THE VALIDATION OF A BIG DATA ANALYTICS CAPABILITY SCALE FOR THE SOUTH AFRICAN CONTEXT

by

RENEE NAICKER

submitted in accordance with the requirements for the degree of

MASTER OF COMMERCE

in the subject

INDUSTRIAL AND ORGANISATIONAL PSYCHOLOGY

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: PROF A-P FLOTMAN

January 2023

DECLARATION

I, Renee Naicker, hereby declare that the thesis titled "The Validation of a Big Data Analytics Capability (BDAC) Scale for the South African Context" is my own work and that all the sources that I have used or have quoted from have been indicated and acknowledged by means of complete references.

I further declare that ethical clearance to conduct the research has been obtained from the Department of Industrial and Organisational Psychology, University of South Africa (UNISA), as well as from the participating individuals. I also declare that the study was carried out in strict accordance with UNISA's Policy for Research Ethics and that I conducted the research with the highest integrity during all phases of the research process, taking into account UNISA's Policy on Copyright Infringement and Plagiarism.

Renee Naicker
Student Number: 46492968

DEDICATION

This dissertation is dedicated to the memory of my beloved father, Anandan Andiappen Naicker. I miss him every day, but I am grateful to him for instilling in me the values and motivation to conquer any obstacles life may present.

To my mother, Ragani Naicker, for constantly demonstrating genuine strength and unwavering love.

To my brother, Lovashin Naicker, for always encouraging me and supporting me through life's challenges.

To my husband, Nisharlin Naidoo, for always supporting my ambitions and helping me throughout this tough journey.

To my son, Avan Naidoo, for being my motivation and inspiration to always be better.

To my uncle and aunt, Mervyn Naicker and Devika Naicker, for always inspiring and supporting me.

ACKNOWLEDGEMENTS

Special mention and thanks go to the following individuals and institutions for their support:

- My supervisor, Dr. Aden-Paul Flotman (University of South Africa), for his continuous expertise, guidance, direction and reassurance throughout this long journey;
- Dr. Jeremy Mitonga-Monga (University of Johannesburg) for his dedication, support, encouragement, and guidance with evaluating statistical analysis;
- All respondents who took part in this study's survey, for their valued input;
- Mr. Richard Thompson, for editing this dissertation;
- My family and friends for their continued support; and
- The University of South Africa, for giving me the opportunity to undertake the Master of Commerce in Industrial Psychology.

SUMMARY

THE VALIDATION OF A BIG DATA ANALYTICS CAPABILITY (BDAC) SCALE FOR THE SOUTH AFRICAN CONTEXT

by

RENEE NAICKER

Supervisor: Dr. A.P. Flotman

Department: Industrial and Organisational Psychology

Degree: M. Com (Industrial and Organisational Psychology)

Literature confirms that few organisations have managed to enhance organisational performance through big data analytics capabilities (BDAC). Therefore, the primary objective of this study was to design and validate a BDAC scale for the South African context, and clarify the nature of the BDAC relationship to organisational performance. The population identified senior managers, executives and data analysts who work in the context of big data (BD) and the BDAC space with organisations or during project implementation. A new scale was designed comprising relevant items based on a comprehensive literature review and items taken from existing literature. Two pilot studies were conducted and data collected from respondents using an online survey provided 239 usable questionnaires. The final scale comprised two primary dimensions (i.e., BDAC and organisational performance) and ten subdimensions. The results confirm that the new BDAC scale is valid for South African organisations and can be used to enhance organisational performance. The study thus contributes a validated BDAC scale for the South African context to benefit academics, researchers and practitioners in the quest to further understand BD and BDAC utilisation in transforming organisations, improving organisational development and enhancing organisational performance.

KEY WORDS:

Big data; big data analytics; big data analytics capabilities; confirmatory factor analysis; exploratory factor analysis; organisational performance; validation studies

ABSTRAK

Die literatuur bevestig dat min organisasies al daarin geslaag het om organisasieprestasie deur grootdata-ontledingsvermoëns (GDOV) te verbeter. Die primêre doel van hierdie studie was dus om ’n geldige GDOV-skaal vir die Suid-Afrikaanse konteks te ontwerp, en om die aard van die GDOV-verhouding tot organisasieprestasie te verduidelik. Die populasie het senior bestuurders en data-ontleders geïdentifiseer wat in die konteks van groot data en die GDOV-ruimte met organisasies of tydens projekimplementering werk. ’n Nuwe skaal is ontwerp wat relevante items bevat gebaseer op ’n omvattende literatuuroorsig en items wat uit bestaande literatuur geneem is. Twee voorondersoeke is uitgevoer en data wat van respondente ingesamel is met behulp van ’n aanlyn opname het 239 bruikbare vraelyste verskaf. Die finale skaal het twee primêre dimensies (i.e. GDOV en organisasieprestasie) en tien subdimensies behels. Die resultate bevestig dat die nuwe GDOV-skaal geldig is vir Suid-Afrikaanse organisasies en gebruik kan word om organisasieprestasie te verbeter. Die studie dra dus ’n geldige GDOV-skaal vir die Suid-Afrikaanse konteks by tot voordeel van akademici, navorsers en praktisyns in die strewe om grootdata- en GDOV-benutting beter te verstaan vir die transformasie van organisasies, en die verbetering van organisasie-ontwikkeling en organisasieprestasie.

SLEUTELWOORDE:

Groot data; grootdata-ontleding; grootdata-ontledingsvermoëns; bevestigende faktorontleding; eksploratiewe faktorontleding; organisasieprestasie; geldigheidstudies

ISIFINQO

Imibhalo iqinisekisa ukuthi izinhlangano ezimbalwa zikwazile ukuthuthukisa ukusebenza kwenhlangano ngamakhono amakhulu okuhlaziya idatha (KAOD). Ngakho-ke, inhloso eyinhloko yalolu cwaningo bekuwukuklama nokuqinisekisa isikali sa- KAOD somongo waseNingizimu Afrika, nokucacisa uhlobo lobudlelwano ba-KAOD nokusebenza kwenhlangano. Inani labantu lihlonze abaphathi abakhulu, abaphathi nabahlaziyi bedatha abasebenza kumongo wedatha enkulu (IE) kanye nesikhala sa- KAOD nezinhlangano noma phakathi nokuqaliswa kwephrojekthi. Kwaklanywa isikali esisha esihlanganisa izinto ezifanele ngokusekelwe ekubuyekezweni okuphelele kwezincwadi nezinto ezithathwe ezincwadini ezikhona. Kwenziwa izifundo zokuhlola ezimbili futhi imininingwane eqoqwe kwabaphendulayo kusetshenziswa inhlolovo ye-inthanethi yanikeza imibuzo engama-239 esebenzisekayo. Isilinganiso sokugcina sasihlanganisa izilinganiso ezingqala ezimbili (okungukuthi, KAOD nokusebenza kwenhlangano) kanye nezingxenye ezingaphansi eziyishumi. Imiphumela iqinisekisa ukuthi isikali esisha sa- KAOD sivumelekile ezinhlanganweni zaseNingizimu Afrika futhi singasetshenziswa ukuthuthukisa ukusebenza kwenhlangano. Ngakho-ke lolu cwaningo lufaka isandla esikalini sa-KAOD esiqinisekisiwe somongo waseNingizimu Afrika ukuze kuzuze izifundiswa, abacwaningi kanye nabasebenzi emzamweni wokuqonda kabanzi ukusetshenziswa kwe- IE na- KAOD ekuguquleni izinhlangano, ukuthuthukiswa kwenhlangano kanye nokuthuthukisa ukusebenza kwenhlangano.

AMAGAMA ABALULEKILE:

Ukuhlaziywa kwedatha enkulu; amakhono amakhulu okuhlaziya idatha; ukuhlaziya isici sokuqinisekisa; ukuhlaziya isici sokuhlola; ukusebenza kwenhlangano; izifundo zokuqinisekisa

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1:
SCIENTIFIC ORIENTATION TO THE RESEARCH

1.1 INTRODUCTION

This is a study of the development and validation of a big data analytics capability (BDAC) scale for the South African context. This chapter outlines the background and motivation of the study. The problem statement is then articulated, the research objectives listed, the paradigm perspective shared, and the disciplinary relationships explained. The meta-theoretical constructs are outlined, and the research design described before the chapter is concluded with an outline of the remaining chapters of the study.

1.2 BACKGROUND OF AND MOTIVATION FOR THE STUDY

In the current economic environment, organisations experience uncertainty and increased competition, and rapid changes, especially with the use of technology, which has made more data available than ever before (Edu, 2022). Big data (BD) and big data analytics capabilities (BDAC) have provided organisations with the opportunity to take advantage of the increased volume, variety, velocity and veracity of data, allowing for increased levels of innovation, proactivity and decision-making and a data-driven culture of evidence-based decision-making (Zheng et al., 2022).

The primary objective of this study was to validate a BDAC scale for the South African context. The study will identify enabling factors of BDAC and how these impact organisations and organisational performance (OP). By establishing these enabling variables, the organisation will be in a position to take corrective action to improve the organisation and its performance (Harris, 2012). It is important to note that apart from existing literature, subjective measures were used. In other words, the perceptions of identified employees and experts have been used to develop and validate the assessment instrument.

Organisations are increasingly challenged by "*Big Data*" (Kubick, 2012, p. 27). The term "Big Data" is described as data sets that grow so large that they become difficult to work with using traditional database management systems. They are data sets whose "size is beyond the ability of commonly used software tools and storage systems to capture, store, manage, as well as process the data within an acceptable time period" (Kubick, 2012, pp. 26–28).

Big data (BD) has emerged as an exciting frontier of productivity and opportunity in the last few years (Rahman & Aldhaban, 2015, p. 479). Table 1.1 is a description of big data characteristics:

Table 1.1

Big Data Characteristics (Rahman & Aldhaban, 2015, p. 479)

Characteristic	Description	Influencer
Volume	Grows from a few terabytes to hundreds of terabytes to petabytes of data that must be captured, processed, stored and analysed.	Data volume keeps growing faster from source.
Velocity	Data flows in today's digital era are being produced real-time and around the clock. Large volumes of data must be captured in real time, stored, processed, and displayed faster for real-time business intelligence (BI) and decision-making.	Data flows in real time and in large volumes. Improved computing, processing, BI & visualisation technologies.
Variety	Originates from a variety of sources with unstructured, semi-structured, and structured data. More than 90% of data are unstructured.	Sensors, social networks, digital pictures, video, transaction records and communication surveillance.
Veracity	In most cases data is unstructured and hence data consistency is an issue. This causes data and findings extracted from subjective comments and opinions difficult to predict.	Data-driven decisions require traceability and justification.

Value	Provides new insights to generate organisational value.	Corporate organisational value

The researcher attempted to understand how organisations use big data analytics capability (BDAC). In today's world people do not just want to collect data, they want to understand the meaning and importance of the data, and then use it to aid them in making decisions (Strong, 2010, p. 731).

Big data analytics capabilities (BDAC) is the process of applying algorithms to analyse sets of data and extract useful and unknown patterns, relationships and information (Adams, 2010, pp.11–19). Further, BDAC is used to extract previously unknown, useful, valid, and hidden patterns and information from large data sets, and to detect important relationships among the stored variables. Therefore, BDAC has had a significant impact on organisational functioning, organisational development, research and technologies, since decision-makers have become more and more interested in learning from previous data, thus gaining a competitive advantage (Song & Kusiak, 2009, pp. 1733–1751).

According to Barton, Davenport, and Harris (2012), BDAC is widely considered to transform the way organisations do business. According to Columbus (2014, p. 2) who reviewed a recent study by Accenture and General Electric, "87% of organisations believe big data analytics capabilities (BDAC) will redefine the competitive landscape of their industries within the next three years. Furthermore, 89% believe that organisations that do not adopt a BDAC strategy in the next few years risk losing market share and momentum".

Motivated by this debate, the researcher aimed to validate a big data analytics capabilities (BDAC) scale in the South African context and the impacts on organisational performance (OP).

In the literature and in practice, it has been and is likely to continue to be difficult to find the organisational fit to information technology as discussed by Strong (2010, p. 731). There are not enough details in the literature about the important roles to be played. This study addresses this critical gap by developing an expanded theoretical understanding of BD and BDAC as key related organisational elements which are then leveraged to understand under what circumstances BDAC use will translate into quick decision-making and ultimately enhanced organisational performance (OP).

