational Culture
FPer	Firm Performance
AI	Artificial Intelligence
BDAC	Big Data Analytics Capability
CDEV	Cultural Development
EINV	Employee Investment
BDA	Big Data Analytics
BDAIF	BDA Infrastructure Flexibility
BDAMC	BDA Management Capability
BDAPE	BDA Personnel Expertise
HMR	Hierarchical multiple regression
KMO	Kaiser-Mayer-Olkin index
RMSEA	Root mean square error of approximation
CFA	Confirmatory Factor Analysis
EFA	Exploratory Factor Analysis
SEM	Structured equation Modelling

Table of Contents

List of Figures

List of tables

Chapter 1: Introduction

1.1. Research problem

The recent rise of Big Data has been fuelled by the exponential growth in the volume of data and the sources available from which data can be gathered. This has resulted in 2.5 Exabytes (1 Exabyte = 1000000 TB) of data creation per day (Markow, Braganza, Taska, Hughes, & Miller, 2017). The popularity of social media, mobiles devices and the ubiquitous nature of sensors through wearable technology, has ensured that data collection is now entrenched into society (Mikalef, Pappas, Krogstie, & Giannakos, 2018).

In order to deliver business value from these various pools of data organizations, Big Data Analytics (BDA) seeks to make sense of the data by creating actionable insights for management to act upon. (Wamba, Gunasekaran, Akter, Ji, Ren, Dubey, et al., 2017). Big data is data with sizes outside of the scope of normal computer software to process, collated from numerous sources and which is intended to be used for large scale problem solving (Kamioka & Tapanainen, 2014). BDA provides invaluable insights to business which has the potential to advance Firm Performance (FPer) if used correctly (Lee, 2017). Through leveraging the resource based view (RBV), FPer is the value created by organisations when they possess or develop resources which are valuable, rare, inimitable and non-substitutable (Akter, Wamba, Gunasekaran, Dubey, & Childe, 2016).

Current research has centred around the application of data techniques through the volume, veracity and variety of data available, however determining the way in which organisations need to change to enhance the development BDA is currently lacking (Mcafee & Brynjolfsson, 2012). This has implications for how data analytics can be utilized to deliver business value and FPer and is further compounded by the under investigated benefits derived from BDA (Mcafee & Brynjolfsson, 2012).

Additionally, the technical nature of the big data research means that although the theoretical benefits and characteristics of big data are well defined, there is little knowledge of how Big Data capabilities are built for organisations (Bharadwaj, 2000). A company's ability to collect Big Data alone does not predispose it to success in the business environment, but rather is seen as necessary to build a capability (Gupta & George, 2016). Thus, the collection of big data alone by an organisation in a competitive industry does not provide for a competitive advantage, as other firms would be collecting data of their own in similar ways (Carr, 2003).

Instead, organisations are looking towards building more complex dynamic capabilities in order to extract the maximum potential from their Big Data. The outward looking artefacts of sensing and seizing opportunities in the external environment are paramount to the success of such technologies, and indeed the justification for further investment once these opportunities have been identified (Teece, Pisano, & Shuen, 1997).

Although challenging for organisations to achieve, seizing business opportunities can be seen as an executive decision requiring senior management to be adept in their understanding of technology and the external environment, as well as the underlying construct of resource scarcity (Mikalef & Pateli, 2017; Teece et al., 1997). However, in a technologically driven market, the challenge lies in the creativity of individuals, where Big data remains a resource, but the leveraging of the resource itself into innovative outputs, results in the development of competitive advantages (Salvato & Vassolo, 2018). Thus, the creation of BDA capabilities cannot be seen as an end in itself, but rather a tool requiring a human element to induce the most favourable results for the organisations.

In organisations built with Big Data at its core, the human element has a natural tendency to positively cohabitate with technology. This however changes when the organisation is based on a workforce aligned to experience related decision making, away from the realm of BDA and data driven decision making.

Central to this idea is the intersection of people, and technology where technological advancements, such as the development of BDA, results in workforce anxiety at the threat of automation replacing jobs (McKinsey Global Institute, 2018). An Organisational Culture (OC) resistant to change and a movement away from experience based decision making, towards data driven decision making, will face difficulties in Big Data project implementation (Lorsch & McTague, 2016; Mcafee & Brynjolfsson, 2012). Recent findings by McKinsey Global Institute (2018) posits that at least 75 million people will need to find new career paths and reskill themselves by the year 2030, or face unemployment. This rapid change in the working environment has sparked a greater resistance to change as workforces are unwilling to openly adopt a data driven culture out of fear of job losses. Furthermore, OC extends its reach to the executive level, where a failure to appreciate the changing business environment contributes to the failure of Big Data projects (Alharthi, Krotov, & Bowman, 2017). With executives entrenched outside of the realm of data driven decision making, cultural shifts which embrace change are required to facilitate healthy adoption and improved financial performance.

Thus, by ignoring OC differences, companies may face a risky future for sustained competitive advantage in a changing business environment. Forcing Big Data projects onto organisations who are unable to swiftly adapt to change, may ultimately result in underwhelming FPer.

In tracking the human impact of technological adoption, the assimilation of Artificial Intelligence (AI) by business is linking the previously disparate worlds of humans and machines. AI can be defined as a system with the ability to understand data , learn from the data , and then apply these lessons to achieve a specific task (Haenlein & Kaplan, 2019). Within the sphere of Big Data, this is paramount to maximising the value generation of Big Data technologies. Supplementing AI into the realm of large data oriented, time consuming tasks, replaces the human element who would otherwise be incapable of task completion in the same time period (Brock & von Wangenheim, 2019). Instead, the skills specialisation of AI related jobs increases, with data scientists required to develop AI technologies and mould them to produce the required business outputs (Brock & von Wangenheim, 2019; Leopald, Ratcheva, & Zahidi, 2018). As reported by Leopald, Ratcheva, & Zahidi (2018) in assessing task hours across 12 industries, showed that current task hours are split between human and machines at 71% and 29%, respectively. This is set to increase dramatically by 2022 with this split revising to 58% to humans and 42% to machines (Leopald et al., 2018). The rate of change is accelerating with companies looking towards leveraging their existing Big Data repositories through the use of AI.

Indeed to understand the impact of Big Data is to understand the further technologies it enables, such as AI, together with the entanglement effect of the workforce , will inevitably propagate the implementation of such projects to begin with. AI cannot exist without the vast banks of Big Data and in turn Big Data needs to be adopted by the workforce to allow for improved FPer to materialise. Thus, there exists an entangled relationship between the constructs of Big Data and its associated capabilities, the organisations culture towards adopting data driven decision making, and further technologies being enabled by Big Data.

Given the proliferation of Big Data in the business environment, and its associated positive effects on FPer, it stands as a beacon for current and future business strategy (Constantiou & Kallinikos, 2015). The challenges however lie in the execution of these projects, thus illuminating the central focus of this study.

1.2. The theoretical need for this study

Through the use of big data analytics, organisations can take advantage of benefits such as reduced costs, improved product quality and more informed decisions being made (Davenport, 2014). In their research, Davenport & Bean (2019), found that 77% of business executives listed business adoption as the number one difficulty in Big Data adoption. These were further expanded to show that only 7.5% of executives believed that implementation of the technology was a challenge, whilst 93% agreed that people and processes were the key resistors to adoption (Davenport & Bean, 2019).

Successful organisations have been found to be twice as likely to use data analytics as there less competitive peers and thus are strongly differentiated in their industry (Lavalle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011). The literature however is still underdeveloped, thus there is a need to understand the underlying relationships under which Big Data can act as an enabler of organisational performance improvement (Akter et al., 2016; Wamba, Akter, Edwards, Chopin, & Gnanzou, 2015). Thus far the study of BDAC has focused on the technical capabilities and their effect on the organisation through enabling IT capabilities (Akter et al., 2016; Garmaki, Boughzala, & Wamba, 2016), with little attention placed on the organisational capabilities. Business however has reiterated that amongst the largest roadblocks to establishing Big Data capabilities within their organisation is the development of a big data strategy, which in turn is based on the deeper levels of OC and the required change needed to raise support for the technologies adoption in the organisation (Vidgen, Shaw, & Grant, 2017). Without the support garnered through an OC shift towards a data driven business environment, managers will be powerless to implement these projects and effectively leverage them for a positive return on investment. Pospiech & Felden (2012) found through their work that 87% of the literature on big data is of a technical nature, leaving just 13% on a non-technical focus. Thus, for business management studies, there is a lack of consensus on the implementation of BDA and its interactions with the organisation, leaving space for the factors which may influence the relationship between BDA and FPer to be analysed.

Furthermore, literature has presented contrasting views on the subject with Côrte-Real, Oliveira, & Ruivo (2017) revealing that OC, amongst its population sample of IT managers and business executives, ranks outside of their top 10 precursors to successful big data implementation. Their research calls for further analysis of the Big Data subject, leveraging a greater degree of theory, to create a movement away from the technical lens through which the topic has thus far been primarily investigated.

With Big Data having been implemented successfully by numerous companies the new frontier has begun to emerge through the development of AI. The difficulties experienced in the past where large enough datasets simply did not exist, combined with the now low costs of data storage, has realised itself in the ubiquity of data (Mikalef et al., 2018). The focus on Big Data is now shifting towards the next frontier in AI (Henke & Kaka, 2018). Literature however has tended to focus on the implementation of AI technologies through strategic challenges facing management from adoption to the considerable effects of these technologies ability to replace the human workforce (Haenlein & Kaplan, 2019; Hosanagar & Saxena, 2017; Morikawa, 2017; Tambe, Cappelli, & Yakubovich, 2019). The unfortunate pattern emerging however is that business is failing to leverage these technologies in the same light as Big Data. This is due to AI being a subset of Big Data technologies, where AI cannot exist without the necessary data for the technology to learn from. AI however has remained a future plan for most organisations with 85% confirming that they have a plan for the use of the technology, but just 20% having gained any substantive benefit from its use (Brock & von Wangenheim, 2019). Although several studies have focused on the interaction of individual strategic constructs on AI implementation and its success, there exists a gap in the literature to confirm the interaction of BDAC on FPer through the concurrent effects of OC and the further development of AI technologies. Achieving an understanding of these interactions provides an initial base for the planning of organisational strategy from developing of BDA capabilities to the ultimate development of the technology in the form of AI.

1.3. The business need for this study

Due to the speed of growth in Big Data and the increasing need for BDA, companies need to continuously keep abreast with the implementation of new BDA technologies or risk becoming uncompetitive (Rao, 2017). These implementations however tend fail often due to the barriers to entry (Alharthi et al., 2017; Tabesh, Mousavidin, & Hasani, 2019). BDA's relationship with organisational factors such as OC needs to be more thoroughly understood for management to decrease the failure rate of these projects, and increase FPer (Alharthi et al., 2017; Lavalle et al., 2011).

Although understanding the relationships to BDAC still challenges businesses worldwide, the accelerated rate of change of technology has meant that the next wave of technologies has arrived in the form of AI (Henke & Kaka, 2018). AI is a necessary technology to leverage Big Data due to the limitations experienced as human beings. The human element of business decision making is anchored on the cognitive limitations of the human brain, where a concentration of information is rationalised as 'white noise' allowing for daily functioning, but limited ability to make sense of the large data sets offered through big data (Colson, 2019). Thus, to develop BDAC in the current business environment is to further plan the development of AI. These technologies however are difficult to implement without the OC which drives the ‘people’ aspect of strategy. Interestingly, South Africa is further behind the development curve both in terms of infrastructure and the workforce readiness to understand the technology, and upskill themselves for the future of work (Moore, 2017). These gaps in knowledge and the intermingled of angst and uncertainty pertaining to AI, provide a research opportunity to remove the veil clouding the true state of affairs for organisations.

This research study thus proposes to gain a further understanding of BDAC and assess its effectiveness in organisations as a tool for enhancing FPer. Furthermore the relationships of OC and its subsequent effects on BDA project implementation will be investigated to derive further insights for business in strategic planning. AI thus forms the third pillar of the study, by assimilating the future prospects of the technology, and investigating the entanglement of relationships between BDAC, OC and AI. Thus, further business insights will be illuminated through these constructs, allowing business to effectively plan a long range technology strategy from BDA through to leveraging their data with AI to enhance FPer.

Chapter 2: Literature review

2.1. Introduction

The purpose of this literature review is to explore the artefacts of Big Data, its relevance in a business context and the challenges experienced by business in successfully deploying Big Data projects. Being the topic of the moment, Big Data is often evangelised as a tool which enables strategic advantages for businesses and thus creates higher financial performance. In reality the relative youth of the literature on the subject remains divided, providing for numerous caveats to enable successful FPer.

To address this issue, this literature review explores the position of Big Data in the business environment today, whilst defining the potential levers of success for the successful implementation of Big Data projects. Beginning with the definition of Big Data, the history of the subject is explored, along with complementary theories from strategy through dynamic capabilities. Likewise, the challenges, as has been highlighted through supporting studies, is taken into account to fully define the environment as it stands today. Of particular interest in this study was the underexplored elements which influence FPer in the Big Data environment, thus the literature takes this exploratory viewpoint, ultimately allowing for the proposition of a research model which forms the focus of this study.

2.2. Definition of Big Data

With the advent of Big Data technologies such as social media and more broadly speaking the internet of things, data collection has reached a point of inflection, whereby organisations view the power of data as necessary to their success (Ylijoki & Porras, 2016). Through the increase in usage of internet-connected devices a resultant movement has emerged i.e. datafication (Lycett, 2013). Defined simply, datafication can be seen as the “sense-making process” which through its data capture abilities, contributes to the spectacle that is Big data (Lycett, 2013; Ylijoki & Porras, 2016). Notably for Big Data, are several definitions of the construct, each derived from a unique perspective. These include the product perspective, process-oriented perspective, cognitive-oriented perspective and the social movement perspective (Ekbia et al., 2015).

The product-oriented perspective views data historically from its inception to growth in recent years, and is based on the size, speed and structure of data (Wang, Zenshui, Fujita, & Shoushen, 2016). Through the product-oriented lens, big data can be defined as datasets with sizes larger than general computer software is capable of processing, storing and analysing (Bharadwaj, 2000; Sun, Chen, & Yu, 2015). The proliferation of social media, IOT and online shopping has led to companies such as Facebook producing over 10 petabytes of data per month through social networking, Google processing hundreds of petabytes of data through its search engine, whilst online shopping giant Baidu, similarly generating tens of petabytes of data per day through online shopping (Chen et al., 2014).

The process-oriented perspective highlights the uniqueness of processes involved in the production with Big Data, usually outlined through the storage, management, aggregation, searching and analysis of such data (Ekbia et al., 2015). This perspective highlights the infrastructure challenges faced by Big Data including, the array of technical tools and programming techniques associated with Big Data (Wang et al., 2016). Under this premise Big Data can be defined as data with a size large enough to force us to seek new methods to deal with its processing, storage and analysis (Bekmamedova & Shanks, 2014).

The cognitive-perspective of Big Data is built around the construct of mental actions or discourse which needs to take place to decipher meaning from data (Wang et al., 2016). It seeks to understanding the cognitive challenges which Big Data places on humans, often exposing our limitations (Ekbia et al., 2015). Kraska (2013) explains this concept with a simple analogy where a terabyte of data would be nothing for the U.S. national security agency to compute, yet for one human being this would be far more challenging. Thus, to interpret Big Data a multidisciplinary approach is needed, which underscored by discourse and supporting technical analyses, allows for more lucid understanding of the phenomena (Ekbia et al., 2015).

Lastly, Ekbia et al. (2015) introduces the social-perspective where the focus is placed on the socio-economic, cultural and political shifts underlined by the Big Data movement. With the desire to move towards a utopian society based in Big Data, several partnerships have begun to emerge, consolidating efforts, and resulting in the furthering of Big Data science. Apache Hardoop and Apache spark are examples of Big data processing programs, developed through strategic partnerships between the universities and independent contributors, thus creating an open source tool (Bello-Orgaz, Jung, & Camacho, 2016). These types of collaborations incite a vision for the future, affecting societal perspectives of how Big Data is imagined and executed.

These four perspectives allow for the understanding of Big Data as a multidimensional field of study, whilst not being mutually exclusive in their scopes, they also provide a basis for the key themes of volume, velocity and variety which was first described by Laney in 2001.

Laney (2001), defined big data as having three dimensions of volume, velocity and variety. This framework has been pervasive throughout academic literature on big data, providing insights into the opportunities and difficulties faced by business (Chen et al., 2014; Lee, 2017; Li, Tao, Cheng, & Zhao, 2015).

Volume refers to both the amount of data generated and collected as well as the minimum limit of one terabyte, to be considered as big data (Chen et al., 2014; Lee, 2017; Ylijoki & Porras, 2016). The volume of big data continues to grow as the world moves through the internet of things (IoT) of web 3.0, with more devices becoming connected, thus increasing the volume of data (Lee, 2017).

Velocity refers to the speed at which data is produced and at which it can be processed (Chen et al., 2014; Wamba et al, 2015; Lee, 2017). This is increasing steadily as with the volume of data due to the new devices being connected each day, sharing and analysing data along the way (Elgendy & Elragal, 2016; Lee, 2017; Tabesh et al., 2019).

Variety refers to the structured and unstructured data which is being shared daily, such as text, audio, video and photo (unstructured) whilst structured data refers to database oriented data such as that offered in Microsoft SQL (Elgendy & Elragal, 2016; Lee, 2017; Tabesh et al., 2019; Ylijoki & Porras, 2016).

A further V, veracity, was added by IBM to account for the unreliability and ambiguity of data (Chen et al., 2014; Lee, 2017; Tabesh et al., 2019). This is prominent on social media where data on consumer opinion and behaviour may not reflect the truth (Lee, 2017; Tabesh et al., 2019). Two more V's have recently been added by technology company SAS, who added variability and Oracle who added value (Lee, 2017). Variability refers to the unpredictability in the data flow rates such as during sudden events, which trigger high activity and stress available resources in their ability to process the data (Elgendy & Elragal, 2016; Lee, 2017). Lastly value is added as a management component to judge the cost to benefit of implementing a big data project (Lee, 2017)

The 5 V's highlight the complexity of big data which grows as the IoT and its associated devices grow in popularity. Tabesh et al. (2019) have described this increase in data as "the new gold rush" which is causing an extraordinary growth in demand for data tools and techniques. Rasmussen & Ulrich (2015) refer to this as "data fetish" however the interest around BDA seems to be valid with a predicted 20 billion in smart devices to be connected by 2020 (Henke, Libarikian, & Wiseman, 2016).

2.3. The rise of Big Data

The current business climate shows a rising number of firms investing in BDA, with the goal of providing further insights which may give them a competitive advantage (Constantiou & Kallinikos, 2015). Whilst this theme continues today, its origins stretch back to the early 1990's with Big Data 1.0.

Lee (2017) posits that Big data 1.0 began in 1994 with the advent of ecommerce. This allowed for web mining to take place where web users actions were monitored to determine behavioural patterns. In 1997, the first mention of the term "Big data" was made in a publication in the ACM digital library, introducing the world to the possibilities of large volume data sets under a specific terminology (Li et al., 2015). Laney (2001) defined big data for the first time stating that challenges and opportunities relating to big data were around its characteristics of volume, velocity and variety, highlighting the importance of Big Data at this early stage. User created content was minimal until the start of web 2.0 in 2005 when technology advanced to a stage where website users could now create their own content, resulting in the birth of social media (Lee, 2017).

The development through web 2.0 and social media continued, with larger volumes of data advocating the need for new types of tools to be developed to analyse the data (Chen et al., 2014; Lee, 2017). This social media revolution created a way for users of data to interact with data for the first time, a marked change from web1.0, and the drive for further data creation through user driven content.

Lastly, web 3.0 began in 2015 with the creation of the internet of things (IoT) where connected devices, such as smart watches, share information about the user without any human involvement (Lee, 2017). This development has changed fundamentally from web 2.0 as data is created passively by users, in comparison to web 2.0 where active content creation drove data growth. Thus, data growth continues on an upward trend, creating further opportunities for business to derive value. Further estimated growth trends show that volumes will continue to grow with 20 billion connected objects by 2020 (Lee, 2017)

As a result of this growth, Big data has cemented itself in decision making for organisations (Hagel, 2015). It is currently viewed as a differentiator of FPer increasing business foresight, whilst reducing costs of deriving new customers by 47% and increasing revenue by 8% (Liu, 2014). Further to this, literature has already predicted the disruption in businesses by the use of Big Data (Davenport, 2014; Manyika et al., 2011) , as it is seen as a high impact technology (Ylijoki & Porras, 2016). Although this provides opportunities for companies who are data enabled, transformational business changes will first be necessary in order to unlock this competitive advantage (Dehning, Richardson, & Zmud, 2003). Therefore, these technologies will provide varying degrees of success of failure for businesses dependant on their ability to change their business model, linking strategy to the digital evolution they now find themselves in (Ylijoki & Porras, 2016).

2.4. Big data and value creation

The value creation of big data has been widely researched in its ability to create competitive advantage and positively impact FPer when implemented successfully (Manyika et al., 2011; Müller, Fay, & Brocke, 2018; Y. Wang, Kung, & Byrd, 2018). FPer has been empirically proven to increase with investments in IT through financial and operational gains or organisational gains (Kim, Shin, Kim, & Lee, 2011; Mikalef et al., 2018). The 5 V model, as eluded to above, provides input into the opportunities and threats of big data, however, this framework is lacking in its practicality for management in assessing the consolidated impact of big data on the organisation. Mazzei & Noble (2017) suggest a three-tier framework to determine the effect of data on strategic management:

Tier 1: Data as a tool
Big data can be seen as a tool to link the organisation to the customer, tracking customer behaviours, preferences and products, allowing the organisation to respond in a way that generates greater FPer (Mazzei & Noble, 2017).

Tier 2: Data as an industry
Implementation of BDA comes with its own set of barriers to entry (Alharthi et al., 2017). Many firms may not wish to generate their own capabilities in BDA due to costs or complexity, thus the industry of data is able to provide this through packaged solutions and outsourcing (Mazzei & Noble, 2017).

Tier 3: Big data as a strategy
Mazzei & Noble (2017) describe the business environment in two ways; the old way where strategy drove data and the new way, where the inverse is true. The strategic decision making process at organisations has changed considerably due to this, with executives such as chief information officer and chief data officer having further input into the organisation’s strategy in order to drive data capability (Côrte-Real et al., 2017) . Effective data strategy is reflected in an organisations ability to understand their data strategy aspirations, i.e. their space in the competitive environment

amongst rivals or developing fields which they wish to explore) and data strategic authentication (the firm's ability to implement data strategies based on available capabilities and resources) (Mazzei & Noble, 2017).

Development and maturity across the three tiers are required for a company to achieve a significant competitive advantage, thus any shortcomings may require a realignment of the business's strategy. Companies either need to emerge as tier three entities ,such as Facebook and Amazon, or be wary of these entities entry into their landscape as they further develop through the tiers (Mazzei & Noble, 2017).
This model is useful in its ability to keep management honest about their data driven aspirations as well as identifying the maturity they currently possess. Value extraction can be maximised per tier if the correct approach is taken, ultimately culminating in a tier three firm.

2.5. Big Data challenges

As with any new technology, there are numerous challenges to be expected in implementation, closing the gap between expectations and reality with regards to FPer. Businesses developing BDA strategies are however experiencing difficulties in leveraging these towards creating business value, with investments in projects showing an estimated 65% success rate via a below average return (Baldwin, 2015).

Research has shown that the challenges of BDA are not merely based in technical fields of deployment, but it is rather skewed towards organisational challenges (Gupta & George, 2016). This concept was explored in detail through a Delphi study by Vidgen, Shaw, & Grant (2017) who found that in order to become a data-driven organisation management have five key areas to focus on, namely; data, technology, processes, people and organisational related challenges. This notion was in response to the management challenges ,originally outlined by Mcafee & Bryjolfsson, 2012 , which still remains a problematic area for organisations today.

Data

Regardless of the organisational challenges Data brings, it has become of paramount importance to business. Recent research has surmised that data, like land, labour and capital, is now a factor of production (Manyika et al., 2011). Lee (2017) sets out data challenges into quality and security each with their own distinct features. Data quality affects decision making in organisations, thus poor quality data leads to incorrect decisions being made and negates the potential competitive advantages that would otherwise have been derived (Lee, 2017).

Further Data complications can arise through the investment of firms into the incorrect type of data. George, Haas, & Pentland, 2014 posit that data can be identified by five distinct categories namely; public data, private data, data exhaust, community data and self-quantification data. Public data typically is generated by government agencies and can be of use to businesses, whilst private data is generated by organisations, securely held away from the public (George et al., 2014). Data exhaust alludes to data that has zero value, unless it is combined with other forms of data which may lead to new insights (George et al., 2014). Community data refers to social media data whilst finally self-quantification data outlines the IOT revolution where data is collected passively through wearable technology such as smart watches (George et al., 2014).

Though a multitude of data types may exist business decision makers lie outside of the sphere of pure data specialists, requiring conceptual understanding of the technology and how-to effective leverage it to their advantage. To this end, a more simplified view of data categorisation places data into internal or external data (Zhao, Fan, & Hu, 2014). Internal data is generated by a firms internal systems, such as inventory systems, in contrast to external data which is generated externally through a wide range of modalities ranging from ecommerce to smart devices (Zhao et al., 2014). One such example may be seen through United airlines, who through the use of both internal and external data have been able to improve the customer experience. Data from internal systems such as seating (internal data), is combined with customer data (external data) such as customer preferences, allowing for a personalised experience (Amankwah-Amoah & Adomako, 2019) . Therefore, high levels of investment into a firms internal data alone, can lower the chances of producing a competitive advantage (Zhao et al., 2014). Instead, companies need to be wary in deciding the scope and mix of their data as only through combining both internal and external data will firms build the capabilities for producing competitive advantage.

Technology

The technology needed to unlock data analytics is vested in the type of data the organisation is using and the nature of the employees (Lee, 2017; Vidgen et al., 2017). Citytrans for instance advocates that their data scientists need a curious mindset in order to be productive; questioning the make-up of the world order and challenging this towards value creation for the organisation (Vidgen et al., 2017). They accomplish this through the use of relational database management systems such as structured query language (SQL), which relate to structured data generated by organisations such as financial data, inventory data and customer data, which may be used by data scientists to generate value, through data manipulation (Vidgen et al., 2017). Data types have however evolved to match the ubiquitous nature of data collection, with the introduction of Big Data, moving from structured data and its associated tools such as SQL, to unstructured data through the use of Hadoop and coding languages such as R (Chen, Chiang, & Storey, 2012; Wamba et al., 2017). These new tools are necessary when considering the growth of Big data, where the traditional technologies of data analysis can no longer cope with the volume of data being generated (Bello-Orgaz et al., 2016). Thus, in order for organisations to leverage BD appropriately, a further investment into the technologies to unlock the potential of BD is needed, as BD in isolation provides for little value to organisations.

Process

The analytics process requires a structured team to be effective in organisations often being composed of data scientists, business analysts and IT professionals (Vidgen et al., 2017). This allows for the effective collection, analysis, integration and interpretation of the available data, allowing for a robust and accurate decision making process by management (Wang et al., 2018). Thus, the analytics processes can be seen as interlinked with both the structure of the organisation’s analytics team, as well as the independent tasks which each follows. Through the analysis of cases related to Big Data integration, research has shown that the process around analytics may be broken into three constituent parts namely; agility, ethics process and the exploration and exploitation of data (Vidgen et al., 2017).

Agility in the Big Data Analytics (BDA) environment rests on the architectural design of the data systems, that is, the software elements, properties and interrelations between them (Bass, Clements, & Kazman, 2003). Without an amalgamated architectural concept, fit for purpose within the organisations BDA aspirations, the artefacts of performance, security and modifiability cannot be achieved (Chen, Kazman, & Haziyev, 2016). Failure to reach said artefacts could result in business not achieving the desired outcomes from Big Data implementations, however chasing these aspirations relentlessly may unleash further ethical quandaries, with their own pitfalls.

Big Data's pace of development comes with certain ethical obligations for organisations, ensuring individual's privacy is upheld and that said organisations do not commit blatant discrimination when using this information for decision making (Zwitter, 2014). To counter these behaviours, responsibility needs to be shared between governments and organisations, where governments should apply certain stringent agreements on ethical practices, whilst organisations should ensure that these are followed correctly (Cao & Duan, 2014; Someh, Davern, Breidbach, & Shanks, 2019). One such way in which this can be achieved is through the establishment of an ethics committee dedicated to the Big Data strategy of the organisation, ensuring correct governance is implemented and followed (Vidgen et al., 2017). The facets of the agreed governance should then be implemented in such a way as to mould the norms, values and beliefs of the organisations to one which aligns to high ethical standards, and thus eliminates risk around Big Data (Someh et al., 2019).

People

People centre challenges of Big Data refer to the correct skills necessary to realise the organisations Big Data strategy. Such skills refer to the knowledge in data enrichment for the production of business centred insights as well as the practical knowledge to correctly leverage associated tools such as Hadoop (Gupta & George, 2016). Vidgen et al., (2017), through their Delphi study on the challenges of Big data, surmised that these skills, although important, were secondary to the characteristics of the data scientist themselves. More importantly a data scientist needs a curious mindset along with the resourcefulness of a “bricoleur” (Vidgen et al., 2017). Combining these attributes creates a unique and rare profile for the people needed by organisations. A recent study into the resources of BDA which lead to high FPer confirmed that these rare technical skills were of great importance to management, due to their rarity and the current difficulty experienced in recruiting them (Mikalef, Boura, Lekakos, & Krogstie, 2019). These skills are currently still being developed and if they are to one day reach the ubiquitous nature that has occurred with IT technical skills, organisations will find an easier path to achieving their Big Data strategies (Gupta & George, 2016). For now, this remains a point of interest, and key space in any organisations Big Data aspirations.

Organisational issues (Culture)

In the case of data-driven decision making, organisations tend to behave in an overconfident manner, often exaggerating their use of data in the business (Mcafee & Brynjolfsson, 2012). This stems from their lack of understanding around the intangible resources required for BDA. OC has been highlighted as one such asset, which is a necessary requirement if organisations expect to derive maximum value from their BDA strategies (Gupta & George, 2016). The central issue around culture is that organisations have historically based decisions on both managerial expertise and intuition, rather than having built their decisions off robust analytics (Provost & Fawcett, 2013). Establishing and strengthening a BDA capability can thus build towards a data driven culture, and a movement away from instinct based decisions (Amankwah-Amoah & Adomako, 2019; Cao & Duan, 2014). Since the constituent parts of building a BDA capability are seen as data, technology, processes and people, organisations need to place an equal importance on each of these to build towards a data-driven culture, and thus realising their Big Data strategies in the future.

These dimensions of big data emphasize the challenges and opportunities for business, however the rate of expansion of Big Data is harming companies on their returns on investment. These Big Data projects are expensive and difficult to implement; therefore, business strategy needs to be clear about the areas in which they are going to play and if these projects will allow the organise to leverage off of BDA to produce the desired rate of return. Management would need to further enhance their understanding of the barriers to entry for business into Big Data, focusing on projects that fit the organisation, reducing the chances of failure.

2.6. Generating Value from Big Data Analytics

Being the subject of the moment, big data has driven the business world to pursue greater firm profits, however few companies have realised the potential of their Big Data projects (Gupta & George, 2016). The has been due to the focus of research around Big Data being of a technical nature, investigating the infrastructure and tools necessary to implement Big Data, resulting in a lack of theoretically focused literature (Gupta & George, 2016; Mikalef & Pateli, 2017). Value however is generated not just through technical means and specific tools, but rather through the consolidated effort of organisations capabilities, from ordinary through to higher order capabilities (Teece, 2007)

Value comes from the successful implementation of Big Data projects by overcoming the barriers to entry (Alharthi et al., 2017), with OC postulated as having a significant impact (Henke et al., 2016). Due to this Gupta & George, (2016) developed a theoretical framework to test the relationship between BDAC and FPer through a lens of resource based theory (RBT) (Gupta & George, 2016). The model grouped variables into three categories according the RBT of tangible resources (data and technology), intangible resource (data driven culture, intensity of organisation learning) and human resources (managerial skills, technical skills) (Gupta & George, 2016). Their findings showed that there was a significant relationship between intangibles and human skills in achieving competitive advantage (Gupta & George, 2016). These results were corroborated by (Vidgen et al., 2017) who investigated a similar concept through a sociomaterialism lens, by identifying the challenges managers face in creating value through business analytics. Their results showed that IT skills were not the only factor in determining FPer through business analytics but rather encompassed human and managerial skills as well. These models however, test their theories on the basis that each firm should be treated in the same way, suggesting that their finding may be universally applied across companies to achieve FPer through BDAC (Abbasi, Sarker, & Chiang, 2018). In reality however, each firm can be differentiated through its norms, values and beliefs, which affect the way it performs in the market.

Through the use of complexity theory, Mikalef, Boura, Lekakos, & Krogstie (2019) were able to remove this limitation whilst testing for FPer using the factors of data (by size and type), process (process allowing for data to produce insights), technology (currently implemented technology), organisation (organisational culture), people (managerial skill and technical skill) and context (Firm size). Their findings showed that in moderately uncertain environments FPer responded positively to increased technological investment, whilst in highly uncertain environments FPer was more responsive to organisational characteristics and managerial skills (Mikalef et al., 2019). Thus, OC affects big data analytics ability to generate performance in a highly uncertain environment more than in a moderately uncertain environment. Since competitive markets are in a constant state of flux, this finding helps frame the challenges that organisations are facing in their Big Data aspirations (Gupta & George, 2016).

Côrte-Real, Ruivo, Oliveira, & Popovič (2019) conducted a Delphi study to verify the drivers of value in firms. The study used a two phase approach highlighting precursors to BDA through a Delphi study, followed by a cross country survey (Côrte-Real et al., 2019). The population sample comprised of business and IT executives across European firms. Their results were in contrast to the findings of Gupta & George (2016), Mikalef et al. (2019) and Vidgen et al. (2017), showing that the organisation culture precursors were in the form of management of exogenous knowledge, management of endogenous knowledge, and analytical decision making culture ranked outside of the top ten precursors in organisations (Côrte-Real et al., 2019). The population sample suggests that IT and business executives do not place as high a value on OC as they do on technical competencies of BDA.

The results of these models present a new set of challenges to organisations. Whereas the focus of business has been on keeping pace with the velocity on Big Data, not only have organisational factors been overlooked, but the limited theory guiding BDAC implementation, excluded the complexity of firms. The common theme however is that OC presents a relationship to FPer and this relationship may be more complex than previously thought. Business leaders however seem to place more value on technical precursors of BDA than non-technical precursors (organisational culture). These results provide a starting point towards understanding these relationships and that of the poor return on investment associated with Big Data projects (Baldwin, 2015).

2.7. Dynamic capabilities in organisations

With over 3000 articles appearing in a google scholar search for the title "Dynamic capabilities", it is clear that the framework is currently at the fore of academic research (Teece, 2018b). Dynamic capabilities theory is today considered a fast developing area of strategic literature, providing the basis for alternative methods of viewing an organisations competitive and strategic management options (Teece et al., 1997). It is developed in the theoretical base of resource based theory, which has its roots in the 1960's (Learned, 1969), however has only come to the fore of academic research in the 1990s with the work of Teece, Rumelt, Dosi, & Winter, (1994) and Teece et al., (1997). The resource based theory surmises that an organisation will need to arrange its resources in such a way as to create a competitive advantage in order to achieve its predetermined goals (Learned, 1969). The RBV interprets organisations as being dissimilar with resources/capital endowments, which are difficult to change in the short term, thus placing a high level of importance on the management decision to enter a particular market (Teece et al., 1997). With this in mind the RBV focuses on creating sustained competitive advantage, which may only be achieved if the resource itself are valuable, rare, inimitable, non-substitutable (VRIN) (Barney, 1991).

In contrasting the RBV against the established strategic theory of competitive forces and strategic conflict (Teece et al., 1997) highlight that RBV allows for management input into producing new capabilities in organisations. These can be highlighted by artefacts such as skills acquisition, the management of knowledge and learning, where learning provides the basis for the development of dynamic capabilities (Teece et al., 1997).

Several definitions of the dynamic capability construct have been developed since its inception, however, since this research takes a position of examining FPer, dynamic capability will be examined through this lens. With this in mind, the dynamic performance of an organisation may be seen as is its ability to create positive performance through its dynamic capabilities under competitive conditions (Iansiti & Clark, 1994). In turn, this positive performance is rooted both in external and internal uncertainties, where external uncertainties may be driven by customer centred challenges such as evolving technology, or internal uncertainties, where changing circumstances for the business such as structure, may affect performance (Iansiti & Clark, 1994). More practically, Iansiti & Clark (1994), use an example of product development whereby problem solving, that is creating the correct product for the market given the changing technological and competitive environment, may be seen as a dynamic capability, not just for product development, but more generally as well.