The research model in this study will better explain the impacts of data analytics use on OP, while also indirectly providing guidance to managers on how they could better leverage such technologies. These findings could be more broadly used to inform organisational development interventions and the effective use of other forms of BD and BDAC in organisations.

1.3 PROBLEM STATEMENT

In "The Global Competitiveness Report 2019" (Schwab, 2019) South Africa was ranked the 60th most business competitive country in the world. This shows the importance of understanding BD and BDAC in organisations and the need to create high value to add benefit to our competitive landscape. It has been reported that organisations fail to utilise BD and BDAC effectively to achieve a competitive advantage (Schwab, 2019). Organisations often struggle to appreciate multiple sources of data at the systems level, and to link information together in its several forms. Organisations also struggle to generate meaningful and actionable insights from BDAC to solve problems or drive broader organisational change (Bean, 2018). Finally, they also struggle to identify the fundamental purpose of the problem or question to be solved that may lead to the specific types of actions to be pursued (Maguire, 2018). However, without a suitable research methodology, the time and financial investment required for successful big data management may not add any value. Organisations thus lack a model to assist with the design, development, and implementation of BDAC, which will enable them to improve performance (Russom, 2013).

In an attempt to take advantage of the potential benefits of BD, a growing number of organisations are attempting to use BDAC to analyse available data and inspire strategic organisational decision-making (Schwab, 2019). For these organisations, it is important to leverage the full potential that BDAC can offer with the aim of enhancing performance. BDAC does have shortcomings regarding value creation and analytics maturity with organisations. This study investigates individual perceptions and understanding related to these changes in terms of opportunities, extent, limitations, challenges, and implications, and the way that organisational performance is measured and managed.

The mindset of managers and decision-makers has a crucial impact on what can be achieved. The success or failure of any organisational change imposed by new information technologies depends, above all, on people's attitudes towards them. Managers' awareness and understanding of BDAC impact the organisation, and the alignment between data and the decision-maker. From an organisational psychology stance, one part of this study is to contribute towards a big data-driven approach for organisational change. A model is needed to address current challenges faced by many organisations in managing large volumes of data. This study has extended existing research by proposing that BDAC enables organisations to generate insight that can strengthen their dynamic capabilities, which, in turn, positively impact their decision-making capabilities.

1.4 RESEARCH AIMS

Given the problem statement discussed above, the aims of the study are formulated below.

1.4.1 General Aim

The general aim of this study is to develop and validate a BDAC scale for the South African context and to examine to what extent big data analytics capabilities (BDAC) impact organisational performance (OP).

1.4.2 Specific Aims

The specific literature aims are as follows:

Literature aim 1: To conceptualise the variables of big data (BD), big data analytics capabilities (BDAC) and organisational performance (OP) from the literature.

Literature aim 2: To report on research studies regarding big data analytics capabilities (BDAC) and organisational performance (OP) from the literature.

Literature aim 3: To determine the effect of big data analytics capability (BDAC) on organisational performance (OP) from the literature.

The specific empirical aims are as follows:

Empirical aim 1: To develop and validate the identified big data analytics capability (BDAC) scale for the South African context.

Empirical aim 2: To make recommendations to the participating organisation, for industrial and organisational psychology, and for future research, based on the results of the study.

Consequently, this study seeks to answer the following closely related research questions:

Does the big data analytics capabilities (BDAC) scale possess acceptable levels of internal consistency and construct validity?

To what extent do big data analytics capabilities (BDAC) impact organisational performance (OP)?

What recommendations can be made to the participating organisations, for industrial and organisational psychology, and for future research, based on the findings of the study?

1.5 PARADIGM PERSPECTIVE

Research strategies are located within the broader frameworks of theoretical or philosophical perspectives, commonly referred to as paradigms (Blaikie, 2007). Creswell and Clark (2011) define a paradigm as a set of generalisations, the philosophy and the values of a community of specialists. Paradigms are social constructions, historically and culturally embedded discourse practices, and therefore, neither inviolate nor unchanging (Teddlie & Tashakkori, 2003).

1.5.1 Theoretical Paradigm: Systems Theory

The theoretical paradigm for this study is the systems theoretical paradigm. Systems theory suggests that all systems are composite things that have interacting components (Checkland, 1999). Accordingly, a system should possess properties that are derived from the interactions among its components (Chesbrough & Bogers, 2014).

Systems theory is particularly suitable for theorising the organisational value of big data (BD) and big data analytics capabilities (BDAC). The central argument is that organisations consist of several interacting systems and sub-systems. Using the combination of BD and BDAC allows an organisation to analyse the gathered data and derive organisational knowledge, which could be beneficial for developing superior organisational performance with the use of data-driven decisions.

Key Assumptions of Systems Theory

- Holism: People who claim to take a system approach probably have most in common with respect to assumptions pertaining to the level of explanation, specifically taking a holistic view rather than a reductionist view (Wallis, 2013).
  - Relationships: Another underlying assumption shared by many systems traditions is that the unit of analysis should be relationships rather than entities. Entities only take on definition when they are interacting with each other (Wallis, 2014).
  - Environment: Another underlying assumption of several systems traditions is that the environment plays a role in the manifestation of the phenomenon (Cabrera et al., 2008). The environment is central to understanding and explaining. Ackoff (1981) suggests that in situations which he refers to as producer-product, any principle or explanation offered must stipulate the conditions under which the principle applies.
  - Indeterminism: The assumption of indeterminism is that at times it is "inherently impossible to determine in advance which direction change will take" (Prigogine & Stengers, 1984, p. xv).
  - Causality: Assumptions about cause and effect, as well as those pertaining to observation and level of explanation (holism or reductionism), give the best indication of worldview (Dent, 1997).

- Self-organisation: The idea that the elements of a system move toward their stable equilibrium states largely independently, versus the assumption that one or a small number of causes affect the elements of a system (Cabrera et al., 2008).

- Observation: A key belief underlying classical science was that observations are independent of the characteristics of the observer. Objectivity was possible if one assumed that very different people looking at the same phenomenon in the same way would create similar descriptions (Cabrera et al., 2008).

- Reflexivity is the system of interest composed of knowing subjects with characteristics such as the following: are they able to generate new states in themselves (think new thoughts, do new things) that they never manifested before? Are people (or machines) best thought of as continually trying to generate such new states? Do they have the property of being able to notice your attempts to theorise about them and model them, and do they modify themselves according to their reaction to this information? (Vaill, 1996, p.117).

1.5.2 Research Paradigm: Post-Positivism

The positivist approach has been selected, due to the nature of the problem, the variables to be explored and the quantitative nature of the study. The positivist approach is followed when researchers believe that reality is stable and can be described from an objective point of view (Remenyi et al., 1998). As explained by Hirschheim and Klein (1989), the positivist method identifies reasons for a problem based on a deductive reasoning process.

In the positivist or deductive method, according to Bryman and Bell (2011) and Creswell (2009), there are three fundamentals explained as constructing the hypothesis model or a relationship and the execution of quantitative methods and value-free explanation provided by the researcher on the research problem. It is understood from Alvesson and Skoldberg (2009), Bryman and Bell (2011) and Creswell (2009) that the essence of the positivistic theoretical approach is to describe variables using quantitative measures, while testing a hypothesis on a sample to generalise to a larger population. The inquiry methods available under positivism are observing, measuring, distributing surveys and questionnaires (Minges, 2003).

1.5.3 The Disciplinary Context

In an increasingly digitised world, big data (BD) and big data analytics capabilities (BDAC) become more prevalent, precise, and available (McAfee et al., 2012). BD and BDAC are quite impactful on industrial and organisational psychology, which is the disciplinary context of this study. Industrial and organisational psychology is defined as the scientific study of human behaviour in organisations and the workplace (McAfee et al., 2012). The speciality focuses on deriving principles of individual, group and organisational behaviour and applying this knowledge to the solution of problems at work (Altman et al., 2008). Organisational psychology has been defined as the area of psychology that applies psychological principles to the workplace, including the structure of organisations, the ways its members work together, and how the organisation attempts to improve itself through motivation, diversity, work attitudes, leadership, culture, and other related processes (Levy, 2006). This study also has implications for other areas of applied practice including organisation development (Church & Dutta, 2013), learning (Saunderson, 2014), and human resources in general (Bersin & Ferrar, 2014). Practitioners need to think more broadly and holistically and treat the dynamic that we call BDAC as an opportunity to make a meaningful impact on organisational applications. What is relevant to this study is that BDAC has the potential to shift the fundamental way in which work is perceived and conducted.

1.5.4 Meta-theoretical Perspective and Variables

The meta-theoretical constructs and perspectives of this study include:

- **Big data (BD)**

  The term "Big Data" is described as data sets that grow so large that they become difficult to work with using traditional database management systems. They are data sets whose size is beyond the ability of commonly used software tools and storage systems to capture, store, manage, and process the data within a tolerable elapsed time (Kubick, 2012, p. 26–28).

- **Big data analytics capabilities (BDAC)**

Big data analytics capabilities (BDAC) is the process of applying algorithms to analyse sets of data and extract useful and unknown patterns, relationships, and information (Adams, 2010, p. 19). Further, data analytics are used to extract previously unknown, useful, valid, and hidden patterns and information from large data sets, and to detect important relationships among the stored variables. Therefore analytics have had a significant impact on organisational functioning, research, and technologies, since decision-makers have become more and more interested in learning from previous data, thus gaining a competitive advantage (Song & Kusiak, 2009, p. 1751).

- **Organisational performance (OP)**

To assess the performance of an organisation, it is essential to choose an organisational model that has certain diagnostic requirements. Martins and Coetzee (2009) mention that such a model must be well researched. The Burke-Litwin model (Burke & Litwin, 1992) is considered in this study.

The Burke-Litwin model of OP is founded on a functional cause-and-effect framework and explains how linkages between elements contribute to OP (Burke & Litwin, 1992). Burke and Litwin (1992) describe OP as the outcome of work performance, effort and achievement. Indicators of this include productivity, customer satisfaction and service quality. According to Jones and Brazzel (2006) and Martins and Coetzee (2009), the Burke-Litwin model highlights two distinct sets of organisational dynamics. One set is primarily associated with the transactional level of human behaviour, or the everyday interactions and exchanges that create the climate of the organisation, while the second set of dynamics is concerned with processes of human conditions that include marketplaces, world financial conditions and political and governmental circumstances.

- **Big data (BD) and big data analytics capabilities (BDAC)**

A growing number of organisations are attempting to use BD and BDAC to analyse available data and inspire strategic organisational decision-making (Schwab, 2019). For these organisations, it is important to leverage the full potential that BD and BDAC can offer, with the aim of gaining significant impact. BDAC has shortcomings in regard to value creation and analytics maturity within organisations. This study investigates individuals' understanding related to these changes in terms of opportunities, extent, limitations, challenges, and implications, as well as the way that organisational performance is measured and managed.

As noted above, in the literature and in practice, it has been and is likely to continue to be difficult to find the organisational fit to information technology as discussed by Strong (2010, p. 731). There is a critical role that is underspecified in the data analytics literature. This study addresses this critical gap by developing an expanded theoretical understanding of people, skills, and data, including these aspects between BD and BDAC as being key related organisational elements, which is then leveraged to understand under what circumstances use of BD and BDAC will translate into agility gains and ultimately organisational impact.

Research Hypothesis

The research hypothesis of this study is that big data analytics capability (BDAC) has a direct and positive impact on organisational performance (OP).

1.6 RESEARCH DESIGN

The researcher's choices of approach, methodology, strategy, and ethical considerations for this study are described below.

1.6.1 Research Approach

A quantitative approach has been employed in this study (Cooper & Schindler, 2006). For the purpose of this research, and in order to answer the research questions and achieve the research objectives, it was appropriate to quantify the opinions, attitudes and behaviour of BI experts and data experts. A quantitative approach is therefore utilised, as it is a systematic way of investigating various phenomena and their relationships with other phenomena (Kothari, 2004). The relationship between the variables was investigated in an objective manner (positivistic stance) for the results to be generalised to larger populations. This approach also allows for statistical inferences to be made and for the results to be replicated in future studies (Kothari, 2004). The survey research method was used to facilitate the logical collection of data from respondents so that a scientific sampling and questionnaire design could be used to measure characteristics of the population sample with statistical precision.