Each organisation has a selection or capabilities which allow it to be successful, however these can be more intensely analysed by breaking this down into two distinct groups, namely ordinary operational and core capabilities and higher dynamic capabilities (Teece, Pisano, & Shuen, 2008) . These ordinary or “zero level” capabilities comprise mainly of administration related tasks, which allow the business to continue its daily operations (Sidney G. Winter, 2003; Teece et al., 2008). The higher order capabilities can be further broken down into micro foundations, the managerial decisions regarding the deployment of resources to meet strategic objectives, and higher order dynamic capabilities (Teece, 2018a; Winter, 2003).

The higher order dynamic capabilities framework was developed from this point of view by Teece et al. (1997) and consists of:

Sensing: In highly competitive environments with changing technology, organisations need to devote resources to searching for new opportunities. Through this iterative process learning takes place improving an organisations ability to locate new opportunities for growth (Mikalef & Pateli, 2017; Teece et al., 2008)

Seizing: Once the opportunity has been identified, it follows that the next decision will be in the investment path to be taken. Since technology often follows several competing development paths in the early stages of any new opportunity, the organisation's managerial team will need to follow stringent paths in order to capitalise on the correct solution (Mikalef & Pateli, 2017; Teece et al., 2008)

Reconfiguring: Once the opportunity has been adopted into a business model and actioned the organisation will be set along a path of profitability. This however creates a path dependency as technology will begin to change and shift, causing the established operational routines once needed for profitability, to become defunct, requiring reconfiguration (Mikalef & Pateli, 2017; Teece et al., 2008).

Although dynamic capability is a relatively well-established area of strategic literature, numerous discrepancies in views litter the field. Indeed for an organisation to effectively develop dynamic capabilities there needs to be a clear distinction between "zero level" capabilities and higher order capabilities, however since these are defined by the organisation itself, some confusion and overlap often occurs (Winter, 2003). Further complexity is then brought to the fore in literature with a disagreement on heterogeneity of the construct, with Teece et al., (1997) suggesting that these are unique to each organisation, whilst others suggest that these are far more homogenous (Eisenhardt & Martin, 2006). The implications of this on organisations who are currently investing exorbitantly into developing their Big Data capabilities are vast, as the aforementioned under performance of these investments, challenges managers decision of capital expenditure (Baldwin, 2015).

Much of the dynamic capability literature is skewed towards the importance of the macro level, emphasising the importance of operational routines in the development of these capabilities in the organisation (Mikalef & Pateli, 2017; Teece, 2007; Teece et al., 2008; Teece et al., 1997; Winter, 2003). In this view, the importance of the creative individuals and their ability to develop new business models in the face of a dynamically changing environment is removed, however organisations need both the methodological approach of routines as well as individual contributions in creativity to develop effective dynamic capability (Salvato & Vassolo, 2018). Thus, drawing on both levels of theory, dynamic capability may be thought of as both the use of operational routines and individual contributions to achieve organisational change in a dynamically changing market.

With dynamic capability being vigorously pursued in academic literature, it stands to reason that the increase in FPer through the use of this construct is guaranteed. This however seems to be a misnomer, as the empirical evidence of performance is divided in the theoretical base, with authors finding positive effects of performance (Fainshmidt, Pezeshkan, Frazier, Nair, & Markowski, 2016; Fang & Zou, 2009; Stadler, Helfat, Verona, & Stadler, 2013), whilst others found a negative impact (Schilke, 2014; Wilden & Gudergan, 2014). Fainshmidt et al. (2016) through their meta-analysis of the topic found that dynamic capabilities, establishes valuable resource structures which indeed contribute to FPer as a general rule, however, is notably dependent on the managerial skill for successful implementation. Furthermore, the type of capability or dynamism of the organisations dynamic capability, affects FPer , with research showing that marketing capabilities produce positive FPer (Wilden & Gudergan, 2014). In contrast , technological dynamic industries produces less of an effect on FPer, due to the relative homogeneity of the dynamic capability's amongst competing organisations (Barrales-Molina, Bustinza, & Gutiérrez-Gutiérrez, 2013; Fainshmidt et al., 2016; Wilden & Gudergan, 2014). The central theme is that dynamic capability is a multi-layered construct, which cannot be viewed from a macro level in isolation, but instead needs the additional creativity of individuals to drive performance (Salvato & Vassolo, 2018). In the case of Big Data this means that the workforce needs to consist of creative individuals who contribute positively to the nurturing of dynamic capability from the bottom up, who are absorbed by their jobs and who are able to be mentored by senior management (Salvato & Vassolo, 2018). Senior executives on the other hand should remain in their roles of developing higher order dynamic capabilities from the top down, driving the restructuring of resources and FPer (Teece et al., 2008). More pragmatically, research has been conducted using the dynamic capabilities framework as a lens for business analytics showing a direct correlation to FPer, and in this way corroborating the current trend towards BDA investments (Torres, Sidorova, & Jones, 2018). The implication for this research shows a valid use of the sensing, seizing and recalibrating framework towards building a dynamic capability of business analytics, and further that dynamic capabilities functions through ordinary capabilities to generate FPer (Bowman & Ambrosini, 2003; Torres et al., 2018).

2.8. Organisation Culture towards new capabilities

With over 4600 articles being published on organisational culture, the topic remains of particular relevance both to academia and organisations in general (Hartnell, Ou, & Kinicki, 2011). In a business context, and with the arrival of Big Data, barriers to implementation of the technologies and projects associated with Big Data have been highlighted to include OC (Alharthi et al., 2017). Authors on the topic have stressed the importance of creating a data-driven culture within an organisation, unlocking the expected performance from big data initiatives, whilst managers in organisations have joined the movement by stressing the practical link between OC and the effective performance of the organisation (Lorsch, McTague, 2016; Mcafee & Brynjolfsson, 2012).

OC has several definitions in academic literature, however most authors seem to agree that it involves the norms, values and beliefs which form the identity of an organisation and define the way in which all employees behave (Deshpande & Webster, 1989; Kotter & Heskett, 1992; Schein, 2010). Others state a simpler definition of “how we do things around here” to express the deeply ingrained and intangible nature of OC(Schein, 2010). Culture is also distinct from concepts of corporate identity or the national culture, and serves as the invisible workings of an organisation which influences the way in which employees behave (Scholz, 1987). These definitions however do not provide one unified theory of the phenomenon, but rather spark more rigorous debate in academia. The reasons for this have been outlined as both an academic and business problem, where the overall organisational interest in the subject to leverage improved performance, has created an opportunity for consulting firms to step in (Chatman & O’Reilly, 2016) . Academia provides a lack of consensus on the basic definition of the construct and the way in which it affects organisations, stalling further research in making conclusive breakthroughs (Chatman & O’Reilly, 2016). Therefore, in order to define organization culture, a view of culture as a whole should be taken from the basic models which have remained pervasive in the field.

Culture as defined by Schein (2010) can be viewed as a three tiered model which highlights the ways in which culture manifests in organisations. At the lowest level sits the values of the organisation, the underlying assumptions and beliefs which have created a sense of “what it’s like to work here” in an organisation (Schein, 2010). The next level shows the norms and values present in the organisation or rather the acceptable behaviour which is necessary to fit in, whilst the upper and similarly most visible layer comprises the artefacts and symbols of organisations reminding the employees of the cultural norms within the organisation (Schein, 2010). Indeed culture can be seen as the guidelines in place to steer the individual interactions between people whilst having an influence on the behaviour of individuals of organisations through the central norms (O’Reilly, 1989). The importance of norms as a measure of control over individuals cannot be understated, as organisations seek to derive innovative capacity to build a higher level of performance (O’Reilly, 1989). In order to achieve this, while meeting the employee requirements for big data analytics, i.e. employees with a certain amount of “bricoleur” or innovative capability to see the world in a different light, whilst creating solutions which business may derive value from Vidgen et al. (2017) , organizations would need to create a culture to support the use of dimensions of organizational culture (Dubey et al., 2003).

These dimensions may be subdivided into the several value generating artefacts which promote innovation, proposed by Dubey et al. (2003), as outlined below:

Success is inherent in any organisation encapsulated in the drive to succeed while creating value from their business strategy. Through the lens of OC this drives performance of employees to function at a higher level, encouraging innovative capacity amongst employees and produces a higher level of performance for the organisation (Vidgen et al., 2017). Employees are further driven to provide innovative solutions to problems, using their skills and distinct “view of the world” to allow further success to be derived by the organisation (Vidgen et al., 2017).

Openness and flexibility to problem solving supports employees in their innovative endeavours, promoting creativity and empowerment, which are key aspects of the “bricolage” in sustaining innovation (Vidgen et al., 2017). Situational learning is also promoted to advance the skills of employees, extending the learning process beyond the theoretical basis, which is paramount to the creation of dynamic capabilities for organisations seeking to be data driven (Gupta & George, 2016; Mikalef & Pateli, 2017; Teece et al., 1997).

Risk taking or the risk appetite of organisations differs greatly, however, to derive value from the use of an emerging technology such as big data, organisations should be ready to encourage calculated risk taking (Mikalef & Pateli, 2017). From dynamic capabilities theory this can be seen through the artefacts of sensing and seizing, inherently checking the environment for opportunities of the next wave of development, dedicating resources and thus taking on the risk of failure in search of higher returns (Mikalef & Pateli, 2017; Teece et al., 1997).

Thus, through the lens of organisational culture, the challenges which firms face in creating a data-driven culture through leveraging Big Data are highlighted. Without a focus on the culture of the organisation through controlling the norms, innovation will be difficult to achieve. Indeed by instilling norms which promote success, openness, flexibility and risk taking, organisations will be far more capable of overcoming the difficulties associated with the implementation of Big Data projects (Alharthi et al., 2017; Dubey et al., 2003; Tabesh et al., 2019) .

Organisations however differ in their cultural fit, ultimately exposing cultural disparities between channel members who are expected to develop mutually beneficial relationships (Prasanna & Haavisto, 2018). These relationships are inevitably the responsibility of employees of the organisation, thus the channel members experience of the OC moderates this relationship in either a positive or negative manner (McAfee, Glassman, & Honeycutt, 2002). Thus, differing from the Schein model of culture, authors have developed their own distinct ways in which in to measure the construct, including relational/transactional culture, where culture is measured in terms of the relationship being either purely transactional to a long term relationship (McAfee et al., 2002). Flexibility/control oriented culture on the other hand is based on organisational structures where flat structures are seen as more accessible versus a control-oriented culture, where organisations deploy a more hierarchical structure, devoid of flexibility (Khazanchi, Lewis, & Boyer, 2007; Srinivasan & Swink, 2018). In their study Prasanna & Haavisto (2018) found that these cultural fits resulted in a different reaction to external shocks by the organisations, and resulted in a different way of interpretation and ultimately reaction by each. Their study used the semi structured interviews to investigate the relationship between humanitarian organisations (buyers) and commercial sellers, showing that less collaboration and product innovation was achieved when these two cultures differed (Prasanna & Haavisto, 2018). Similarly, Dubey et al. (2019), found that OC and the structure of the organisation itself influences the degree of collaboration possible between two organisations. Thus, for organisations facing cultural barriers to the implementation of big data projects, particular attention should be placed on the internal and external culture of the business. This suggests that managers should focus their efforts on creating a culture of innovation to derive value from their employees, whilst instilling a culture of success to match channel members to advance their business strategies more effectively.

2.9. AI towards new capabilities

AI (AI) has recently come to the fore of business strategy, often hyped as the key to success and digital transformation (Overgoor, Chica, Rand, & Weishampel, 2019). AI leverages machine learning, together with big data, to provide businesses with faster decision making capabilities (Metcalf, Askay, & Rosenberg, 2019). The technology itself ,however, is not new, with its roots being traced back to the 1940's and popularised by the now famous world war two code breaker, Alan Turing, through the use of his code breaking computer "the Bombe" (Haenlein & Kaplan, 2019). Disappointingly, this was then followed by a relative trough in development in the coming decades due to the lack of significant computational power and an easily available repository of data (Haenlein & Kaplan, 2019; Overgoor et al., 2019). With the recent rise of Big Data however, barriers to the development of AI technologies have significantly declined, allowing for effective technological diffusion into the business environment (Mikalef et al., 2018).

Given that organisations are now attempting to deliver digital transformation strategies to improve their performance, they are faced with several technologies which aid in their strategy (Hess, Matt, Benlian, & Wiesböck, 2019) . Pragmatically, however, this leaves management with further decisions on the technologies to leverage to achieve their digital goals. AI (AI) is one such technology, but with the current scarcity of empirical evidence to show the benefits of such implementations, it becomes difficult to validate investment into this area (Brock & von Wangenheim, 2019).

Through their study, Brock & von Wangenheim (2019), surveyed over 3000 executives and managers worldwide to provide an understanding of the diffusion of AI. Their findings showed that implementation of AI technologies had become pervasion across the world, with 85% of organisations currently having AI plans and 20% having derived value from these implementations (Brock & von Wangenheim, 2019).

Digital transformation projects were implemented in a pragmatic fashion, attempting a symbiotic relationship with other technologies such as big data analytics and IOT, to achieve a greater degree of process efficiency and automation (Brock & von Wangenheim, 2019). Airbus is one such company to implement this, by the use of AI algorithms and stored data, the company was able to match problems experienced on the manufacturing floor with prior solutions at an effective rate of 70% (Ransbotham, Kiron, Gerbert, & Reeves, 2017). Not only did this technology allows for an amalgamation of process breakdowns with solutions, but the real time advise also shortened the time to solutions being implemented, improving manufacturing efficiency through decreased factory downtime (Ransbotham et al., 2017).

Whilst several theoretical lenses highlighted across this study corroborate with Brock & von Wangenheim (2019) in encouraging organisations to focus on business agility, innovation , process and data management (Dubey et al., 2003; Ekbia et al., 2015; Gupta & George, 2016; Vidgen et al., 2017), their study goes further to take a macro view on the phenomenon of digital transformation, introducing their DIGITAL framework as outlined below:

Digital is paramount to the success of AI as it is based on the data collected by the organisation. The algorithms on which AI is based is essentially developed as untrained, thus requiring a Big Data and its associated system to "learn" to provide value to the organisation (Ransbotham et al., 2017). Thus, data, along with the necessary skills to interpolate, manipulate and manage these datasets correctly, is a requirement for any successful AI technology implementation (Brock & von Wangenheim, 2019).

Be Intelligent highlights the necessary skills required in the organisation to effectively develop and deploy AI technologies. Indeed, this requirement is ubiquitous across the digital sphere and provides a challenge for managers to recruit to correct human resources for success (Brock & von Wangenheim, 2019; Vidgen et al., 2017).

Be Grounded in the planned strategic use of these AI technologies. Leaders in the field have used these technologies in the development of current offerings, improving processes and thus deriving value. Ambitious projects expose organisations to high levels of risk and ultimately could result in failure, thus smaller projects should be favoured (Brock & von Wangenheim, 2019).

Be Integral in the approach to AI projects, as these traverse across the organisations business, encompassing areas from strategy to culture. Without a firm handle on these constituent parts, managers will struggle to achieve success in these projects (Brock & von Wangenheim, 2019).

Be Teaming suggests the active partnering with technology companies in order to leverage these partnerships to success. External to partnerships within the business, the organisation should also partner across the value chain, developing the capabilities of the firm (Brock & von Wangenheim, 2019; Ransbotham et al., 2017).

Be Agile highlights the business requirement of adaptability in the changing business environment. Managers should ensure that the organisation can respond swiftly to external changes, leveraging the artefacts of dynamic capabilities in sensing and seizing new opportunities in the market (Brock & von Wangenheim, 2019; Teece et al., 1997).

Lead in the diffusion of a digital strategy in the organisation as executive management should be actively involved in the supporting and driving of these projects (Brock & von Wangenheim, 2019).

AI implementation however is not without its challenges. Taking into context the organisation's departmental segmentation, unique problems arise. In the field of human resources for example, the replacement of management judgement with AI can pose difficulties in the hiring process as decisions are based on historical data (Tambe et al., 2019) . Amazon failed in their attempt at this technology in 2018, as the AI learned from its historic data that the best employees to hire were white men (owing to the workforce currently being comprised in majority of the same demographic) , thus created a host of legal issues for the firm, and swiftly ending the project (Tambe et al., 2019). This case, sparks a debate on the integration of AI with the human workforce, questioning the boundaries of the technology and indeed how far should the decision-making capabilities be allowed to go. Mechanistic tasks such as data mining for instance have proved to be a forte of algorithmic based technologies with the field of marketing at the forefront of deriving business value (Tirunillai & Tellis, 2014). Through electronic responses given by customers, algorithms are able to understand customer satisfaction from large data sets and quickly assimilate a corrective outcome (Tirunillai & Tellis, 2014). In the pursuit of a digital strategy, managers are thus faced with decisions not only relating to the current implementation through extrinsic factors, but furthermore on the intersection of AI and intrinsic factors which subjugate the workforce's future. Indeed, the implementation of AI projects possess a risk to the organisational culture, in a similar fashion as general Big Data projects, suggesting a new way in which the workforce should be viewed, negating a breakdown in future performance.

In taking this concept further, Huang, Rust, & Maksimovic (2019) separate out the thinking tasks which humans have thus far prized in the work space, stating that the workforce is moving towards a "feeling economy"; that is a workforce centred around performing tasks requiring high levels of empathy, relationship building skills and communication. Similarly Metcalf et al., (2019), build on this concept by providing a clear distinction between explicit knowledge, knowledge that is available and which is easily codified, and tacit knowledge such as feelings and human intuition, which is thus far difficult to replicate in the field of AI. Their study suggests the creation of a symbiotic swarmed super intelligence amongst humans, linked through the use of AI which would offer greater organisational decision making capabilities, thus aiding rather than replacing humans in the workforce (Metcalf et al., 2019). Articulating these concepts into strategy poses a difficult challenge to managers, with well entrenched concepts, such as organisational decision-making structures, set to change. With an increase in decision making responsibilities passed from human to AI technologies managers should be cognisant of their position via actioning on the DIGITAL artefacts of “being integral” and “agile” (Brock & von Wangenheim, 2019). In search of a resolution to this quandary, Shrestha, Ben-Menahem, & von Krogh (2019) postulate a framework with three structural categories for decision making in an AI integrated organisation, as outlined below:

Full human to AI delegation which comprises a situation where decisions are made entirely by AI, however the human still remains responsible for the decision taken, as evidenced by traffic congestion systems (Shrestha et al., 2019). Ultimately this level of decision making in favour of AI is desired in time sensitive situations, where little value can be derived from human intervention (Shrestha et al., 2019).

Hybrid sequential decision-making structures is a symbiotic system where decisions are made sequential by humans and AI, with either starting the process. This approach provides the enhancing factor of each substituting for the others weaknesses, thus improving the overall efficacy of the decision making process and its outcomes (Shrestha et al., 2019).

Aggregated human-AI decision making structures consolidates the decisions made by both human and AI decision makers, providing each with a weighting towards the ultimate decision to be made (Shrestha et al., 2019).

Thus, to be “integral”, managers are required to shift focus to the implications of successful AI implementation and what this means for the organisation, its decision-making structures, and the workforce at large. With the rapid pace of development of digital technologies, the agility of the organisation remains paramount to its success and ability to derive value from their digital strategy (Tambe et al., 2019). Several artefacts of implementation in AI overlap with Big Data implementation due to the ultimate relationship to digital transformation, thus, to enable AI successfully, is to first implement Big Data (Brock & von Wangenheim, 2019; Gupta & George, 2016). These technologies can be seen as complementary, Big data analytics with

AI, Big Data with AI, with organisations aiming to develop their Big Data technologies to ultimately evolve into AI technologies (Metcalf et al., 2019). As the pinnacle of digital transformation, managers therefore should adhere to AI adoption principles, starting small and leveraging existing tools to develop the capabilities required to achieve their digital strategy.

2.10. Conclusion

With the current technological pace of development highlighted through IoT and social media, Big Data has come to the forefront of business, positioning itself as the key to unlocking financial performance (Ylijoki & Porras, 2016). Through leveraging data mining ,and the insights therein, these vast banks of data provide businesses the opportunity to develop new innovative products and services for consumers (Constantiou & Kallinikos, 2015). The process of reaching such success however, is marred with pitfalls ranging from executives' understanding of how to strategic leverage the available technologies, to using these technologies in a responsible manner (Gupta & George, 2016; Zwitter, 2014). Delving into these challenges presents OC as a prominent artefact in the success of an organisations ability to foster successful Big Data projects, wherein the current workforce remains at odds with the uncertainty presented in innovative technologies (Dubey et al., 2003; Gupta & George, 2016). Unfortunately for organisations still struggling to implement Big Data projects, newer technologies have now emerged evidenced by the rise of AI. By way of leveraging the repositories of Big Data in organisations, AI attempts to develop automated decision making, providing strategic advantages for organisations who can deploy these technologies with aplomb (Metcalf et al., 2019).

This literature review reiterated the importance of the BDAC model in a technology driven society, its constituent factors, and its importance in creating a competitive advantage for organisations. Further to this the researcher delved into the challenges and the current developments in the technology space, uncovering the symbiotic relationship of people, through organisational culture, and the technologies changing their work environment. Through the leveraging the BDAC model, this study sought to understand this relationship of OC and AI on FPer, delving into the entangled relationships between these artefacts.

Chapter 3: Research Questions

3.1. Research overview

Although numerous studies have been conducted on the value drivers of BDA projects, there is still a lack of consensus on how these artefacts interact with organisations to ultimately derive value for the firm (Côrte-Real et al., 2019; Gupta & George, 2016; Mikalef & Pateli, 2017). Some confusion exists between the technical precursors for value generation versus the organisational precursors such as OC (Côrte-Real et al., 2019). This study is aimed towards investigating the interplay of business value drivers for BDA through the base model suggested by Akter, (2016). Through the lens of the BDAC model, this study aims to understand the characteristics which affect the deployment Big Data capabilities, and thereby contribute to the understanding of the low success rate of these project implementations.

The study includes two research constructs, entangled with the constructs presented in the BDAC model (Figure 1 below) derived by Akter, (2016), which will be assessed. Taking a pragmatic approach to business value, the two constructs will be measure through FPer via financial return/market share. Thus, through the assessment of these constructs the study aims to provide practical recommendations for business to consider. The original model as defined by Akter, (2016) (Figure 1) derives BDAC as an aggregate of two sub dimensions conceived in second order and first order constructs. Using a base of sociomaterialism and RBV, the first order constructs present at the lowest and most voluminous level, with eleven constructs, These are subsequently aggregated into the three second order constructs, BDA infrastructure capability, management capability, and personnel capability (Akter et al., 2016). Finally, by aggregated these second order constructs into BDAC, the model encompasses organisational capability into this third order construct. To provide a lucid effect on business, the model shows BDAC as directly affecting FPer (Figure 1), at which point the research constructs of OC and AI are placed.

Thus, it is the aim of this research to investigate if the constructs of OC and AI influence the successful deployment of BDAC projects. The influence of each individual construct will therefore be analysed as well as the convergent relationships of these constructs, and moderating relationships.

3.2. Research Questions

The research questions posed in this study were separated into four distinct assessments, each with its own hypothesis as outlined below.

Research Question 1

Is there a direct positive relationship between BDAC and FPer?

This research question serves as a confirmatory analysis of the original relationship outlined by the BDAC model, which several studies have subsequently tested (Wamba et al., 2017; Y. Wang et al., 2018). This assessment was done through H_1

H_1: BDAC has a significant positive relationship with FPer

Research question 2

Is there a direct positive relationship between AI, organisational performance respectively (independent variables) and FPer (dependent variable)?

Big Data has been widely accepted as increasing FPer through value creation for business (Stanek, 2017; Wamba et al., 2015b; Wamba et al., 2017). Thus far the literature implies that OC is a barrier to entry for digital transformation projects with several authors stressing the importance of organisational acceptance of both Big Data and AI projects (Alharthi et al., 2017; Brock & von Wangenheim, 2019; Müller et al., 2018). Likewise ,through practical experience, managers have highlighted the link between OC and effective performance, attempting to leverage this to build increased performance in the firm (Lorsch & McTague, 2016; Mcafee & Brynjolfsson, 2012). In a similar context, AI has been earmarked as the key to unlocking FPer in the digital transformation space (Overgoor et al., 2019), however just 20% of AI projects have shown any value generation at this point. Thus, this question serves to investigate these constructs' relationships with FPer, under the pretext that BDA is present, but excluding BDA itself. This was addressed through **H_2** and **H_3**.

H_2: AI has a significant positive relationship with FPer

H_3: OC has a significant positive relationship with FPer

Research question 3

Is there a direct positive relationship between AI and OC(independent variables) on BDAC (dependent variable)?

The BDAC model highlights both the organisational and technological influences in its causal link in generating such capabilities (Akter et al., 2016; Wamba et al., 2015). AI requires sufficient success in a firms Big Data technologies as it learns from these technologies in its operation, however once implemented can form a complementary relationship to data analytics (Brock & von Wangenheim, 2019). Therefore, this research questions seeks to understand the influence AI and OC has on BDAC. Drawing from the original model this study tests the relationships of OC and AI on BDAC through **H_4** and **H_5**.

H_4: AI has a significant influence on BDAC

H_5: OC has a significant influence on BDAC

Research question 4

What are the combined impacts of OC and AI (independent variables) on the BDAC-FPer (dependent variable) relationship and do these independent variables moderate that relationship?

As evidenced by several studies, OC can be seen as a barrier to entry for the implementation of digital transformation projects, requiring management to shift focus onto this construct if project success rates in this field are to improve (Alharthi et al., 2017; Brock & von Wangenheim, 2019; Müller et al., 2018). Interestingly, since

AI is one such digital project which is enabled by robust Big Data capabilities in organisations, it stands as a technology which through algorithmic automation, can enhance BDA and ultimately decision making(Brock & von Wangenheim, 2019; Metcalf et al., 2019). This study investigating these relationships through the moderating effect of OC and AI on FPer using **H_6** and **H_7**.

H_6: OC has a significant moderation effect on BDAC enabling FPer

H_7: AI has a significant moderation effect on BDAC enabling FPer

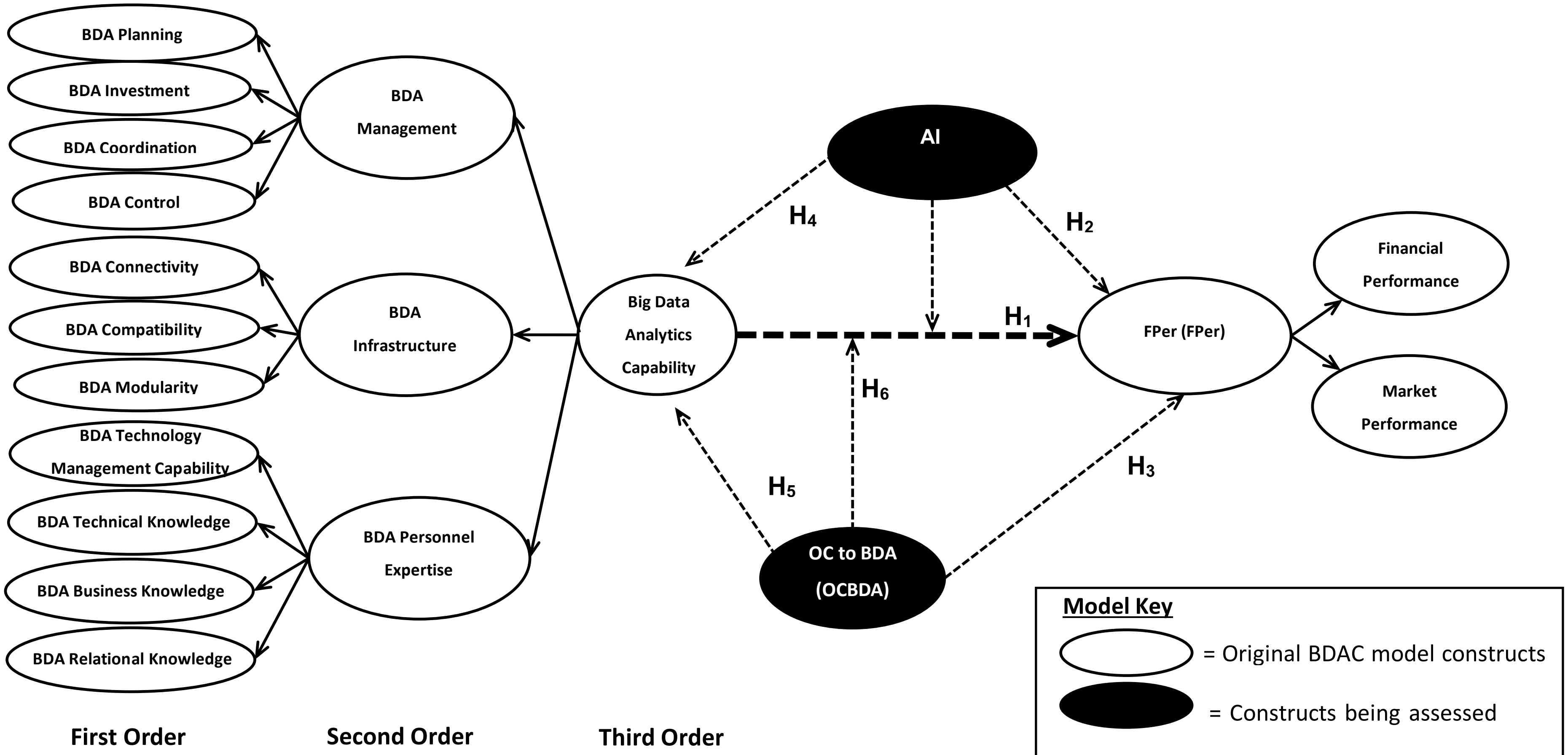


Figure 1: Research model applied in the study

Chapter 4: Research methodology

4.1. Research methodology and design

The objective of the study was to empirically examine the relationship between BDA and FPER as well as the moderating ability of OC and AI. The theoretical background provided in chapter 2 outlined the basis for development of the research questions presented in chapter 3. These questions established the research view to be used for evaluating the relationships between organisational culture, AI, BDAC and FPer. In addition, the base model's theoretical underpinnings were testing via BDAC's influence on FPer, as well as to ascertain moderating effects of OC and AI. Thus, the research employed a philosophy of positivism, where positivism relates to the way of life of a scientist, encompassing all philosophical ideals held by a scientist (Saunders & Lewis, 2018; Wamba et al., 2017). Using science as a grounding, the outcome of a positivist philosophy is defined as being "unambiguous and accurate knowledge using methods designed to yield pure data and facts uninfluenced by human interpretation" (Saunders & Lewis, 2018). Thus in this philosophical ideal, the world may be seen as based on lucid cause and effect laws, from which empirical testing may be conducting with logical and verifiable outcomes (Muijs, 2011).

Owing to the study applying well established theoretical lenses and constructs to test the relationships between variables, the research conducted may be described as deductive (Saunders & Lewis, 2018). The BDA-FPer relationship and the complexity involved in the entanglement of constructs as outlined by the BDAC model, was assumed (Akter et al., 2016; Wamba et al., 2017). Thus, owing to this entanglement view, and the numerous variables involved, not all of the parameters could be addressed in the study.

The study conducted structured research which was objective in nature, using measurable variables as is the assumption underlying the positivist philosophy. In turn this philosophy forms the basis of quantitative research, guiding the choice of research for the study. The currently accepted view that there is no difference in BDA implementations across organisations further negated the need for a qualitative study to be done, instead supporting the path of quantitative research (Gupta & George, 2016; Mikalef et al., 2019; Vidgen et al., 2017).

The quantitative approach has been described as “God’s view” in that it attempts to study things as they are in reality (Slevitch, 2011). Thus, it suggests one truth exists, and is scientific in its nature. Since quantitative methods adhere to a positivist philosophy, there are facts to be determined with no human intervention or bias to skew the results (Saunders & Lewis, 2018). Furthermore, the study sought to investigate ‘the facts’ through the relationships between variables, adhering to determinism, an ideal of the chosen philosophy. Determinism seeks to examine the interlinking relationships between constructs through the use of surveys or experiments (Cresswell, 2009). Saunders & Lewis (2018) further advise that the chosen philosophy dictates the research instrument to be used. Thus, adhering to this principle and the evidence given through prior studies in the field, a survey was chosen as the research instrument (Côrte-Real et al., 2017; Wamba et al., 2017). Adopting a survey strategy allows for population sampling, which placed the study as a non-experimental and explanatory research methodology (Zikmund, Babin, Carr, & Griffin, 2010). Unlike Wamba et al. (2017), who presented a two phase approach to administering a survey, this study used a single phase approach ,in part due to the constructs being tested through the BDAC model having a prior basis in literature, and also due to the time limitations for the study to be completed. To test the core constructs of the model, attitudinal predetermined questions were administered in a self-administered online survey (Toepoel, 2016). Due to the aforementioned time considerations of the study, and for the largest possible sample size to be achieved, this method of sampling was chosen. Due to the above considerations, the population was sampled cross-sectionally to gather insight into the constructs while implementing BDA and its effect of FPer, thus defining this study as using a single measurement analysis (Zikmund et al., 2010). This is outlined by Saunders & Lewis (2018) as an effective method to use as it limits bias in the results. Corresponding to the previous research done in the field, the constructs and model validity have been thoroughly tested and verified, thus this approach to analysis was deemed effective.

4.2. Population

Big Data in recent years has experienced intense growth in the business environment. Through the use of smart devices the process of ‘datafication’ has become ubiquitous, and thus defined the business landscape by an adoption of a digital strategy to gain competitive advantage (Constantiou & Kallinikos, 2015; Lycett, 2013). In choosing a population, the diffusion of Big Data was taken into account, as Zikmund et al., (2010) defines the population as “any complete group of entities that share some common set of characteristics” (p.387). Therefore, the target for the study was to reach businesses who actively used Big Data in their decision making. To target these businesses, however, would require the use of a census of the target population. This was deemed unfeasible in light of its cost and time to conduct, thus generalisation around the target sample was done. The study instead actively targeted users of Big Data who varied in their seniority and exposure to the field, from analysts to senior management. Although the product-oriented lens offers a definition of big data based on size, the rapid growth being experienced in the industry, along with advances in computing power, means that size limitations of Big Data will change (Bharadwaj, 2000; Sun et al., 2015). Thus, no limitation of Big Data based on the minimum data requirements were used, whilst the size of the company by employees was also not taken into account as the access to data by companies was now done through several channels, allowing for large storage capacity to be independent of company size (Bello-Orgaz et al., 2016; Lee, 2017). Akter (2016) limited his study to e-commerce and m-commerce industry in China, focusing only on the IT industry. This study widens the horizon of the population as Big Data has developed in more recent years to become pervasion across industries and businesses. Limiting the sample by industry would bias results towards historically technological capable businesses, excluding business who are attempting to implement Big Data today. Thus, to determine the relationships between organisational culture, AI, BDAC and FPer this generalised population sample was used, which would allow for a wider sample to be extracted and more realistic conclusion to be drawn.

4.3. Unit of Analysis

Zikmund et al. (2010) defines a unit of analysis as “what or who should provide the data and at what level of aggregation it should be analysed” (p.660). This is to determine the levels at which the data will be analysed, as is the case in this study, one could report the data at an individual or organisational level. However, since the study looked toward investigating organisational traits of OC and AI, from several distinct organisations, an organisational viewpoint was taken. Therefore, the unit of analysis for this study was proposed as the organisations which utilise Big Data.

4.4. Sampling method

In choosing the sampling method for the study certain considerations were made due to limitations imposed on cost and time to complete the research. Furthermore, since the proliferation of Big Data has grown in business, and continues to grow, the total population size could not be determined with any level of accuracy. Zikmund et al. (2010) confirms that researchers are bound by these conditions, which in turn present a trade off in accuracy. Probability sampling was therefore not considered in this study, as this technique requires the researcher to know the total population size and the number of individuals using Big Data (Zikmund et al., 2010). Based on the field of study chosen, and the constructs being evaluated, the researcher chose to approach the sampling method in two parts, both instructed by non-probability sampling.

Firstly, purposive sampling was used. This was due to the researcher’s background and professional network possessing numerous individuals who showed congruence to the Big Data environment, from analysts to senior management. An element of convenience sampling also occurred as the researcher entered a working environment in the technology field during the course of the study, thus enabling easy access to several individuals (Etikan, Musa, & Alkassim, 2016; Zikmund et al., 2010). Each individual was selected through their experience in the field of Big Data, whilst through professional networks, numerous large multinational organisations were represented.