1.6.2 Research Strategy and Method

A cross-sectional research design has been used in this study. Cross-sectional studies examine data at one point in a time, that is, data is collected on only one occasion with the same subjects rather than on the same subjects at several time points (LoBiondo-Wood & Haber, 2002; Polit & Beck, 2008). Accordingly, Bowling and Ebrahim (2006) explain that a cross-sectional study describes the frequency (or level) of a particular attribute. Cross-sectional studies can be descriptive or may include an analytical component.

The aim of cross-sectional studies is to obtain reliable data that make it possible to generate robust conclusions, and create new hypotheses that can be investigated with new research (Katz, 2006). Analytical studies seek to establish relationships and associations between two or more phenomena (called variables), and descriptive studies are only about the detailed and organised description of one or more phenomena. The systematisation and standardisation of the data collection methods, and also the strategy adopted to obtain them, is called the design or, more correctly, the study design.

The main characteristic of cross-sectional studies is that the observation of variables, whether they are cases, individuals, or other types of data, is performed in a single moment (the same), when the researcher records a “photograph” of the facts (variables) of interest and not the “movie” of their evolution (Katz, 2006). These characteristics make cross-sectional studies particularly useful for studying the prevalence of a particular phenomenon, whether it is assumed to be the cause or the consequence, or both, in a defined population. These studies, even if purely observational and descriptive, are very useful in the field of organisational development (Joubert & Ehrlich, 2009).

1.6.2.1 Population and Sample Size

According to Lind, Marchal and Wathen (2008), a survey population is defined as the entire set of individuals or organisations from which survey data can be collected, which is used as a basis for research. Two phases have been used in this study:

Phase 1 - Questionnaire development (new items to be collected)
Phase 2 - Questionnaire validation

To collect the required data, organisations in the South African context with a population size of 400 or more were considered, and a single organisation from South Africa was selected. Electronic questionnaires were distributed to potential respondents. Business intelligence (BI) experts and data experts are the target population. Conclusions has been drawn about the entire population. A list of BI professionals (business intelligence consultants, technical business architects, project managers, product managers, data analytics experts, business analysts, system analysts, executives, operational users and technology specialists) has been obtained.

This study has used non-probability, convenience sampling (Battaglia, 2008). In this type of sampling, researchers prefer participants as per their own convenience (Andale, 2015). Subjects who are readily accessible or available to the researcher are selected. In other words, in this type of non-probability sampling method, whoever meets the researcher’s criteria qualifies to be the part of the sample. In this instance, the best sampling method to use is using a sample of the total population (Sarstedt & Mooi, 2019). Sampling is a technique where the entire population that meet the criteria (e.g., specific skill set, experience, etc) are included in the research being conducted (Walliman, 2011). Sampling is more commonly used where the number of cases being investigated is relatively large.

Sampling criteria were:

a) Permanent employees.
b) Employees with knowledge of BD and BDAC.
c) Employees willing to participate in this study.

Table 1.2

Research Sample

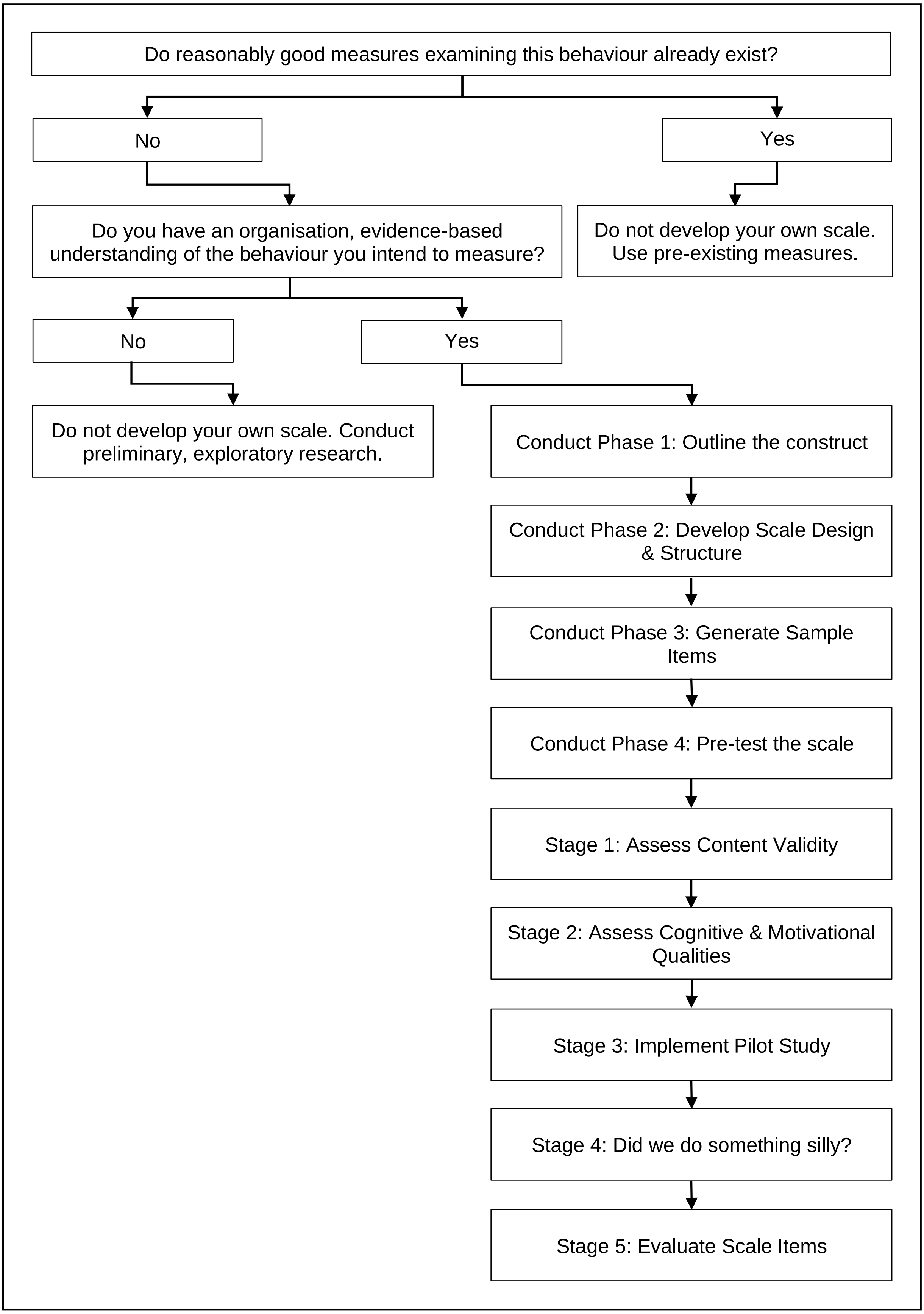
Designation (BI Experts and BI Users)	Sample Size(N)
Business intelligence consultant	19
Technical business architect	24
Project manager	31
Product manager	18
Data analytics expert	37
Business analyst	16
System analyst	15
Executives	9
Technology specialist	59
Other	11
Total	**239**

Scale Development

A methodological map is provided that has guided this research through a scale development protocol and assist with the developing of an effective survey instrument.

Figure 1.1

Scale Development (Barry et al. 2011, p. 98)



1.6.2.2 Research Measures

The researcher has used a questionnaire for the South African context. A questionnaire-based survey method enables generalisability of outcomes, allows for easy replication, and facilitates the simultaneous investigation of many factors (Check & Schutt, 2012). Additionally, survey-based research is a well-documented way of accurately capturing the general tendency and identifying associations between variables in a sample. In this study, all measurement items were taken from the existing literature and were adapted to fit the big data (BD), big data analytics capability (BDAC) and organisational performance (OP) for the South African context. Questions were customised to fit the context of the study to ensure that they were applicable to the population. Refer to Appendix A for the survey design and response items.

A pilot study was conducted to test the questionnaire's content validity. The following steps were taken to conduct the pilot study (as in Saunders et al., 2007):

Step 1: Confirm research objectives.

Step 2: Confirm types of information to be collected. For example, attributes, behaviours and experiences, knowledge or awareness, attitudes, and opinions.

Step 3: Confirm type of data to be collected. For example, nominal, ordinal, interval, and ratio.

Step 4: Collect or formulate relevant questionnaire items and design preliminary questionnaire.

Step 5: Build up face validity by utilising specialists to assess the reaction to things as indicated by the research objectives.

Step 6: Run a pilot test – identify a sub-set (sample 20-30), of intended participants and run a pilot test of the questionnaire to weed out problematic questions and have the items reviewed by an expert on question construction.

Step 7: Clean collected data.

Step 8: Use, for example, principal component analysis to identify underlying constructs being measured – factor loadings.

Step 9: Check internal consistency, for example, Cronbach's Alpha (CA).

Step 10: Revise the survey.

The above steps have ensured that:

- The questionnaire's face validity has been confirmed by experts;
- The questionnaire was pilot tested on a sample; and
- Pre-launch statistical analysis has been included (Collingridge, 2014).

The pilot study has been conducted within a small study of 27 individuals, within a number of organisations to identify the statistical properties to be measured.

Table 1.3

Focus Areas: Indicates the factors used within the questionnaire.

Number	Factors
1	Demographics
2	Big data analytics planning
3	Big data analytics resources
4	Data and analytic investment
5	Organisational Performance
6	Connectivity
7	Big data analytics capability
8	System design
9	Technical knowledge
10	Technology management knowledge
11	Organisational knowledge
12	Relation knowledge

1.6.2.3 Research Procedures

Permission to conduct the research has been obtained from the University of South Africa (UNISA) Research Committee and a single organisation from South Africa, with reference number: 2021/CEMS/IOP/025. Each potential member of the sample has received a link containing the purpose of the study, the approval of the study, confirmation of the safekeeping and confidentiality of the responses, a consent form explaining that participation in this research is voluntary, and the instructions for supplying the socio-demographic information and completing the questionnaire. Each participant has submitted all completed questionnaires, in a link, which has been submitted to the researcher. The survey platform that has been used is Lime Survey.

1.6.3 Data Analysis

The following section provides a broad overview of the statistical analysis, descriptive analysis and inferential analysis that has been used to analyse the data as described in Table 1.4. In addition, the section describes which statistical methods have been used to analyse the study's hypotheses. The following statistical analyses have been conducted:

Research Validation Process

Step 1:

Item analysis for variable 1 (Cronbach's Alpha).

Step 2:

Exploratory factor analysis for variable 2.
Determine factorial structure (construct validity).

Step 3:

Measurements (AVE and CR – discriminant and convergent validity).

Step 4:

First and second order relationships and structural equation modelling.

Step 5:

Repeat the same process above for variable 2 (organisational performance).

Step 6:

Determine the effect of BDAC on OP.

Step 7:

Summary of discriminant validity.

Table 1.4

Statistical Analysis Stages

Stage	Description
Finalise data entry	Data will be ready for analysis after all data entries were included in storage tables. This will occur after the responses that were obtained from the online survey, soft or hard copies, were closed.
Analyse data	Once the data is ready for analysis after the preparation process, I will then analyse and present the research findings.
Present results	For the purpose of analysis, the Likert scale items were grouped to obtain a majority score comprising responses in respect of the “Strongly agree and Agree”, and “Disagree and Strongly disagree” options. All the results will be presented quantitatively.

1.6.4 Ethical Considerations

According to Cooper and Schindler (2006), ethical clearance and considerations ensure that no respondent is abused or suffers adverse consequences from research activities. The below ethical considerations have been used:

- Participants have been recruited to participate voluntarily and are able to withdraw at any time without facing any penalties;
- An informed consent letter has ensured that potential participants are fully aware of the procedure of the proposed survey, primarily the right to privacy, confidentiality, and anonymity; and
- Ethical clearance has been obtained from UNISA and the participating organisation in South Africa to ensure the legitimacy of the research conducted.

All these ethical considerations ensured that the respondents did not suffer from physical harm, discomfort, pain, embarrassment, or loss of privacy when they participated in the research.

1.7 RESULTS OF THE STUDY

The study reports on the results of the research, pertaining to the factor analysis, reliability analysis, descriptive analysis, and inferential statistics**:** correlations and regression.

1.8 DISCUSSION OF THE STUDY

The discussion of the study has been guided by the general aim, the specific literature aims, the specific empirical aims and research objectives of the study.

1.9 CONCLUSION, LIMITATIONS AND RECOMMENDATIONS

The conclusions of the study have been guided by the specific literature and empirical aims of the study.