Due to these professional networks possessing easy access to further individuals in the field, snowball sampling was introduced as the survey was shared. This was done through email and messages sent via networking applications allowing for improved dispersion of the and diversification of the sample by both organisation and industry. The field of Big Data is highly specialised with limited individuals being able to successfully participate in the study, thus through deep investigation of these networks, like minded and skilled individuals were able to participate (Saunders & Lewis, 2018). Unfortunately, through snowball sampling, the survey may reach individuals outside of the Big Data environment, who may in turn generate random sampling error through their participation (Etikan et al., 2016; Zikmund et al., 2010). To account for this the researcher added disqualifying questions to the survey, thus ensuring the sample stayed within the confines of Big Data.

Non-probability sampling however does have its drawbacks as snowball sampling introduces sampling bias via a lack of independence in the individuals and their networks (Zikmund et al., 2010). Like minded individuals tend to share the same traits and therefore the results of the survey may be skewed. This however was not a large concern as the researcher additionally made used of purposive sampling to mitigate this bias.

4.5. Sample Size

Zikmund et al. (2010) purports that the population size has no bearing on the minimum required sampled size of the study, but instead the variance inherent in the population has the largest effect. To achieve the desired accuracy, a larger sample size is required to generate a smaller confidence interval to be used in statistical testing (Bonett & Wright, 2015). A larger sample size directly reduces the sampling error, although this is often overlooked in business research, it is paramount to achieving a reliable understanding of the phenomenon under investigation (Zikmund et al., 2010) . Although statistical methods exist for calculating the required sample size for a study, these required accurate measures of the population itself, which in the case of Big Data and its growing proliferation in business, was deemed impossible to measure (Toepoel, 2016).

Tabachnick & Fidell (2013) suggest a formula for calculating the minimum sample size as N > 50 + 8M (where N is the sample size and M is the number of independent variables). This formula was chosen due to its significance in using multiple regression testing, which is one of the statistical tests relevant to this study (Section 4.X). Using the proposed model's independent variables of OC and AI, the minimum sample size may be calculated as 64. Section X.X provides an overview of the sample size obtained

4.6. Measurement Instrument

The chosen research philosophy, and previous studies, informed the research instrument used in this study, being a predetermined, self-administered, cross sectional survey (Akter et al., 2016; Saunders & Lewis, 2018; Wamba et al., 2017). The survey questions were informed by the research directive eluded to in chapter 3, and the prior research studies conducted on BDAC (Wamba et al., 2017). Owing to the study providing a confirmatory analysis of the BDAC model, prior studies by Akter et al. (2016); Dubey et al. (2003) and Wamba et al. (2017) were leveraged to includes questions which investigate the base model. Additional questions were posed for the constructs of OC and AI which are being tested in this study. A survey is used to generalise findings about the research population, whilst providing the data to test the constructs of the proposed model (Fricker & Schonlau, 2002; Steinmetz, 2016). Furthermore, surveys have numerous advantages due to geographic reach, speed, convenience, and low costs (Zikmund et al., 2010).

In order to evaluate the chosen constructs of BDAC, FPer, OC and AI, previously published and tested research questions were used. This was done in order to ensure objectivity in the testing, as well as congruence in the scales being used to measure each construct. As shown in Appendix A, the questionnaire used in the study was divided into seven sections, each collecting data for a distinct aspect of the study. The first section of the survey is titled "Context and respondent", which seeks to collect the demographic data of each respondent in the sample population. Demographics are essential as essential component to survey strategy as they allow for the collection of descriptive statistics as well as enabling analysis of the diversity in the population sample, thereby highlighting any possible sampling bias. As mention above, a screening question was also introduced to limit the population sample to individuals who are aware of and understand Big Data. This is turn controls for the quality of data received for the study, which will be statistically analysed in Chapter 5. Sections two to four then measure for the underlying BDAC model based on the research questions originally proposed by Akter (2016). Due to the aforementioned 11 first order constructs of the BDAC model namely, BDACon, BDAComp, BDAMod, BDAP, BDADM, BDACoord, BDACont, BDATK, BDATMK, BDABK and BDARK (Figure 1), these sections contained 43 questions. Sections five and six measured the constructs, added through this study, of OC and AI. Questions were derived by the researcher through the use of literature and to ensure an easy understanding of the constructs by the sample population. These sections contained a total of 15 questions, 7 relating to measuring OC and 8 towards measuring AI. This is in accordance with literature on the subject of survey strategy, specifying a minimum of 5 questions to data information on a construct (Zikmund et al., 2010). Likewise, section seven contained six questions relating to firm and market performance which were based on the aforementioned research on BDAC. In accordance with the scales used in studies by Akter et al. (2016) and Wamba et al. (2017), the researcher chose a seven point Likert scale across all sections, in the interests of congruence across the survey. These ranged from “strongly disagree” to “strongly agree” as well as “very positive to very negative”. It should be noted that section one did not use a Likert scale as this was deemed inappropriate, due to the data collection being of a descriptive nature. The total survey length was 65 questions was considered to be of an appropriate length to incite a good response rate. The questionnaire was then transcribed into SurveyMonkey, a popular online survey tool, for distribution ease and the researchers experience in using this tool beforehand. The tool features numerous analytic tools for administration purposes, and the ability to download the data into popular spreadsheet software programs.

4.7. Pre-testing

Survey pre-testing was done in accordance with typical survey strategies outlined in literature (Saunders & Lewis, 2018; Toepoel, 2016; Zikmund et al., 2010). Pretesting is a screening procedure using a test group of individuals to determine if the survey is appropriate, clearly specifies the purpose of the research, avoids ambiguity and generally misinterpretation of questions (Zikmund et al., 2010). There exists no general consensus on the appropriate size of a pre-test sample, however ranges consist of between 12 and 30 participants (Hunt, Sparkman, & Wilcox, 1982). The researcher thus chose a sample size of 15 individuals, as the study itself was not deemed to be to be "large". These individuals were professionals from the field of Big Data, having practical work experience in the field, with five participants also being MBA students to add further academic opinion.

Feedback was received through email, highlighting grammatical errors which were subsequently fixed. One notable comment was the length of the survey was "too long". This was noted, and the survey re-inspected, however each question was based on the constructs contained within the BDAC model and deemed to be of significance to the reliability of the data and statistical tests to follow. Further feedback was received on the construct definitions of Big Data and AI, which were simplified for the reader to easily understand and removing any misinterpretation.

4.8. Data gathering process

Using the aforementioned online survey tool, SurveyMonkey, allowed for administration to take place by separating campaigns for collection. Two such campaigns were created, one via a general electronic link, and another via a social media link. These were subsequently shared via email and WhatsApp to the researcher's networks, whilst the social media link was shared through Facebook and LinkedIn. Additionally, the researcher used a third-party provider, iFeedback, to collect data through a targeted database. This was done due to the time limitations of posed by the study, the rarity of Big Data specialists and the use of this method in a similar study by Wamba et al. (2017). The chosen database was the ICT database, isolating individuals in the field on technology in South Africa. Emails were sent out with links to the SurveyMonkey survey to allow for administration to be centralised in the online tool. A total of 1000 individuals were emailed the survey, with two follow up emails subsequently sent out on week two and four of the data collection campaign.

Additionally, to this the researcher sent out the survey to targeted Big Data groups on LinkedIn to further isolate professionals in the field. The data was gathered cross sectionally over the period 8th August 2019 to the 11th October 2019. The survey attracted a total of 190 responses over this period, with the average response taking

12 minutes. The data was then exported in XLS format into Microsoft Excel for further analysis to take place.

4.9. Analysis Approach

As stated above the data was collected through SurveyMonkey which provides a data file for further data interpolation. The data was quantitative, ordinal and categorical in nature. In congruence with Zikmund et al. (2010), the approach taken to data analysis followed the four step process of data editing, coding, data file preparation and data analysis.

4.9.1. Data Coding

The data collected throughout the survey was in symbol form thus, in accordance with Pallant, (2005) the data was coded to convert the symbols into numeric data. This was done to allow for statistical analysis to be completed.

4.9.2. Data Editing

Missing data is a common experience in quantitative research, where surveys are deployed as the collection instrument resulting in a reduced total number of participants (N) (Allison, 2002). In survey data collection, respondents may refuse to answer questions which they find sensitive (such as income), or may overlook questions due to human error (Allison, 2002). Missing data can be categorised into broadly two patterns of nonresponse, unit nonresponse when the respondent refuses to participate in the survey, and item nonresponse, where data collected is partially complete (Schafer & Graham, 2002). Furthermore, data can be deemed missing at random or non-random seen through three distinct methods (Holmes, 2014; Little & Rubin, 2014)

1. Missing completely at random (MCAR) – where the missingness of the data happens at random and has no reliance on the value itself. If this condition holds true, then the missing entries may be deleted without affecting the outcome (if the dataset is not significantly reduced)
2. Missing at random (MAR) – where the missingness of data is determined by the observed data and not the data which is missing i.e. this is also known as

an ignorable response. MAR is a characteristic which poses a risk to the propensity of the data.

3. Missing not at random (MNAR) – The missingness of data is dependent on the missing data itself, thus this affects the dataset and is nonignorable in its nature. These are the most likely to skew the dataset and affect results.

Due to the length of the survey being a key consideration during the pretesting phase, the researcher assumed that the missing data was due to MAR. Other considerations due to the relative complexity of the BDAC model was that certain respondents were not able to answer all of the questions posed, or simply did not have the time to complete the survey in one attempt, neglecting to return at a later time.
In deploying the MAR technique, the researcher considered five types of data imputation, each with their own characteristics which affect the overall accuracy of the dataset (Holmes, 2014).

1. Hot deck imputation – this involves substituting a missing variable with an observed variable within the same dataset.
2. Cold deck imputation - which involves substituting for the missing variable using an observed value from a different dataset.
3. Single imputation (SI) - which requires simulation of a single variable related to the missing value.
4. Multiple imputation (MI) – MI deploys simulation in a similar fashion as SI however involves averaging multiple values for the missing variable.
5. Plausible imputation (PI) - which uses homogenous propensity strata to compute the missing variable

In order to keep the uncertainty within the dataset at a minimum the researcher chose to use MI due to it providing the benefits of SI, whilst through using a larger sample set towards estimation of inputs and ease of deployment, allowed for flexibility in integration(Allison, 2002). Imputation can be defined as an estimate of missing data chosen to be the best representation of the respondent (Zikmund et al., 2010). Due to MI using a representative sample of the respondent from the dataset, not only is the best guess chosen, but the statistical noise is also retained, removing bias and thus achieving a robust sample for hypothesis testing (Allison, 2002) MAR was thus computed using MI by using respondents with a response rate above 50% and below 100% (Allison, 2002; Schafer & Graham, 2002).

In accordance with MI, the researcher divided the sample by industry into sub-groups, calculating the mean values of each question. Imputation was then done to each respondent which met the minimum criteria, thereby increasing the sample size for statistical testing (Holmes, 2014).

4.9.3. Statistical analysis background

The data collected through the survey was categorical for section 1, whilst the remaining sections provided ordinal data. Therefore section one was employed for descriptive analysis of the population sample, whilst sections two to seven provided for inferential statistics to be done by treating the data as continuous due to the use of a seven point scale (Cresswell, 2009; Pallant, 2005; Zikmund et al., 2010). As described above, missing data was handled using MI ensuring that the final data used was complete. Following the data editing using Microsoft Excel, the data file was then saved in XLS format and uploaded to SPSS.

4.9.4. Descriptive statistics

Descriptive statistics summarize the characteristics of the data and set out the data in an easily consumable format (Zikmund et al., 2010). The researcher has included the raw data and results post imputation into the descriptive statistics as outlined in chapter 5. This consisted of reporting the initial total respondents, the percentage complete prior to imputation, and the contribution to the dataset applied. The descriptive statistics analysis was done on the post imputation results which contained 100 participants. As the data collected was both categorical and ordinal in nature, these were analysed separately. Categorical data collected in section 1 of the survey was analysed using frequency and percentage frequency to describe the characteristics of the data. Likewise, the ordinal data from section 2 to section 7 was analysed through the mean, standard deviation, skewness and kurtosis of the data. All descriptive statistics were done through IBM SPSS.

4.9.5. Shapiro Wilk test for normality

In order to conduct statistical analysis on the dataset, first a clear understanding of the nature of the data is needed. Normality is central to this decision as normally distribute data may undergo parametric statistical testing, whilst skewed data which is not normally distributed will undergo non-parametric testing (Singh, 2007). Although this study conducted normality tests through skewness and Kurtosis, the Shapiro Wilk test for normality was deployed to confirm the results as well as provide a statistical verification to the assumption of normal distribution (Shapiro & Wilk, 1965). Thus, through the Shapiro Wilk test, using $p<0.05$, if the value generated was less than 0.05, the null hypothesis would be confirmed, and the data would therefore not be normally distributed (Singh, 2007).

4.9.6. Cronbach's Alpha – Internal reliability of data and constructs

Central to a research study which deploys a research instrument is the test for validity and reliability of said instrument (Tavakol & Dennick, 2011). As one of the most widely used measures for reliability, Cronbach's Alpha describes the reliability of a sum of measures, such as the measures deployed in the survey instrument in this study (Bonett & Wright, 2015). It tests the consistency of measurement from one measure to the next where this internal consistency purports the degree of error is the measurement instrument (such as the survey in this study) inability to return the same results when all other conditions held constant (Cronbach & Shavelson, 2004). Thus, the questions in the measurement instrument should correlate to each other where the same underlying concept is being measured (Bonett & Wright, 2015).
This internal consistency provides validation that the measurement instrument used was effective and is represented by the following equation;

$$a = \frac{k}{k-1}\left(1 - \frac{\sum s_i^2}{s_T^2}\right)$$

Where k represents the number of conditions being assessed from the measurement instrument, s_i^2 is the standard deviation of each of these conditions and s_T^2 is the total standard deviation of all conditions (Cronbach & Shavelson, 2004). Cronbach's Alpha is most typical used in behavioural and organisational studies where questions are posed in survey's by the use of Likert scales, as was the case in this study (Bonett & Wright, 2015; Cronbach & Shavelson, 2004).

The researcher has thus approached determining the internal consistency of the constructs by firstly using the calculated Cronbach's Alpha as correlation coefficient and examining the results, and secondly by confirming these results via the use of factor analysis. As Cronbach's Alpha is suited to Likert scale data (ordinal data) and not categorical data, section 1 was excluded from these tests. Section 2 to section 7 was however examined with results shown in Chapter 5 table 3. This allowed for questions to be examined in the consistency where alpha's < 0.7 would be deleted resulting in an iterative analysis process.

4.9.7. Internal validity of data and constructs

According to Zikmund et al., 2010, internal validity can be determined by the extent to which the independent variable creates any variance in the dependent variable. Validity itself is both an internal and external construct where external validity refers to the world outside of the experiment, whilst internal validity is internal to the experiment (Druckman, Green, Kuklinski, & Lupia, 2011). For research to be extended beyond the confines of the study being conducted, it needs to first show internal validity, that is manipulations were carried out successfully to show a difference in independent variables (for a Likert scale these would be answers which differed from the average value) (Druckman et al., 2011; Zikmund et al., 2010). The importance of internal validity cannot be understated as without sound validity the researcher cannot make verifiable conclusions to the research and thus cannot extrapolate the findings to the external environment (Zikmund et al., 2010). The researcher thus chose to conduct validity testing in two ways, firstly by using confirmatory factor analysis and secondly by conducting bi-variate correlations between survey questions and the construct mean value for which they are measuring as shown in Appendix C, table 17.

4.9.8. Factor Analysis

Factor analysis is a group of statistical methods which aim to reduce the observed variables of a study into a smaller group of factors through data reduction (Nunnally & Bernstein, 1994). This reduction is done by leveraging statistics to quantitatively identify the most prominent factors from the larger group, allowing for the emergence of latent constructs for the researcher to examine (Zikmund et al., 2010). The factors themselves are distinct from one another allowing for the assessment of construct validity to arrive at a set of parsimonious factors through simplification (Pett, Lackey, & Sullivan, 2003). Once the reduction in variables occurs the researcher is faced with a far simpler task in examining the interrelated constructs, as variables with the greatest variances are left to be analysed (Pett et al., 2003).

Factor analysis contains two distinct statistical methods namely, confirmatory factor analysis (CFA) and exploratory factor analysis (EFA). EFA is performed when the number of factors which may exist in a given set of variables, is unknown, thus the researcher uses this method to explore the underlying factors and thereby ascertaining the interrelation between the underlying variables (Pett et al., 2003; Zikmund et al., 2010). Conversely, CFA is used when, through the researcher's understanding of the theory, the number of factors are known (Nunnally & Bernstein, 1994; Pett et al., 2003; Zikmund et al., 2010).

4.9.9. Assessment by CFA

Through the use of Cronbach's alpha, internal reliability of the constructs were tested. CFA was then used as a confirmatory method by deploying AMOS and building a standardised estimate model for assessment (Shek & Yu, 2014). Each first and second order construct was tested individually in independent models, allowing for the confirmation of factors and the internal reliability assessed in Cronbach's alpha section 4.9.8. Through this process, questions which did not load sufficiently were removed from the dataset, allowing for a parsimonious model to be developed in conjunction with the research objectives of this study.

4.9.10. Assessment by EFA

To deploy EFA a few basic assumptions need to be met. Firstly the dataset itself is assumed to have underlying factors which are less than the number of observed variables (Nunnally & Bernstein, 1994). Due to EFA additionally leveraging the Pearson product moment correlations, the underlying assumptions of this need to be met. This includes continuous distributions, which the study meets through its use of Likert scales, linear relationships, which was shown through scatterplots in Appendix E and Appendix I, holds true (Nunnally & Bernstein, 1994). Finally a large dataset is required, of which this particular study could only gather 100 participants, thus forming a limitation (Nunnally & Bernstein, 1994).

The researcher in accordance with Zikmund et al., 2010 and Nunnally & Bernstein, 1994 applied EFA per section of the survey, thereby investigating the configuration of the underlying factors. Following this assessment the factor loadings were taken into account, identifying which factors did not load sufficiently and thus could be removed (Nunnally & Bernstein, 1994; Zikmund et al., 2010). Varimax rotation was then applied with a maximum convergence set to 25, followed by principle component analysis with the eigenvalues set to greater than 1 to identify the correct factor reduction (Nunnally & Bernstein, 1994).

4.10. Examining the relationships between constructs

To examine the relationships between constructs, the researcher employed correlation coefficients. A simple correlation coefficient can be defined as measuring the covariance between two variables (Zikmund et al., 2010). There are several types of correlations, namely; intraclass correlations, product moment correlations and rank correlations, however for the purposes of the linear regression being applied in this study, the Pearson product moment correlation was chosen (Kraemer & Blasey, 2016). In order to deploy this correlation the assumption of bi-variate normal distribution needs to be met between the X and Y variables (Kraemer & Blasey, 2016).

The Pearson product moment correlation is represented by the following equation:

$$r = \frac{n(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{\sqrt{[n\sum x_i^2 - (\sum x_i)^2][n\sum y_i^2 - (\sum y_i)^2]}}$$

Where r is the Pearson correlation coefficient, n is the sample size and x_i and y_i are the sample variables.

When the variables are continuous in nature, as is the case with this study, then the Pearson correlation or simple correlation is appropriate (Zikmund et al., 2010). The coefficient is measured between values of -1.0 and +1.0 where +1.0 denotes a perfect positive relationship between variables and -1.0 represents a perfectly negative relationship. If the correlation reaches a value of 0 no correlation exists between the variables (Zikmund et al., 2010).

In order to effectively test the hypotheses proposed on Chapter 3 of this study, Pearson correlations were run to indicate the relationships between variables. The study conducted all tests at a 95% confidence interval.

4.11. Hierarchical Multiple regression

Although correlation and regression are mathematically similar, correlation measures interdependence of variables whilst regression measure the dependence of variables (Schumacker, 2015; Zikmund et al., 2010). Regression measure the linear relationship between the independent and dependent variables, as it attempts to predict the values of dependent variable using values from the independent variable (Schumacker, 2015; Zikmund et al., 2010).
Linear regression is represented by the following equation:

$$Y = a + bX + e$$

Where Y represents the dependent variable, X represents the independent variable, b is the coefficient and e is the prediction error.

For this study however, the researcher aimed to predict one dependent variable (FPer) from the data provided by two independent variables (OC and AI), thus multiple regression testing was used. This type of regression analysis would allow for two or more independent variables to be tested against the single dependent variable (Schumacker, 2015; Zikmund et al., 2010).
Multiple regression testing is similar in its mathematical equation to linear regression and is represented as follows:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + e$$

All variables hold the same context as the linear regression equation, however with multiple regression there may be two or more independent variables.

The researcher chose to use multiple regression analysis through hierarchical multiple regression (HMR) due to the control provided over a stepwise regression (Pallant, 2005). Instead of the statistical program deciding the variable order, HMR allowed the researcher the freedom to develop an ordered model, adding or removing variables in each subsequent model to reach the maximum possible total variance explained (Pallant, 2005). Additional variance explained by each independent variable can be interpreted as the moderating effect these have on the BDAC-FPer relationship.

The statistical program SPSS was used to calculate the HMR with $p > 0.05$ and FPer set as the dependent variable.

4.12. Limitations

Notably for this study is the use of correlation coefficients to answer the first three research questions. Correlation coefficients assist in identifying if a relationship exists between two or more variables, and the direction, positive of negative, but cannot predict the causation therein (Hair, Black, Babin, & Anderson, 2014; Pallant, 2005; Tabachnick & Fidell, 2013). Thus, inferences can be made about the possible reasons for the strength of the relationships, or lack thereof, but no conclusive insight regarding causality may be given.

Secondly, the dynamic business environments which organisations are currently competing in are in a state of fluctuation, with new technologies arriving at an accelerated pace (Markow et al., 2017; Mikalef et al., 2018). Ideally this research should investigate the BDAC environment through a longitudinal study, examining the changes in the effects OC and AI have on FPer. This would allow the changes in the business environment to be encapsulated within the study, however, due to the time constraint under which this research is required to be completed, a cross-sectional study was pursued.

Lastly, the construct of AI and the subsequent questions posed to the respondents were conceived in the assumption of a base level of understanding of the technology. Although a definition of the construct was supplied in the survey instrument, the construct itself and its related technologies, are a specialised field, with far less prominence than big data. With this in mind, the limited understanding of AI poses a limitation in the quality of data which could be gathered.

Chapter 5: Research results

5.1. Introduction

This chapter presents the results from the data collection and statistically analysis completed, as outlined in chapter four. Firstly, the descriptive statistics will be outlined, providing context for the sample and describing the data from the survey. This will then be followed by the inferential statistical testing, addressing the hypotheses outlined in chapter three.

5.2. Descriptive statistics

5.2.1. The research sample

From the seminal work done in BDAC by Akter et al. (2016) and Wamba et al. (2017), the researcher targeted the data sample of 250 participants in an attempt to achieve congruency between these studies. Tabachnick & Fidell (2013) however suggest a formula for calculating the minimum sample size , as outlined in chapter four, where this was calculated as being a 64 participants.

The researcher however was unable to meet these requirements due to the specificity of the specialisation of Big Data, where relevant respondents where difficult to find. The timeline given for completion of the study also proved to be a limiting factor, as data gathering required extension time in searching for the required participants.

5.2.2. The response rate

Due to the chosen methodology, the response rate of the survey could not be accurately determined. The survey was distributed online through social media platforms, with particular interest placed on “Big Data” related social media groups. Furthermore emails were sent out to the ICT database as provided by a third party research provider who assisted with administration of the SurveyMonkey survey. Since several methods of survey distribution were used, no scientifically quantifiable number for response rate can be reliably stated, however the researcher believes that the overall response rate could be characterised as “Low”. A total raw sample size of 190 participants was achieved, however this was reduced due to the qualifying question posed in the survey of “*Are you aware of or associated with a BDA Capability within the organisation being described in this questionnaire?”.* As outlined in chapter four, and to follow below, MI was used to extend the usable sample size for statistical testing. Ultimately the sample reached 100 participants who completed a satisfactory amount of the questionnaire for statistical testing to be done.

The geography of the study was centred around South Africa, focusing on technology related industries. This took a similar approach, as taken by Akter (2016), who looked for deployed a research company to administer the survey targeted at individuals who belonged to specific groups including ‘business analysts’, ‘IT professionals’ and ‘big data analytics’.

5.2.3. The total sample

With every effort made in pursuit of completed responses, this study’s total sample size was 190 respondents, far below the studies on which it is based. The usable sample size however was deemed to be 100 respondents after MI was used. This was above the targeted sample of 64 participants as seen in table 1.

Table 1: Summary of data collected and imputed

Respondent breakdown	**Number (Total Data Set)**	**Percentage Total Data**
Total respondents	190	100.00%
Total respondents with over 50% questions completed	100	53%
Total respondents’ data Imputed	12	12%
Number of Total Potential Answers	6500	100.00%
Total number of Answers Imputed	308	4,7%

The data includes the use of MI which allowed for the completion of questionnaires which were between 51% and 99%. The total number of potential answers for the sample population was taken as 6500 data points (65 questions X 100 participants).

Of this, 12 respondents were deemed necessary to use MI on, which brought the total sample size from 88 participants to 100 participants. A total of 308 answers were imputed, which represents 4.7% of all answers from the 100 participants. Two participants answered 15 and 16 questions respectively and were therefore below the 51% completion rate necessary for MI and were therefore left out of the testable sample.

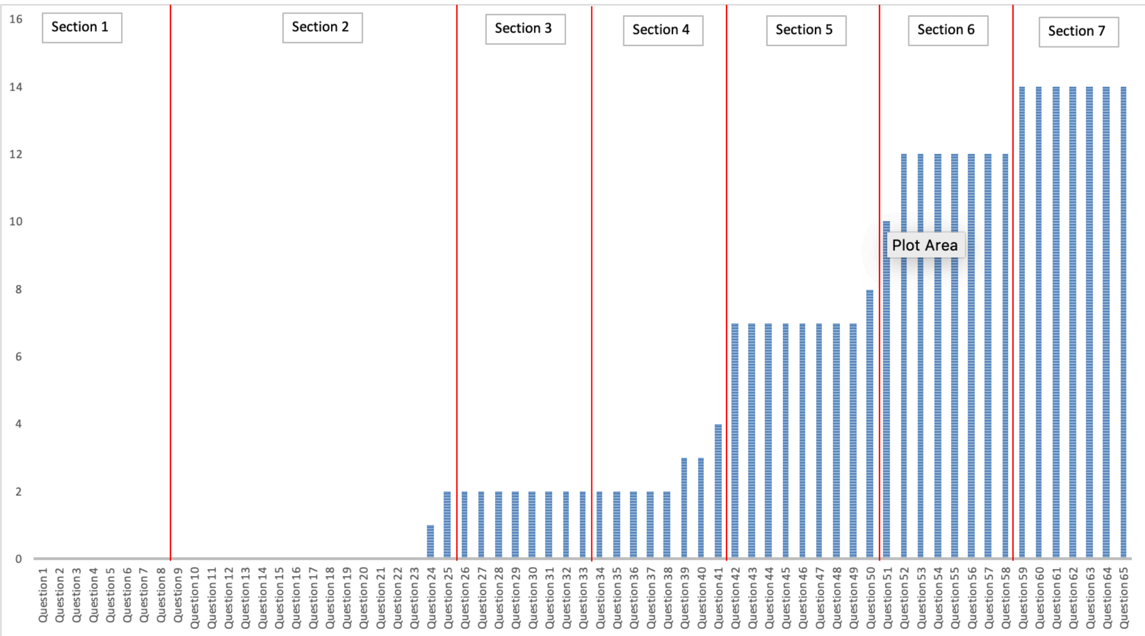


Figure 2: Number of missing responses per question and section of survey

Figure 2 depicts the same imputation information via the number of missing responses per question and section within the survey. Section one was fully answered with zero missing data. The first missing answer appears in section two question 24 and continues to grow through to section seven which holds the highest number of unanswered questions at 14. This is made up off the aforementioned 12 participants whose questionnaires were completed through the use of MI, whilst the two remaining participants were deemed to have answered too few questions to be imputed. Feedback from the survey testing as outlined in chapter four, showed that several individuals commented on the survey length, stating that the survey was too long. This seems to have realised itself in the steady increase in the number of missing participants through the sections of the survey. Participants seemed to stop completing the questionnaire, the further they proceeded, which could be directly attributable to the 65 questions posed.

Table 2: Summary of missing responses per section

Section Detail	**Question Range**	**Missing/Imputed Responses**	
		Number	**Average**
Section 1: Context of Organisation and Respondent	1-8	0	0.00
Section 2: BDAMC	9-25	3	0.176
Section 3: OC	26-33	16	2
Section 4: BDAIF	34-41	20	2.5
Section 5: BDAPEC	42-50	64	7.1
Section 6: AI	51-58	94	11.5
Section 7: FPer	59-65	98	14

Table 2 summarises the number of missing responses per section of the participants who were imputed. Section seven has an average of 14 missing responses, drawing the largest number of imputation. The missing data grew steadily through section 2 to seven, showing no significant spikes, therefore it can be surmised that no single construct caused the incompletion of surveys.

5.3. Descriptive statistics of the respondents

Section one of the survey contained a total of 8 descriptive questions to be answered allowing for the analyses of the sample via groupings for subsequent testing to be understood. The initial question also acted as a disqualifying question by way of checking the respondents awareness of the topic of Big Data. Thus all 100 respondents being used in this study answered “Yes” to this question, with a further 88 answering “No”. Further descriptive questions interrogated the respondent’s for their demographic information such as age and gender, whilst descriptions of their primary industry, organisation’s size, seniority, and geography were also pursued.

5.3.1. Age of the population

The respondents were primarily distributed between three age groups namely, 25 – 34, 35 – 44 and 45 – 54, each making up 35%, 31% and 24% of the sample respectively. This is depicted in figure 3 below

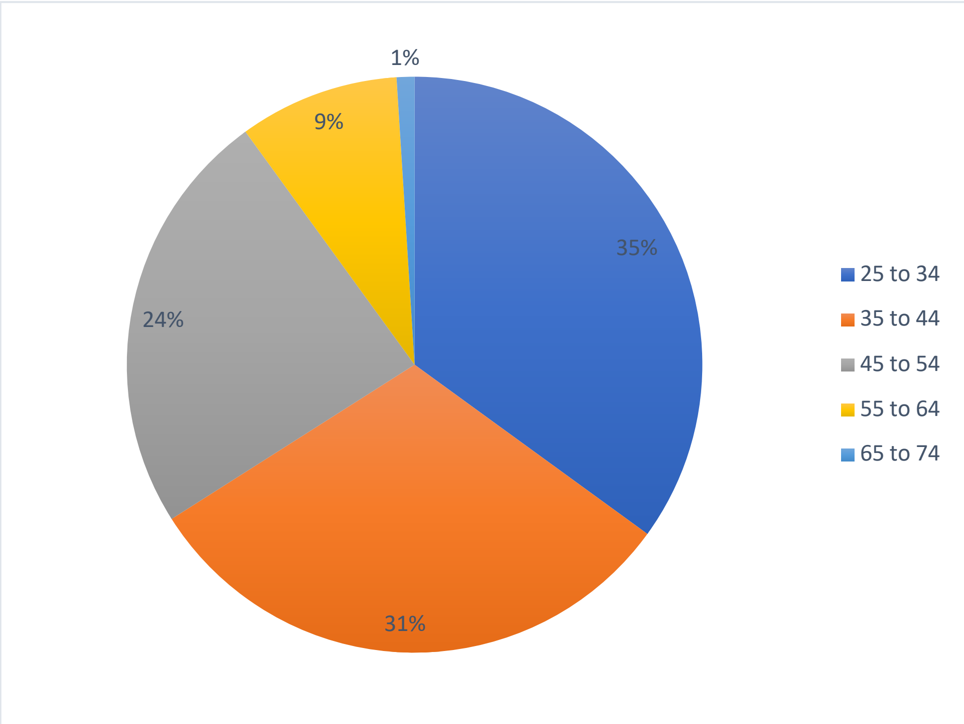


Figure 3: Age of the population

5.3.2. Gender

The survey was dominated by male respondents as shown in figure 4 below. Male respondents made up 69% of the sample, whilst females accounted for 31%.

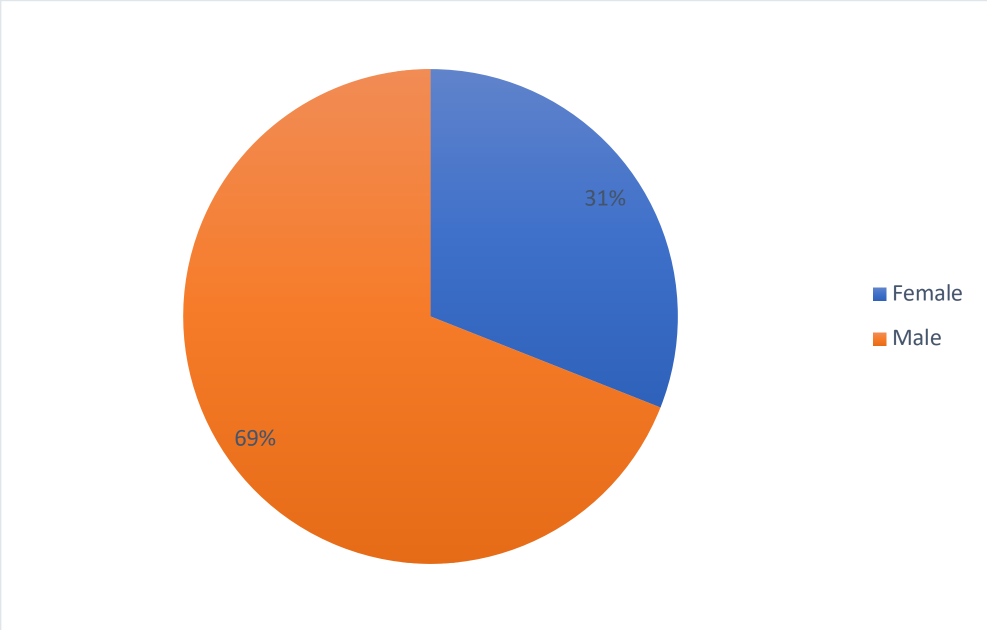


Figure 4: Respondents by gender

5.3.3. Geography

This study was centred in South Africa, with the researchers primary networks being placed in country. Thus, the survey itself was unsurprisingly answered primarily by individuals in South Africa, regardless of the use of social media groups in an attempt to achieve distribution beyond the country.

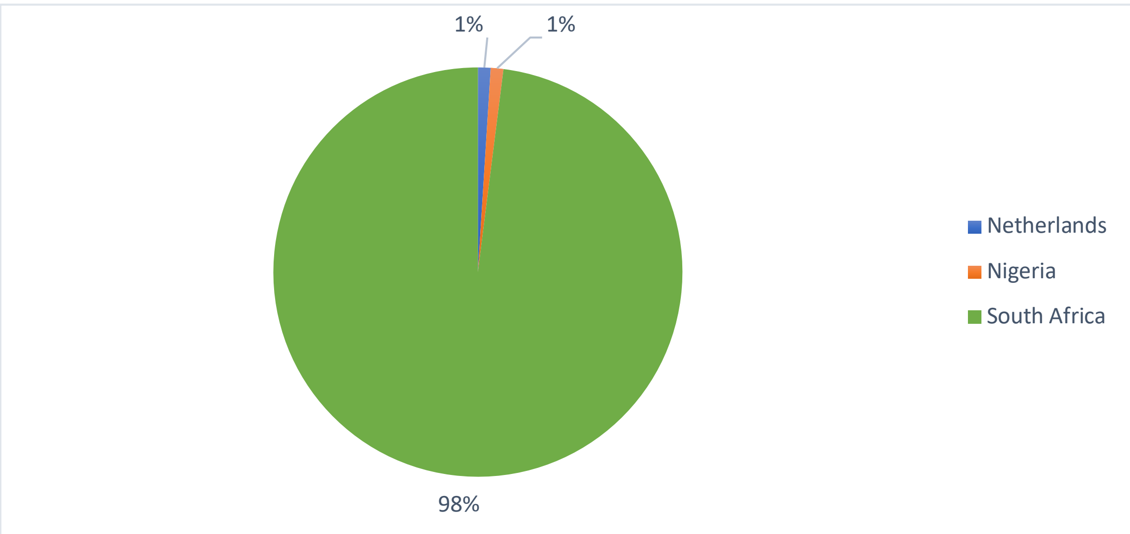


Figure 5: Participants by geographical location

Figure 5 shows the distribution of the participants by geography. South Africa remains the dominant country of origin with 98% of the respondents residing there. A further 1% of respondents are based in Nigeria and the Netherlands respectively.

5.3.4. Occupation level of respondents

The occupation level of the respondents consisted predominantly of owner/executive/c-suite level individuals representing 35% of the survey as depicted in figure 6. This was closely followed by middle management and senior management with 32% and 28% of the sample respectively. The higher level of seniority of the organisations shows the survey requirements of awareness and experience in the Big Data environment attracted more a more senior sample.

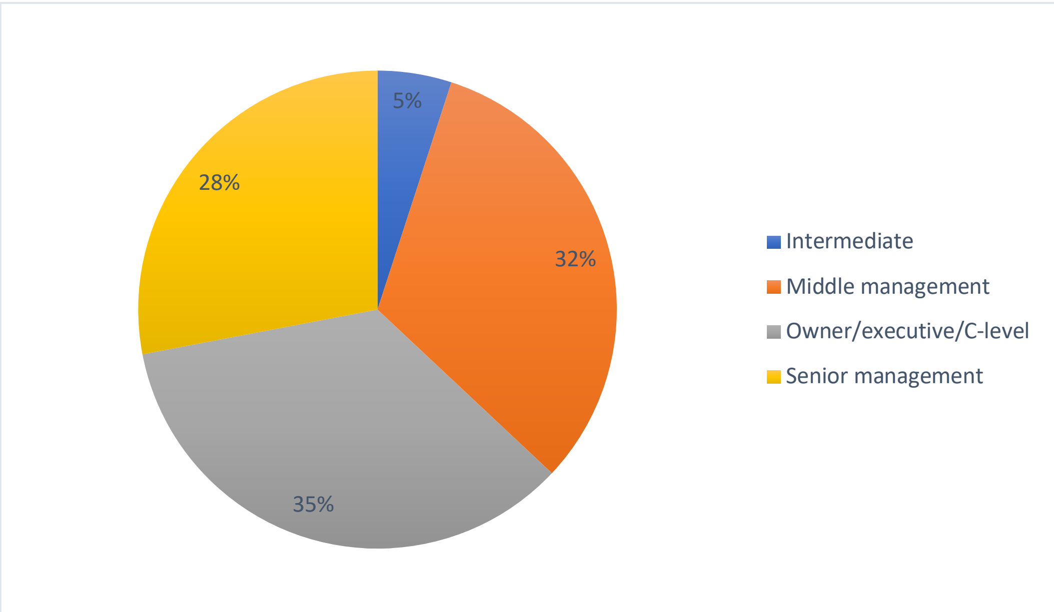


Figure 6: Summary of occupation level

5.3.5. Association with Big Data

Participants showed an overwhelming association to Big Data via the category of user of analytics within business representing 47% of the sample. IT systems or infrastructure represented 23% of the survey whilst participants who did not fit into the predetermined descriptions assigned themselves to other.

As the study focused on the business implications of Big Data, with strategic decisions being made to derive value from the technology, it is interesting to note to the high percentage of users of analytics which inevitable drive data driven decision making. Coupled with the results from respondent's occupancy levels being dominated by executive/owner/c-suite level employees, participants of this study are skewed towards senior employees, who make use of Big Data through business analytics.

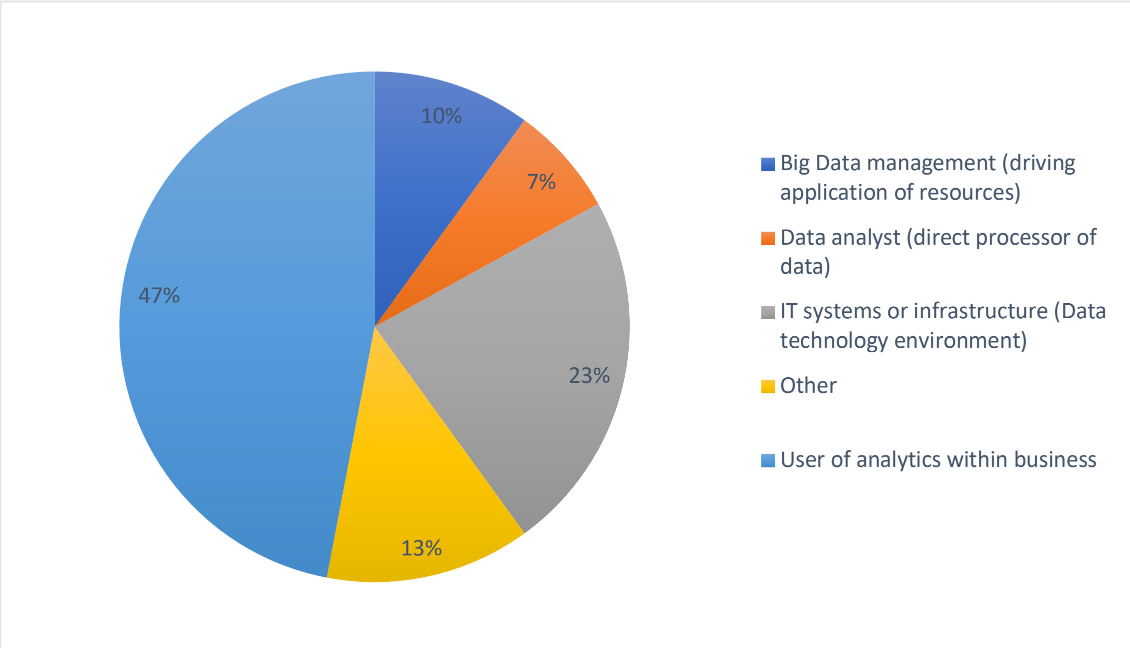


Figure 7: Summary of respondent's association with Big Data

5.3.6. Industries contained within the survey

The industries which dominated the survey, highlighted by figure 8 below, were telecommunications, technology, internet and electronics as well as finance and financial services, representing 42% and 19 % of the sample respectively. This was expected through the similar studies done by Akter et al.(2016) and Wamba et al. (2017) who made similar findings, where their highest industry was the 'information and communication' industry. For South Africa however, the second most dominant industry differs from these studies, being represented by the finance industry. Due to South Africa having a well-developed financial sector, the results are expected, whilst further studies do advocate the use of Big Data in the financial sector as having become pervasive since the introduction of the technology. Interestingly the third highest industry was represented by the business support and logistics industry, manufacturing and utilities and energy making up the fourth most dominant industries with just 4% of the sample each.

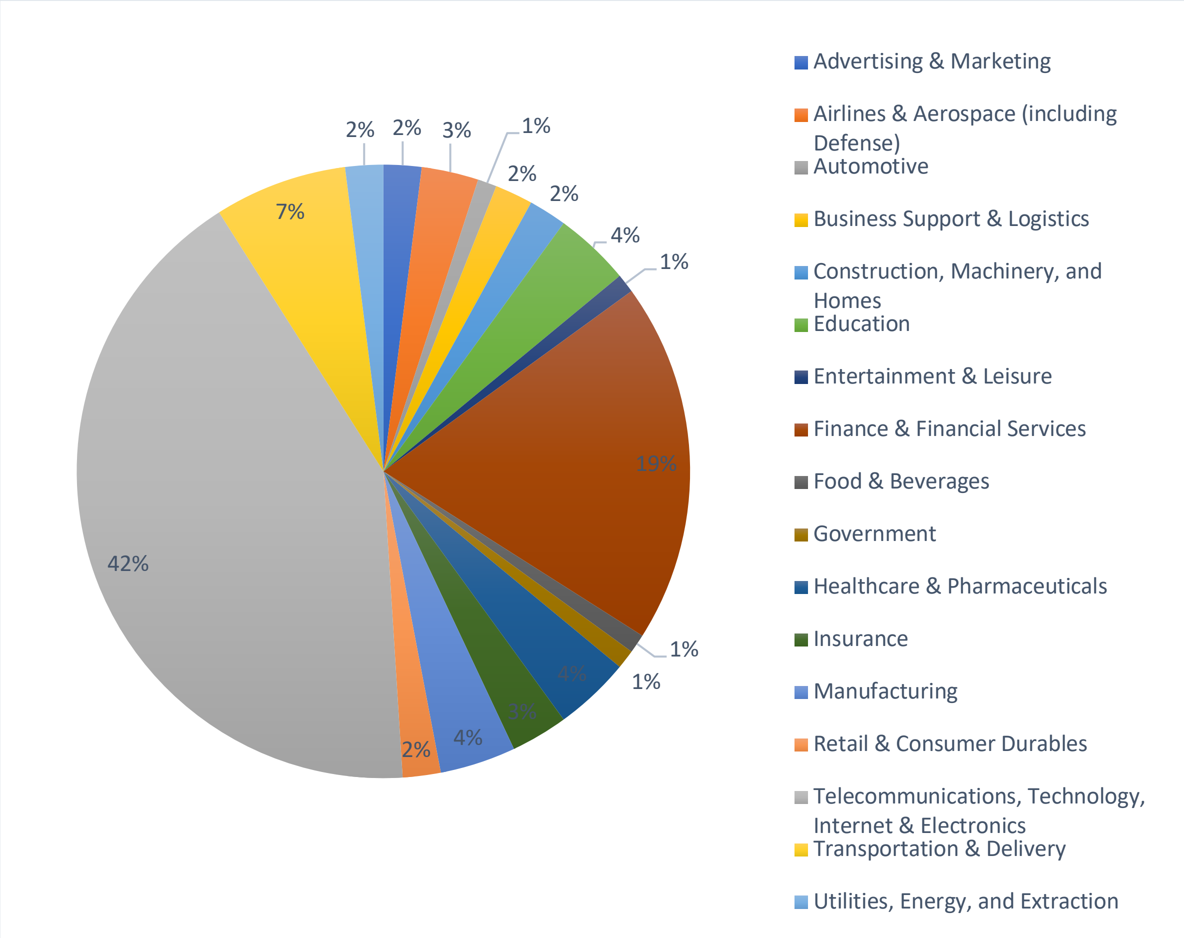


Figure 8: Participants by industry

5.3.7. Organisational size

The organisations represented by the sample data were predominantly represented by larger organisations, with 47% of the sample containing 1000 or more employees as illustrated in figure 9. A further 2% had 500 – 999 employees whilst smaller organisations with headcounts of 1 - 99 and 100 - 499 represented 29% and 22% of the sample respectively. Due to the size of data necessary to constitute Big Data, it was expected that the sample would largely encompass organisations with higher headcounts. The high investment and further resources needed in deploying these Big Data technologies also poses a limitation on smaller organisations, thus leaving larger organisations to lead the way in Big Data.

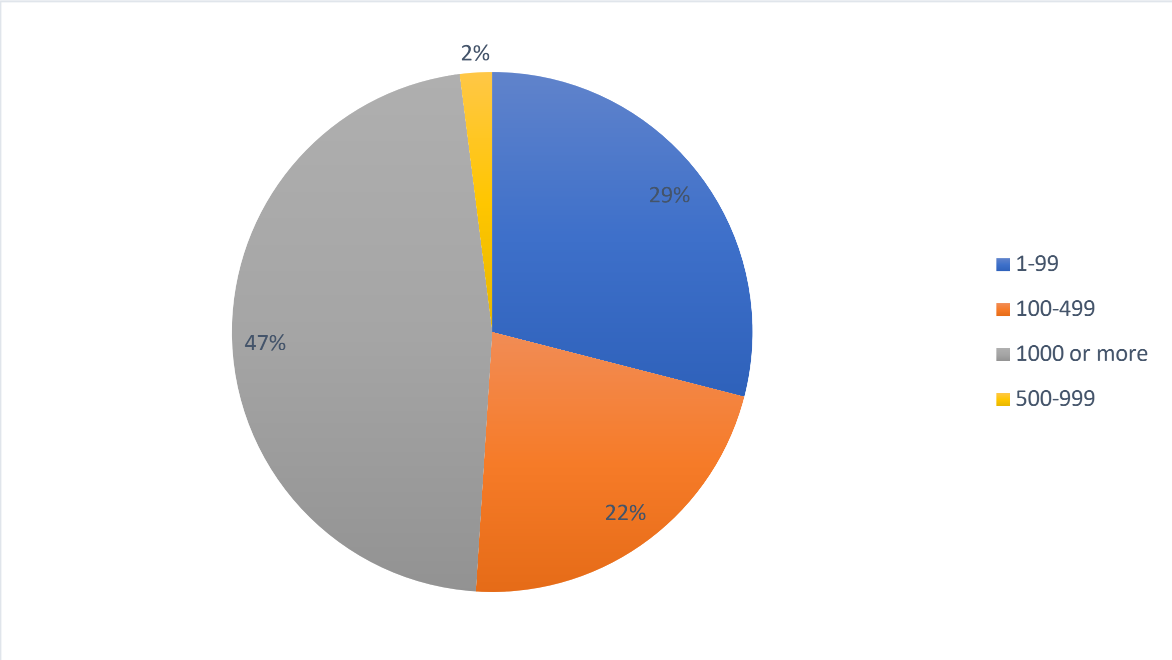


Figure 9: Summary of organisation size

5.4. Validity

Validity seeks to test the differences in scores achieved by respondents. This was tested in the study by Pearson correlations between each question and the construct total representative of that question. The results may be seen in Appendix C. All correlations were found to be significant, each varying between 0.355 and 0.930, all of which were above the recommended value of 0.3.

5.5. Reliability

Cronbach's alpha was used at a per question level, resulting in all sections being deemed a reliable measure except for AI (table 3). AI showed an initial Cronbach's Alpha of .336. The individual questions were then assessed and item AI1, AI5, AI7 and AI8 presented Cronbach's Alpha values below the threshold of 0.7.

Although it would be correct to remove these items from the construct and retest using Cronbach’s Alpha for the improvement in the measure, the researcher decided to further test the reliability through the development of CFA models. This would allow for verification of the Cronbach’s Alpha results, and the isolation of the questions which may need to be removed to improve the validity of the model.

Table 3: Reliability of first order constructs

Constructs	Cronbach's Alpha
Planning	0.894
Investment Decision Making	0.903
Coordination	0.896
Control	0.885
Organisational Culture	0.891
Connectivity	0.890
Modularity	0.906
Technical Knowledge	0.897
Business Knowledge	0.894
Relational Knowledge	0.894
AI	0.911
FPER	0.905

5.6. Factor Analysis

Factor analysis provides a more reliable measurement of scores through its measurement at a factor level, in contrast to those achieved at an individual item level (Tabachnick & Fidell, 2013).

As outlined in section 4.9.8 factor analysis consists of two principal techniques, namely; confirmatory factor analysis and exploratory factor analysis. Due to this study achieving a relatively low number of responses (N=100), it is pertinent in the minimising of measurement error to use EFA as a complementary test in conjunction with CFA (Hair et al., 2014). Thus, this study first deployed CFA to decipher the factors discriminant and convergent reliability (Hair et al., 2014).

5.6.1. Confirmatory factor analysis

CFA was conducted across all BDAC second order constructs along with the additional constructs posed in this study of AI and FPer, as displayed in table 4 below, with full results including the standardised factor loadings presented in Appendix D.

Table 4: Summary of CFA constructs

Scale	IFI	CFI	RMSEA
OC	0,888	0,886	0,151
FPer	0,943	0,942	0,16
AI	0,905	0,894	0,067
BDAMC	0,869	0,865	0,104
BDAIF	0,993	0,993	0,035
BDAPEC	0,911	0,909	0,165

Each second order construct was represented by an individual CFA model, to further understand the validity achieved in section through the use of Cronbach's Alpha in section 5.4.

The incremental fit index (IFI) of each model was assessed against the boundary value of 0.9, where models >0.9 are assessed as having a good model fit. With this in mind, the models of OC and BDAMC were deemed to not have a good model fit with value of 0.888 and 0.869 respectively. Good model fits were found for FPer, AI, BDAIF and BDAPEC with values of 0.943, 0.905, 0.993 and 0.911 respectively.

Similarly, examining the comparative fit index (CFI) values, good model fits were found for FPer, BDAIF, and BDAPEC with values of 0.942, 0.993, and 0.909 respectively. Interestingly using CFI, AI was now deemed to be below the lower bound of a good model fit.

Each model presented an acceptable model fit utilising the root mean square measure (RMSEA) with the exception of AI and BDAIF which presented a good model fit with values of 0.067 and 0.035 respectively. Interpreting RMSEA with the standard low bound of good model fit < 0.1 shows that OC, FPer, BDAMC and BDAPEC are tending above the minimum bound. This limit however, has been debated in literature to show that it could be set too low (Brown, 2015). Furthermore, the relatively low number of responses collected for this study provide a limiting factor in achieving consistent goodness of fit statistics through the use of CFA (Hinkin, 1965). The overall use of CFI, which is comparatively less sensitive to sample size than RMSEA, showed that two out the three second order constructs of achieved a good model fit (Hair et al., 2014). Due to this, and the ultimate propagation of these constructs into a third tier BDAC construct, model fit was deemed to not be a concern.

Assessing the CFA results shows the factor loadings per construct, where a generally theme of acceptable values was achieved. The majority of factors presented above the lower bound of λ = 0.5, with the exception being AI, which corroborated with the low Cronbach's Alpha values outlined in section 4.9.8.

Table 5: Summary of standardised factors loadings for AI

AI question	Standardised Factor loadings	Factor
AI1	0,402	AI2
AI2	**0,519**	AI1
AI3	**0,839**	AI1
AI4	**0,560**	AI1
AI5	0,448	AI2
AI6	0,419	AI1
AI7	**0,539**	AI2
AI8	-0,081	AI2

Through the CFA model for AI the standard factor loadings and separation of individual questions into two separate factors was achieved. AI2, AI3, AI4 and AI7 loaded sufficiently (table 5), with λ values of 0.519, 0.839, 0.560 and 0.539 respectively. AI7 however loaded independently onto a separate construct and was thus disregarded from the model. For reliability to be sufficiently the general consensus regarding studies focusing on organisations, a minimum of 3 questions per construct is recommended (Hinkin, 1965). Thus, AI was reduced to consist of

AI2, AI3 and AI4. These corroborate with the Cronbach's Alphas achieved for AI, thus removal of these items was done prior to conducting EFA.

5.6.2. Exploratory factor analysis

Exploratory factor analysis is a useful statistical technique for summarising information from large datasets which are interdepended (Hair et al., 2014). It allows for the testing of correlated data for discovering of additional factors, thereby defining a final structure of the data (Hair et al., 2014). This model, is primarily based on the well-established and tested BDAC model, however additional variables have been introduced in the form of the theoretical constructs of AI and OC. Thus, EFA is used in this study as a confirmatory test of the CFA results, and will generate the necessary structure for the additional variables unique to this study.

In order to establish if factor analysis could be applied to the model, two statistical tests in the form of the Kaiser-Meyer-Olkin test and the Bartlett's test of sphericity were applied. Furthermore, a visual inspection of the correlation matrices was done to confirm correlations of at least 0.3 per variable with another variable. Principal Component Analysis (PCA) was deployed as the dimension reduction technique.

Visual inspection of the correlation matrices, as presented in Appendix C, table 17 shows that all variables possessed suitably large correlations with at least one other factor, to allow for factor analysis to take place.

The summarised results of the KMO test and Bartlett's test of sphericity are shown below in table 6 with full result available in Appendix E. As per the KMO results, the lowest sampling adequacy attained was 0.5, which is on the threshold of acceptable, however resulted in a 'Miserable' descriptive result for the first order constructs of modularity, technological management knowledge, business knowledge and relational knowledge. The 'Mediocre' constructs of this model were represented by Investment-decision making, connectivity and AI with KMO values of 0.675, 0.692 and 0.635 respectively. Only the constructs of control, OC and compatibility received 'meritorious' rating with KMO values between 0.811 and 0.868. The constructs of planning coordination and technical knowledge produced KMO results of 0.772, 0.758 and 0.715 which translates into a descriptive value of 'Middling'. Finally, FPer was the only construct to achieve a ‘marvellous’ descriptive value, with a KMO measure of 0.903. Thus, although several constructs achieved ‘miserable’ value of 0.5 this is deemed to be sufficient for factor analysis to be performed.

Additionally, Bartlett’s test of sphericity produced results of p = 0.000 across all of the tested constructs. Thus, significant results were achieved for all construct, therefore meaning that all constructs were deemed to be factorizable.

Table 6: KMO and Bartlett's test of sphericity

Construct	Kaiser-Meyer-Olkin Measure of sampling adequacy	Meaning	Bartlett’s test of Sphericity		
			Chi-Square	df	Sig.
Planning	0.772	Middling	114.53	6	0.000
Investment-Decision making	0.675	Mediocre	150.166	6	0.000
Coordination	0.758	Middling	140.784	6	0.000
Control	0.811	Meritorious	231.290	10	0.000
Organisation Culture	0.868	Meritorious	407.131	28	0.000
Connectivity	0.692	Mediocre	66.875	3	0.000
Compatibility	0.818	Meritorious	50.406	1	0.000
Modularity	0.500	Miserable	13.992	1	0.000
Technical Knowledge	0.715	Middling	105.439	3	0.000
Technological management Knowledge	0.500	Miserable	50.304	1	0.000
Business Knowledge	0.500	Miserable	70.881	1	0.000
Relational Knowledge	0.500	Miserable	32.498	1	0.000
AI	0.625	Mediocre	45.393	3	0.000
FPer	0.903	Marvellous	611.649	21	0.000

PCA was then conducted on the data after factorizability was determined to be achievable. PCA is a test which seeks to capture the greatest explanation of the construct, using the least number of variables possible (Hair et al., 2014). This was performed using the statistical program SPSS, with Kaiser’s criteria where eigenvalues of greater than 1 are acceptable, with item convergence set to a maximum of 25 iterations. Orthogonal rotation was also used through varimax rotation, allowing for the reduction of ambiguity in the data, increasing accuracy (Hair et al., 2014).

PCA was conducted across all first order constructs with the results illustrated in table 7 with further details in Appendix E. All constructs loaded onto one factor with the exception of organisational culture, which loaded onto two constructs.

Table 7: Summary of PCA

Construct	Number of items	Number of components extracted	Cumulative % of variance
Planning	4	1	61.434
Investment-Decision making	4	1	64.631
Coordination	4	1	63.070
Control	5	1	63.954
Organisational Culture	8	2	70.793
Connectivity	3	1	66.555
Compatibility	2	1	81.768
Modularity	2	1	68.282
Technical Knowledge	3	1	73.759
Technological management Knowledge	2	1	81.743
Business Knowledge	2	1	85.939
Relational Knowledge	2	1	76.620
AI	8	1	59.945
FPer	7	1	72.629

Due to this the researcher subdivided OC into the separate factors of CDEV and EINV as per the results shown in table 8 and thematic analysis of the survey instrument at a per question level. This was necessary as the construct of OC will be further tested in this study, in answering the research questions as posed in Chapter three.

Table 8: Summary of PCA for organisational culture

Construct	Included items	Cumulative % of variance explained
CDEV	OC1, OC2, OC3, OC4, OC5, OC8	43.994
EINV	OC6, OC7	26.799
Total		70.793

Questions OC1, OC2, OC3, OC4, OC5 and OC8 were amalgamated to the form CDEV construct, whilst OC6 and OC7, were amalgamated to form the EINV construct. The cumulative variance achieved showed that both constructs cumulatively explained 70.793% of the variance in organisational culture.

5.7. Normality

An important aspect of multivariate testing lies in the normality of the variables, where normally distributed samples are considered preferential to skewed data (Tabachnick & Fidell, 2013). The assumption of normality is a necessary prerequisite in the application of parametric testing to be applied (Tabachnick & Fidell, 2013). Normality can be assessed through the skewness of data, that is the deviation away from a bell shaped curve, and kurtosis of the frequency of data, which should be aligned towards the middling of the largest frequency of scores (Pallant, 2005). Additional testing can however be done using the Shapiro- Wilk and Kolmogorov-Smirnov tests for normality, as has been done in this study.

Firstly, normality testing was done of first order constructs, the results of which are displayed in Appendix B, table 16. These results showed that just one construct, modularity, was normally distributed. To further understand the normality, the researcher then re-ran the testing using second order constructs.

The results shown in table 9 highlights the normality of the second order constructs applied within this study. These include Big Data Analytics Management Capability (BDAMC), Big Data Infrastructure Flexibility (BDIF), Big Data Analytics Personnel Expertise Capability (BDAPEC). BDAMC and BDIF were found to have significant values through the Kolmogorov-Smirnov test, thus showing a normal distribution of data. This was the same result from the Shapiro-Wilk test where these two constructs were the only two to be found to have a normal distribution.

Table 9: Normality of second order constructs

Second-order constructs	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
BDAMC	0.061	100	.200*	0.98	100	0.14
BDIF	0.08	100	0.12	0.978	100	0.1
BDAPEC	0.138	100	0	0.912	100	0

Following these tests, the researcher used the Histograms and QQ plots of the variables to further evaluate the skewness of the data as show in figure 10 with full results in Appendix G. Regardless of the violation of the assumption of normality Maxwell, Delaney, & Kelley (2017) suggest that analyses can continue without transformations.

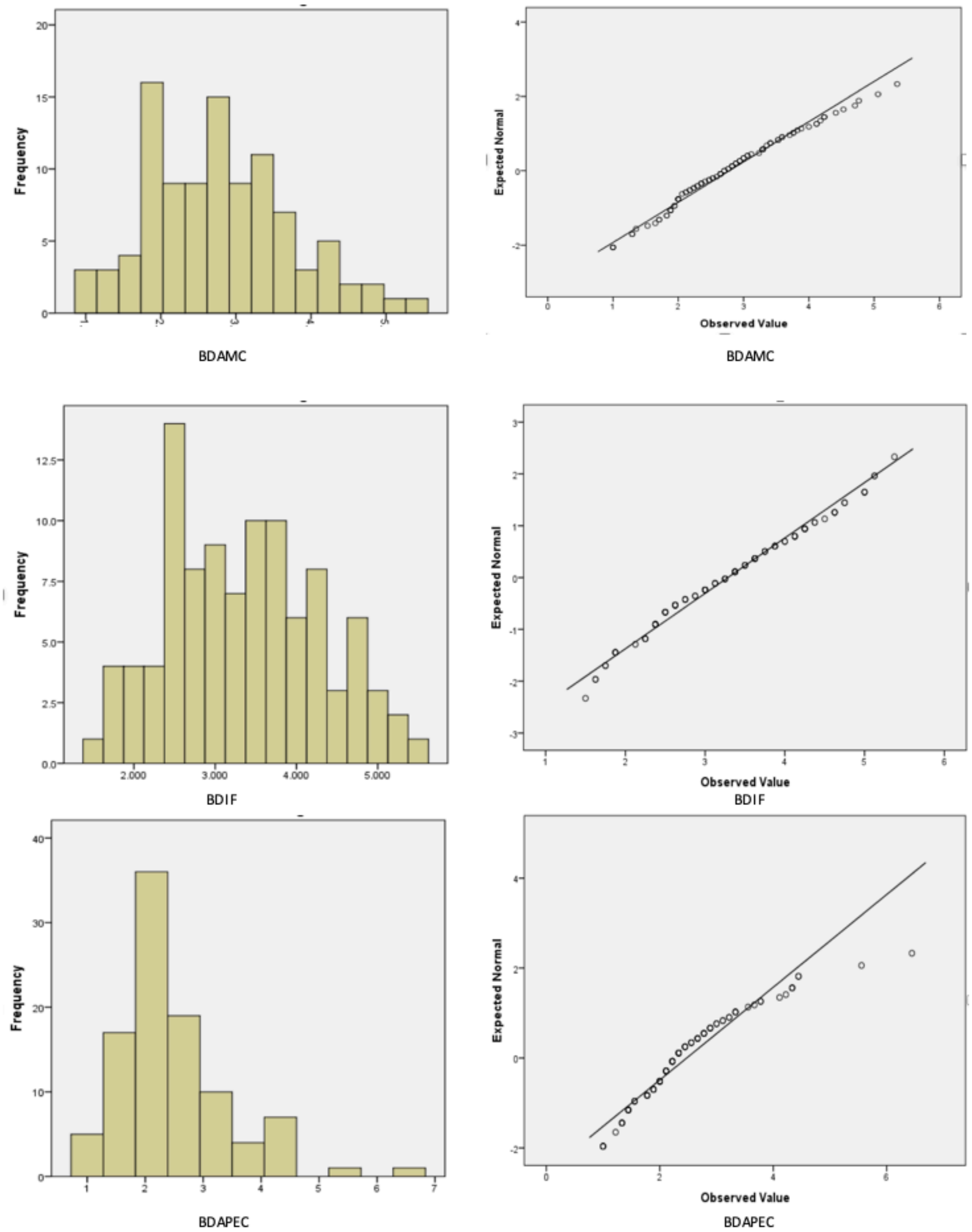


Figure 10: Histograms and QQ Plots of second order constructs

5.8. Research hypotheses

5.8.1. Direct relationships

The following two questions assess the direct relationships between constructs and thus leverage inferential statistical testing through correlation coefficients. The constructs are assessed in the direct relationship with FPer, establishing the basic relationships held within the study.

5.8.2. Research Hypothesis 1

The first research question sought to understand the relationship between BDAC and FPer as originally proposed by Akter (2016). This was done through H_1 which suggested a positive relationship between the two constructs with BDAC as the independent variable and FPer as the dependant variable. Through the use of CFA, the researcher was able to confirm the use of BDAC as a parent construct to the child constructs of BDAMC, BDAIF and BDAPE. Thus, BDAC was applied as directly as a construct in testing the relationship to FPer. Table 10 highlights the Pearson's for all constructs being assessed in this study.

Table 10: Summary of correlation coefficients

FPer	Pearson Correlation Coefficient	p (2-tailed)	N
BDAC	**.456**	**0.000**	100
Planning	.365	0.000	100
Investment-Decision making	.237	0.018	100
Coordination	.271	0.006	100
Control	.454	0.000	100
Connectivity	.509	0.000	100
Compatibility	.455	0.000	100
Modularity	.474	0.000	100
Technical knowledge	.137	0.174	100
Technical Management Knowledge	.253	0.011	100
Business Knowledge	.302	0.002	100
Relational Knowledge	.328	0.001	100

As seen in table 4 BDAC has a moderate correlation with FPer of 0.456 using the Pearson's correlation coefficient with a p-value of 0.00. Therefore, the null hypothesis is rejected in favour of the alternate hypothesis H_1. Thus, this relationship shows that with an increase in BDAC, there is an increase in FPer. The relationship is positive, which corroborates with the initial findings posed by Akter et al., (2016) and Wamba et al. (2017). Further evidence of this relationship is shown through the relationships of the first order constructs which consolidate to form BDAC. Connectivity, control and compatibility exhibit the strongest relationships with FPer. Again, these are moderate positive relationships, where an increase in any of the first order constructs will have a positive effect on FPer. Technical Knowledge showed the weakest relationship with a Pearson's correlation coefficient value of 0.193 at p = 0.54. Overall these results show that BDAC along with its first order constructs present a positive relationship with FPer.

5.8.3. Research Hypothesis 2

The second research hypothesis sought to investigate if there existed a relationship between the proposed variables in this model of OC and AI. Previous studies have highlighted the importance of organisational capabilities such as OC as well as technological capabilities as is represented by AI, however these constructs have not been tested in this configuration previously (Akter et al., 2016; Overgoor et al., 2019). The research question is assessed through hypotheses H_2, H_3.

Table 11: Summary of correlation coefficients for research question 2

FPer	Pearson Correlation Coefficient	p (2-tailed)	N
Cultural Development	.451	0.000	100
Employee Investment	.347	0.001	100
AI	.292	0.003	100

For the testing, the Pearson's correlation coefficient was used, due to the assumption of normality being deployed in this study. As outlined in table 11, OC was found to have two distinct factors, thus this correlation was achieved using the EFA derived constructs of Cultural Development (CDEV) and Employee Investment (EINV). Both constructs show a positive relationship with FPer where CDEV is represented by the Pearson's correlation coefficient of 0.451 at p = 0.00, whilst EINV achieved a correlation coefficient of 0.347 at p = 0.001. Therefore, the null hypothesis for H_2 has been rejected in favour of the alternative hypothesis. This confirms a moderate relationship between CDEV and FPer valid at the 1% significance level. EINV however achieves a weak relationship with FPer at valid at the 1% significance level.

AI achieved a Pearson's correlation coefficient of 0.318 at p = .001 allowing the null hypothesis to be rejected in favour of the alternative hypothesis for $H_{3.}$ This shows a weak direct relationship between AI and FPer.

5.8.4. Research Hypothesis 3

Research hypothesis 3 sought to establish if there existed a relationship between OC, AI and BDAC, in a similar manner as the relationship investigated with FPer. This was assessed through hypotheses $H_{3\ \text{and}}$ $H_{4.}$ As eluded to in section 5.8.3 and table 10, the two factors of OC are represented below in place of the single OC construct originally described in Chapter three. Table 12 below shows the Pearson's correlation coefficients for these constructs.

Table 12: Summary of correlation coefficients for research question 3

BDAC	Pearson Correlation Coefficient	p (2-tailed)	N
CDEV	.714	0.000	100
EINV	.667	0.000	100
AI	.075	.458	100

CDEV and EINV show a strong positive relationship with BDAC with a Pearson's correlation coefficient of 0.714 (p = 0.000) and 0.667 (p = 0.000) respectively. Thus, the null hypothesis was rejected in favour of the alternative hypothesis, confirming that there exists a direct positive relationship between OC and BDAC.

AI produced a Pearson's correlation coefficient value of 0.75 at p = 0.458. Thus, AI fails the hypothesis test at the 5% significance level and the null hypothesis cannot be rejected. Therefore, AI does not have a significant direct relationship with BDAC, measured at the 5% significance level.

5.8.5. Research Hypothesis 4

Research hypothesis 4 represents the main research question of the study encompassing OC, AI and their combined moderating impact on the BDAC-FPer relationship. These are represented by H_5 and H_6 as originally shown in chapter three. Due to the factorising of the data, OC was reconstituted into two new variables, namely CDEV (CDEV) and EINV (EINV).

A hierarchical multiple regression model was thus created to assess the moderating effects of CDEV, EINV and AI. The results are shown in table 13 below with further details in Appendix E.

Table 13: Summary of HMR results

Model	**R**	**Entered**	**Removed**	R^2	**Adj. R^2**	**Std. Err.**	**Change Statistics**				
							ΔR^2	**F Change**	**df1**	**df2**	**Sig. F Change**
1	0,46	BDAC		0,21	0,20	0,96	0,21	25,72	1	98	0,00
2	0,52	MOD CDEV, CDEV		0,27	0,25	0,93	0,07	4,44	2	96	0,01
3	0,53	MOD EINV, MOD EINV		0,28	0,24	0,94	0.00	0,04	2	94	0,96
4	0,46		MOD CDEV, CDEV, MOD EINV, MOD EINV	0,21	0,20	0,96	-0,07	2,20	4	94	0,08
5	0,52		MOD OC, OC	0,27	0,25	0,93	0,06	3,97	2	96	0,02
6	0,6		MOD AI, AI	0,35	0,32	0,89	0,09	6,29	2	94	0,00

The model was based on the dependant variable of FPer and consisted of 4 iterations. Model 1 assessed the basic relationship between BDAC and FPer resulting in an R^2 value of 0.21, thus 21% of the variance in FPer can be explained with BDAC. Each new iteration of the model added independent variables and was assessed for the change in R^2 and the F test significance level. Thus, when the F test failed the effect of adding said variable would not improve the model's explanatory power.

The initial culture moderator of CDEV, entered in model 2, shows an adjusted R^2 of 0.25 with an F test significance level of 0.01. In contrast, the second moderator based on OC, the EINV variable and its subsequent moderator MOD EINV, reduced the R^2. The moderating effect of EINV through MOD EINV lowered the adjusted R^2 from 0.25

to 0.24 with an F test significance level of 0.96. Therefore, EINV is not an effective moderator of the relationship between BDAC and FPer. Subsequently, model EINV and its moderator variable were removed for model 4, raising the Adjusted R^2 to prior levels. This was done in order to facilitate the testing of model 5 which added in the propagated variable OC and its moderating counterpart.

Model 5 thus resulted in a change in R^2 value of 0.06 with an f test significance level of 0.02 showing that OC does show moderating effects on the BDAC-FPer relationship, significant at a 5% confidence interval.

Model 6 presented the moderation effect of AI showing a change in R^2 of 0.09, resulting in an overall model Adjusted R^2 value of 0.32. The F-test resulted in a significance level of 0.000, thus by increasing the model's ability to explain the variance in the dependent variable FPer, AI exhibits moderating effects of the BDAC – FPer relationship.