Some potential limitations of the study are:

- The cross-sectional design of the study means that the findings do not imply any causality among the variables;
- The results of the study could be adversely impacted, depending on the number of usable completed questionnaires; and
- The outcome of the pilot study could have a direct impact on the success of the study.

Recommendations for IOP and for future research are to be made, based on the findings of the study.

1.10 CHAPTER LAYOUT

This study has been presented in detail in the following chapters:

Chapter 1: Scientific orientation to the research.

Chapter 2: Literature review: Achieving competitive advantage and organisational performance (OP) through big data (BD) and big data analytics capabilities (BDAC).

Chapter 3: Literature review: Research studies on validation and the effect of big data analytics capability (BDAC) on organisational performance (OP).

Chapter 4: Research methodology.

Chapter 5: Results and discussion.

Chapter 6: Conclusions, limitations, and recommendations.

1.11 CHAPTER SUMMARY

In chapter one, the scientific orientation to the research is discussed. This chapter contains the background and motivation, the research problem and aims, the paradigm perspective, the research design and method, and the chapter layout.

CHAPTER 2:
LITERATURE REVIEW

ACHIEVING COMPETITIVE ADVANTAGE AND ORGANISATIONAL PERFORMANCE THROUGH BIG DATA (BD) AND BIG DATA ANALYTICS CAPABILITIES (BDAC)

2.1 INTRODUCTION

The previous chapter offered the motivation for conducting this research study, its purpose, and the primary research question and objectives. The methodology, preliminary literature evaluation, and other essential parts of this research study that would assure its success were highlighted.

The purpose of this chapter is to provide an understanding of big data (BD), big data analytics capability (BDAC) and organisational performance (OP), their uses and the associated organisational value, as reflected in the literature. It covers in detail the key concepts and terminologies that are needed and/or are commonly used to understand the research objectives.

2.2 UNDERSTANDING BIG DATA (BD)

In the next section, the concept of BD is discussed, and its characteristics are explored. This is followed by a discussion on the relevance and rationale of BD in the global and South African context.

2.2.1 Defining and Characterising Big Data (BD)

BD is a collection of data that is massive in volume and continues to increase exponentially in size over time. It is data that is so massive and complex that none of the usual data management methods can efficiently store or process it. BD is also data, but it is enormous in size (Ronda-Pupo et al., 2012). According to Bigelow (2020), BD also encompasses a wide variety of data types, including:

- Structured data, such as transactions and financial records;
- Unstructured data, such as text, documents, and multimedia files; and

- Semi structured data, such as web server logs and streaming data from sensors.

When it comes to BD's characteristics, most of the literature relies on the four V's method (Marr, 2021). McAfee et al. (2012) define BD according to its characteristics, namely, "volume, variety, velocity, and veracity". This can be explained as follows:

- Volume – there is a huge increase in data volume. Google processes about a billion requests per day. Every day, 24,000 terabytes (24 million gigabytes) of data are generated (Davenport et al., 2017).

- Variety – new data types have emerged. Traditional data types, for example, text or digits, have traditionally been used to record information. Present-day systems allow data to be recorded in photos, films, and places (Alharthi et al., 2017).

- Veracity – this refers to the data's accuracy; for example, certain data may be inaccurate. Social media, for example, generates a lot of noise and ought to be avoided (Goes, 2014; Pigni et al., 2016).

- Velocity – the rate at which data is updated has accelerated to the point where that near-real-time data analysis is possible (Pigni et al., 2016).

2.2.2 The Relevance and Rationale of Big Data (BD)

Although the BD phenomenon has just recently gained prominence (Chen et al., 2015), the knowledge that organisational intelligence contributes to decision-making has been around for a long time (Chen et al., 2014). The difference between BD and organisational intelligence is that BD is endless and real-time, as opposed to organisational intelligence, which is finite, offline, and unstructured (Alharthi et al., 2017). This ever-expanding range of large data sources is relevant because it offers a more comprehensive perspective of events, whether they are in person, online, or on a mobile device. Internet clicks, social media posts, and sensor interactions are examples of transactions on a network (Akbay, 2015; George et al., 2014).

Data-driven decision-making directly benefits organisations and operations (Brynjolfsson et al., 2011; Lavalle et al., 2011; McAfee & Brynjolfsson, 2012; Nannetti, 2012), which indicates that BD should be considered a strategic competence to the smooth operations of an organisation. Attaining this appears to be difficult for large organisations. Many organisations rely on intuition rather than database-based or data-driven decision-making choices. On average, only a tenth of a percent of an organisation's BD is available to it and is analysed for insights (Bradley et al., 2013), and almost half of this statistic in huge organisations (Comuzzi & Patel, 2016).

Another relevant point is that BD can be classed as technical or non-technical in nature (Akter et al., 2016; Chen et al., 2015, Dutta & Bose, 2015; Frizzo-Barker et al., 2016, Wah et al., 2015; Frizzo-Barker et al., 2016). These authors discuss system infrastructure and architecture, data storage and data management curation, complexity of data, security, compliance, analytic tools and approaches, resources, computational and processing complexity, visualisation, and scalability. Several non-technical issues include access to data, expertise, decision-making, resource allocation, executive support, an organisation's grasp of outputs and capabilities of individuals, organisational structures, strategy alignment, legislation, assisting analysts with complicated challenges, stakeholder buy-in, system and capability implementation, and transformation management.

Frizzo-Barker et al. (2016) conducted a comprehensive empirical examination of the state of BDAC. These authors also conducted a review of the literature and discovered that three concurrent primary areas of concern have been identified. The first and most obvious of these was a lack of necessary skill sets and tools to exploit the extraordinary volume of data accessible in such a way that deliverables for BDAC are possible. The second most frequently expressed issue was around data privacy and regulation. This area is growing in importance and increasing in popularity as a result of the amount of personally identifiable information exchanged in the current period, customers' enhanced knowledge, and the proactive creation of legislation, such as South Africa's POPI act (Alharthi et al., 2017; Kambatla et al., 2014).

By definition, BDAC projects span multiple disciplines, as data is generated at all levels of organisations. To enhance the value of BDAC, the outputs of BDAC initiatives should be distributed among multiple organisation divisions and degrees of data-driven decision-making in an organisation (Chen et al., 2014; Comuzzi & Patel, 2016). This necessitates an awareness for and acceptance of the indigenous culture on the organisation's platform (Alharthi et al., 2017) in order to produce adequate results and avoid silo-based execution. This is consistent with the assertion of Lavalle et al. (2011) that a poor result in implementation of BD is a manifestation of an organisation's lack of comprehension of how to use statistics to improve organisation's and management's focus on competing and on priorities. Poor processes for data collection and implementation, and an unwillingness to change the organisational culture also play a role in poor implementation of BD. These are not uncommon enterprise-wide difficulties with deployments of technology, such as ERP (enterprise resource planning) systems, which are notoriously difficult to regulate (Comuzzi & Patel, 2016), and they should therefore be addressed accordingly.

BDAC initiatives are often poorly executed. Despite the obvious links between organisational success and widespread data-driven analytics in making decisions and developing strategies, organisations do not implement these strategies effectively to create value (Akter et al., 2016; Lavalle et al., 2011; McAfee & Brynjolfsson, 2012; Monino, 2016; Power, 2015). The most repeated issue in the literature is that enhanced skills (e.g., machine learning and artificial intelligence) continue to be an obstacle that can only be overcome via fear of organisations being irrelevant. There are too few experienced individuals equipped with the necessary system resources, which are usually the base capabilities for BDAC projects. BDAC projects often fail because organisations do not understand BD and the implementation of BDAC projects, and not as a result of technological or organisational skill deficiencies.

McAbee et al. (2017) say that the literature on organisations puts excessive emphasis on deductive and theoretical areas. Thus, there is a demand to substantiate previously established causal linkages between the use of BD, BDAC and OP. The next sections will examine the theoretical foundations of OP in order to determine its influence from BD and BDAC.

The literature is replete with subjective evidence of BDAC's capacity to provide enormously significant and influential findings and insights (Inamdar et al., 2020 & Lunde et al., 2019). BDAC uses available information with insight development and knowledge management and execution to ensure an organisation's success. This point has been demonstrated directly, by increasing operational efficiencies, and indirectly, by improving data-driven decision-making, selecting optimal product development or placement options, and individually valuable knowledge or information (Chen et al., 2012; Lee, 2017; Pospiech, 2017).

Numerous market evaluations have been conducted on the applicability of BD (Walker & Brown, 2019; Kiron et al., 2012; Lavalle et al., 2011; Nannetti, 2012). These studies demonstrate appreciation for BD's potential to enhance organisational success, but show that BD is often ineffective in execution (Bradley et al., 2013; Wang et al., 2016). While significant benefits are mentioned, they are frequently achieved by isolated projects or pockets of activity within organisations and thus cannot accurately describe mature BD deployments that provide sustainable competitive advantage and contribute to organisational success (Elgendy & Elragal, 2016; Espinosa & Armour, 2016; Lee, 2017; Mazzei & Nobel, 2017). Not all organisations are equally capable of deriving value from their data. There are, nevertheless, countless informal ways for great value to be added. Value to organisations ranges from direct influence on competitive capability to indirect impact on competitive capability (Espinosa & Armour, Özköse et al., 2015; Kubina, Varmus, & Kubinova, 2015; Mikalef & Pateli, 2017). BD can assist in overcoming contemporary organisational difficulties (Kubina et al., 2015; Perrons & Jensen, 2015) and establishing new organisational capacities (Addo-Tenkorang, & Helo, 2016), providing an enabling environment novelty (Walker & Brown, 2019; Gobble, 2013; Mazzei & Noble, 2017; Vargo & Lusch, 2004), enhancement of services, marketing personalisation, all significant developments in organisations.

Given the range of use and the magnitude of the impact that BD may have, the potential for value creation is staggeringly high across practically all organisations. Mazzei and Noble (2017) define the value provided by BD in terms of a three-tiered paradigm in an organisation (Table 2.1), where data becomes a greater driver of success as one advances through the stages than strategy. They do so by encapsulating McAfee and Brynjolfsson's (2012) findings, which highlight increased value as organisations focus on data-driven decision-making.

Table 2.1

Three tiers of value creation as a result of Big Data. Modified from Mazzei and Noble (2017).

Tier	Category	Description
1	Data as a tool	Allows for traditional or core functions and value chains to be improved. Outputs include improved efficiently and effectively using the data to make key decisions. This is where most organisations will find themselves, as it is an operational view that is relatively easy to achieve with improvements through real-time or customised data-driven decision-making and individualised consumer focus.
2	Data as an industry	As the majority of organisations do not possess the in-house capability to appropriately access and leverage their data, many insource specialised entities focused on acquisition, storage, construction of infrastructure, processing and analysis of data, and development of software devoted to managing BD.
3	Data as a strategy	Few organisations are dedicated to building data resources which can direct development of radical or innovative organisation models that link traditional and modern strategic thought. This relied on organisational leaders being evidence-based or data-led in their data-driven decision-making and strategy development. They focus on data flow rather than data stocks to create leverage through insight that allow them to create competitive advantage.

The relevance of BD is reflected in the first tier, which employs BD as a means of enhancing and facilitating the efficiency of existing systems and internal competencies of the organisation (Mazzei & Noble, 2017). The overwhelming majority of cases of the above-mentioned value-adding applications fit in this tier, owing to their combined considerations, including: first, organisational leaders view data and analytics as a valuable resource positioned to fix or improve existing paradigms in a particular organisation, it is a significantly simpler solution proposition. Second, achieving applications at the subsequent tiers necessitates adequate resources, scarce skills, organisational alignment, and not only the willingness but also the ability to achieve actions that are redirected based on data insights (Akter et al., 2016).

The second tier defines BD as a self-sustaining stimulus for industry. Unlike BDAC, which relies on highly specialised skill sets and infrastructure for information systems, it permits the development of infrastructure-focused organisations (Mazzei & Noble, 2017). Amazon, Cloudera, Hortonworks, Capgemini, Microsoft, and Palantir are among the companies that have begun to operate in this field.

The third layer contains a smaller number of organisations, but it is also the sphere where the biggest potential for BD applications exists. These are organisations that not only allow data to impact their operations, but also actively contribute to the development of ecosystems that allow for the direction of information and resources, and the use of devices as a result of data access. Mazzei and Noble (2017) cited Apple, Google, Amazon, and Microsoft as examples of organisations that have been able to disassociate themselves from numerous competing situations through their commitment to data-driven insights and actions, which have a direct impact on their products and services.