Thus, due to these results the hypotheses H_5 and H_6 are confirmed due to the rejection of the respective null hypotheses in each case. In conclusion, both AI and OC present moderating effects on the BDAC-FPer relationship of the established theoretical model specified by Akter et al. (2016) and Wamba et al. (2017).

Chapter 6: Discussion of results

6.1. Introduction

This research study set out to investigate the relationship of the BDAC model with the independent variables of OC and AI, as defined in the model shown in chapter 3. The influencing behaviour of these variables and their impact on the BDAC – FPer relationship were also tested. In line with these objectives, research question one was a retest of the BDAC-FPer relationship already establish by Akter et al. (2016) and Wamba et al. (2017) through the base model of BDAC. Research question two additionally independently tested the relationship of OC- FPer and AI-FPer. This was done to establish the relationships of these constructs as a means for business performance improvement in a Big Data enabled environment. Research question three was applied in a similar manner, setting out to establish the independent relationships of OC-BDAC and AI-BDAC. This allows for the investigation of how business may improve their BDAC, if at all. Lastly, the study sought to understand and characterise the effect on the Big Data environment of the independent variables, including the moderating effects that could be expected. In doing so, this study adds to the insights which business can effectively leverage to develop their BDAC and assess future investment in AI technologies as part of their digital strategy.

6.2. Research question 1

Is there a direct positive relationship between Big Data Analytics Capability (BDAC) and FPer (FPer)?

Research question one sought to identify if their existed a relationship between BDAC and FPer. As this study is based on a survey, the respondents opinions of the relationship is examined through inferential statistical testing to outline this relationship. This represented the retest of the base BDAC model and thus was used to establish the BDA environment on which further research questions are based. In establishing the causal link between these variables further vindication of the BDA

environment as an important area of academic literature and further study is also established.
Big Data has become ubiquitous in the business environment, now being classified as a factor of production (Manyika et al., 2011). With the accelerated rate of growth of Big Data due to the lowering of storage costs, businesses are paying further attention to the technology and how best to leverage the maximum benefit for their organisations (Müller et al., 2018). These benefits materialise through the enablement of IT capabilities and has been deemed to improve FPer (Kim et al., 2011; Mikalef et al., 2018). The BDA environment has been outlined by Akter et al. (2016) and Wamba et al. (2017) to consist of several first and second order constructs which aggregate to form BDAC. It is these first order capabilities which enable superior FPer, through both operational efficiencies and strategic potential which may be realised by businesses. The IT capabilities, as posed in the original BDAC model, are bound by the resource based view and dynamic capabilities (Akter et al., 2016). Through the RBV, it is envisioned that firms are capable of producing sustained competitive advantage when the resources which they deploy are valuable, rare, inimitable and non-substitutable. The several first order constructs are thus bound together through an entangled relationship to form the second order constructs of BDA management, BDA infrastructure, and BDA personnel expertise and finally materialising as a single third order construct of BDAC, representing the amalgamation of these constructs (Wamba et al., 2017). Thus, BDAC seeks to leverage dynamic capabilities through its artefacts of sensing and seizing, to unlock latent potential within a business. The ability to integrate the numerous constructs effectively, allows the maximised leveraging of these capabilities, as businesses seeks to gain the highest ROI on their digital strategies (Akter et al., 2016; Mcafee & Brynjolfsson, 2012; Wamba et al., 2017).

As the study was based on a Likert scale derived research survey, and BDAC was found to be normally distributed, outlined in section 5.6, the mean can be used as an acceptable measure of central tendency. The mean value for Big Data Analytics Capability was 2.80 with a standard deviation of 0.82. The Likert scale used was based on 7 points thus a mean value of 2.80 can be seen as low, indicating that the respondents viewed their organisations Big Data Analytics Capability as negative. In the business context this infers that organisations represented in this study are underdeveloped with regards to their Big Data environment. Regarding the general low success rate of Big Data projects, this stands to represent the continuation of the difficulties in management’s ability to leverage the technologies to gain a competitive advantage in the marketplace (Baldwin, 2015).

Delving further into the into the second order constructs of Bid Data Analytics Capability shows that Big Data Management capability represents a mean value of 2.78 with a standard deviation of 0.92. Thus, it can be inferred that the respondents view management’s ability to integrate Big Data into the business environment as sub-par. The first order constructs planning, Investment-Decision making, coordination and control range in mean values from 2.72 to 2.914. Although the range remains small between the three first order constructs, it is interesting to note that the lowest perception of management involves planning and investment-decision making abilities, whilst the highest value is seen through control. Respondents overwhelmingly view the management skills in their organisations negatively, attributing to the ineffective leveraging of Big Data.

As posited by Akter et al. (2016) and Wamba et al. (2017) Big Data Analytics Capability is based on the RBV and dynamic capabilities, therefore the artefacts of sensing , seizing and reconfiguring resources by management is pertinent to the successful implementation of Big Data (Teece et al., 1997). With the results from this study, the represented organisations are composed of management teams who are ineffective at establishing these dynamic capabilities within the organisation due to their negatively viewed planning, investment-decision making, coordination and control capabilities.

Comparatively, table 14 highlights the means achieved across the seminal Big Data Analytics Capability studies done by Akter et al. (2016) and Wamba et al. (2017) and the current study. This is comparable due to the congruency of the questions and scales across the three studies.

Table 14: Comparative means across studies for BDAMC

First order construct	Current study	Akter et al., (2016)	Wamba et al., (2017)
BDA Planning	2.72	4.90	5.03
BDA Investment-Decision making	2.72	4.85	5.13
BDA Coordination	2.82	4.60	5.01
BDA Control	2.91	4.58	5.29

The means show disparate values between this study and the seminal studies, where Akter et al. (2016) and Wamba et al. (2017) found values above average for management capabilities. Owing to each study's use of a survey research instrument, the population sample differences clearly exhibit dissimilar views on their organisations ability to foster the development of Big Data Analytics Capabilities. This study's sample was represented in majority by 'users of big data analytics', with a professional career level dominated by 'owner, executive, c-suite level' employees. Thus, the results are surprising due to the majority of respondents being in the executive level decision making roles within organisations who would ultimately drive the implementation of Big Data projects. The majority of 'owner, executive, c-suite level' respondents exhibiting 'user of big data' statuses within organisations provides some understanding towards the structure of these organisations, where executives are not directly making decisions towards the adoption of Big Data.

Big Data Infrastructure Flexibility posed the highest means in the study, with a mean value of 3.29 and a standard deviation of 0.94. This value is closer to the mid-point of the scale, which can be considered 'Neutral' in terms of how respondents view their organisations infrastructure flexibility. In order for organisations to successfully implement Big Data solutions into their organisations, sufficient investment into the required technologies are necessary (Bello-Orgaz et al., 2016). Not only is investment required, but this investment needs to be tailored to the type of data needed for the organisation (Lee, 2017; Vidgen et al., 2017). Currently, the rate of change of technology places further stress on organisations, as in their attempt to leverage one particular type of technology, path dependencies may be produced through the establishment of dynamic capabilities, resulting in inflexible infrastructure in a dynamic market environment (Mikalef & Pateli, 2017; Teece et al., 2008).

Table 15: Comparative means across studies for BDAIF

First order construct	Current study	Akter et al., (2016)	Wamba et al., (2017)
BDA Connectivity	3.14	4.53	5.09
BDA Compatibility	3.24	4.54	5.10
BDA Modularity	3.46	4.47	5.17

Table 15 highlights the first order constructs between the studies for Big Data Analytics Infrastructure Flexibility. Once again, the means across all three studies are disparate, however for each study the mean values of each first order constructs do exhibit a minimal range. This suggests that much of the variability in the three studies could be attributed to the sample population and organisations featured in each study. With the highest value achieved for the first order constructs of Big Data Infrastructure Flexibility being represented by BDA modularity, with a mean value of 3.46, the respondents clearly view modular flexibility of their IT infrastructure neutrally. Interestingly, the tending of BDA connectivity towards a negative view with a mean value of 3.14, suggesting that respondents viewed the organisations connectivity as below average. By way of the three V's which define Big Data, Volume, Velocity and Variety, the organisations infrastructure should allow for swiftness in change and development as new business challenges emerge (Elgendy & Elragal, 2016; Lee, 2017; Tabesh et al., 2019). These infrastructures foster the organisations ability to adapt to such change, build innovative solutions and respond to a dynamic market environment. The 'neutral' mean values seen in this study could thus be attributed to an organisations slow development of the technologies and skills necessary to build a Big Data environment capable of sustained competitive advantage.

The proposed relationship between BDAC and FPer was tested using the Pearson's correlation coefficient as outlined in table 4. The correlation coefficient was found to be .456, representing a positive moderate relationship between BDAC and FPer at $p = 0.000$. This indicates that there is a positive relationship between Big Data Analytics Capabilities and FPer, where an increase in Big Data Analytics capability results in an increase in FPer. In the context of this study being derived from a survey, respondents foresee that when an organisation invests in Big Data Analytics capabilities, there is an increase in monetary gain by the firm. The correlation coefficient however does not propose causality between the variables, but rather explains the relationship as moderate, i.e. they change in the same direction. Further inferences made cannot be conclusively stated but are rather suggestions for the results, based on theory.

This finding is in agreement with Akter et al. (2016) and Wamba et al. (2017), however it should be noted that these studies found that Big Data Analytics capability had a strong and significant impact on FPer. The differences between the findings to this study can be attributed to the variation in data and the population sample itself.

6.3. Research question 2

Is there a direct positive relationship between AI, organisational performance respectively (independent variables) and FPer (dependent variable)?

In order to leverage an organisations Big Data Capabilities to maximise FPer, several authors have highlighted the lack of attention placed on the social aspects of culture (Lorsch, McTague, 2016; Mcafee & Brynjolfsson, 2012). Currently, in the assimilation of Big Data projects within organisations, executives have emphasised the acceptance of change by ‘people’, that is its employees, as the largest roadblock to the success of these projects (Davenport & Bean, 2019). Given that successful organisations are twice as likely to effectively assimilate Big Data technologies, the focus of firms who are searching for improved FPer should be mindful of the resistance a non-data driven OC will provide (Lavalle et al., 2011).

Furthermore, since Big Data has become ubiquitous in the current market environment through datafication, combined with the accelerated pace of technological development, has meant that non data driven organisations have fallen further behind (Lycett, 2013). Organisations who have effectively deployed Big Data strategies, are seeking further leveraging of their banks of data and knowledge, through the deployment of new frontier technologies in AI (Henke & Kaka, 2018). AI provides organisations with the ability to enhance data driven decision making, since these repositories of data have grown to sizes that the human mind simply cannot process fully, leaving vast arrays of learnings and insights to be uncovered (Colson, 2019). Indeed, to establish FPer through Big Data in today's dynamic market environment, is to leverage the available technologies, whilst balancing the interaction of the workforce. In organisations where a data centric culture does not exist, the establishment of Big Data and any subsequently advanced technology such as AI, could result in a breakdown in organisational culture. Data centric tasks will be taken over by these technologies at a greater pace, resulting in workforce angst, in fear of being replaced by machines (Leopald et al., 2018). Thus, for organisations to effectively deploy Big Data projects, plan for the use of advanced technologies such as AI, and return the maximum ROI, a deeper understanding of the relationship between AI and OC is needed. To this end the study tested the relationships between AI, OC and FPer, verifying the theorised increase in performance of these independent variables.

As outlined in section 5.8.3, Organisational Culture, through the process of EFA, was determined to constitute two distinct factors, namely; Cultural Development and Employee Investment. Thus, the researcher has conducted the correlations at this level, to delve into the deeper understandings of the construct itself.

CDEV showed a moderate correlation with FPer with a Pearson's correlation coefficient of 0.451. This represents a moderate direct and positive relationship with FPer as shown in table 11. CDEV also presented a mean value of 2.08 with a standard deviation of 0.93, which is notably below the mid-point of the Likert scale deployed in this study. As per the questions posed in this study relating to this variable, CDEV specifies the amount of attention organisations place on instituting a data driven culture within the organisation. As this mean value may be interpreted as the respondents viewing the organisations development of the inherent culture negatively, it is evident that there is a lack of focus on creating a data driven culture. This is a worrying outcome for the organisations represented in the study, as the theoretical underpinnings with regards to the implementation of Big Data projects is centred around OC(Lorsch & McTague, 2016; Mcafee & Brynjolfsson, 2012). Without driving the organisations culture towards creating a fertile data centric culture, the risk of failure on Big Data projects increases (Mcafee & Brynjolfsson, 2012). Furthermore, an essential aspect of OC in Big Data environments, is the development of business flexibility, where flatter structures promotes cross-functional collaboration, an essential aspect of innovation (Vidgen et al., 2017). Delving deeper into the CDEV variable highlights this point further, with the statement posed (OC5) of “Our organisation can be described as visionary and flexible”. This statement achieved a mean value of 2.37 with a standard deviation of 1,315, which may be interpreted as respondents viewing their organisations negatively with regards to organisational flexibility and the subsequent vision needed to design a forward-thinking environment.

Essentially, organisations are leveraging Big Data technologies to understand and solve complex market and business problems, then leveraging this knowledge to create innovative solutions for customers (Prasanna & Haavisto, 2018). Without leveraging these structures and fostering innovative collaboration, as posed through the DIGITAL framework, organisations will inevitably meet with resistance in their digital strategy aspirations (Brock & von Wangenheim, 2019).

As the lowest scoring questions posed through CDEV, OC1 and OC3 presented a mean score of 1.69 each with standard deviation of 1.002 and 0.761 respectively as shown in Appendix A. As purported to by the respondents in this study, the represented organisations are viewed negatively in their ability to establish innovative cultures or deploying the latest technology to deliver FPer. This could elude to poor FPer for organisations who are conducting business as usual, instead of sensing the opportunities being leveraged in the marketplace. Furthermore, according to the three tier framework presented by Mazzei & Noble (2017), the represented organisations in this study can be considered pre-tier 1, that is currently failing to see Big Data at its most basic use, as a tool for generating greater FPer. This does not constitute to future fit high-performance organisational development, as reaching the third tier of this framework where data drives strategy, is a basic premise of leveraging data driven decision making and developing the associated culture therein. Thus, it could be inferred that due to the slow uptake of technology by these organisations, development of a data driven culture has not been identified as a focus area.

In contrast to CDEV, the second Organisation Culture variable of EINV presented a weak correlation to FPer with a Pearson's correlation coefficient of 0.347 at p = 0.001. From the results it can be inferred that EINV has a direct positive relationship with FPer, however this relationship is weaker than that achieved by Cultural Investment.

As EINV is an amalgamation of two questions posed by the research instrument, OC6 and OC7, these were investigated at an individual level. The first question, OC6, presented the statement of "There is an extensive employee orientation program for new employees to ensure employees share the corporate vision and purpose", which resulted in a mean value of 2,72 with a standard deviation of 1.6. This could infer that the culture of the organisation is not being transferred effectively through the explicit sharing of the norms, values and beliefs of the organisation (Schein, 2010). According the cultural theory and the three-tiered model of culture, sharing of the culture to employees communicate the norms which in turn provides a measure of control for the organisation. Without the controls, organisations will find it difficult to build innovative capacity to deliver organisational flexibility and ultimately drive FPer through new products and services (O'Reilly, 1989; Vidgen et al., 2017).

Similarly, the second question (OC7) posed the statement of "We invest in targeted training and support at all levels of our organisation to assist our organisation to understand or know how to use data that is available". This question achieved a mean value of 2.79 and a standard deviation of 1.365. The respondents therefore viewed the targeted training offered by their organisations in a negative manner. To be future fit, organisations are required to invest in the necessary skills for their workforce to be effective with new technologies. The skills required for Big Data are distinct and rare, thus fostering development within the organisation expands the concentration of said skills, creating an easier path for organisations to deploy effective Big data strategies (Gupta & George, 2016). From the results it could be inferred that the organisations represented in this study, have a low tendency to providing this training to their workforce, thus placing a skills limitation on any Big Data strategies they wish to deploy.

AI's relationship with FPer presented a Pearson's correlation coefficient of 0.292 at p = 0.003. Thus, this relationship can be characterised as significant, direct and positive with FPer. Therefore, it can be inferred that the opinion derived survey instrument used in this study, encapsulated the respondents view that AI, does directly affect FPer. Through recent studies it has been noted that 20% of companies who have implemented AI related projects reached a point of inflection, allowing for value to be created from the technology (Brock & von Wangenheim, 2019). The relatively low number of projects reaching this value creation point, reverberates with the low correlation coefficient achieved in the AI-Fper relationship.
Since the questions composing AI were reduced to AI2, AI3 and AI4, further analysis was done on more detailed question level. AI2 posed the statement of "My organisation is currently going through change due to automation of tasks" which achieved a mean value of 2,74 and a standard deviation of 1.4. This value represents a low mean in relation to the Likert scale used. This infers that respondents view their organisations negatively with regards to the automation of tasks. Accordingly, these smaller task related automations are referred to as the basis for establishing and building of AI capabilities within organisations (Brock & von Wangenheim, 2019). Once a company has committed to extending their Big Data capabilities into new technologies such as AI, larger projects should be avoided as the risk of failure rises. As these technologies are seen as complementary to Big Data technologies, the organisations in this study could be inferred as not taking full advantage of the data currently at their disposal.

The second and third items for AI, AI3 and AI4, leveraged AI in the form of machine learning by posing statements regarding the current and future adoption of AI through machine learning in the organisation. The mean value achieved for AI3 were 2.28, with a standard deviation of 1.319, whilst AI4 subsequently achieved a higher mean of 2,71 with a standard deviation of 1.37. This is a surprising result based on the market environment and pace of technological development, with evidence showing that this statement is possible today, and will become a reality in a short period of time (Leopald et al., 2018). The differences in the latest studies being evidenced and the current study however lies with the population sample. The demographics show that 99% of the sample were South Africans, which could infer that South Africa is underdeveloped in terms of AI related technologies and the subsequent workforce necessary to deploy said technologies (Moore, 2017). Organisations however, need to be agile and innovative for leveraging these technologies, thus requiring foresight and engagement of the workforce, or risk being surpassed by competitors (Brock & von Wangenheim, 2019).

Research question 3

Is there a direct positive relationship between AI and OC (independent variables) on BDAC (dependent variable)?

Although a host of challenges exists in the implementation of Big Data, these challenges are not merely based on the technical aspects of the technology, but are rather weighted towards the organisational challenges (Gupta & George, 2016). Current research has documented the interdependence of OC and BDAC, with several authors highlighting the importance of creating a data driven culture to allow for the effective leveraging of BDAC projects (Davenport & Bean, 2019; Lorsch & McTague, 2016; Mcafee & Brynjolfsson, 2012; Vidgen et al., 2017). Furthermore, As BDAC is based on the RBV and dynamic capabilities, the organisational environment ,determined by the firm's own culture, establishes the norms inherent in 'how things get done around here' (Akter et al., 2016; Schein, 2010; Teece et al., 1997). Without these norms, organisations risk losing control over the workforce and the ability to steer the trajectory of development towards the envisioned strategic goals of the firm (O'Reilly, 1989; Vidgen et al., 2017). Thus, to effectively enable the development of BDAC, organisations are required to invest in the culture of their organisation.

BDAC's constituent parts of BDAMC, BDAIF and BDAPEC facilitate the overall effect of BDAC as a parent construct, however these variables are further amalgamations of deeper first order constructs. These constructs and their associated entanglements with the inherent workforce which deploy them are inseparable, held together by sociomaterialism (Akter et al., 2016). As such there exists an expectation in examining the relationship of OC to BDAC, which this study objectively set out to quantify.

In a similar manner to the previous research question, but changing dependent variables, this question seeks to understand the relationship of AI and OC on BDAC. The results however were in contrast with those achieved when examining the relationship with FPer. Instead, the correlations of the OC variables, CDEV and EINV, showed strong correlations with BDAC presenting values of 0.714 and 0.667 respectively at $p = 0.000$.
Firstly, CDEV seeks to create a unified culture of the organisation, taking in account the goal of the organisation, to achieve higher FPer through the leveraging of BDAC. These capabilities however are driven on certain prerequisites as posed through theoretical evidence, to include innovation, agility, flexibility and a certain degree of measured risk taking (Dubey et al., 2003). In turn these cultural artefacts are not achievable if the cultural alignment has not been reached. In the context of Big Data and the creation of a data driven culture, organisations will need to drive these cultural dynamics to exhibits these proposed behaviours. If one unified culture is not presented, channel disparity may become evident driving opposing forces both on internal relationships and cross organisational relationships (Prasanna & Haavisto, 2018). Whilst not directly observable disparate channel cultures with external partners, portrays a poor image on the organisation, ultimately slowing the flow of information, slowing or degrading the quality of output from these symbiotic relationships, and ultimately driving poorer BDAC (Prasanna & Haavisto, 2018).

An interesting point to note is the correlation coefficients achieved by both CDEV and EINV are higher for BDAC than for FPer. This result is inherent in the construct as OC is predisposed to affect the internal environment of the organisation through BDAC and its first and second order constructs than it is to externally focused FPer. Thus, BDA which is reliant on OC would produce an enhancing effect on FPer through both strategic and operational capabilities, and the inherent improvements on FPer with an unified data driven culture (Lavalle et al., 2011; Lee, 2017; Lycett, 2013).

The second construct under OC assessed in this study is EINV. From the research survey, and the statements posed to respondents, this encapsulates the direct training and assistance provided to the employee in adjusting and understanding the culture of the organisation. The global business environment is characterised by the deployment of these data related technologies with businesses both new and established taking the leap of high investment spend, banking their future potential growth on the Big Data movement (Alharthi et al., 2017; Gupta & George, 2016; Lee, 2017; Leopald et al., 2018; Moore, 2017). Establishing said technologies is innate in established a distinct and sustainable competitive advantage for organisations, however in order to produce such advantages, possession of technology resource in the form of infrastructure, and workforce capabilities in data scientist is not enough. The current environment demands a further investment into the cultural aspects of the workforce, stemming from the need to drive a flexibility/control-oriented culture. Without flexibility, skilled professionals such as data scientist will not be allowed the control and space to perform 'bricolage' where experimentation leads to development of new ideas and thus products for the organisation (Dubey et al., 2003; Vidgen et al., 2017).

AI produced unsurprising results to BDAC owing to its position as a subset of Big Data. Indeed, the technology is at the frontier of business innovation, enabled by the relatively low cost of data storage and ubiquitous nature of data through the development of IoT. AI showed a Pearson's correlation coefficient of 0.75 at $p = .458$. Therefore, no significant relationship was found between AI and BDAC. Since the nature of BDAC involves the technology and sociomaterialism constructs which enabled it, AI poses no relationship outside of these bounds. This is contrast with the FPer relationship under which AI was found to have a weak correlation and thus a significant relationship with.

The relationships shown by correlations however do not predict causality therefore no causal relationship can be concluded, however inferences can be made in both directions of possible causation. AI should be looked at as a construct based in Big Data where the primary elements of effective implementation provide congruence between Big Data and AI. The DIGITAL framework posed by Brock & von Wangenheim (2019) draws on similar bases of implementation success as Big Data through the need to deploy highly skilled individuals, experiment with the technology, start small and grow the size of projects, enhancing teamwork and creating an agile organisation (Brock & von Wangenheim, 2019; Dubey et al., 2003). Thus, through these similarities, both BDAC and AI are technologies enhanced by the organisational environment, requiring OC to foster their development. Furthermore, the results seem surprising as Big Data is an input into AI, therefore BDAC would allow for higher quality data inputs into AI and improved outcomes for FPer. As this result was not expected, it provides for further postulation around the artefacts for AI to function within an organisation. The population sample itself could also explain part of this result as the geographically skewed nature of the data provides for respondents opinions in a country where AI is in its infancy, limiting the exposure of respondents and therefore accounting for the findings (Moore, 2017).

The results of OC and its constituent parts, AI and the relationships established through statistical testing with BDAC and FPer highlight the complicated entanglement of relationships between the constructs. These require further testing to understand the effects of the established BDAC-FPer relationship as AI increase FPer but not BDAC.

Research question 4

What are the combined impacts of OC and AI (independent variables) on the BDAC-FPer (dependent variable) relationship and do these independent variables moderate that relationship?

Much has been said around the accelerated growth of Big Data and the enablement of improved FPer, resulting in a race amongst organisations to achieve the elusive tier 3 status, joining the likes of companies such as Google and Amazon (Lee, 2017; Mazzei & Noble, 2017). The reality is somewhat removed with companies dedicating significant financial resources, without the accompanying ROI (Mcafee & Brynjolfsson, 2012). The established relationship between BDAC and FPer has further evangelised the use of Big Data by organisations, however due to the failure of such projects, research has focused on the effects of the organisational environment in recent times. There however still exists some disparity in the findings of studies, with certain papers advocating for the importance of OC(Gupta & George, 2016; Mikalef & Pateli, 2017; Vidgen et al., 2017), whilst others claim the importance of culture is not ranked amongst the top 10 precursors to Big Data success (Côrte-Real et al., 2019).

Given these interactions, an HMR model was used to investigate the moderating effects of OC and AI on the BDAC-FPer relationship. As has been established thus far in this study, there exists an entanglement of relationships between these variables, when correlation evaluation is performed. HMR however, seeks to delve deeper into the moderation effects through the use of variance explained as an identifier of the impacts these variables may pose. The model assessed OC, its constituent parts CDEV and EINV, AI and BDAC as independent variables and FPer as the dependent variable.

The results of the HMR evaluation are presented in table 13 each construct was added in subsequent models evaluating their impact on the model allowing for the ΔR^2 to be evaluated. The model culminates with a total variance explained (R^2) by the constructs of 35%. Thus, the model allows for the assessment of independent effects of each of the constructs being tested in relation to each other, and as a total model to explain the BDAC-FPer relationships with the variables added in this study.

Model 1 evaluated the BDAC-FPer relationship as tested in research question 1, with an R^2 value of 21%. This established a baseline for the original model being posed, however the total variance explained can be considered quite low. Due to the survey instrument being used, and the corresponding low mean values seen throughout this study, this result is not surprising. The population sample views the BDAC construct overall in a negative manner for the organisations represented in this model.

After the baseline model was established, the CDEV variable was added as the first portion of the overall OC construct. The results showed a ΔR^2 value of 7%, meaning that CDEV in the organisation increased the total variance explained by the BDAC-FPer relationship by 7%. Assessing the F test values shows that this impact is significant at the 5% level. Owing to the findings in the model from CDEV's correlation with both FPer and BDAC, these findings are not surprising, verifying the theoretical underpinnings of the sociomaterialism view and the practical experiences reported from business (Gupta & George, 2016; Leopald et al., 2018; Vidgen et al., 2017).

To evaluate the impact of the all independent variables of OC, EINV was added in model 3. The results showed a ΔR^2 of 0.00, with a total variance explained increasing by 1%. The F test however ruled out this variable with a p = 0.96, thus the impact of EINV was not significant. The EINV variable did pose lower correlation coefficients than its OC sibling when measured against FPer and BDAC respectively. Overall the means for the questions posed for this variable were amongst the lowest achieved in the survey, thus the resultant ΔR^2 achieved was not surprising. This however does pose a challenge to the workforce, who are a necessary component of the effective deployment of BDAC and ultimately generate FPer. Without sufficient CDEV, the challenges of such tasks will be lost to the obscurity of divergent cultural channels and moreover detrimental to the organisations Big Data aspirations (Gupta & George, 2016).

Consequently, the EINV and CDEV variables were removed from the model, in its fourth iteration. This was done first to negate the negative effects of EINV in relation to the ultimate goal of achieving a maximum total variance explained. Secondly, the original research question posed the investigation of the moderating effects of OC as a single construct, thus with the removal on one component of OC, the researcher needed to test the effects of the construct as a whole. As expected, the effect returned to the R^2 value to its model 1 levels.

With the addition of OC in model 5, the total variance explained increased to 27% with an F test value of p = 0.02, thus OC as a total construct presented significant effects at a 5% significance level. Although the ΔR^2 of 0.06 did not match the removal effects from CDEV and EINV of -0.07, the results were significant, and thus posed a better overall fit.

The final iteration of the model added AI in model 6, which increased the total variance explained to 35%. The ΔR^2 value was the highest of the constructs added in this study at 9% achieved with an F test value of p = 0.00. The relatively high moderating impact of AI can be contrasted to the correlation coefficients, where a weak direct relationship with FPer was established and no significant relationship with BDAC was noted. Thus, the result is surprising as the weighted variance explained is geared towards the AI constructs, related to technology, more than towards organisational culture. This goes against the findings of numerous studies in which claims were made of OC being the primary enabler of Big Data, and ultimately unlocking FPer (Alharthi et al., 2017; Davenport & Bean, 2019; Gupta & George, 2016; Vidgen et al., 2017). From the results it can be purported that due to the negligible differences of the ΔR^2 values between AI and OC, both constructs moderate the BDAC-FPer relationship similarly. The respondents being geographically based in South Africa could provide some explanation to this finding, due to the country being relatively slow in the adoption of Big Data and AI (Moore, 2017). A distinctly data disengaged organisation would place less emphasis on the development of OC in a data driven context, when little to no development of advanced analytics technologies to being implemented. Furthermore, the lack of exposure to these technologies, as identified through the low mean values of each of the constructs, shows little in the way of opinion differences between the constructs. In essence the population sample seem to be at an early stage of their usage of AI, a likely reasoning provided by the relative infancy of the technology in a business environment (Haenlein & Kaplan, 2019). The final model total variance explained reached 35% thus may be considered low, requiring further variables to explain the depth of the relationships.

Conclusion

The moderation effects identified in this study pose a furthering of the theoretical understanding of the entangled relationships of OC, AI, BDAC and FPer. Ultimately through the investigation of the independent and moderating effects of these constructs, a clear theme arises. OC and its constituent components of CDEV and EINV exhibit significant direct and positive relationships with BDAC, which in turn are greater than the relationship with FPer. In contrast AI as a technological outcome of BDAC, exhibits no significant relationship with said construct, but rather holds a significant and positive relationship with FPer. Thus, in a business context firms should derive value from BDAC through instituting strong data driven cultural reforms, ultimately increasing FPer. AI thus forms part of a long-term planning process, once BDAC has been enabled, the Big Data repositories may be effectively leveraged to derive further FPer for the business. Since this study uses a survey instrument, effectively leveraging the opinions of a sample population on the constructs herein, the study poses a further possible rationale for the low AI to FPer result through South African business sluggish deployment of the technology.

Chapter 7: Conclusion

7.1 Introduction

This research study set out to investigate the relationships of the BDAC model, compounded with the constructs of OC and AI. This was done to further examine the resistance to BDAC enablement and the new frontier of AI technologies which have begun their resurgence in the technical field of information systems. Big Data has been widely evangelised as the technology to provide businesses with the strategic advantage over its rivals through more accurate data driven decision making (Akter et al., 2016; Mcafee & Brynjolfsson, 2012). In the pursuit of achieving Big Data superiority within their industries, organisations have invested vast resources into the technology, with somewhat disappointing results (Baldwin, 2015). Achieving tier 3 status, as evidenced by global corporates such as Amazon and Google, for established companies has meant a complete upheaval of their current business strategies.

Digital technologies such as Big Data have become a prerequisite in the pursuit of greater FPer (Constantiou & Kallinikos, 2015). Unfortunately for the vast majority of organisations, the challenge of implementing said technologies was seen as an operational challenge, skewed towards the technical aspects of infrastructure and technical skills, whilst ignoring the organisational environment in the form of culture (Gupta & George, 2016). Further complications have begun to arise with early adopters of Big Data, now in a position to leverage their expertise in the field with more ambitious technologies in the space of AI (Haenlein & Kaplan, 2019). A new precedent has thus been created for digital strategy, tasking AI to work alongside human counterparts in the pursuit of accelerated learning, and creating dynamic products capable of expanding firm profitability and meet customers evolving demand (Henke & Kaka, 2018). Indeed, to enable the subset of technologies such as AI, the parent technology of Big Data will first be required to show a positive ROI. With OC showing that it is undeniably a key to unlocking this performance and implementation success, their exists an entanglement of relationships between BDAC, FPer, OC and AI, which this study has investigated. These relationships have been posed in figure 1 of this study, which informed the development of further hypotheses as set out in chapter 3. This chapter therefore summarise the key findings posed by the proposed model to pursuit of furthering the theoretical and business understandings of the constructs herein.

7.2 Theoretical contributions

This study makes theoretical contributions to the seminal BDAC studies conducted by Akter et al. (2016) and Wamba et al. (2017) in the field of information systems and strategic management. The specific influence of the organisational environment is built on the work of Gupta & George (2016) where organisational resistance to BDAC project implementation was theorised. Thus, the study had a particular focus which is congruent with these seminal works, where additional constructs and their influences on BDAC were hypothesised and tested. Furthermore, this study introduced a longer-range technological construct of AI, asserting that the accelerated pace of development of the technology changed the frontier of digital strategy.

BDAC ultimately remains a well-established model built on the IT capabilities of the firm viewed through a sociomaterialism lens, and based on the RBV together with dynamic capabilities theory (Akter et al., 2016; Wamba et al., 2017). However, the model itself is still contested in its efficacy, with no real consensus on BDAC provided by these sociomaterialism views over the pure technological challenges that these projects assert (Côrte-Real et al., 2019; Lee, 2017). Thus, this study set out to quantify the BDAC-FPer relationship and further assess the efficacy of OC.

Initial findings presented a strong positive and direct relationship between BDAC and FPer, reasserting the findings of Akter et al. (2016) and Wamba et al. (2017) and further affirming that the relationship is based on the sociomaterialism view. This is in contrast to findings from Lee (2017), where IT capabilities was posited to being the main obstacle to effective Big Data project implementation. Thus, this study adds to the theoretical base of this model vindicating its use in further research as a base for further organisational constructs and their associated influence to be examined.

Further findings showed that OC has significant, direct and positive relationships with both BDAC and FPer. The constructs have been shown to allow for the leveraging of OC in the pursuit of unlocking enablement for BDAC technology. More specifically, although there does exist a relationship between OC and FPer, this relationship was found to be significantly weaker than the OC-BDAC relationship. This finding therefore reasserts the theoretical underpinnings of similar studies (Gupta & George, 2016; Vidgen et al., 2017), while reaffirming the established findings of the BDAC models sociomaterialism view . Divergent cultures within organisations can be seen as detrimental to the organisation's pursuit of sustained competitive advantage. Furthermore, findings from Côrte-Real et al., (2019), found that IT and business executives did not view OC as a precursor to the success of Big Data technology implementation, which is in contrast to the findings of this study. Instead a suggestion can be inferred from the results herein that technological and infrastructure challenges are secondary to the organisational challenges posed by Big Data project implementations. In a similar manner OC's moderating effect on the BDAC-FPer relationship, reaffirms this theoretical understanding of the effect OC may have in an organisation. A unified data driven culture will therefore serve to enhance the BDAC-FPer relationship, assisting in overcoming challenges of Big Data project implementations.

A clear defining line in research between IT capabilities and organisational environmental capabilities importance in the BDAC sphere has developed, with this study adding to the theoretical base with its reaffirming of OC as a key component for the enablement of BDAC within organisations.

Further to the organisation environment, this study sought to contribute to the long rand strategy of organisations, by examining the impact AI has to both BDAC and FPer. AI showed a weak correlation to FPer, that is respondents view of their organisations ability to leverage the technology towards increasing FPer was viewed negatively. This result was expected due to the technology being on the forefront of business capabilities, due to its reliance on Big Data repositories to 'learn' and ultimately assist with the human decision making function (Haenlein & Kaplan, 2019). Regardless on the infancy of the technology, studies have shown that significant value is currently being derived from AI through a multitude of applications ranging from assisted decision making, to marketing and human resources (Brock & von Wangenheim, 2019; Metcalf et al., 2019; Overgoor et al., 2019; Tambe et al., 2019). Significant in this study, is the low correlation to FPer, which was an unexpected result, but reaffirms the theoretical understanding of the lag in technological development by organisations who are currently leveraging Big Data, with just 20% of companies having achieved value out of their AI projects (Brock & von Wangenheim, 2019).

Furthermore, since the participants of the study held a 99% geographical alignment to South Africa, this study reaffirms the view that the country is lagging behind in technological development, where businesses are not leveraging Big Data and AI at a competitive pace (Moore, 2017). Notably AI and BDAC were found to have no significant relationship, however due to correlation analysis being used, no directional relationship could be surmised. This present a further finding as AI is a subset of Big Data, and therefore BDAC, providing data for the technology to function. Hence, the higher the quality of data provided through the BDAC process would effectively improve the inputs to AI, and ultimately the output of improved FPer (Haenlein & Kaplan, 2019). As this relationship was not found in the study, it suggests that other constructs could explain the relationship in more depth, however this is outside of the confines of the current study.

The concurrent effects of AI, together with OC, showed that both variables achieved a moderating effect on the BDAC-FPer relationship, with AI adding the second highest variance explained to the hypothesised model. The notable effect of AI as a moderating variable provides validation of the hypothesised model, and draws congruency with established theory on the value creation effects of the technology (Brock & von Wangenheim, 2019; Metcalf et al., 2019; Overgoor et al., 2019; Tambe et al., 2019). Thus, this research contributes to the understanding of organisations current leveraging of OC and AI towards the development of FPer in a dynamic market environment.

7.3 Implications for business

In the current dynamic competitive environments, which organisations have to derive maximum value and ROI on their investments, this study takes the stance of the sociomaterialism view in enabling a sustained competitive advantage. Yet, with the advancing of technology it encapsulates the intersection of forward planning of Big Data technologies through AI as the next step in the technology evolutionary chain. A powerful concept was revealed in the assimilation of BDAC enablement and the development of competitive advantage through the entanglement of the BDAC-FPer relationship, with the moderating effects of AI and OC. Indeed, as firms invest heavily in the development of their technological capabilities, the cultural alignment of the workforce to their strategy is paramount to their success. Organisations who invest further in cultural development will inevitable leveraged BDAC further and develop dynamic capabilities to entrench innovation and maintain higher levels of competitive advantage. To develop deeply entrenched dynamic capabilities involves more than just the macro level of resource structures, but rather a microcosm of individual contributors (Salvato & Vassolo, 2018). Through creative individuals who are allowed to develop in a flexible and agile organisation, the release of their inner bricoleur is achieved, a necessary resource in the field of data science (Dubey et al., 2003; Vidgen et al., 2017). These constructs are driven by the culture instilled by the organisation, therefore, to allow for a significant increase in innovative capacity, and thus improved FPer, organisations should drive a data driven, open culture which invests in its workforce.

The benefits of improved BDAC however, are further enhanced with long term planning, with organisations assessing the next step in the data driven journey. Towards this end, AI has found popularity through the significant improvement in Big Data in terms of its artefacts of variety, velocity and veracity (Brock & von Wangenheim, 2019; Henke & Kaka, 2018; Laney, 2001). This study showed a significant moderating relationship of AI on FPer, encapsulating the promise of improved competitive advantage through the leveraging the technology on the base of Big Data. As management are the decision makers of the organisations, a general understanding of the possible investment or frontier technology is necessary to seize the opportunity at an early stage (Mikalef & Pateli, 2017; Teece et al., 2008). The surprising finding from the sample population in this study is that respondents viewed management's investment decision making capability as sub-par, a prerequisite skill to establishing dynamic capabilities (Teece, 2018a; Winter, 2003). Thus, in order for management to establish a greater level of FPer through leveraging advanced technologies such as AI, significant investment into management's the decision making ability should be made. Knowledge of the market and sufficient understanding of frontier technology remains a necessary condition for sustainable competitive advantage (Mikalef & Pateli, 2017; Teece et al., 2008). Since technology generally follows path dependencies, failure to understand the to gather this knowledge may result in further incorrect technological investments and unacceptable ROI's.

7.4 Recommendations for future research

This study focused on the reassertion of the BDAC-FPer relationship whilst investigating a configuration of constructs which has thus far not been used in literature. Although OC has been extensively covered in the context of data driven decision making and Big Data, the topic itself remains broad and complex. Further exploration of the workforce's motivation in the face of developing AI technology, as well as the threat of job losses could be asserted through the model.

The BDAC-AI relationship presented no significant relationship between variables, the theoretical base on the subject suggest the contrary (Brock & von Wangenheim, 2019; Haenlein & Kaplan, 2019; Henke & Kaka, 2018). Therefore, with the combined efforts of a deeper sample size and the use of SEM, further evidence can be presented on the relationship and its causal effects.

Furthermore, as AI is a subset of Big Data, BDAC and its first order constructs should be reconfigured in light of the DIGITAL framework, to understand the differences in leveraging each technology (Brock & von Wangenheim, 2019). The relative infancy of the technology itself has meant that the workforce remains underdeveloped towards the skill requirement, and subsequent understanding of the future of work (Haenlein & Kaplan, 2019; Metcalf et al., 2019). To this end, Swarm intelligence presents a vast sphere of opportunity for business, as multiple configurations of the hypothesised technology are being pursued (Metcalf et al., 2019). Thus, the effects of OC are magnified, as organisational ignorance towards the subject may cause further resistance to change, ultimately harming FPer. Future studies should delve deeper into the possible uses of AI, examining the proposed effects of these technologies on employee motivation and FPer.

As dynamic capabilities are both operationally and strategically based within a business environment, they are paramount to the success of a technologically driven organisation. Technological dynamic capabilities, however, are postulated as providing less of an effect on FPer due to the homogeneity of technology amongst firms (Barrales-Molina et al., 2013; Fainshmidt et al., 2016; Wilden & Gudergan, 2014). In contrast AI remains a technology that polarises organisations as few have been able to generate business value from it. Thus, future research should investigate the effects of technological dynamic capabilities and their effect of FPer in a dynamic market environment where AI is not pervasive. Quantification of the differences in FPer amongst organisations who are both leveraging AI and are yet to implement the technology will provide an important contribution to theory.

7.5 Limitations

This study offered numerous limitations extending across methodological and theoretical boundaries. The limitation placed on data collection has been noted, with the sample size reaching 100 participants. Thus, with the low number of responses, the statistical analyses was limited. A retest of the model leveraging structured equation modelling would be ideal to deepen the statistical rigour provided in this study. Structured equation modelling is highly suited to business research, due to its ability to predict causal dependent relationships if the assumptions if theoretical support, sequence, covariation and nonspurious covariation are evident (Hair et al., 2014). Thus, through SEM further conclusions may be drawn about the causality of relationships, as opposed to this study which has been limited to possible inferences from the results (Hair et al., 2014).

The study also achieved a relatively lower total variance explained in the proposed model of 35%. Although the BDAC-FPer relationship is well established in literature, the construct of AI could be quantified in greater depth. This construct was used at a high level in this study to account for the level of knowledge amongst respondents. The researcher thus notes that a more granular conceptualisation of the AI construct could allow for a greater understanding of the influence on the BDAC-FPer relationship. Furthermore, with a low total variance explained, their exists further variables which should be added to the model to increase the model's efficacy.

Although great effort was placed on widening the reach of the study to diversify the opinion driven survey research instrument, 99% of the sample was represented by South African's. This largely could account for the low mean scores seen throughout the survey as South Africa lags behind the world in terms of AI advancement (Moore, 2017).

Lastly, the researcher is noted as being inexperienced in the field of research, with this study being his first. Due to this inexperience, the researcher may have made errors in judgement with regards to the design and execution of the research.

References

Abbasi, A., Sarker, S., & Chiang, R. (2018). Big Data Research in Information Systems: Toward an Inclusive Research Agenda. *Journal of the Association for Information Systems*, *17*(2), I–XXXII. https://doi.org/10.17705/1jais.00423

Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. (2016). How To Improve Firm Performance Using Big Data Analytics Capability. *International Journal of Production Economics*, *182*(December), 113–131. https://doi.org/10.1016/j.ijpe.2016.08.018

Alharthi, A., Krotov, V., & Bowman, M. (2017). Addressing barriers to big data. *Business Horizons*, *60*(3), 285–292. https://doi.org/10.1016/j.bushor.2017.01.002

Allison, P. (2002). *Missing Data* (Third edit). California: SAGE Publications.

Amankwah-Amoah, J., & Adomako, S. (2019). Big data analytics and business failures in data-Rich environments: An organizing framework. *Computers in Industry*, *105*(2), 204–212. https://doi.org/10.1016/j.compind.2018.12.015

Baldwin, H. (2015). When Big Data Projects Go Wrong. Retrieved August 23, 2019, from Forbes website: https://www.forbes.com/sites/howardbaldwin/2015/01/22/when-big-data-projects-go-wrong/#18272ea46427

Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, *17*(1), 99–120. https://doi.org/10.1177/014920639101700108

Barrales-Molina, V., Bustinza, Ó. F., & Gutiérrez-Gutiérrez, L. J. (2013). Explaining the Causes and Effects of Dynamic Capabilities Generation: A Multiple-Indicator Multiple-Cause Modelling Approach. *British Journal of*

Management, *24*(4), 571–591. https://doi.org/10.1111/j.1467-8551.2012.00829.x

Bass, L., Clements, P., & Kazman, R. (2003). Software Architecture in Practice , Second Edition. In *Software Architecture* (Second). Boston: Addison-Wesley Longman.

Bekmamedova, N., & Shanks, G. (2014). Social media analytics and business value: A theoretical framework and case study. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 3728–3737. https://doi.org/10.1109/HICSS.2014.464

Bello-Orgaz, G., Jung, J. J., & Camacho, D. (2016). Social big data: Recent achievements and new challenges. *Information Fusion*, *28*(3), 45–59. https://doi.org/10.1016/j.inffus.2015.08.005

Bharadwaj, A. (2000). A resource-based perspective on information technology capability and firm performance: An empirical investigation. *MIS Quarterly*, *24*(1), 169–196. Retrieved from https://www.jstor.org/stable/3250983?seq=1#page_scan_tab_contents

Bonett, D. G., & Wright, T. A. (2015). Cronbach's alpha reliability: Interval estimation, hypothesis testing, and sample size planning. *Journal of Organizational Behavior*, *36*(1), 3–15. https://doi.org/10.1002/job.1960

Bowman, C., & Ambrosini, V. (2003). How the Resource-based and the Dynamic Capability Views of the Firm Inform Corporate-level Strategy. *British Journal of Management*, *14*(4), 289–303. https://doi.org/10.1111/j.1467-8551.2003.00380.x

Brock, J. K.-U., & von Wangenheim, F. (2019). Demystifying AI: What Digital Transformation Leaders Can Teach You about Realistic Artificial Intelligence. *California Management Review*, *61*(4), 110–134. https://doi.org/10.1177/1536504219865226

Brown, T. (2015). *Confirmatory Factor Analysis for applied research* (Second Edi). Boston: The Guildford Press.

Cao, G., & Duan, Y. (2014). Gaining competitive advantage from analytics through the mediation of decision-making effectiveness: An empirical study of UK manufacturing companies. *Proceedings - Pacific Asia Conference on Information Systems, PACIS 2014*, (June), 377. https://doi.org/10.13140/2.1.3505.5360

Carr, N. G. (2003). IT Doesn't Matter Leadership Development : *Harvard Business Review*, 5–17. Retrieved from https://www.classes.cs.uchicago.edu/archive/2014/fall/51210-1/required.reading/ITDoesntMatter.pdf

Chatman, J. A., & O'Reilly, C. A. (2016). Paradigm lost: Reinvigorating the study of organizational culture. *Research in Organizational Behavior*, *36*, 199–224. https://doi.org/10.1016/j.riob.2016.11.004

Chen, H.-M., Kazman, R., & Haziyev, S. (2016). Agile Big Data Analytics for Web-Based Systems: An Architecture-Centric Approach. *IEEE Transactions on Big Data*, *2*(3), 234–248. https://doi.org/10.1109/tbdata.2016.2564982

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business Intelligence and analytics. *MIS Quarterly*, *36*(4), 1165–1188. Retrieved from https://pdfs.semanticscholar.org/f5fe/b79e04b2e7b61d17a6df79a44faf358e60cd.pdf

Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, *19*(2), 171–209. https://doi.org/10.1007/s11036-013-0489-0

Colson, E. (2019). What AI-Driven Decision Making Looks Like. Retrieved from Harvard Business Review Digital Articles website: https://hbr.org/2019/07/what-ai-driven-decision-making-looks-like

Constantiou, I. D., & Kallinikos, J. (2015). New games, new rules: Big data and the changing context of strategy. *Journal of Information Technology*, *30*(1), 44–57. https://doi.org/10.1057/jit.2014.17

Côrte-Real, N., Oliveira, T., & Ruivo, P. (2017). Assessing business value of Big Data Analytics in European firms. *Journal of Business Research*, *70*, 379–

390. https://doi.org/10.1016/j.jbusres.2016.08.011

Côrte-Real, N., Ruivo, P., Oliveira, T., & Popovič, A. (2019). Unlocking the drivers of big data analytics value in firms. *Journal of Business Research*, *97*(12), 160–173. https://doi.org/10.1016/j.jbusres.2018.12.072

Cresswell, J. W. (2009). Research design: Qualitative, quantitative, and mixed methods approaches. In *Muqarnas* (3rd ed.). Washington DC: Sage publications.

Cronbach, L. J., & Shavelson, R. J. (2004). My Current Thoughts on Coefficient Alpha and Successor Procedures. *Educational and Psychological Measurement*, *64*(3), 391–418. https://doi.org/10.1177/0013164404266386

Davenport, T. (2014). *Big data at work: Dispelling the myths, uncovering the opportunities* (2nd ed.). Boston: Harvard Business Review Press.

Davenport, T., & Bean, R. (2019). Companies Are Failing in Their Efforts to Become Data-Driven. Retrieved from Harvard Business Review Digital Articles website: https://hbr.org/2019/02/companies-are-failing-in-their-efforts-to-become-data-driven

Dehning, B., Richardson, V. J., & Zmud, R. W. (2003). The value relevance of announcements of transformational information technology investments. *MIS Quarterly: Management Information Systems*, *27*(4), 637–656. Retrieved from https://misq.org/the-value-relevance-of-announcements-of-transformational-information-technology-investments.html

Deshpande, R., & Webster, F. E. J. (1989). Organizational Culture and Marketing: Defining the research agenda. *The Journal of Marketing*, *53*(1), 3–15. https://doi.org/10.2307/1251521

Druckman, J., Green, D., Kuklinski, J., & Lupia, A. (2011). *Cambridge Handbook of Experimental Political Science* (First Edit). New York: Cambridge University Press.

Dubey, Rameshwar;, Gunasekaran, A., Childe, S., Roubaud, D., Fosso Wamba, S., Giannakis, M., & Foropon, C. (2019). Big data analytics and organizational culture as complements to swift trust and collaborative performance in the humanitarian supply chain. *International Journal of Production Economics*, *210*(9), 120–136. https://doi.org/10.1016/j.ijpe.2019.01.023

Dubey, Rameshwar, Gunasekaran, A., Childe, S., Blome, C., Papadopoulos, Hanos, … Terblanche, F. (2003). Building organisational culture that stimulates creativity and innovation. *European Journal of Innovation Management*, *6*(1), 55–78. https://doi.org/10.1111/1467-8551.12355

Eisenhardt, K. M., & Martin, J. A. (2006). Dynamic Capabilities : What Are They ? *Strategic Management Journal*, *21*(10), 1105–1121. Retrieved from https://www.jstor.org/stable/3094429

Ekbia, H., Mattioli, M., Kouper, I., Arave, G., Ghazinejad, A., Bowman, T., … Sugimoto, C. R. (2015). Big data, bigger dilemmas: A critical review. *Journal of the Association for Information Science and Technology*, *66*(8), 1523–1545. https://doi.org/10.1002/asi.23294

Elgendy, N., & Elragal, A. (2016). Big Data Analytics in Support of the Decision Making Process. *Procedia Computer Science*, *100*, 1071–1084. https://doi.org/10.1016/j.procs.2016.09.251

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of Convenience Sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1–4. https://doi.org/10.11648/j.ajtas.20160501.11

Fainshmidt, S., Pezeshkan, A., Lance Frazier, M., Nair, A., & Markowski, E. (2016). Dynamic Capabilities and Organizational Performance: A Meta-Analytic Evaluation and Extension. *Journal of Management Studies*, *53*(8), 1348–1380. https://doi.org/10.1111/joms.12213

Fang, E., & Zou, S. (2009). Antecedents and consequences of marketing dynamic capabilities in international joint ventures. *Journal of International Business Studies*, *40*(5), 742–761. https://doi.org/10.1057/jibs.2008.96

Fricker, R. D., & Schonlau, M. (2002). Advantages and Disadvantages of Internet

Research Surveys: Evidence from the Literature. *Field Methods*, *14*(4), 347–367. https://doi.org/10.1177/152582202237725

Garmaki, M., Boughzala, I., & Wamba, S. F. (2016). The effect of Big Data Analytics Capability on firm performance. *Proceedings - Pacific Asia Conference on Information Systems, PACIS 2016*, 301. Retrieved from http://aisel.aisnet.org/pacis2016/301

George, G., Haas, M. R., & Pentland, A. (2014). From the editors: Big data and management. *Academy of Management Journal*, *57*(2), 321–326. https://doi.org/10.5465/amj.2014.4002

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information and Management*, *53*(8), 1049–1064. https://doi.org/10.1016/j.im.2016.07.004

Haenlein, M., & Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*, *61*(4), 5–14. https://doi.org/10.1177/0008125619864925

Hagel, J. (2015). Bringing analytics to life. *Journal of Accountancy*, *219*(2), 24–25. Retrieved from https://www.questia.com/library/journal/1G1-400529124/bringing-analytics-to-life

Hair, J., Black, W., Babin, B., & Anderson, R. (2014). *Multivariate data analysis* (Seventh Ed). New York: Pearson.

Hartnell, C. A., Ou, A. Y., & Kinicki, A. (2011). Organizational Culture and Organizational Effectiveness: A Meta-Analytic Investigation of the Competing Values Framework's Theoretical Suppositions. *Journal of Applied Psychology*, *96*(4), 677–694. https://doi.org/10.1037/a0021987

Henke, N., & Kaka, N. (2018). Analytics comes of age. In *McKinsey Analytics*. Retrieved from https://www.mckinsey.com/~/media/McKinsey/Business Functions/McKinsey Analytics/Our Insights/Analytics comes of age/Analytics-comes-of-age.ashx

Henke, N., Libarikian, A., & Wiseman, B. (2016). Straight talk about big data.

McKinsey Quarterly, *10*(4), 42–51. Retrieved from https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/straight-talk-about-big-data

Hess, T., Matt, C., Benlian, A., & Wiesböck, F. (2019). Options for Formulating a Digital Transformation Strategy. *MIS Quarterly E*, *15*(2), 123–125. https://doi.org/10.1108/10878571211209314

Hinkin, T. (1965). A review of scale development practices in the study of organisations. *Journal of Management*, *21*(5), 967–988. https://doi.org/10.1111/j.1745-4603.1989.tb00456.x

Holmes, W. M. (2014). *Using Propensity Scores in Quasi-Experimental Designs* (Second edi). London: SAGE Publications, Ltd.

Hosanagar, K., & Saxena, A. (2017). The First Wave of Corporate AI Is Doomed to Fail. Retrieved from Harvard Business Review Digital Articles website: http://ezproxy.leedsbeckett.ac.uk/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=124260109&site=eds-live&scope=site

Huang, M.-H., Rust, R., & Maksimovic, V. (2019). The Feeling Economy: Managing in the Next Generation of Artificial Intelligence (AI). *California Management Review*, *61*(4), 43–65. https://doi.org/10.1177/0008125619863436

Hunt, S. D., Sparkman, R. D., & Wilcox, J. B. (1982). The Pretest in Survey Research: Issues and Preliminary Findings. *Journal of Marketing Research*, *19*(2), 269. https://doi.org/10.2307/3151627

Iansiti, M., & Clark, K. B. (1994). Integration and dynamic capability: Evidence from product development in automobiles and mainframe computers. *Industrial and Corporate Change*, *3*(3), 557–605. https://doi.org/10.1093/icc/3.3.557

Kamioka, T., & Tapanainen, T. (2014). Organizational use of big data and competitive advantage—exploration of antecedents. *Pacific Asia Conference on Information Systems*, 372. Retrieved from http://aisel.aisnet.org/pacis2014http://aisel.aisnet.org/pacis2014/372

Khazanchi, S., Lewis, M. W., & Boyer, K. K. (2007). Innovation-supportive culture: The impact of organizational values on process innovation. *Journal of Operations Management*, *25*(4), 871–884. https://doi.org/10.1016/j.jom.2006.08.003

Kim, G., Shin, B., Kim, K. K., & Lee, H. G. (2011). IT Capabilities, Process-Oriented Dynamic Capabilities, and Firm Financial Performance. *Journal of the Association for Information Systems*, *12*(7), 487–517. https://doi.org/10.17705/1jais.00270

Kotter, J., & Heskett, J. (1992). *Corporate culture and performance* (Second edi). New York: Free Press.

Kraemer, H. C., & Blasey, C. (2016). *How Many Subjects?: Statistical Power Analysis in Research* (Second Edi). London: SAGE Publications, Ltd.

Kraska, T. (2013). Finding the needle in the big data systems haystack. *IEEE Internet Computing*, *17*(1), 84–86. https://doi.org/10.1109/MIC.2013.10

Laney, D. (2001). 3D data management: Contolling data volume, velocity and variety. In *Meta group research note*.

Lavalle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big Data , Analytics and the Path From Insights to Value Big Data , Analytics and the Path From Insights to Value. *MIT Sloan Management Review*, *52*(2), 21–31. Retrieved from http://foresight.ifmo.ru/ict/shared/files/201309/1_9.pdf

Learned, E. P. (1969). *Business policy: Text and cases* (First edit). Homewood: R.D Irwin.

Lee, I. (2017). Big data: Dimensions, evolution, impacts, and challenges. *Business Horizons*, *60*(3), 293–303. https://doi.org/10.1016/j.bushor.2017.01.004

Leopald, T. A., Ratcheva, V., & Zahidi, S. (2018). The future of jobs report 2018.

In *World Economic Forum* (Vol. 2). https://doi.org/10.1177/0891242417690604

Li, J., Tao, F., Cheng, Y., & Zhao, L. (2015). Big Data in product lifecycle management. *International Journal of Advanced Manufacturing Technology*, *81*(1–4), 667–684. https://doi.org/10.1007/s00170-015-7151-x

Little, R. J. A., & Rubin, D. B. (2014). *Mixed Normal and Non-Normal Data with Missing Values, Ignoring the Missing-Data Mechanism* (Second edi). New York: John Wiley & Sons.

Liu, Y. (2014). Big Data and predictive business analytics. *Journal of Business Forecasting*, *33*(4), 40–42. Retrieved from https://search-proquest-com.uplib.idm.oclc.org/docview/1657419180?accountid=14717

Lorsch, J., & McTague, E. (2016). Culture Is Not the Culprit: When Organizations Are in Crisis, It's Usually Because the Business Is Broken. *Harvard Business Review*, *94*(4), 96–105. Retrieved from https://www.hbs.edu/faculty/pages/item.aspx?num=50894

Lycett, M. (2013). "Datafication": Making sense of (big) data in a complex world. *European Journal of Information Systems*, *22*(4), 381–386. https://doi.org/10.1057/ejis.2013.10

Manyika, J., Chui Brown, M., B. J., B., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). Big data: The next frontier for innovation, competition and productivity. *McKinsey Global Institute*, (May). Retrieved from file:///Users/graceqiu/Library/Application Support/Mendeley Desktop/Downloaded/Manyika et al. - 2011 - Big data The next frontier for innovation, competition and productivity.pdf

Markow, W., Braganza, S., Taska, B., Hughes, D., & Miller, S. (2017). *The Quant Crunch*. Retrieved from https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=IML14576USEN

Maxwell, S., Delaney, H., & Kelley, K. (2017). *Designing Experiments and Analyzing Data* (Third edit). New York: Routledge.

Mazzei, M. J., & Noble, D. (2017). Big data dreams: A framework for corporate strategy. *Business Horizons*, *60*(3), 405–414. https://doi.org/10.1016/j.bushor.2017.01.010

Mcafee, A., & Brynjolfsson, E. (2012). Big Data: The Management Revolution. *Harvard Business Review*, (10), 1–9. Retrieved from http://tarjomefa.com/wp-content/uploads/2017/04/6539-English-TarjomeFa-1.pdf

McAfee, R. B., Glassman, M., & Honeycutt, E. D. (2002). The Effects of Culture and Human Resource Management Policies. *Journal of Business Logistics*, *23*(1), 1–18. https://doi.org/https://doi.org/10.1002/j.2158-1592.2002.tb00013.x

McKinsey Global Institute. (2018). How will automation affect jobs, skills, and wages? Retrieved from McKinsey & Company website: https://www.mckinsey.com/featured-insights/future-of-work/how-will-automation-affect-jobs-skills-and-wages

Metcalf, L., Askay, D. A., & Rosenberg, L. B. (2019). Keeping Humans in the Loop: Pooling Knowledge through Artificial Swarm Intelligence to Improve Business Decision Making. *California Management Review*, *61*(4), 84–109. https://doi.org/10.1177/0008125619862256

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research*, *98*(7), 261–276. https://doi.org/10.1016/j.jbusres.2019.01.044

Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2018). Big data analytics capabilities: a systematic literature review and research agenda. *Information Systems and E-Business Management*, *16*(3), 547–578. https://doi.org/10.1007/s10257-017-0362-y

Mikalef, P., & Pateli, A. (2017). Information technology-enabled dynamic capabilities and their indirect effect on competitive performance: Findings from PLS-SEM and fsQCA. *Journal of Business Research*, *70*(1), 1–16.

https://doi.org/https://doi.org/10.1016/j.jbusres.2016.09.004

Moore, R. (2017). Artificial intelligence: Is South Africa ready? Retrieved from ee publishers website: https://www.ee.co.za/article/artificial-intelligence-south-africa-ready.html

Morikawa, M. (2017). Firms' Expectations About the Impact of Ai and Robotics: Evidence From a Survey. *Economic Inquiry*, *55*(2), 1054–1063. https://doi.org/10.1111/ecin.12412

Muijs, D. (2011). *Doing Quantitative Research in Education with SPSS* (Second Edi). London: Sage publications.

Müller, O., Fay, M., & Brocke, J. (2018). The Effect of Big Data and Analytics on Firm Performance : An Econometric Analysis Considering Industry Characteristics. *Journal of Management Information Systems*, *35*(2), 488–509. https://doi.org/https://doi.org/10.1080/07421222.2018.1451955

Nunnally, J., & Bernstein, I. (1994). *Psychometric Theory* (3rd Editio). New York: McGraw-Hill.

O'Reilly, C. (1989). Corporations, Culture, and Commitment: Motivation and Social Control in Organizations. *California Management Review*, *31*(4), 9–25. https://doi.org/10.2307/41166580

Overgoor, G., Chica, M., Rand, W., & Weishampel, A. (2019). Letting the Computers Take Over: Using AI to Solve Marketing Problems. *California Management Review*, *61*(4), 156–185. https://doi.org/10.1177/0008125619859318

Pallant, J. (2005). *SPSS survival manual* (Second edi). Sydney: Allen and Unwin.

Pett, M., Lackey, N., & Sullivan, J. (2003). *Making Sense of Factor Analysis* (Third). California: SAGE Publications, Inc.

Pospiech, M., & Felden, C. (2012). Big Data A state pof the art. *AMCIS 2012 Proceedings*, 1–11. Retrieved from http://aisel.aisnet.org/amcis2012/proceedings/DecisionSupport/22/

Prasanna, S. R., & Haavisto, I. (2018). Collaboration in humanitarian supply chains: an organisational culture framework. *International Journal of Production Research*, *56*(17), 5611–5625. https://doi.org/10.1080/00207543.2018.1475762

Provost, F., & Fawcett, T. (2013). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big Data*, *1*(1), 51–59. https://doi.org/10.1089/big.2013.1508

Ransbotham, S., Kiron, D., Gerbert, P., & Reeves, M. (2017). Reshaping business with artificial intelligence. *MIT Sloan Mangement Review*, *59*(1), 1–17. https://doi.org/10.3934/dcdsb.2008.10.925

Rao, B. Y. A. (2017). A Strategist ' s Guide to Artificial Intelligence. *Tech and Innovation*, *4*(87), 46–50. Retrieved from https://www.strategy-business.com/article/A-Strategists-Guide-to-Artificial-Intelligence?gko=d2e2b

Rasmussen, T., & Ulrich, D. (2015). Learning from practice: How HR analytics avoids being a management fad. *Organizational Dynamics*, *44*(3), 236–242. https://doi.org/10.1016/j.orgdyn.2015.05.008

Salvato, C., & Vassolo, R. (2018). The sources of dynamism in dynamic capabilities. *Strategic Management Journal*, *39*(6), 1728–1752. https://doi.org/10.1002/smj.2703

Saunders, M., & Lewis, P. (2018). *Doing research in business and management* (2nd ed.). Harlow: Pearson.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147–177. https://doi.org/10.1037/1082-989X.7.2.147

Schein, E. H. (2010). *Organizational culture and leadership* (Fourth edi). San Francisco: Jossey-Bass.

Schilke, O. (2014). Second-Order Dynamic Capabilities : *Academy of Management Perspectives*, *28*(4), 368–380.

https://doi.org/10.5465/amp.2013.0093

Scholz, C. (1987). Corporate culture and strategy- The problem of strategic fit. *Long Range Planning*, *20*(4), 78–87. https://doi.org/10.1016/0024-6301(87)90158-0

Schumacker, R. E. (2015). *Learning Statistics Using R* (Second Edi). London: SAGE Publications, Inc.

Shapiro, S., & Wilk, M. (1965). An analysis of variance test for normality. *Biometrika*, *52*(3), 21. https://doi.org/10.2307/2333709

Shek, D. T. L., & Yu, L. (2014). Confirmatory factor analysis using AMOS: A demonstration. *Human Developmental Research: Experience from Research in Hong Kong*, *13*(2), 65–86. https://doi.org/10.1515/ijdhd-2014-0305

Shrestha, Y. R., Ben-Menahem, S. M., & von Krogh, G. (2019). Organizational Decision-Making Structures in the Age of Artificial Intelligence. *California Management Review*, *61*(4), 66–83. https://doi.org/10.1177/0008125619862257

Singh, K. (2007). *Quantitative Social Research Methods* (First edit). New Delhi: SAGE Publications.

Slevitch, L. (2011). Qualitative and quantitative methodologies compared: Ontological and epistemological perspectives. *Journal of Quality Assurance in Hospitality and Tourism*, *12*(1), 73–81. https://doi.org/10.1080/1528008X.2011.541810

Someh, I., Davern, M., Breidbach, C. F., & Shanks, G. (2019). Ethical Issues in Big Data Analytics: A Stakeholder Perspective. *Communications of the Association for Information Systems*, *44*(5), 718–747. https://doi.org/10.17705/1cais.04434

Srinivasan, R., & Swink, M. (2018). An Investigation of Visibility and Flexibility as Complements to Supply Chain Analytics: An Organizational Information Processing Theory Perspective. *Production and Operations Management*, *27*(10), 1849–1867. https://doi.org/10.1111/poms.12746

Stadler, C., Helfat, C. E., Verona, G., & Stadler, C. (2013). The impact of dynamic capabilities on resource access and development. *Organiasational Science*, *24*(6), 1601–1869. https://doi.org/https://doi.org/10.1287/orsc.1120.0810

Stanek, R. (2017). Getting Real Business Value From Artificial Intelligence. Retrieved July 20, 2019, from Big Data Quarterly website: http://www.dbta.com/BigDataQuarterly/Articles/Getting-Real-Business-Value--From-Artificial-Intelligence-118226.aspx

Steinmetz, S. (2016). Non-probability sampling. In *The SAGE Handbook of Survey Methodology* (First edit, pp. 329–343). London: Sage publications.

Sun, E. W., Chen, Y. T., & Yu, M. T. (2015). Generalized optimal wavelet decomposing algorithm for big financial data. *International Journal of Production Economics*, *165*(8), 194–214. https://doi.org/10.1016/j.ijpe.2014.12.033

Tabachnick, B., & Fidell, L. (2013). *Using Multivariate Statistics* (Sixth edit). New York: Pearson.

Tabesh, P., Mousavidin, E., & Hasani, S. (2019). Implementing big data strategies : A managerial perspective. *Business Horizons*, *62*(3), 347–358. https://doi.org/10.1016/j.bushor.2019.02.001

Tambe, P., Cappelli, P., & Yakubovich, V. (2019). Artificial Intelligence in Human Resources Management: Challenges and a Path Forward. *California Management Review*, *61*(4), 15–42. https://doi.org/10.1177/0008125619867910

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*(6), 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Teece, D. (2007). Explicating dynamic capabilities: The nature and microfoundations of(sustainable) enterprise performance. *Strategic Management Journal*, *28*(13), 1319–1350. https://doi.org/10.1002/smj.640

Teece, D. (2018a). Business models and dynamic capabilities. *Long Range Planning*, *51*(1), 40–49. https://doi.org/10.1016/j.lrp.2017.06.007

Teece, D. (2018b). Dynamic capabilities as (workable) management systems theory. *Journal of Management and Organization*, *24*(3), 359–368. https://doi.org/10.1017/jmo.2017.75

Teece, D., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Technological Know-How, Organizational Capabilities, and Strategic Management: Business Strategy and Enterprise Development in Competitive Environments*, *18*(7), 509–533. https://doi.org/10.1142/9789812834478_0002

Teece, D., Rumelt, R., Dosi, G., & Winter, S. (1994). Understanding corporate coherence. Theory and evidence. *Journal of Economic Behavior and Organization*, *23*(1), 1–30. https://doi.org/10.1016/0167-2681(94)90094-9

Tirunillai, S., & Tellis, G. J. (2014). Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent dirichlet allocation. *Journal of Marketing Research*, *51*(4), 463–479. https://doi.org/10.1509/jmr.12.0106

Toepoel, V. (2016). *Doing Surveys Online* (Second Edi). London: SAGE Publications Ltd.

Torres, R., Sidorova, A., & Jones, M. C. (2018). Enabling firm performance through business intelligence and analytics: A dynamic capabilities perspective. *Information and Management*, *55*(7), 822–839. https://doi.org/10.1016/j.im.2018.03.010

Vidgen, R., Shaw, S., & Grant, D. B. (2017). Management challenges in creating value from business analytics. *European Journal of Operational Research*, *261*(2), 626–639. https://doi.org/10.1016/j.ejor.2017.02.023

Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How “big data” can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, *165*(7), 234–246. https://doi.org/10.1016/j.ijpe.2014.12.031

Wamba, S. F., Gunasekaran, A., Akter, S., Ji, S., Ren, F., & Dubey, R. (2017). Big data analytics and firm performance: effects of dynamic capabilities. *Journal of Business Research*, *70*(219), 24–25. Retrieved from http://ro.uow.edu.au/cgi/viewcontent.cgi?article=1983&context=buspapers

Wamba, S. F., Gunasekaran, A., Akter, S., Ji, S., Ren, F., Dubey, R., … Papadopoulos, T. (2017). Social big data: Recent achievements and new challenges. *Journal of Business Research*, *70*(3), 263–286. https://doi.org/10.1016/j.techfore.2015.12.019

Wang, H., Zenshui, X., Fujita, H., & Shoushen, L. (2016). Towards felicitious decision making: An overview on challenges and trends of Big Data. *Information Sciences*, *367–368*(11), 747–765. https://doi.org/http://dx.doi.org/10.1016/j.ins.2016.07.007

Wang, Y., Kung, L. A., & Byrd, T. A. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change*, *126*(1), 3–13. https://doi.org/10.1016/j.techfore.2015.12.019

Wilden, R., & Gudergan, S. P. (2014). The impact of dynamic capabilities on operational marketing and technological capabilities: investigating the role of environmental turbulence. *Journal of the Academy of Marketing Science*, *43*(2), 181–199. https://doi.org/10.1007/s11747-014-0380-y

Winter, S. (2003). Understanding dynamic capabilities. *Strategic Management Journal*, *24*(10), 991–995. Retrieved from http://dx.doi.org/10.1002/smj.318

Ylijoki, O., & Porras, J. (2016). Perspectives to definition of Big Data: A mapping study and discussion. *Journal of Innovation Management*, *4*(1), 69. https://doi.org/10.24840/2183-0606_004.001_0006

Zhao, J. L., Fan, S., & Hu, D. (2014). Business challenges and research directions of management analytics in the big data era. *Journal of Management Analytics*, *1*(3), 169–174. https://doi.org/10.1080/23270012.2014.968643

Zikmund, W. G., Babin, B. J., Carr, J. C., & Griffin, M. (2010). *Business research methods* (8th editio). Canada: South-Western: CENGAGE Learning.