Though the vast majority of organisations limit their operations to tier 1 activities, there are numerous examples of how data can be used to influence organisational strategy. As previously noted this is not always easy to achieve and appears to be the basis of most organisations' inefficient transfer from potential to actualised value.

The relevance of BD is accentuated by the future scope of BD. Global data growth began quickly a decade ago and has not slowed. These are the major drivers of the worldwide BD market, which has already reached $49 billion (US dollars) (Mazzei &

Elragal, 2017). The world is fuelled by BD, pushing organisations to hire data analysts who can handle complex data processing. But will it continue?

Five predictions have been made regarding the future of BD (Marr, 2021):

1. Data volumes will grow and transfer to the cloud.

2. Most BD specialists agree that data generation will continue to expand tremendously in the future. For Seagate, IDC (International Data Corporation) predicts the global datasphere will hit 175 zettabytes by 2025. This stack would have covered two-thirds of the Earth-Moon distance, and this would have grown 26 times by 2025.

3. Machine learning will continue to evolve.
   Machine learning (ML) is another emerging technology that will have a major impact on our future. Every year, machine learning improves. That potential goes beyond self-driving cars, fraud detection, and retail trend analysis. Automated machine learning is a rapidly evolving technology. In 2020, ML initiatives received more funding than all other artificial intelligence (AI) systems combined.

4. Demand for data scientists and chief data officers increases.
   Although data scientists and chief data officers (CDO) are relatively new jobs, demand for these individuals is already considerable. The gap between demand and supply for data specialists is high. In 2020, KPMG surveyed 3,600 CIOs and IT leaders from 108 countries and found that 67% faced skill shortages (an all-time high since 2008), with big data/analytics, security, and AI being the most in demand.

5. Privacy will be an important topic.
   Data security and privacy have long been great concerns. Data protection cannot keep up with the data growth rates, therefore safeguarding it from invasions and hacks becomes more difficult.

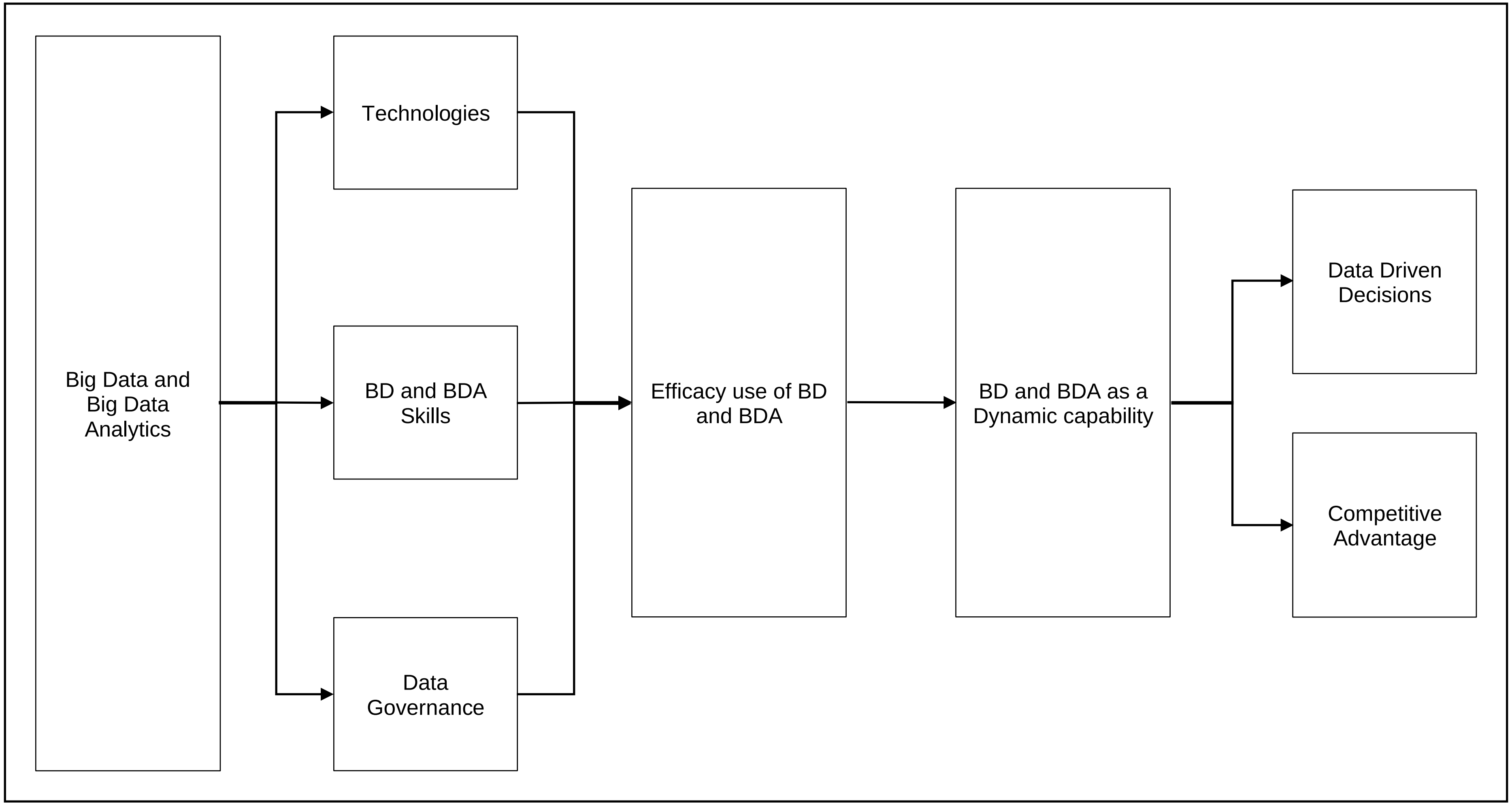
2.2.3 Applicable Models and Frameworks of Big Data (BD) and Big Data Analytics Capability (BDAC)

The literature reflects a number of models and frameworks applicable to BD. The ones that will be discussed are BDAC abilities as a dynamic capability and predictive modelling. These models or frameworks have been selected because they create multiple windows of opportunity to enhance organisational performance (OP). The first model for BDAC as a dynamic capability is depicted in Figure 2.1 below. The model was devised as a result of exposure to several BD and BDAC implementations at organisations, as well as research into the existing literature on the subject.

The **BD and BDAC as a dynamic capability model** is primarily concerned with the interaction between BD and BDAC, and how this contributes to OP, which has an impact on an organisation's ability to gain a competitive edge. Al-Sai (2019) describes the role of BD and BDAC in organisations and elements such as the data governance process, technological enablers, and investment in BD and BDAC skills. The model also shows how BDAC can be utilised for predictive modelling and deep learning (Kumari et al., 2018), demonstrating that BDAC can be a dynamic capability that informs data-driven decisions in organisations and subsequently creates competitive advantage.

Figure 2.1

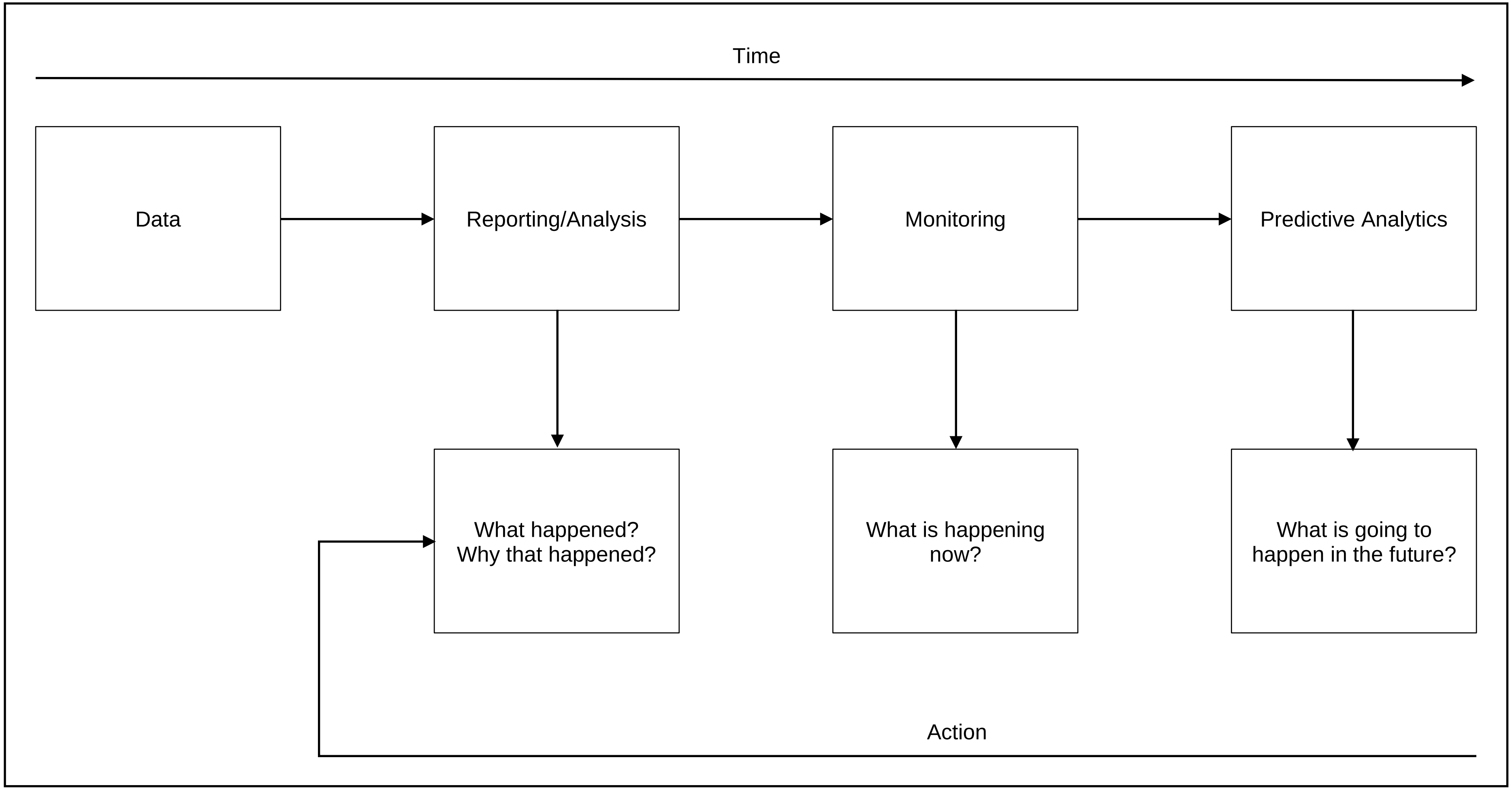
BD and BDAC as a dynamic capability model (Mikalef et al., 2020).



In the second model, **predictive modelling** with BDAC opens a new window for organisational performance (Kumari, Patil, & Jeble, 2018). Predictive analytics uses data, algorithms, and machine learning to predict possible trends based on real data (Diaz-Aviles et al., 2015). Cloud computing and predictive analytics combine to give organisations with massive amounts of data the ability to do complex analysis (Fiedler et al., 2016). This implies organisations may employ data-driven analytics to monitor hardware, predict loss patterns, and handle communication infrastructure issues more efficiently (Moe & Kallin, 2011). BDAC would help identify root causes and foresee problems, as shown in Figure 2.2. These developments will foster long-term competitive strategies that focus on building consumer value and brand recognition. Given that BDAC is seen to be effective when embedded in customer centric initiatives, it is reasonable to assume that BDAC is a dynamic capability.

Figure 2.2

Predictive Modelling (Moe & Kallin, 2011).



Lastly, according to Cohn and Marshall (2014), **organisational framework** strives to integrate BDAC and BD, organisational knowledge, and information technology (IT). This is discussed in the following focus areas,

2.1 Does the organisation consider BD and BDAC to be a critical function, like finance, IT, sales and marketing, product development, etc?