Zwitter, A. (2014). Big Data ethics. *Big Data & Society*, *1*(2), 1–6. https://doi.org/10.1177/2053951714559253

Appendices

Appendix A

Research instrument

Section 1: Context and Respondent

1. Are you aware of or associated with Big Data Analytics within the organisation being described in this questionnaire?	Yes/No
2. What is you gender?	Male/Female
3. What is your age?	• <20 years • 20-30 years • 31-40 years • 41-50 years • 51-60 years • >60 years
4. Which of the following best describes the principle industry of your organisation?	• Standard SurveyMonkeyTM drop down menu of industries
5. What is your main association with the data analytics capability?	• User of analytics within business • Data analyst (direct processor of data) • IT systems or infrastructure (Data technology environment) • Big Data management (driving application of resources)
6. What is the approximate total number of employees in the organisation?	• 1-99 • 100-499

	• 500-999 • 1000 or more
7. Which of the following best describes your current job level?	• Owner/executive/C-level • Senior management • Middle management • Intermediate • Entry level
8. In what country do you work?	• Standard SurveyMonkeyTM drop down menu of industries

<u>Section 2: Big Data Analytics Management Capabilities</u>

9. We continuously examine the innovative opportunities for the strategic use of big data analytics	7 point Likert scale
10. When we make business analytics investment decisions, we estimate the time managers will need to spend overseeing the change	7 point Likert scale
11. When we make big data analytics investment decisions, we project about how much these options will help end-users make quicker decisions	7 point Likert scale
12. Our analytics personnel work closely with customers and maintain productive user/client relationships	7 point Likert scale
13. We enforce adequate plans for the introduction and utilization of big data analytics	7 point Likert scale

14. We perform big data analytics planning processes in systematic and formalized ways	7 point Likert scale
15. We frequently adjust big data analytics plans to better adapt to changing conditions	7 point Likert scale
16. When we make big data analytics investment decisions, we think about and estimate the effect they will have on the productivity of the employees' work	7 point Likert scale
17. When we make big data analytics investment decisions, we think about and estimate the cost of training that end-users will need	7 point Likert scale
18. In our organisation, business analysts and line people meet regularly to discuss important issues	7 point Likert scale
19. In our organisation, business analysts and line people from various departments regularly attend cross-functional meetings	7 point Likert scale
20. In our organisation, information is widely shared between business analysts and line people so that those who make decisions or perform jobs have access to all available know-how	7 point Likert scale
21. In our organisation, the responsibility for big data analytics development is clear	7 point Likert scale

22. We are confident that big data analytics project proposals are properly appraised	7 point Likert scale
23. Our analytics department is clear about its performance criteria	7 point Likert scale
24. Our company is better than competitors in connecting (e.g., communication and information sharing) parties within a business process	7 point Likert scale
25. Our company is better than competitors in bringing detailed information into a business process	7 point Likert scale

Section 3: OC to Big Data Analytics Capability

26. Our organisation has a widely held belief that innovation is an absolute necessity for the organisation's future	7 point Likert scale
27. Our organisation enables learning, accumulation and application of new knowledge better than our competitors	7 point Likert scale
28. We believe it is important to adopt new and cutting-edge practices to continuously improve product or service delivery	7 point Likert scale
29. People in our organisation are continuously encouraged to expand their capacities to achieve more and apply new capabilities	7 point Likert scale

30. Our organisation can be described as visionary and flexible	7 point Likert scale
31. There is an extensive employee orientation program for new employees to ensure employees share the corporate vision and purpose	7 point Likert scale
32. We invest in targeted training and support at all levels of our organisation to assist our organisation to understand or know how to use data that is available	7 point Likert scale
33. Our executive level actively and visibly supports our big data analytics capability	7 point Likert scale

Section 4: Big Data Analytics Infrastructure Flexibility

34. Compared to rivals within our industry, our organisation has the foremost available analytics systems 7-point Likert scale)	7 point Likert scale
35. All other (e.g., remote, branch, and mobile) offices are connected to the central office for analytics	7 point Likert scale
36. There are no identifiable communications bottlenecks within our organisation when sharing analytics insights	7 point Likert scale
37. Software applications can be easily transported and used across multiple analytics platforms	7 point Likert scale

38. Our user interfaces provide transparent access to all platforms and applications	7 point Likert scale
39. Reusable software modules are widely used in new analytics model development	7 point Likert scale
40. The legacy system within our organisation restricts the development of new applications	7 point Likert scale
41. End-users can use software provided to create their own analytics applications	7 point Likert scale

Section 5: Big data analytics Personnel Expertise Capability

42. Our analytics personnel are very capable in terms of programming skills (e.g., structured programming, web-based application, CASE, tools, SQL etc.)	7 point Likert scale
43. Our analytics personnel are very capable in terms of managing project life cycles	7 point Likert scale
44. Our analytics personnel are very capable in the areas of data and network management and maintenance	7 point Likert scale
45. Our analytics personnel are very capable in data decision support systems (e.g., expert systems, AI, warehousing, mining, marts, etc.)	7 point Likert scale

46. Our analytics personnel are very knowledgeable about the critical factors for the success of our organisation	7 point Likert scale
47. Our analytics personnel are very capable in interpreting business problems and developing appropriate technical solutions	7 point Likert scale
48. Our analytics personnel are very knowledgeable about our business environment	7 point Likert scale
49. Our analytics personnel are very capable in terms of planning and executing work in a collective environment	7 point Likert scale
50. Our analytics personnel are very capable in terms of teaching others in our business	7 point Likert scale

Section 6: AI to Big Data Analytics Capability

51. My organisation is safe from automation	7 point Likert scale
52. My organisation is currently going through change due to automation of tasks	7 point Likert scale
53. Machine learning will replace tasks performed by humans at the workplace in the next 10 years	7 point Likert scale
54. Our organisation is planning to or currently making use of machine learning to create business value	7 point Likert scale

55. Our organisation does not have a good understanding of the potential uses of machine learning to increase FPer	7 point Likert scale
56. We believe that using the latest technology will assist in developing our Big Data Analytics Capabilities	7 point Likert scale
57. Our organisation can be described as adverse to technological change	7 point Likert scale
58. How positive or negative are the consequences of human level machine learning likely to be in your organisation	7 point Likert scale

Section 7: Firm Financial and Market Performance

59. Using big data analytics improved customer retention during the last 3 years relative to competitors	7 point Likert scale
60. Using big data analytics improved Sales Growth during the last 3 years relative to competitors	7 point Likert scale
61. Using big data analytics improved Profitability during the last 3 years relative to competitors	7 point Likert scale
62. Using big data analytics improved Return on Investment (ROI) during the last 3 years relative to competitors	7 point Likert scale

63. Using big data analytics improved overall financial performance during the last 3 years relative to competitors	7 point Likert scale
64. Our success rate of new products or services has been higher than our competitors	7 point Likert scale
65. Using analytics our market share has exceeded that of our competitors	7 point Likert scale

Appendix B

Table 16: Tests for normality

First Order Constructs	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Planning	0.237	100	0	0.893	100	0
Investment Decision Making	0.11	100	0	0.956	100	0
Coordination	0.097	100	0.02	0.971	100	0.03
Control	0.11	100	0	0.959	100	0
Organisational Culture	0.133	100	0	0.915	100	0
Connectivity	0.135	100	0	0.945	100	0
Modularity	0.088	100	0.06	0.976	100	0.06
Technical Knowledge	0.156	100	0	0.914	100	0
Business Knowledge	0.236	100	0	0.857	100	0
Relational Knowledge	0.171	100	0	0.922	100	0
AI	0.105	100	0.01	0.955	100	0
FPER	0.116	100	0	0.971	100	0.03

Appendix C

Table 17: Assessing validity

First order constructs and construct total		
	Sig	Pearson's Correlation
		Planning Total
BDAMC1	0.00	0,629
BDAMC5	0.00	0,795
BDAMC6	0.00	0,842
BDAMC7	0.00	0,756
		Investment Decision making
BDAMC2	0.00	0,818
BDAMC3	0.00	0,784
BDAMC8	0.00	0,808
BDAMC9	0.00	0,801
		Coordination
BDAMC10	0.00	0,811
BDAMC11	0.00	0,869
BDAMC12	0.00	0,819
		Control
BDAMC13	0.00	0,746
BDAMC14	0.00	0,745
BDAMC15	0.00	0,822
BDAMC16	0.00	0,837
BDAMC17	0.00	0,842
		Connectivity
BDAIF1	0.00	0,820
BDAIF2	0.00	0,805
BDAIF3	0.00	0,822
		Compatibility
BDAIF4	0.00	0,908
BDAIF5	0.00	0,901
		Modularity
BDAIF6	0.00	0,567
BDAIF7	0.00	0,355
BDAIF8	0.00	0,648
		Technical Knowledge
BDASPEC1	0.00	0,897
BDASPEC2	0.00	0,884
BDASPEC4	0.00	0,693
		Technological Mangement Knowledge
BDASPEC3	0.00	0,920
BDASPEC5	0.00	0,887
		Business Knowledge
BDASPEC6	0.00	0,930
BDASPEC7	0.00	0,924
		Relational Knowledge
BDASPEC8	0.00	0,916
BDASPEC9	0.00	0,828
		Artificial Intelligence
MODAI1	0.00	0,345
MODAI2	0.00	0,388
MODAI3	0.00	0,504
MODAI4	0.00	0,454
MODAI5	0.00	0,401
MODAI6	0.00	0,428
MODAI7	0.00	0,434
MODAI8	0.00	0,337

		Organisational Culture
OC1	0.00	0,781
OC2	0.00	0,718
OC3	0.00	0,679
OC4	0.00	0,767
OC5	0.00	0,797
OC6	0.00	0,651
OC7	0.00	0,753
OC8	0.00	0,805
		Firm Performance
FPER1	0.00	0,732
FPER2	0.00	0,897
FPER3	0.00	0,885
FPER4	0.00	0,907
FPER5	0.00	0,911
FPER6	0.00	0,775
FPER7	0.00	0,837

Appendix D

AMOS Models

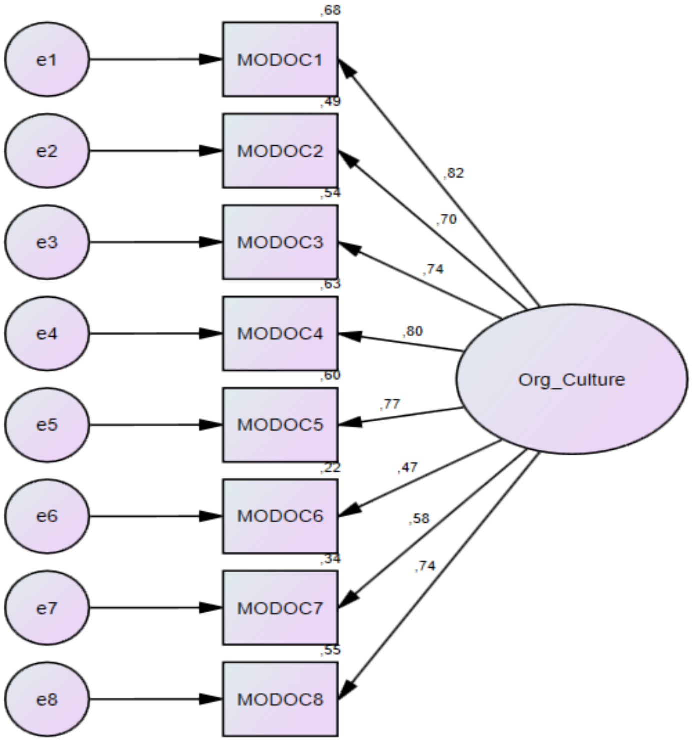


Figure 11: CFA of OC

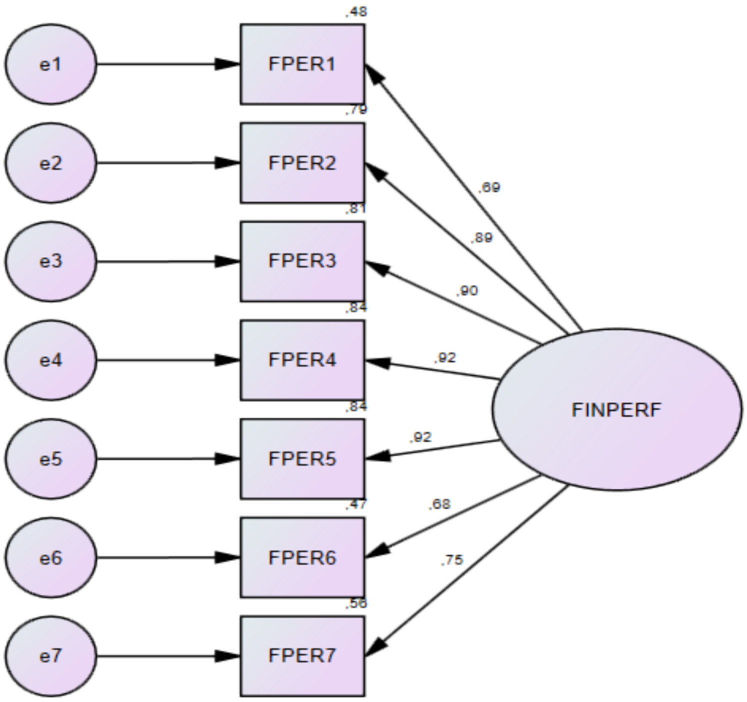


Figure 12: CFA of FPer

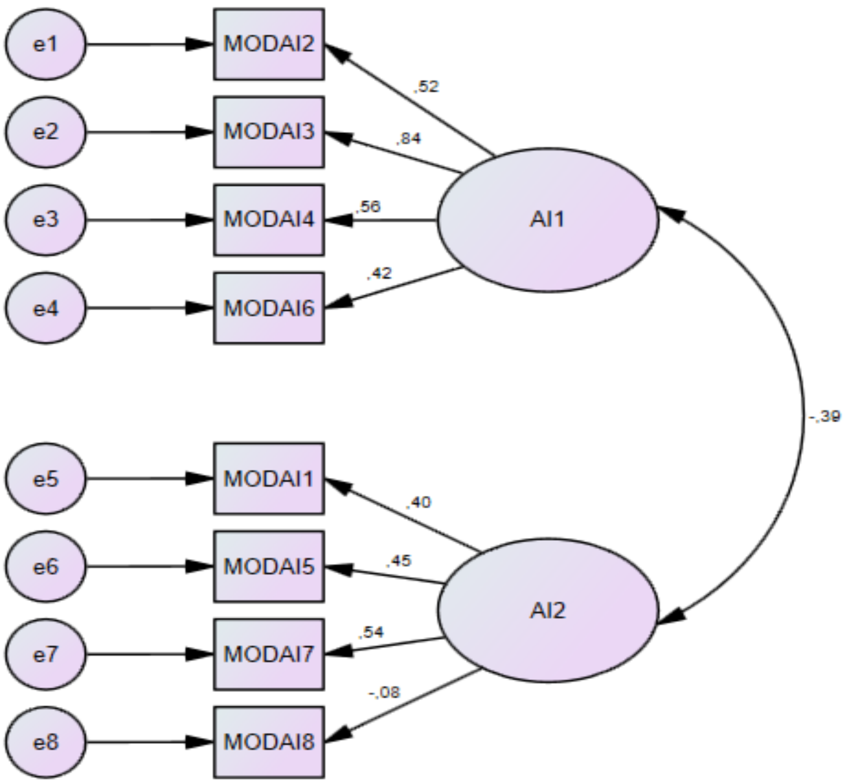


Figure 13: CFA of AI

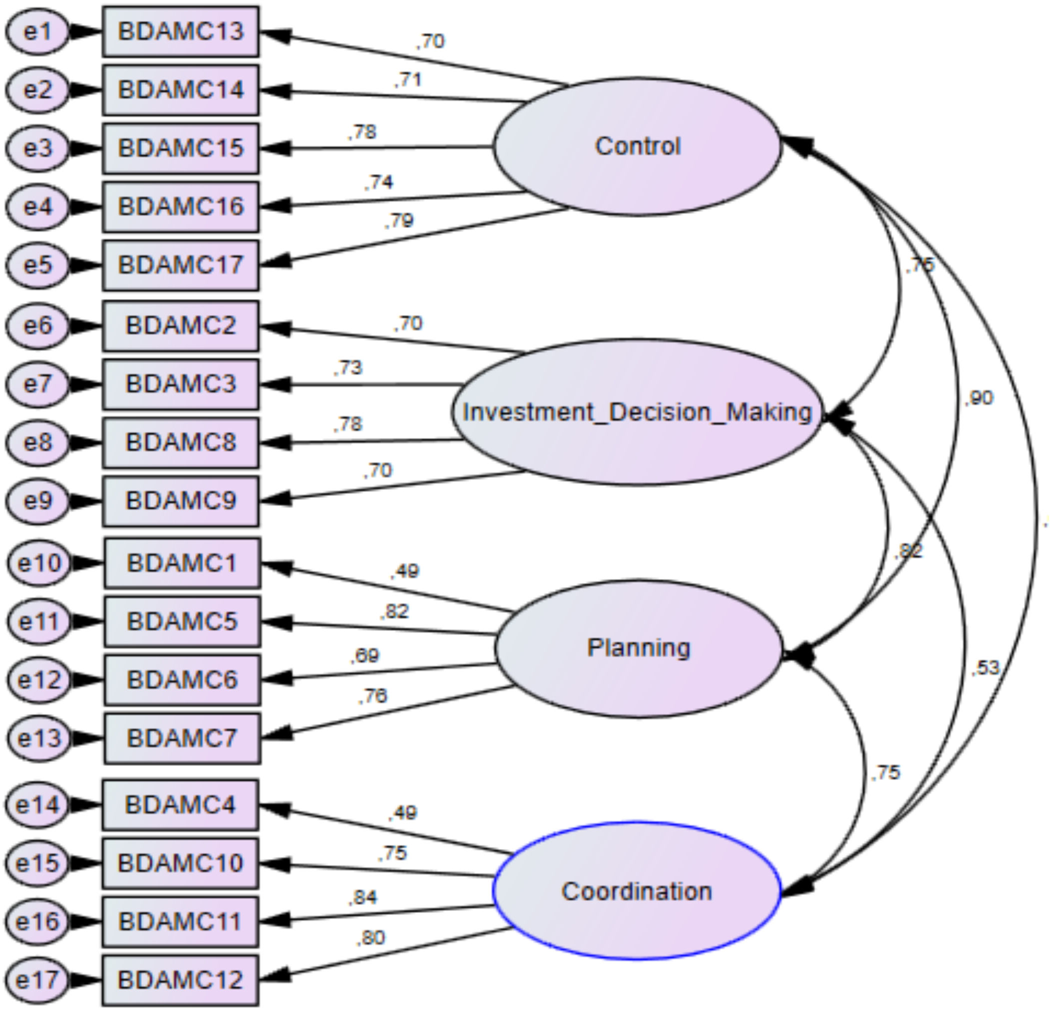


Figure 14: CFA of BDAMC

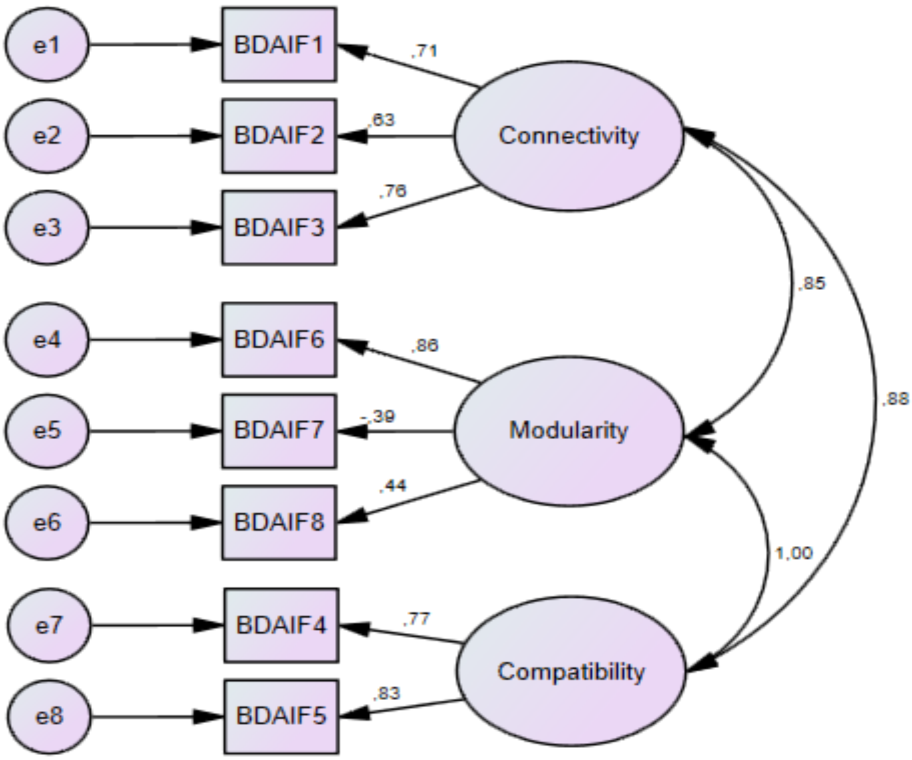


Figure 15: CFA of BDAIF

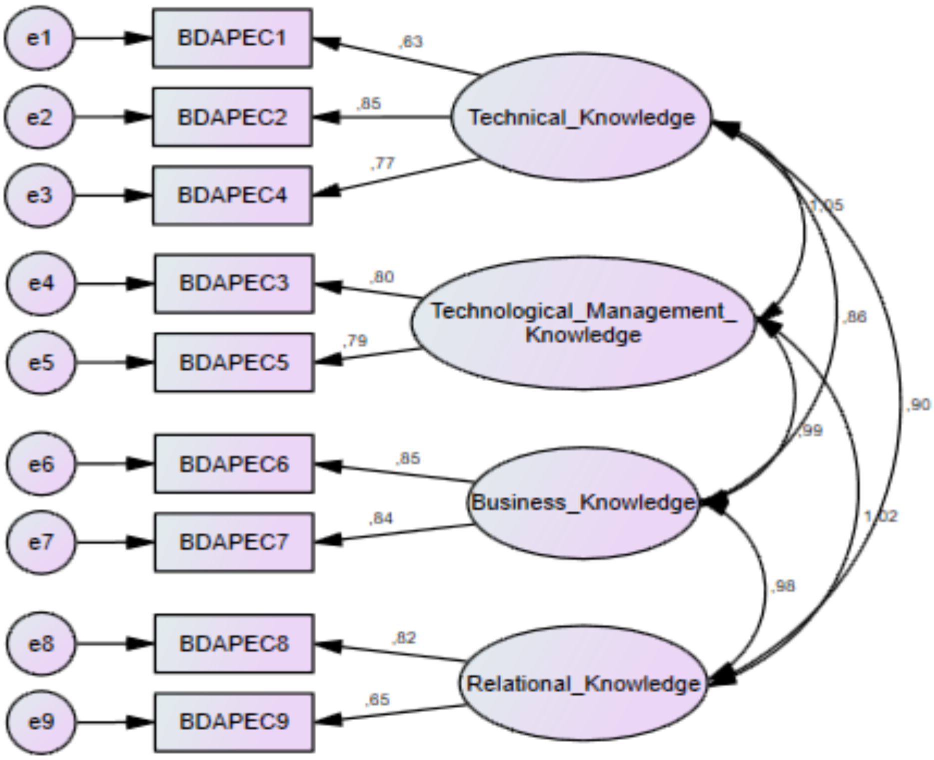


Figure 16: CFA of BDAPEC

Appendix E

Table 18: KMO and Barlett's test for sphericity

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0,868
Bartlett's Test of Sphericity	Approx. Chi-Square	407,131
	df	28
	Sig.	0,000

Table 19: PCA extraction of total variance explained for OC

Total Variance Explained						
	Initial Eigenvalues			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
OC1	4,52	56,48	56,48	3,52	43,99	43,99
OC2	1,15	14,31	70,79	2,14	26,80	70,79
OC3	0,58	7,22	78,01			
OC4	0,48	5,94	83,95			
OC5	0,40	5,05	89,01			
OC6	0,34	4,20	93,21			
OC7	0,30	3,76	96,97			
OC8	0,24	3,03	100,00			
Extraction Method: Principal Component Analysis.						

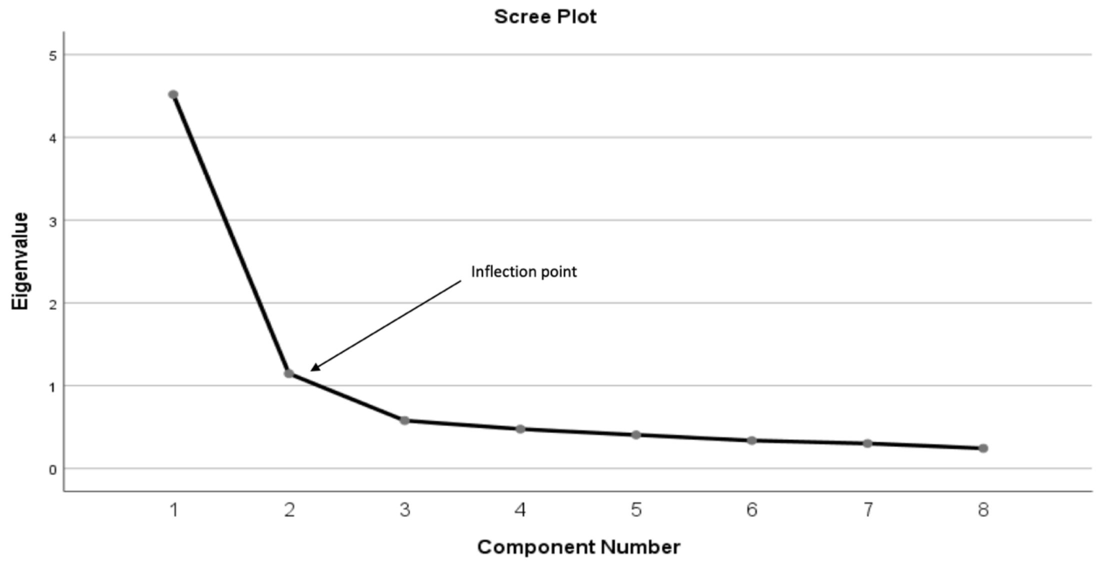


Figure 17: Scree plot of OC with inflection point

Table 20: Rotated component matric for EFA of OC

Rotated Component Matrix[a]		
	Component	
	1	2
OC1	0,83	0,25
OC2	0,74	0,22
OC3	0,84	0,08
OC4	0,83	0,21
OC5	0,69	0,43
OC6	0,10	0,87
OC7	0,27	0,84
OC8	0,58	0,57

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

Appendix F

Table 21: Descriptive statistics per question and variable

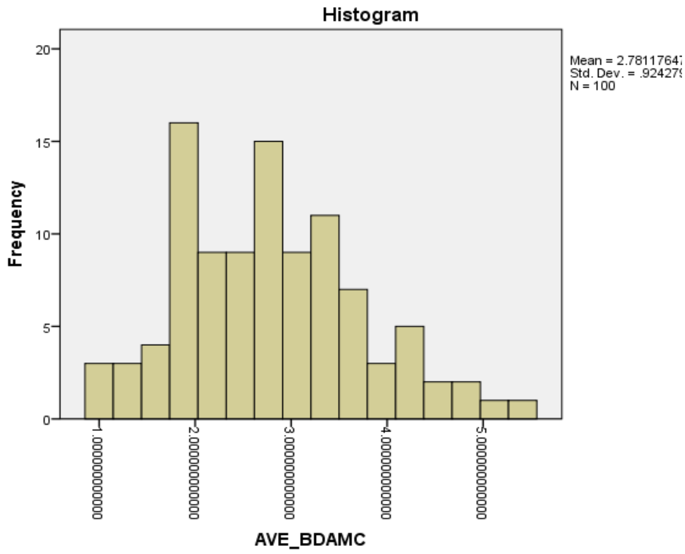
Descriptive Statistics

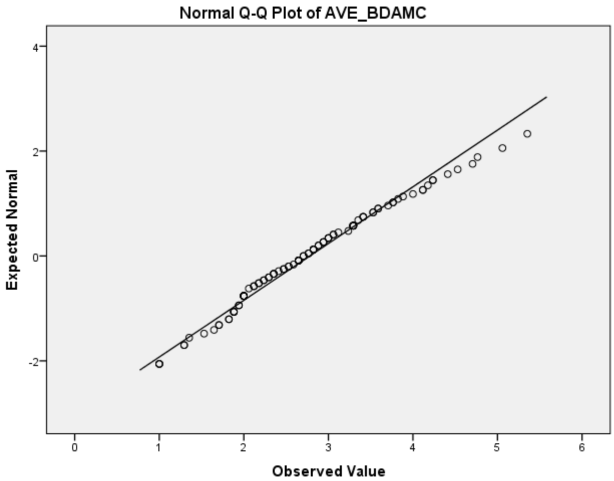
Question	N	Mean	Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
BDAMC1	100	1,98	1,08	1,17	1,46	0,24	2,34	0,48
BDAMC2	100	3,14	1,55	2,40	0,39	0,24	(0,31)	0,48
BDAMC3	100	2,29	1,21	1,46	0,61	0,24	(0,53)	0,48
BDAMC4	100	2,38	1,36	1,85	1,14	0,24	0,79	0,48
BDAMC5	100	3,04	1,30	1,70	0,29	0,24	(0,04)	0,48
BDAMC6	100	2,90	1,44	2,07	1,07	0,24	0,80	0,48
BDAMC7	100	2,63	1,26	1,59	1,29	0,24	1,92	0,48
BDAMC8	100	2,55	1,31	1,72	0,91	0,24	0,95	0,48
BDAMC9	100	2,89	1,50	2,24	0,74	0,24	0,30	0,48
BDAMC10	100	2,88	1,34	1,78	0,56	0,24	(0,19)	0,48
BDAMC11	100	3,05	1,39	1,93	0,63	0,24	0,24	0,48
BDAMC12	100	2,98	1,33	1,76	0,46	0,24	(0,12)	0,48
BDAMC13	100	2,77	1,35	1,84	0,85	0,24	0,10	0,48
BDAMC14	100	2,98	1,39	1,94	0,98	0,24	0,77	0,48
BDAMC15	100	2,82	1,42	2,03	0,80	0,24	(0,05)	0,48
BDAMC16	100	3,01	1,55	2,41	0,66	0,24	(0,31)	0,48
BDAMC17	100	2,99	1,35	1,83	0,54	0,24	0,39	0,48
Ave_Planning	100	2,72	1,04	1,07	0,59	0,24	0,18	0,48
Ave_Investment-Decision making	100	2,72	1,12	1,25	0,65	0,24	0,42	0,48
Ave_Coordination	100	2,82	1,07	1,14	0,43	0,24	(0,29)	0,48
Ave_Control	100	2,91	1,13	1,28	0,73	0,24	0,43	0,48
AVE_BDAMC	100	2,78	0,92	0,85	0,44	0,24	(0,06)	0,48

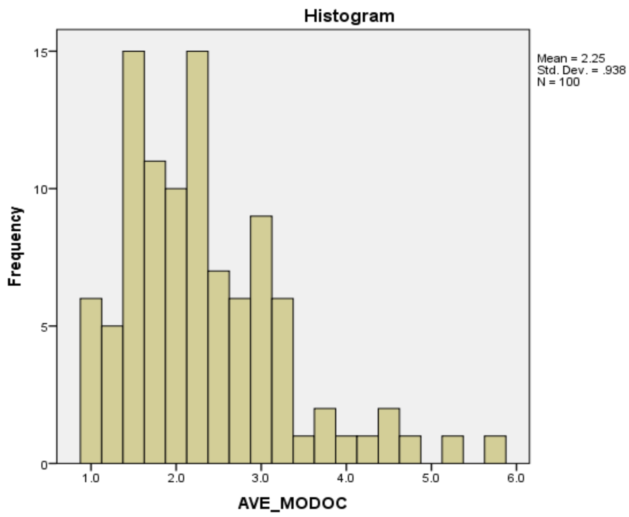
MODOC1	100	1,69	1,00	1,00	1,95	0,24	4,43	0,48
MODOC2	100	2,31	1,28	1,65	1,15	0,24	1,26	0,48
MODOC3	100	1,69	0,76	0,58	1,15	0,24	1,45	0,48
MODOC4	100	2,14	1,17	1,37	0,76	0,24	(0,25)	0,48
MODOC5	100	2,37	1,32	1,73	1,54	0,24	2,70	0,48
MODOC6	100	2,72	1,64	2,69	1,11	0,24	0,41	0,48
MODOC7	100	2,79	1,55	2,39	0,98	0,24	0,34	0,48
MODOC8	100	2,29	1,37	1,86	1,45	0,24	2,15	0,48
Ave_Cultural_Development	100	2,08	0,93	0,86	1,41	0,24	2,22	0,48
AVE_Employee_investment	100	2,76	1,43	2,04	1,07	0,24	0,86	0,48
AVE_MODOC	100	2,25	0,94	0,88	1,18	0,24	1,62	0,48
BDAIF1	100	3,13	1,60	2,56	0,64	0,24	(0,28)	0,48
BDAIF2	100	2,77	1,50	2,24	0,81	0,24	(0,24)	0,48
BDAIF3	100	3,53	1,60	2,55	0,49	0,24	(0,73)	0,48
BDAIF4	100	3,24	1,54	2,37	0,28	0,24	(0,70)	0,48
BDAIF5	100	3,23	1,48	2,20	0,39	0,24	(0,23)	0,48
BDAIF6	100	3,19	1,43	2,03	0,28	0,24	(0,56)	0,48
BDAIF7	100	3,50	1,82	3,32	0,44	0,24	(1,00)	0,48
BDAIF8	100	3,69	1,62	2,64	0,33	0,24	(0,60)	0,48
Ave_Connectivity	100	3,14	1,28	1,63	0,70	0,24	(0,06)	0,48
Ave_Compatibility	100	3,24	1,37	1,87	0,33	0,24	(0,50)	0,48
Ave_Modularity	100	3,44	1,26	1,59	0,33	0,24	(0,31)	0,48
AVE_BDIF	100	3,29	0,94	0,87	0,19	0,24	(0,72)	0,48
BDAPEC1	100	2,31	1,38	1,89	1,11	0,24	0,48	0,48
BDAPEC2	100	2,58	1,30	1,70	1,33	0,24	2,00	0,48
BDAPEC3	100	2,55	1,31	1,72	1,32	0,24	1,92	0,48
BDAPEC4	100	2,63	1,33	1,77	1,13	0,24	1,24	0,48
BDAPEC5	100	2,41	1,12	1,25	1,33	0,24	2,82	0,48

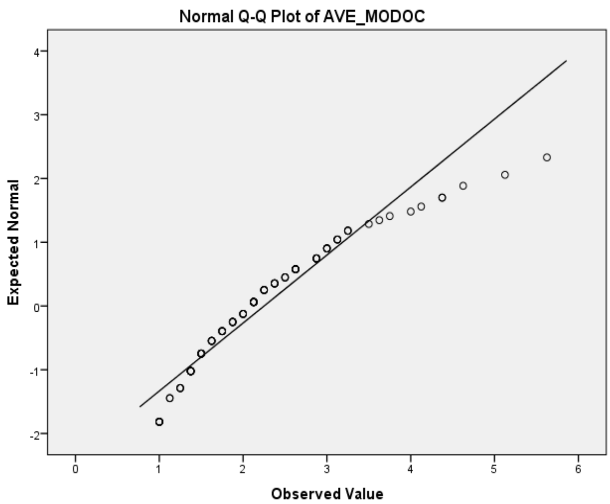
BDAPEC6	100	2,54	1,27	1,62	1,07	0,24	1,10	0,48
BDAPEC7	100	2,38	1,23	1,51	1,70	0,24	3,56	0,48
BDAPEC8	100	2,50	1,17	1,36	1,30	0,24	2,27	0,48
BDAPEC9	100	2,37	0,84	0,70	0,69	0,24	0,90	0,48
Ave_Technical_knowledge	100	2,45	1,19	1,42	1,01	0,24	0,65	0,48
Ave_Technical_man_Knowledge	100	2,48	1,10	1,21	1,30	0,24	2,87	0,48
Ave_Business_Knowledge	100	2,46	1,16	1,35	1,40	0,24	2,31	0,48
Ave_Relational_Knowledge	100	2,44	0,88	0,78	0,88	0,24	1,95	0,48
AVE_BDAPEC	100	2,47	0,97	0,94	1,30	0,24	2,64	0,48
MODAI1	100	4,37	1,69	2,84	(0,19)	0,24	(0,83)	0,48
MODAI2	100	2,74	1,45	2,09	1,10	0,24	0,55	0,48
MODAI3	100	2,28	1,32	1,74	1,73	0,24	3,39	0,48
MODAI4	100	2,71	1,37	1,86	0,86	0,24	0,60	0,48
MODAI5	100	4,23	1,64	2,70	(0,14)	0,24	(0,96)	0,48
MODAI6	100	2,02	0,88	0,77	1,25	0,24	2,55	0,48
MODAI7	100	5,09	1,78	3,17	(0,82)	0,24	(0,53)	0,48
MODAI8	100	2,44	1,03	1,06	0,68	0,24	(0,05)	0,48
AVE_MODAI	100	2,58	1,06	1,13	0,95	0,24	1,24	0,48
FPER1	100	2,96	1,19	1,41	0,63	0,24	1,29	0,48
FPER2	100	2,91	1,18	1,40	0,48	0,24	0,43	0,48
FPER3	100	2,90	1,20	1,44	0,52	0,24	0,19	0,48
FPER4	100	3,02	1,26	1,60	0,42	0,24	(0,18)	0,48
FPER5	100	3,01	1,24	1,55	0,56	0,24	0,16	0,48
FPER6	100	3,20	1,41	2,00	0,45	0,24	(0,27)	0,48
FPER7	100	3,49	1,38	1,91	0,28	0,24	(0,37)	0,48
AVE_FPER	100	3,07	1,07	1,16	0,46	0,24	0,08	0,48
Ave_BDAC	100	2,80	0,82	0,67	0,51	0,24	0,14	0,48