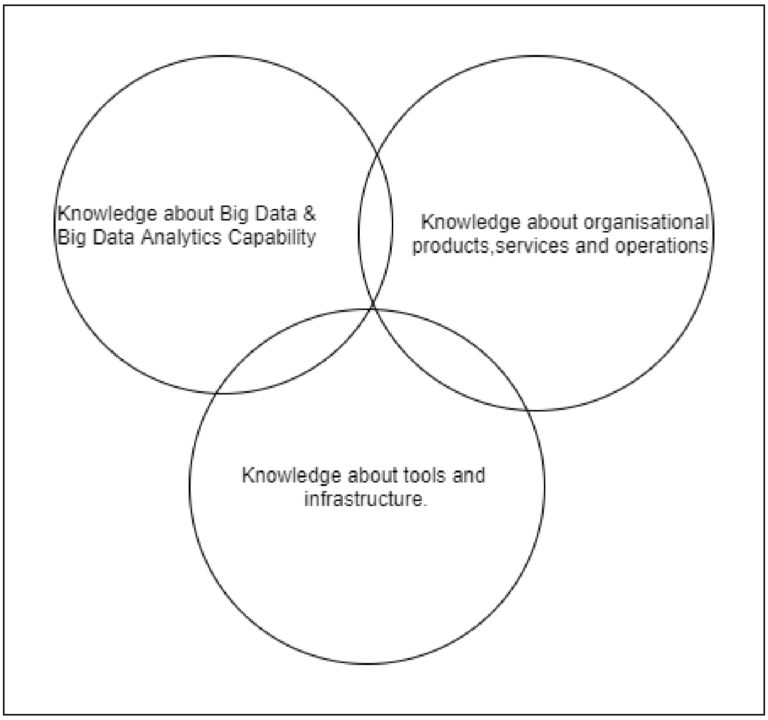
2.2 Are there enough data scientists? Without a critical mass of data scientists, subject matter experts' (SMEs') knowledge is insufficient to address all issues. Also, the analytics infrastructure is not understood well enough to get or create the required data and manage it. Finally, operational use of statistical and data mining methods may be lacking.

2.3 Are there data scientists familiar with the organisation's unit domains? Making valuable organisational models is tough without this understanding, and complex organisational problems tend to create specialisations. A data science team should have a mix of generalists and specialists.

2.4 Is the analytics governance framework adequate? A governance structure aids stakeholders in prioritising BD and BDAC opportunities, obtaining data, deploying analytical models, and assessing their organisational impact.

Figure 2.3

The knowledge required by data scientists (Cohn & Marshall, 2014)



This framework is known as CSPG (culture, staffing, processes and governance). The CSPG framework guides the organisation designer in hiring, educating, and organising a group of BD and BDAC experts. Establishing analytics staff, procedures, and systems government structure starts with culture, which organisational leaders must acknowledge. An organisation's BDAC is as a broad organisational function similar to other major functions.

2.2.4 The Measures of Big Data (BD) and Big Data Analytics Capability (BDAC)

BDAC indicators can help analyse the organisation's data and the impact within which BD grows and innovates for the economy (Mazzei & Elragal, 2017). But there are no longer any widely approved BD measurement systems, even though BD ecosystems is a key component. In 2012, the Massachusetts BDAC Initiative built a competitive BDAC ecosystem that was measured. The ecosystem was measured in terms of organisations, technology, talent, and capital. The Massachusetts Technology Collaborative Innovation Institute established eight critical metrics that characterise Massachusetts' competitive position in BD and the BDAC ecosystem expansion.

Table 2.2

Big Data indications and measures (Mazzei & Elragal, 2017).

Indicator	Factors
İnformation infrastructure	A number of Internet users.
	Worldwide per capita information.
	A number of devices connected to the Internet.
	A number of mobile devices connected to the Internet.
	A number of mobile phone users.
	A number of Internet-connected enterprises.
	Productivity of large data-related devices (storage, computing, generating).
	Input speed of broadband internet network (mbit / sec).
	A number of data centres.
Innovative	A number of research centres associated with BD.
	A number of large data-related studies.
	A number of scientific publications related to BD.
	A number of patents associated with BD.
	A number of invested BD projects.
	Volume of investment in BD projects.
Human capital factor	A number of enterprises (universities, colleges) teaching BD.
	A number of data science programs.
	A number of graduates in science, technology, engineering, and mathematics related to data.
	Courses related to data.
Economic factors	Number of organisations offering BD products and services.
	Volume of investment in BD related to organisations.
	Volume of revenues of organisations associated with BD.
	BD market volume (software, hardware, services).
	Number of jobs associated with BD.

Organisations measure BDAC to find new opportunities, cut expenses, improve decision-making, and ultimately raise customer happiness (Alexander, 2017). So, based on the hypothesis generated, the value of BDAC will be judged by the type of data usually needed in organisations, organisational functions that employ BD and BDAC tools.

The list of indicators (Table 2.2) is not exhaustive, but it hints to possible classifications. To achieve this, it may be necessary to focus on harmonising the various techniques. It is appropriate to establish a more formal framework.

In terms of the survey that has been used in this study the measured factors are:

1. Present State of investment in BDAC.
2. Rate of investment in BDAC.
3. Organisations preparing to capitalize on BDAC investments.
4. Challenges establishing a data culture.
5. Organisations becoming more data-driven.
6. Organisations using BDAC to alter themselves.

2.2.5 Research Studies on Big Data (BD)

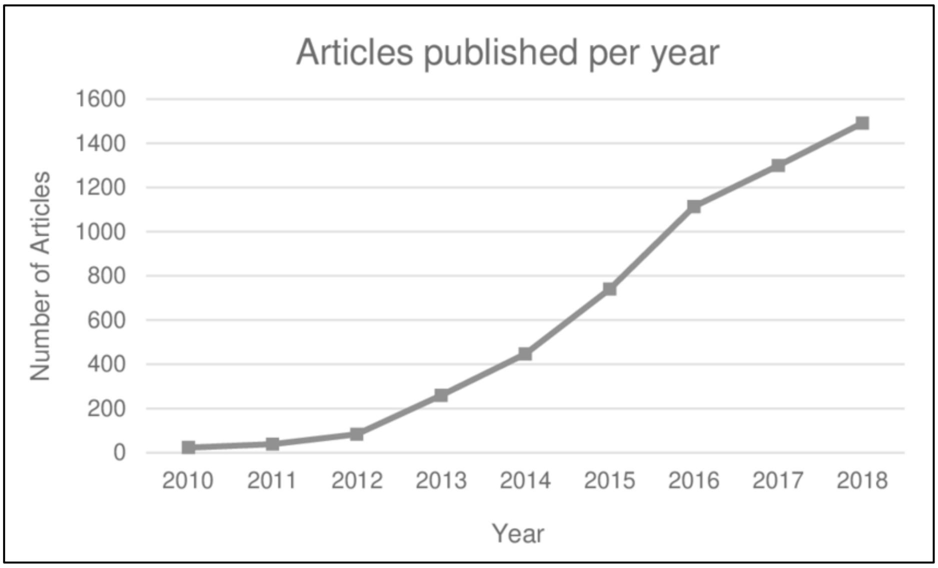
The concept of BD has been researched by various writers (Mikalef et al., 2019; Popovic et al., 2012). Although the BD phenomenon has only recently acquired popularity from 2007, Chen et al. (2012) suggest that the notion that BD and BDAC adds to data-driven decision-making has been around for much longer. BD and BDAC are distinguished by the fact that BD is unlimited, real-time, and unstructured, whereas BDAC is finite, offline, and structured (Alharthi et al., 2017). This ever-expanding range of BD sources enables a more comprehensive perspective of events, whether they are physical, online, or mobile transactions; or interactions such as internet clicks, social media posts, or sensor network interactions (Akbay, 2015; George et al., 2014).

What have we learned about BD since the first significant articles describing BD's promise to change the way we work, live, and do business? The interest in BD has increased dramatically (Chen, et al. 2014; LaValle et al., 2011; McAfee et al., 2012; Walker & Brown, 2019, 2011). Since then, experts have tried to understand how organisations may produce and capture value from their data resources. Figure 2.4

below depicts the number of articles published and found in the academic database Scopus. The researcher used a query to grab publications from 2010 to 2018, when BD and BDAC emerged; 5 495 items were found in total. This figure illustrates that the field is continually increasing and that more researchers are focusing on BD. It is becoming increasingly important to summarise significant research streams and determine what underlying assumptions should be challenged. Examining the broader area and the research issues helps identify important streams that have received insufficient attention, despite pressing practical needs.

Figure 2.4

Articles published per year for the period 2010-2018.



While BD has been in the spotlight for over a decade, they are not new ideas and have been offered under various labels in the past. While these ideas have resurfaced throughout history, they have managed to capture the attention of academics and practitioners since 2010. This can be credited to a number of factors that aided in the emergence of BD. First, storage costs and capacity have steadily decreased, allowing massive data gathering at minimal cost (Ji et al., 2012). Second, modern computers' processing capacity has improved dramatically, while the cost of delivering it has decreased (Sagiroglu & Sinanc, 2013). Third, the creation of sensors and connected devices in physical and digital objects has enabled organisations to acquire previously inaccessible data in real-time (Atzori & Morabito, 2010). As network infrastructures mature and cloud computing organisational models evolve, more enterprises may access scalable services, transmit data, and generate insights practically instantly and cheaply (Agrawal et al., 2011). The above elements together have made it possible to cost-effectively operationalise advanced analytics techniques that demand large amounts of data.

2.3 UNDERSTANDING BIG DATA ANALYTICS CAPABILITIES (BDAC)

This section provides an understanding of big data analytic capabilities (BDAC). It covers in detail the terminology, relevance, applicable models and frameworks, measures and tools, and previous studies, so as to explain the research objectives.

2.3.1 Defining and Characterising Big Data Analytics Capabilities (BDAC)

Big data analytics capability (BDAC) is defined as "a new generation of technologies and architectures, designed to economically extract values from very large volumes of a wide range of data, by enabling high-velocity capture, discovery, and analysis" (Mikalef et al., 2020, p. 242). In addition, Wamba et al. (2017) described BDAC as a holistic strategy for managing, processing, and analysing volume, variety, velocity, and value in order to provide actionable ideas for generating long-term value, monitoring performance, and establishing competitive advantages. Chen et al. (2015), Mikalef et al. (2018) and Sheng et al. (2017) all agree that BDAC is a complicated technique that is used to uncover insightful information through the use of structured and unstructured data by revealing hidden patterns (Gandomi & Haider, 2015; Lee, 2017; Najafabadi et al., 2015). As a result, organisations are increasingly adopting BDAC for the goal of making operations and data-driven decision-making processes simpler and faster. Therefore, BDAC, which is powered by complicated software and algorithms, provides a variety of organisational benefits. These include:

1. New revenue possibilities (Banerjee, 2013a).

2. Marketing that is both robust and effective (Kitchens et al., 2018; Sheng et al., 2017).

3. Improved customer service and real-time functionality (Aluri et al., 2019; Hung et al., 2016).

4. Improved operational efficiency (Wamba et al., 2017).

5. Long-term competitive advantage against competitors (Chen et al., 2015; Kitchens et al., 2018; Mikalef et al., 2018; Opresnik & Taisch, 2015b).

The use of BDAC for compelling organisational strategies has gotten a lot of attention in recent years (Mikalef et al., 2020), but there have been few research studies on the competitive potential of BD.

2.3.2 The Relevance and Rationale of Big Data Analytics Capability (BDAC)

Big data analytics capability (BDAC) is considered a key differentiator between organisations that perform well and those that underperform (Mikalef et al., 2020), especially since it improves efficiency and effectiveness, because of its high operational and strategic potential (Wamba et al., 2017). Studies such as those conducted by Hung et al., 2006, Tsai et al., 2015, Vidgen et al., 2017, and Wamba et al., 2017, show that the use of BDAC as dynamic capability, enables managers, statisticians, and trend analysts, among other professionals, to systematically analyse rapidly incoming data.

Capabilities "facilitate the most efficient, effective, and competitive use of an organisation's assets, whether tangible or intangible" (Mikalef et al., 2020, p. 3). This study further indicates that capabilities operate differently from organisation to organisation; hence, they result in varying levels of organisational impact and even competitive advantage. However, from this definition, it is apparent that dynamic capabilities are needed to create value.

In the current environment, organisations collect large volumes of data from customer interaction, financial information, and social network insights as well as location-based information (Mikalef et al., 2020). Thus, the use of BDAC to make sense of such vast amounts of data which in turn can help organisations reconfigure their strategies based on observed trends, within their competitive environment, and thereby improve their competitive advantage (Mikalef et al., 2020). Further, even though there is sufficient potential for BDAC as a source of organisational impact (Chen et al., 2015; Mikalef et al., 2020; Verhoef & Lemon, 2013; Wamba et al., 2017), there are indications that this could prove futile without the needed organisational capabilities. Hence, BDAC as a capability could enable organisations to realign themselves in changing environments to address customer needs (Mikalef et al., 2020).

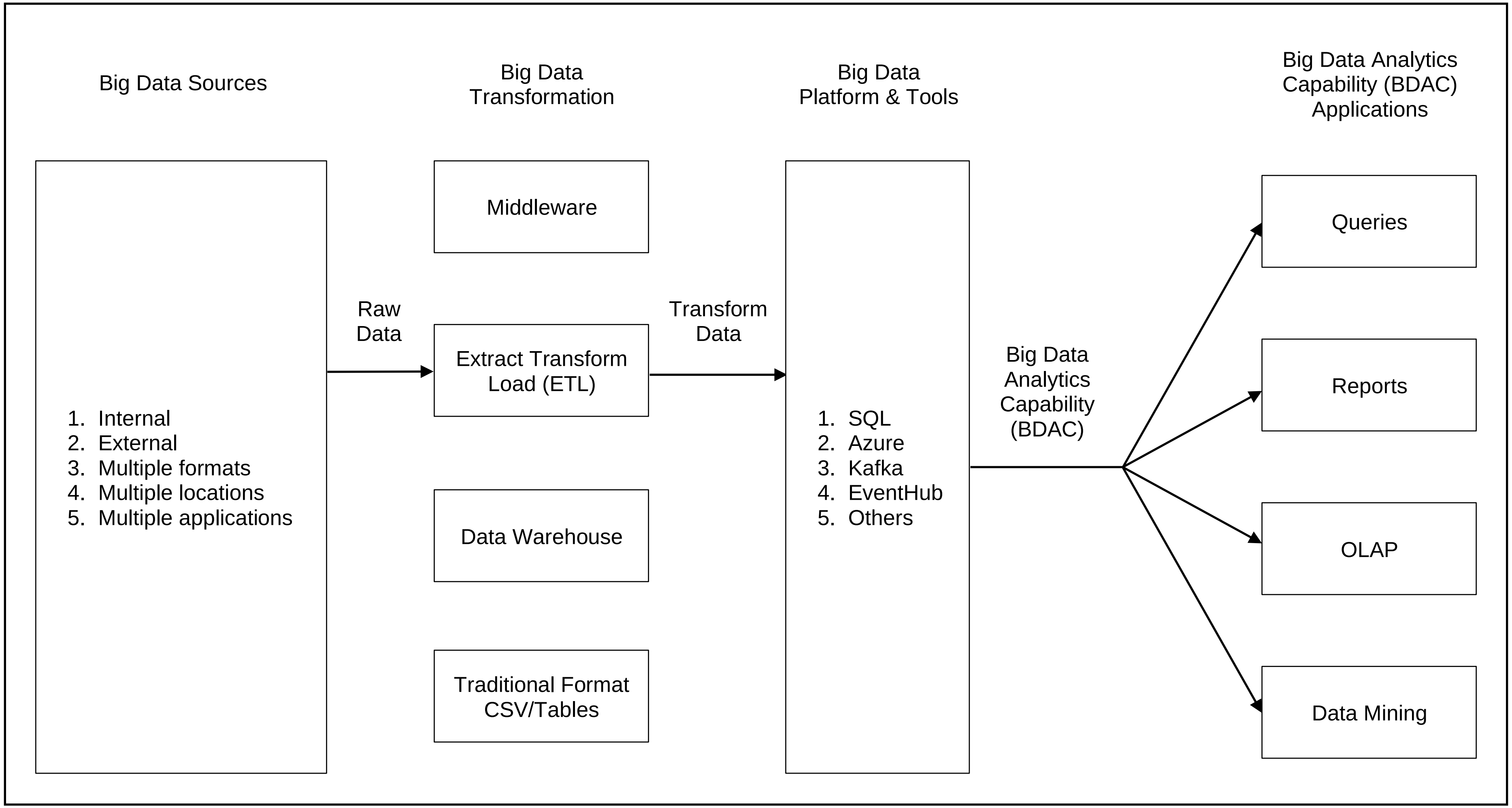
2.3.3 Applicable Models and Frameworks of Big Data Analytics Capabilities (BDAC)

A BDAC project follows the same conceptual framework as an analytics project. The main distinction is in the processing. In a typical analytics project, a business intelligence product (SQL) installed on a desktop or laptop is used to analyse data. Because BD is enormous by definition, it is split up and processed by several nodes (Hevner et al., 2004). While distributed processing techniques have been around for decades, their use in analysing very big data sets is relatively new as organisations begin to use data to make data-driven decisions. The cloud availability of open-source technologies like Azure/SQL has further boosted the use of BDAC in different fields. While the algorithms and models are identical, the user interfaces are not. Classical organisational analytics tools are now visible and easy to use, however, BDAC technologies are exceedingly sophisticated, need coding, and require a wide range of expertise (Becker et al., 2009).

As seen in Figure 2.5 below, data is a fundamental component. The data can come from internal and external sources and be stored in multiple legacy and other applications. This data must be gathered for analytics. The data is raw and must be processed. There are various alternatives. One option is a service-oriented architecture using web services (middleware) where the data remains unchanged, and services are used to call, retrieve, and process it (Venable et al., 2012). On the other side, data warehousing is a method of aggregating and processing data from various sources. But the information is not real-time. Extraction, transformation, and loading (ETL) of data can be seen in Figure 2.5 reflected as an applied conceptual architecture of BDAC.

Figure 2.5

An applied conceptual architecture of Big Data Analytics Capabilities (Raghunathan, 2014)



Several data formats can be used to enter structured or unstructured data to Hadoop/MapReduce (Sathi, 2012; Zikopoulos et al., 2013). Several decisions are made about data intake, distributed design, tool selection, and analytics models in this conceptual framework stage. Finally, the far right shows four common BDAC applications. These are searches, reports, and data mining. Visualisation is a common theme in all four apps. BD can be aggregated, manipulated, analysed, and visualised using numerous techniques and technologies. These methods and tools are based on statistics, computer science, applied mathematics, and economics (Courtney, 2013).

2.3.4 The Measures and Tools of Big Data Analytics Capability (BDAC)

It is worthwhile to examine some of the tools used to assist organisations in comprehending the role of BDAC in the development of useful data-driven insights. The four-dimensional paradigm of analytics is one such instrument. This measuring tool is presented and discussed below, by reflecting on its primary characteristics.

1. **Descriptive: What is happening?**

   The most popular form. It shows the analyst vital key performance indicators and measures within the organisation. A monthly profit and loss statement is one example. Similarly, an analyst may have data on a wide consumer base. Understanding consumer demographics (e.g., 30% of our customers are self-employed) falls under the category of descriptive analytics. Using good visualisation tools strengthens descriptive analytics messaging.

2. **Diagnostic: Why is it happening?**

   Descriptive analytics is the next level of data analytics sophistication. Using descriptive data, an analyst can apply diagnostic analytical methods to go down and find the problem's fundamental cause. Such analysis is possible with well-designed dashboards (Power BI, Tableau) that read time-series data (data over time) and allow filtering and drilling down.

3. **Predictive: What is likely to happen?**

   Predictive analytics is forecasting. Predictive models can estimate the possibility of an event occurring in the future, a specific amount, or a specific time frame. To produce a prediction, predictive models usually use a multitude of variables, for example, a person's age is related to their risk of having a heart attack – we would say age has a linear association with heart attack risk. After that, a score or forecast is generated. In an uncertain world, being able to forecast helps make smarter decisions. Predictive models are widely used in many fields.

4. **Prescriptive: What do I need to do?**

   The prescriptive model adds value and complexity. The prescriptive model helps the user select the best course of action by understanding what happened, why it happened, and what might happen in the future. Prescriptive analysis usually involves multiple activities, not simply one. Despite the fact that different types of analytics may bring varying levels of value to different organisations, they all have a place in the world.

BDAC includes data and a procedure. The procedure handles data from acquisition to interpretation. Several investigations have identified two BDAC components. A hybrid model analyses both data and process (Serhani et al., 2016), but in quality terms. Accuracy, throughput, and reaction time were used to evaluate pre-processing and processing analyses. Their analysis mentioned several data quality metrics. Villalpando (2014) presented the large data performance analysis model. Performance efficiency and reliability were tested using ISO 25010 software quality concepts using the "devil's quadrangle", a process redesign framework. The framework integrates time, quality, cost, and flexibility (Dumas et al., 2013). The framework includes financial (cost) and non-financial (time, quality, and flexibility) measurements.

More process performance measuring frameworks and models are found in the literature, mostly in organisational process and manufacturing perspectives. One is the TOPP system framework. The TOPP system analyses performance in three areas: efficiency, effectiveness, and changeability (Brownell et al., 1997). The TOPP system is similar to the devil's quadrangle in that cost and time are measurable concepts. Customer satisfaction defines TOPP system effectiveness. Comparatively, the quality dimension in the devil's quadrangle aims to satisfy both customers and process participants (the staff). Changeability and flexibility have a lot in common. Both frameworks can be used to make this trade-off. In manufacturing, TOPP uses questionnaires to analyse process and organisational performance. Table 2.3 summarises existing frameworks and models.

Table 2.3

Existing Models and Frameworks (Serhani et al., 2016)

Model/ Framework	Measures/Metrics
A hybrid model for assessing quality of BDAC value chain	Accuracy, throughput, and response time.
Performance analysis model for BDAC applications	Performance efficiency (time behaviour) resource utilisation, and capacity), reliability (maturity, availability, fault tolerance, and recoverability).
Process performance dimensions (the devil's quadrangle)	Time, quality, cost, and flexibility.
TOPP system	Efficiency, effectiveness, changeability.
Process performance model	Overall, cost, quality, service, and time as four classes of indicators with three stakeholders, namely: customers, operators, and management.
Process performance evaluation methodology	Time, quality, service, efficiency, cost, and importance.
A model of process performance	Customer satisfaction, product development time and the cost for product design and manufacturing.

2.3.5 Research Studies on Big Data Analytics Capabilities (BDAC)

Numerous studies have been conducted on BDAC. According to Mikalef et al. (2020), BDAC is considered as a tactical advantage because it may be used to restructure organisational plans, to gain a better understanding of consumer preferences, to facilitate data-driven decisions and, as a result, optimal strategic operations. Decision-making that is effective enables data-driven organisations to be robust and adaptable in the face of change (Wamba et al., 2017).

The dynamics of organisations are rapidly shifting. There is a lack of predictability in unexpected circumstances (Mikalef et al., 2020), and of understanding of how organisations should use BDAC to attain their goals and use this as an advantage in the marketplace. There is also a widespread lack of awareness of how BDAC works (Aluri et al., 2019).

Studies also found that BDAC can be used as a dynamic capability by organisations (Tsai et al., 2015; Verhoef et al., 2007; Verhoef et al., 2015). Despite the fact that many organisations are intending to participate, there is little information regarding BDAC if you want to invest in it or have already done so. Some studies suggest that it is still not clear how to go about building this capability in organisations. For example, questions such as "How should this be built as a dynamic capability?" (Gupta & George, 2016). Further, even if client value is considered a source of long-term revenue, BDAC is not being properly used as a competitive advantage (Wamba et al., 2018).

Finally, it has also been found that conventional analytical tools such as data mining and dynamic capabilities inside organisations like warehousing are still frequently employed (Hung et al., 2006; Tsai et al., 2015).

The next section discusses organisational performance (OP) in relation to BD and BDAC.

2.4 ORGANISATIONAL PERFORMANCE (OP)

This section provides an understanding of organisational performance (OP). It covers in detail the terminology, relevance, applicable models and frameworks, measures and tools, and previous research studies in order to explain the research objectives.

2.4.1 Defining Organisational Performance (OP)

Organisational performance (OP) refers to an organisation's ability to meet its objectives and meet the expectations of its stakeholders, and to stay afloat in the economy (Griffin et al., 2003). It may also be defined as the process of examining and measuring an organisation's performance in relation to its objectives and goals, which includes a comparison of actual and planned outcomes (Richard et al., 2009). The actual productivity or outcomes of the organisation are contrasted to the desired outcome or objectives in terms of OP. Higher performing organisations, according to Teece (2019), are able to deal with innovation, safeguard, and employ intangible knowledge assets beneficially. Further, OP can be defined as the process of ensuring that organisational resources are properly used, and it encompasses all actions or activities undertaken by managers at various levels in order to assess the extent to which an organisation has met its goals (Teece, 2000). In this study, the researcher uses Teece's (2019) definition. Teece’s definition has relevance in our current economic environment and the need for organisations not just to stay afloat in the economy, but rather needing to be relevant and competitive.