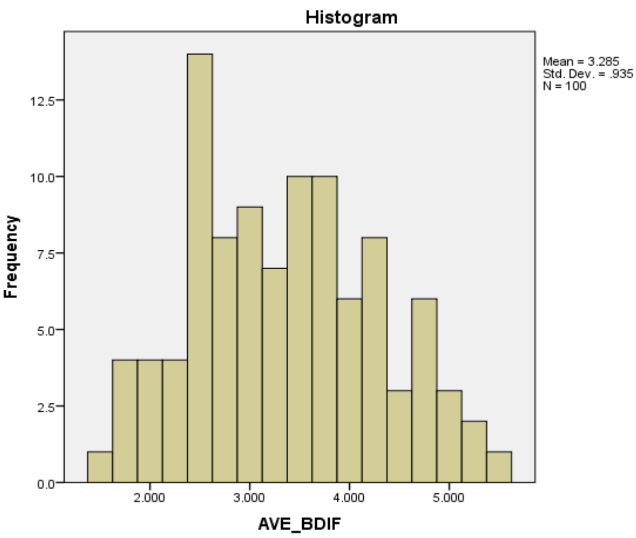
Appendix G

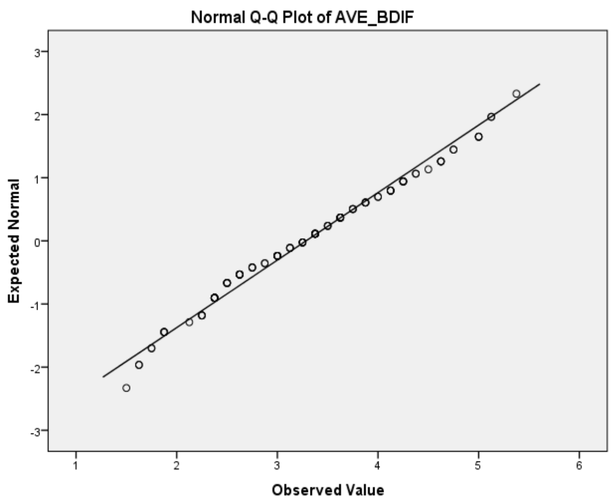

Histogram
Frequency
Mean = 2.7811764
Std. Dev. = .92427
N = 100
AVE_BDAMC


Normal Q-Q Plot of AVE_BDAMC
Expected Normal
Observed Value


Histogram
Frequency
Mean = 2.25
Std. Dev. = .938
N = 100
AVE_MODOC


Normal Q-Q Plot of AVE_MODOC
Expected Normal
Observed Value


Histogram
Frequency
Mean = 3.285
Std. Dev. = .935
N = 100
AVE_BDIF


Normal Q-Q Plot of AVE_BDIF
Expected Normal
Observed Value

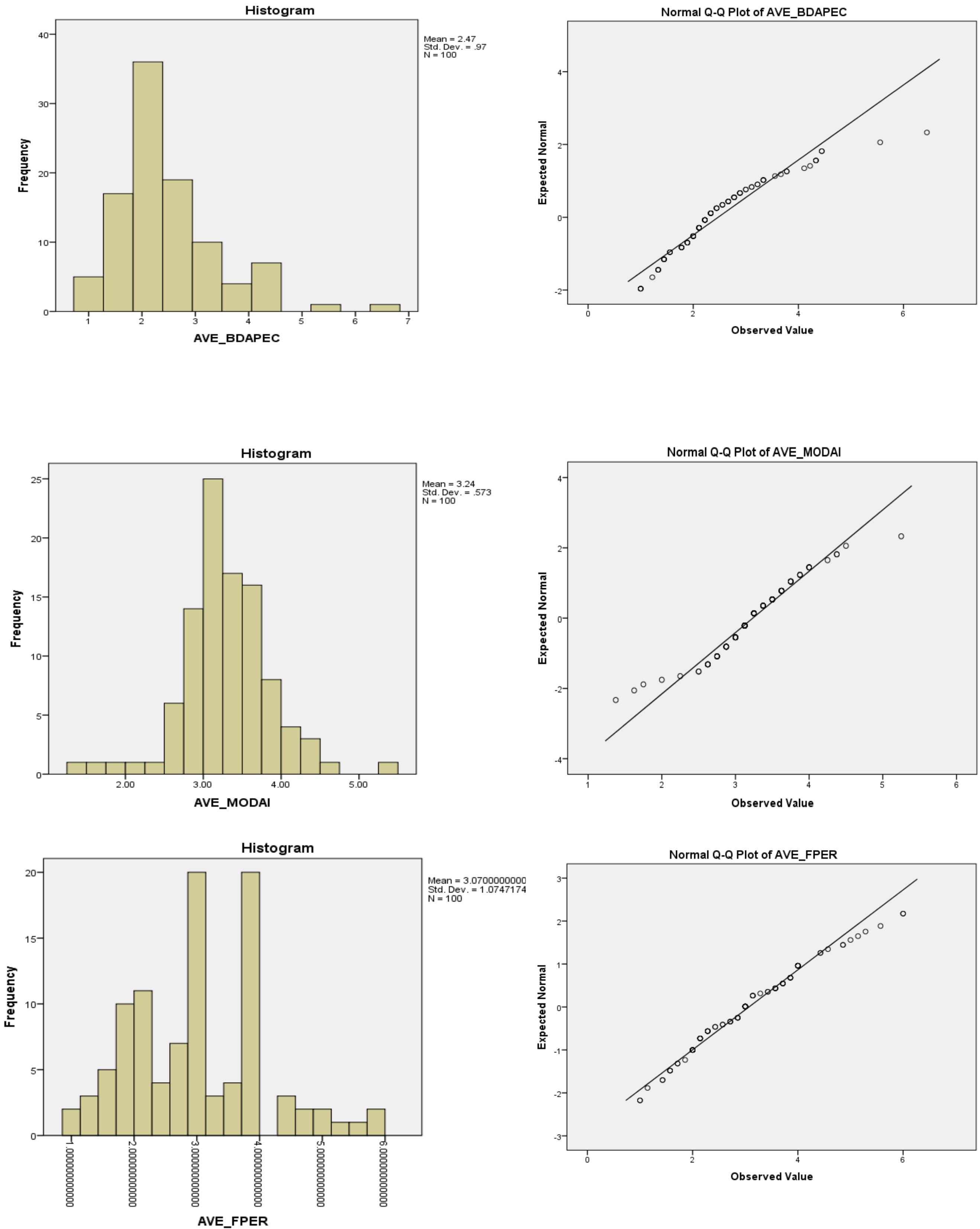


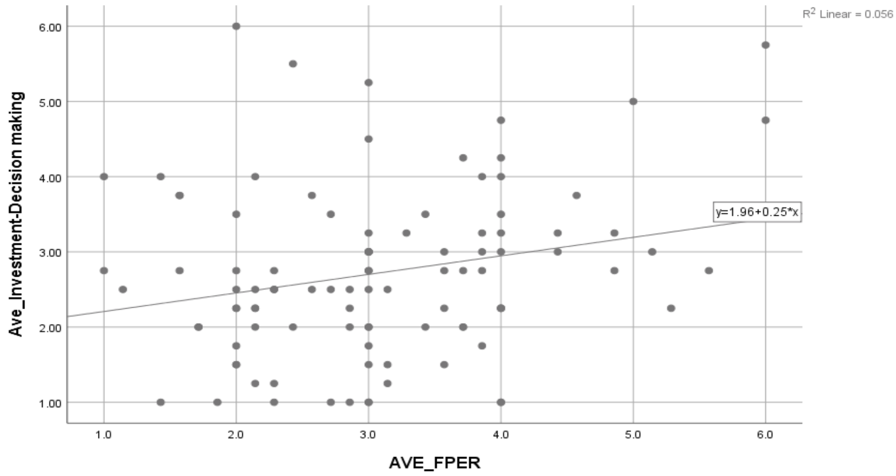
Figure 18: Histograms and QQ plots of Variables

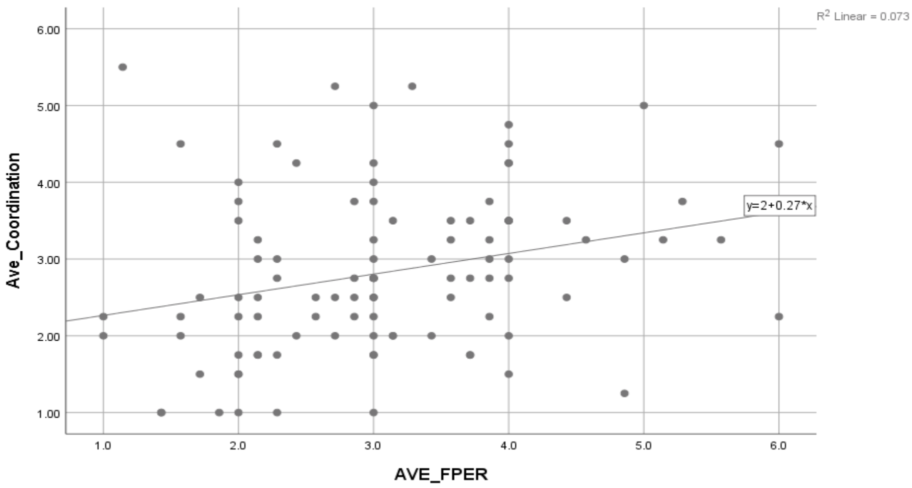
Appendix H

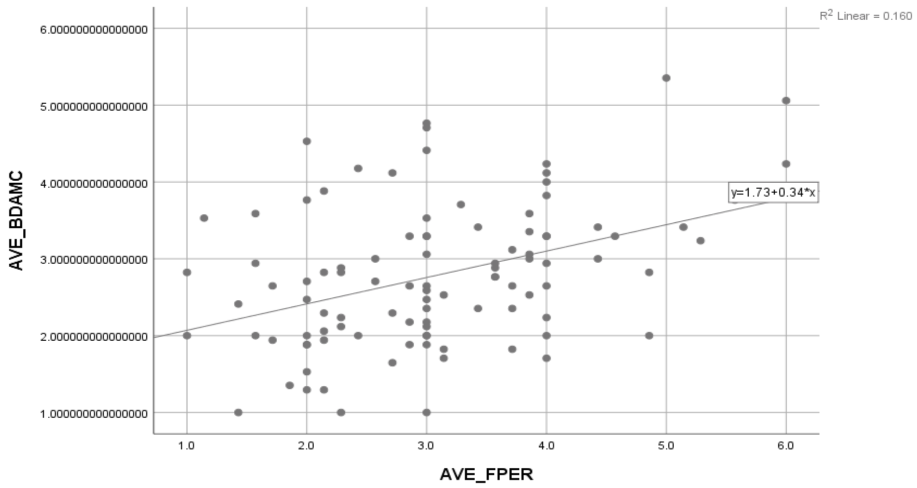
Table 22: Hierarchical multiple regression results of proposed model

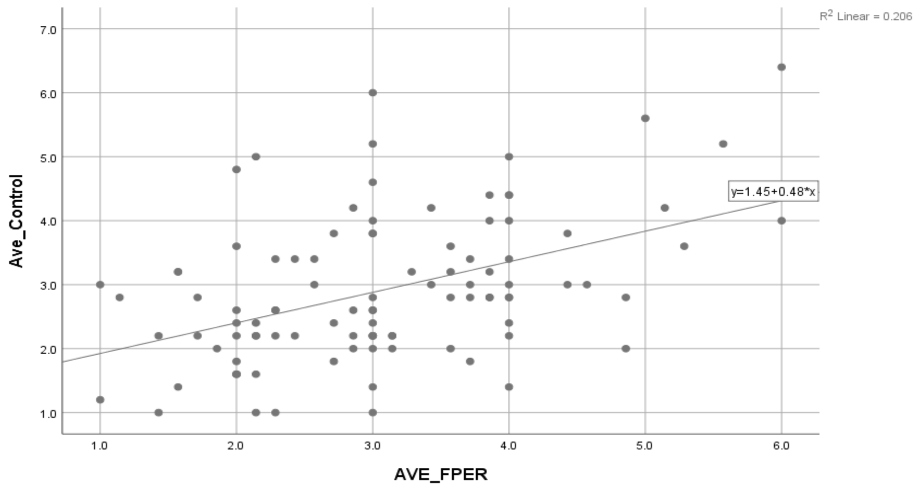
Model Summary									
					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.456[a]	0,21	0,20	0,96	0,21	25,72	1	98	0,00
2	.524[b]	0,27	0,25	0,93	0,07	4,44	2	96	0,01
3	.525[c]	0,28	0,24	0,94	0,00	0,04	2	94	0,96
4	.597[d]	0,36	0,31	0,89	0,08	5,77	2	92	0,00
a. Predictors: (Constant), Ave_BDAC									
b. Predictors: (Constant), Ave_BDAC, MOD_BDACXOCCD, Ave_Cultural_Development									
c. Predictors: (Constant), Ave_BDAC, MOD_BDACXOCCD, Ave_Cultural_Development, AVE_Employee_investment, MOD_BDACXOCEI									
d. Predictors: (Constant), Ave_BDAC, MOD_BDACXOCCD, Ave_Cultural_Development, AVE_Employee_investment, MOD_BDACXOCEI, MOD_BDACXAI, AVE_MODAI									

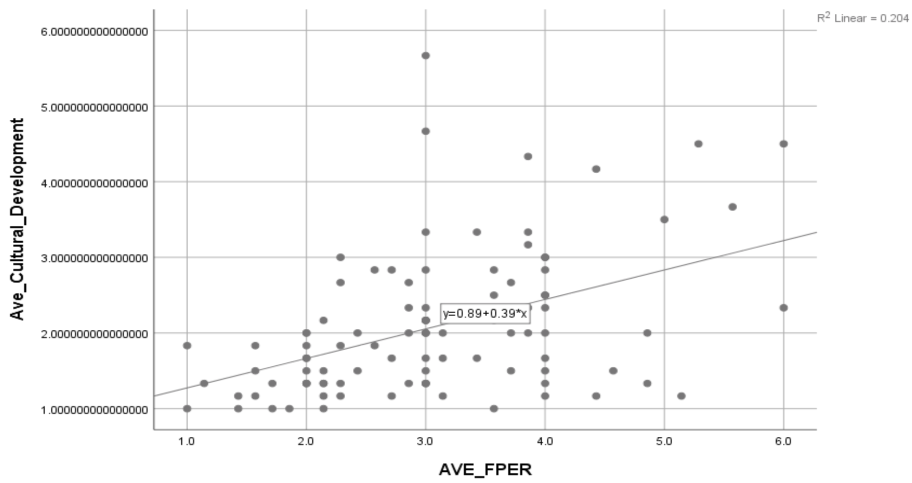
Appendix I

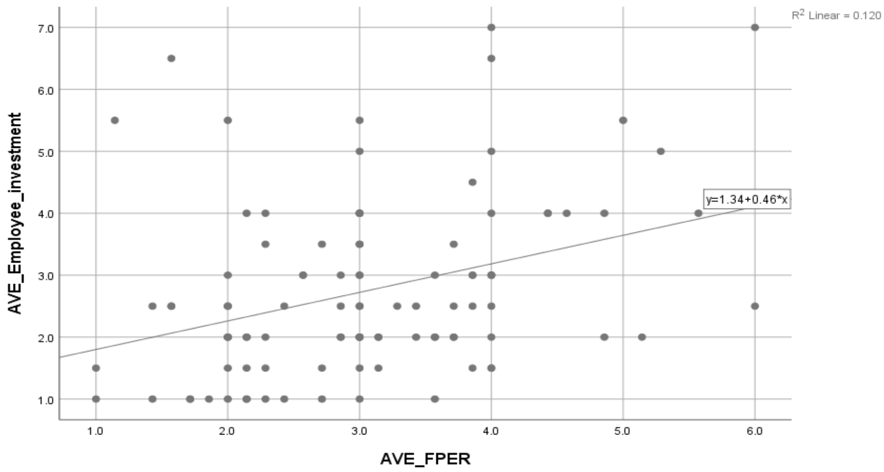

R^2 Linear = 0.056
y=1.96+0.25*x
Ave_Investment-Decision making
AVE_FPER
6.00
5.00
4.00
3.00
2.00
1.00
1.0
2.0
3.0
4.0
5.0
6.0

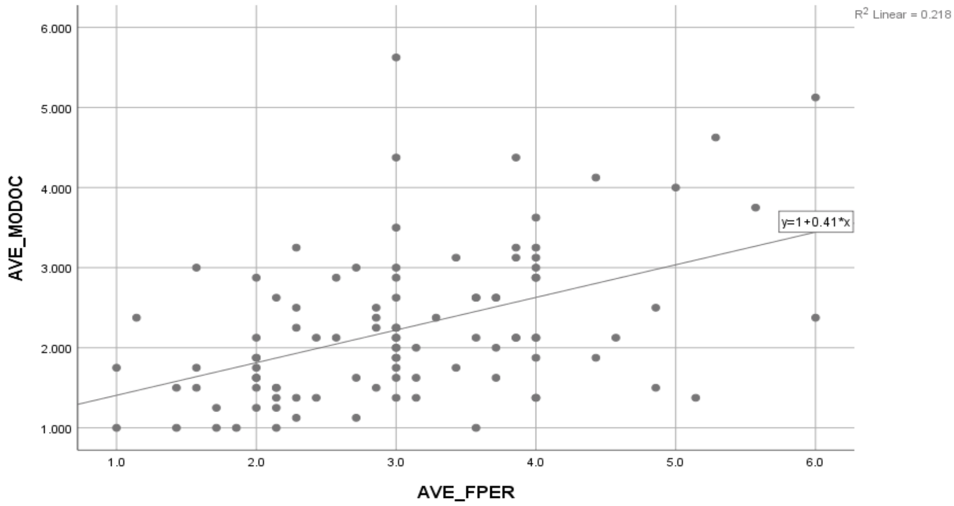

R^2 Linear = 0.073
y=2+0.27*x
Ave_Coordination
AVE_FPER
6.00
5.00
4.00
3.00
2.00
1.00
1.0
2.0
3.0
4.0
5.0
6.0

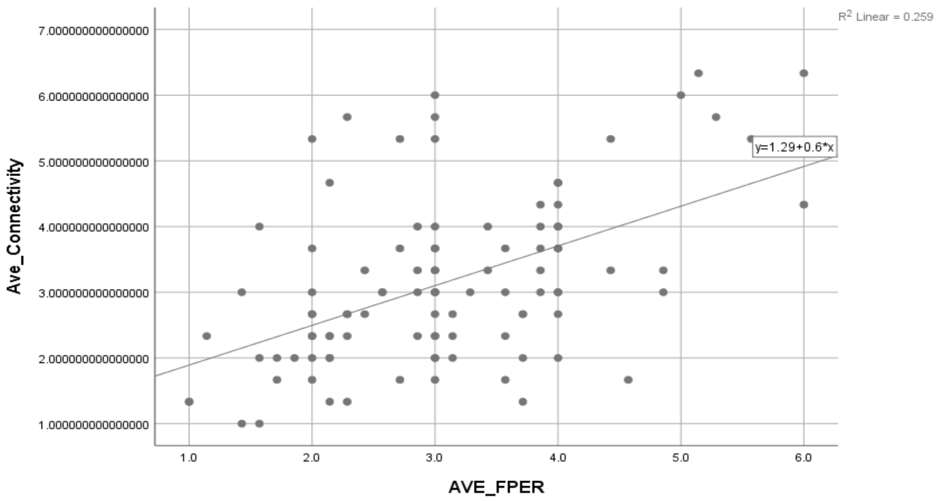

R2 Linear = 0.160
y=1.73+0.34*x
AVE_BDAMC
AVE_FPER
6.000000000000000
5.000000000000000
4.000000000000000
3.000000000000000
2.000000000000000
1.000000000000000
1.0
2.0
3.0
4.0
5.0
6.0

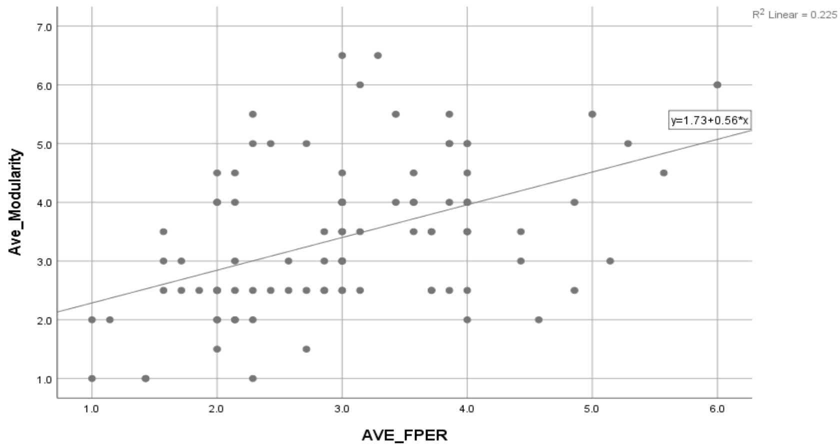

R2 Linear = 0.206
y=1.45+0.48*x
Ave_Control
AVE_FPER
7.0
6.0
5.0
4.0
3.0
2.0
1.0
1.0
2.0
3.0
4.0
5.0
6.0

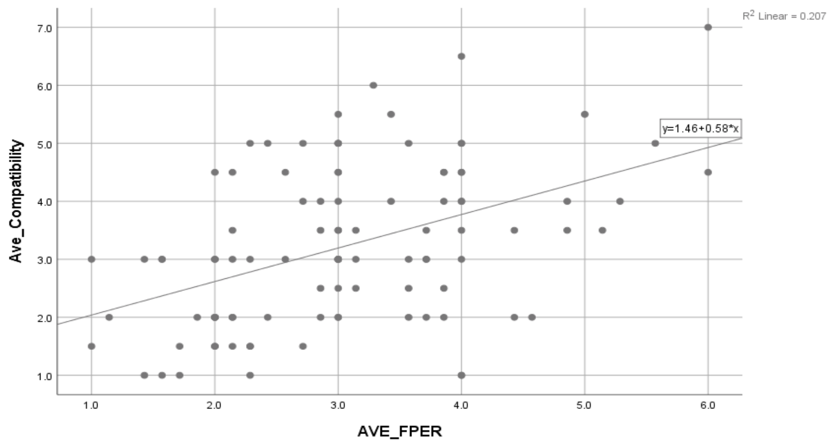

R2 Linear = 0.204
y=0.89+0.39*x
Ave_Cultural_Development
AVE_FPER
6.000000000000000
5.000000000000000
4.000000000000000
3.000000000000000
2.000000000000000
1.000000000000000
1.0
2.0
3.0
4.0
5.0
6.0

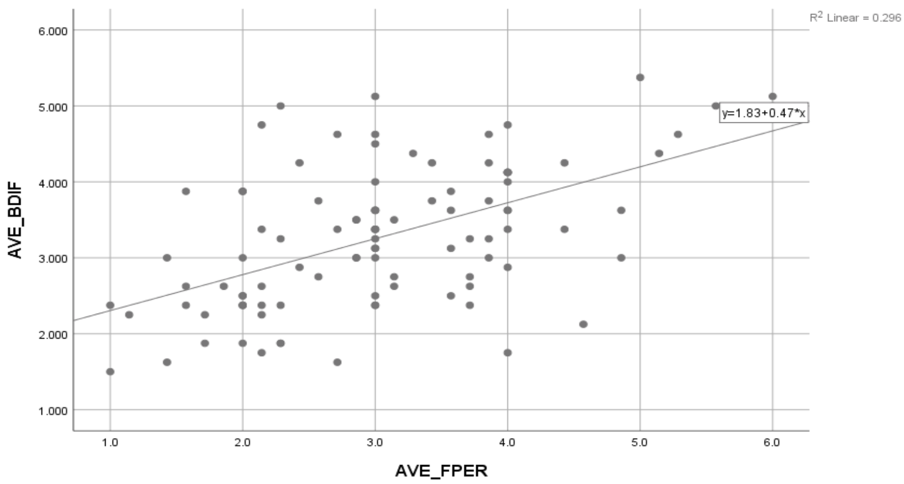

R2 Linear = 0.120
y=1.34+0.46*x
AVE_Employee_investment
AVE_FPER

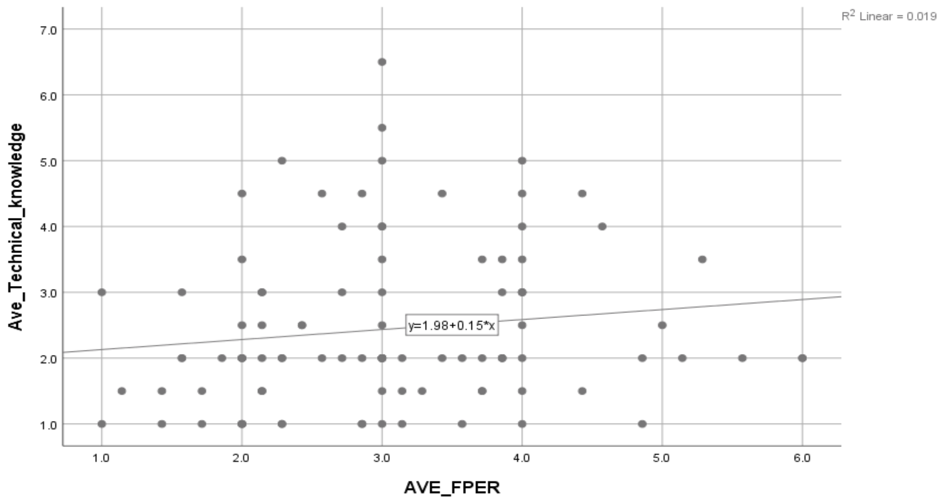

R2 Linear = 0.218
y=1+0.41*x
AVE_MODOC
AVE_FPER

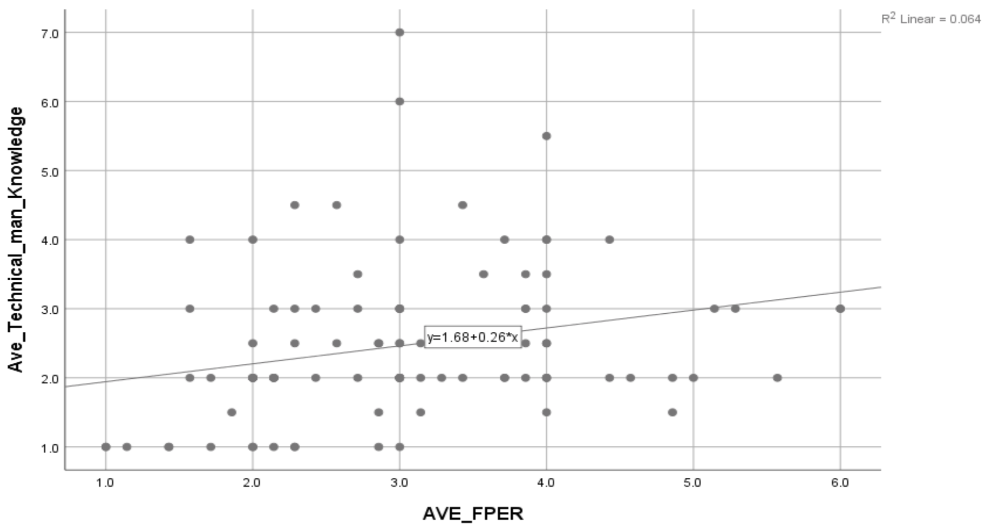

R2 Linear = 0.259
y=1.29+0.6*x
Ave_Connectivity
AVE_FPER

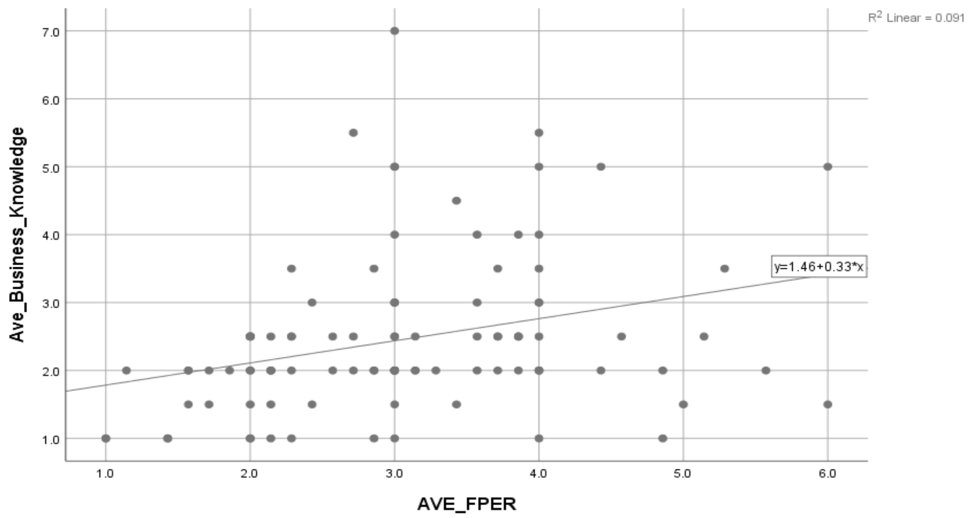

R2 Linear = 0.225
y=1.73+0.56*x
Ave_Modularity
AVE_FPER

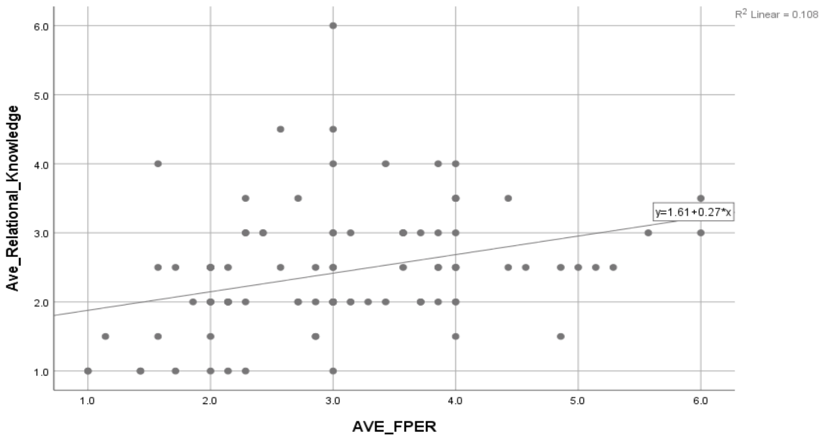

R2 Linear = 0.207
y=1.46+0.58*x
Ave_Compatibility
AVE_FPER

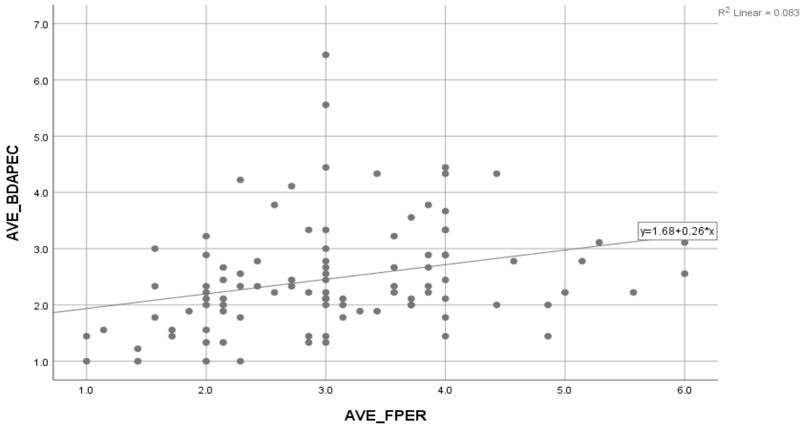

R2 Linear = 0.296
y=1.83+0.47*x
AVE_BDIF
AVE_FPER

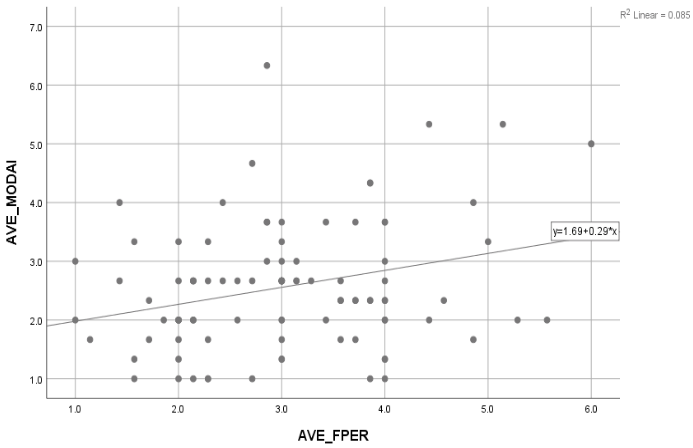

R2 Linear = 0.019
y=1.98+0.15*x
Ave_Technical_knowledge
AVE_FPER


R2 Linear = 0.064
y=1.68+0.26*x
Ave_Technical_man_Knowledge
AVE_FPER


R2 Linear = 0.091
y=1.46+0.33*x
Ave_Business_Knowledge
AVE_FPER


R2 Linear = 0.108
y=1.61+0.27*x
Ave_Relational_Knowledge
AVE_FPER
1.0
2.0
3.0
4.0
5.0
6.0


R2 Linear = 0.083
y=1.68+0.26*x
AVE_BDAPEC
AVE_FPER
1.0
2.0
3.0
4.0
5.0
6.0
7.0


R2 Linear = 0.085
y=1.69+0.29*x
AVE_MODAI
AVE_FPER
1.0
2.0
3.0
4.0
5.0
6.0
7.0

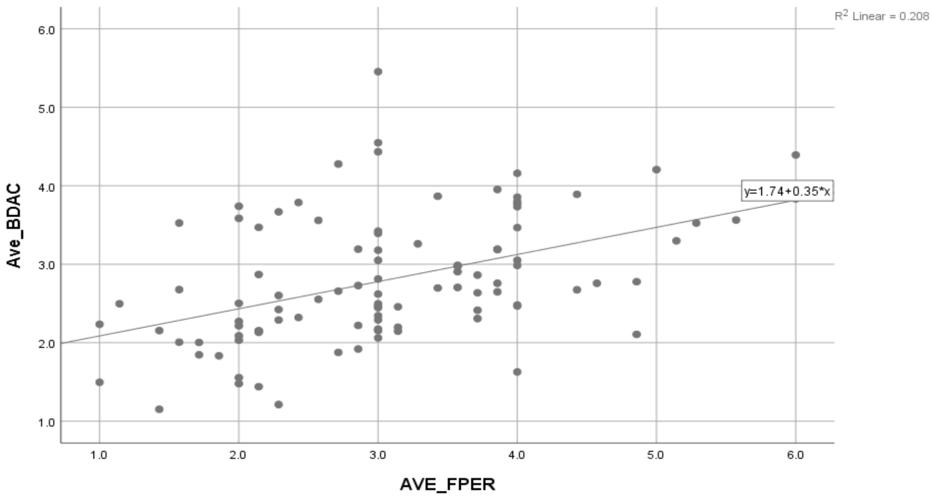


Figure 19: Scatterplots of constructs

Appendix J

Gordon Institute of Business Science
University of Pretoria

21 August 2019

Chetty Tiveshen

Dear Tiveshen

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

You are therefore allowed to continue collecting your data.

Please note that approval is granted based on the methodology and research instruments provided in the application. If there is any deviation change or addition to the research method or tools, a supplementary application for approval must be obtained

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