2.4.2 The Relevance and Rationale of Organisational Performance (OP)

The ability of an organisation to adapt to changes in the external environment is critical to its success. Every organisation strives to improve over time, and OP is linked to the individual performance of team members at the organisational level. Achieving success necessitates three factors (Horga, 2012):

1. Economic efficiency: any organisation wishes to achieve its goals with few resources.

2. Customer satisfaction: superior results with few resources, but in a way that exceeds customer expectations.

3. Leadership effectiveness: the effectiveness of the leadership process depends on the leader's ability to enthusiastically lead the working team by meeting the personal needs of each member; "Leadership builds a relationship between individual performance and organisational efficiency," (Horga, 2012, p. 274).

In this context, the leadership style used to achieve organisational performance should be given special consideration, because the attainment of goals that lead the entire team to success depends on how it is practiced (Tseng, 2014). To support the attainment of organisational goals, effective leaders must create a favourable internal environment. It's critical not to overlook the BD and BDAC component in this strategy.

The link between leadership and OP is becoming more apparent, with the leader impacting the organisation's performance in a variety of ways. As a result, according to Horga (2012), leadership ensures that attempts are made to attain the proposed objectives, coordinates employees' efforts to achieve the goals, drawing on his or her personal expertise rather than formal procedures and motivates people to achieve the results. Creating a pleasant work environment, healthy connections, and communication inside organisations, and emphasising a positive sense of work are all examples of performance levels (Horga, 2012).

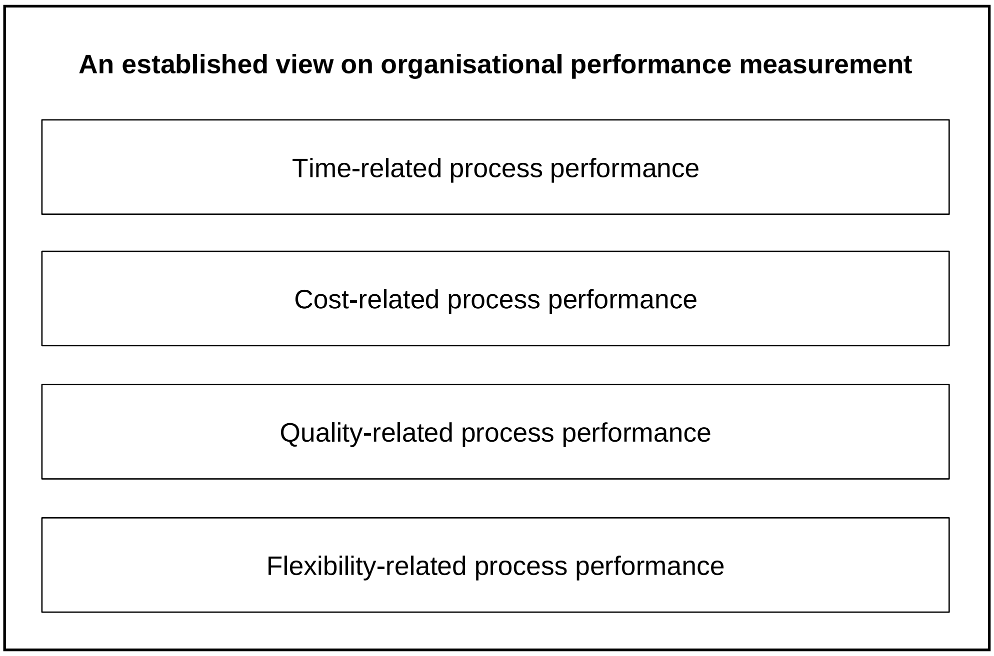
Leadership encourages positive emotions in the workplace, and compassion and thankfulness are promoted, which contribute to team members' jobs going smoothly, and in turn impact OP. Additionally, establishing a good feeling of work at the organisational level, emphasising the work of the entire team, leads to OP through engaging the leader on human relationships (Masa'deh et al., 2015). Positive leadership communication, on the other hand, motivates the team while also encouraging critical feedback (Horga, 2012). At the individual level, attention is focused on the individual leader and the relationships he or she establishes with individual followers. At the organisational level, attention is focused on the individual leader and the relationships he or she builds with individual followers. As a result, a leader and a follower affect each other through time, with leadership being defined as the mutual impact of a leader and his or her followers (Horga, 2012). At this time, the key to success for every organisation that fulfils its objectives and solves difficulties in a creative manner is the leader. Leadership analysis at the organisational level entails defining the role of the leader within the many teams that make up an organisation. What is therefore expected of leadership in the context of this study is to create the necessary BD and BDAC that will enable OP and efficiency.

2.4.3 Applicable Models and Frameworks of Organisation Performance (OP)

Performance measurement can also be focused on specific organisational processes, such as statistical process control, workflow-based monitoring, or process performance measurement systems, in addition to organisational models (Kueng 2000; Neely et al., 2000). Process performance measurement takes a less holistic approach. For example, Dumas et al. (2013) define time, cost, quality, and flexibility as common performance views of organisational process performance measurement. See Figure 2.6.

Figure 2.6

An overview of the performance perspectives in Dumas et al. (2013)



Concrete performance measures or indicators should be created for each process performance perspective, just as they are for organisational performance assessment. In this regard, Dumas et al. (2013) developed views to define internal organisational process performance perspective. On the other side, a performance measure's major goal is to deliver comprehensive and timely information on organisational process performance. Early warning signals, diagnosing process weaknesses, deciding whether corrective actions are required and assessing the impact of actions taken are all possible uses of this data (Kueng, 2000).

According to Kueng (2000), performance measure is a stakeholder-driven measurement that focuses on those who have an interest in the organisational process. Process stakeholders must be identified. Each stakeholder or group of stakeholders must have process-relevant goals. He defines process performance as the degree of stakeholder satisfaction (Kueng & Krahn, 2000).

Stakeholders include financiers, staff, consumers (suppliers and buyers), and society. Each stakeholder group is represented by a performance dimension (Kueng, Meier, & Wettstein, 2001):

- **Financial view**

  Financial view is the term used to describe the process of evaluating financial performance by focusing on metrics such as profit, return on investment (ROI), return on sales (ROS), and profit per unit of production.

- **Customer view**

  A customer view has a customer satisfaction index that is used as a performance indicator to measure the objective "increase customer satisfaction". An indicator may not always be directly related to an objective, as Kueng points out. In this case, a refinement may help.

- **Employee view**

  Employee morale, training, promotional growth, and individual performance ALL impact organisations. Employee satisfaction helps achieve high performing individuals who show more dedication to the organisational objectives and performance.

- **Societal view**

  Organisations are confronted with increased competition, which compels them to cut costs and improve the value they provide to their customers. To adapt to this new environment, organisations shifted their strategic priorities and adopted new management philosophies.

- **Innovation view**

  Innovation and learning metrics focus on an organisation's ability to develop and introduce new products and services, as well as improve the performance of internal and external processes.

Hence, Kueng (2000) proposes a more comprehensive approach to process performance with the use of leading indicators. Leading indications are important. It is critical to measure not just the financial results of yesterday's decisions, but also the signs of tomorrow's performance. That is, identify and measure the performance drivers. Early warning signs impact performance measure system (PMS), says Bitici (2005), should help comprehend the patterns and relationships among measures.

Multiple performance metrics must be used. McNair et al., (2015) advocated monitoring performance at three levels: work unit, organisational operating system, and work central level. According to Fitzgerald et al., (2001), strategic organisational areas should be incorporated. Kueng (2000) claims organisational process performance should be monitored – not only on an organisation's functional level.

Aspects that generated consensus:

- Changes in the organisation's surroundings should be monitored. If the environment changes significantly, organisational objectives and strategy must adapt. Changes in the environment may affect the performance metrics.

- Some authors (Fitzgerald et al., 2001) recommend a PMS with an external monitor.

- Examining IT talents. Only a few studies explore IT's role in performance measurement (Bitici, 2017).

- Real-time performance data. Several authors address the issue of delay, i.e., the time between occurrence (of good or bad performance) and communication, which requires the reporting period be as brief as possible.

Overall, the many performance measurement frameworks proposed during the last decade emphasised dimensions and performance measurements (performance indicators).

2.4.4 The Measures of Organisational Performance (OP)

Organisational performance measurement models aim to provide a holistic view of an organisation's performance by taking into account several performance perspectives.

There are four viewpoints through which objectives and performance indicators ensure that plans and operations are in sync (Kaplan & Norton 1996, 2001). Figure 2.7 below outlines the elements of organisational performance.

Figure 2.7

Balanced Scorecard (Organisational Performance Measure)

An established view on organisational performance measurement
Financial performance
Customer performance
Internal Business Process Performance
Performance related to learning and growth

Kaplan and Norton (2001) emphasised the financial perspective, noting that financial performance is a lag indicator that defines an organisation's success. This strategy often describes how an organisation aims to produce long-term growth, profitability, and shareholder value (Kaplan & Norton, 2001). Due to the nature of most profit-making organisations, this one statistic encapsulates the core motivation for their existence. This viewpoint asks, "How do our shareholders see us?" and ties organisational performance to financial goals (Du Plessis et al., 2001, pp. 424-427). Customer satisfaction is key to enhanced financial performance (Kaplan & Norton, 2001). In addition to lagging indicators, Kaplan and Norton (2004) suggest tracking customer success factors like satisfaction, retention, and growth. Choosing the customer value proposition as the core aspect of this approach asks, "How should our customers regard us?" (Du Plessis et al., 2001, p. 424). When a non-profit organisation uses the balanced scorecard, this perspective becomes the driving force (Kaplan & Norton, 2001).

Internal organisational processes produce and give value to customers (Kaplan & Norton, 2004). Innovation, operations, and after-sales support are processes that managers should excel in (Du Plessis et al., 2001: p. 428). Kaplan and Norton (2004) claim that improving internal processes leads to better customer and financial outcomes.

Finally, intangible assets are the ultimate source of long-term wealth generation (Kaplan & Norton, 2004). Lessons learned and goals for growth describe how the people, technology, and environment support the plan. Indicators of internal process, customer, and financial performance include "learning and growth" measurements. They emphasise the need for investing in non-traditional areas for the future (Du Plessis et al., 2001; Kaplan & Norton, 2004).

The balance scorecard (Figure 2.7) seeks to achieve success from all four viewpoints. Nonetheless, one perspective may be given more weight to address organisational areas that require development (Brackertz & Kenley, 2002). Prioritizing strategic goals is usually done by senior management and normalised. However, Hasan and Tibbits (2000) warn against prioritising one viewpoint above another. They cite examples of organisations emphasising the financial perspective because it is the most easily measured. Moreover, Hasan and Tibbits (2000) say the balanced scorecard is designed to measure what counts, not what is simple. The weighting should reflect the importance of each perspective in relation to the organisation's goals (Brackertz & Kenley, 2002).

Similar insights can be found in other organisational performance evaluation approaches. Cross and Lynch (1988), for example, propose a four-level performance pyramid:

1. A top-level vision.
2. A second level set of market and financial objectives for each organisational unit.
3. A third level with customer satisfaction, flexibility, and productivity objectives for each of the organisation's operating systems.
4. An operational level with quality, delivery, process time, and cost targets.

Despite the differing views, this study uses the balanced scorecard since it helps approach on how to evaluate organisational performance (Seang, 2003). Furthermore, an organisation's performance measuring system, and proper implementation and maintenance of models are critical.

2.4.5 Research Studies on Organisational Performance (OP)

According to Mdluli and Makhupe (2017), the current organisational environment is fast-paced, volatile, and unpredictable, necessitating dynamic analytic capabilities (Eisenhardt & Martin, 2000; Helfat & Peteraf, 2014; Pisano, 2017; Teece et al., 1997; Wang & Ahmed, 2007). Pisano (2017) defines dynamic capabilities as an organisation's "capacity to reconfigure and extend its competencies" (p. 748). According to Erevelles et al. (2016), organisations that use BD to analyse and respond to changing consumer needs are more dynamic. The relationship between dynamic skills and competitive advantage has also grown in importance. Many studies suggest that dynamic capabilities contribute to organisational success and long-term competitive advantage (Li & Liu, 2014; Lin & Wu, 2014; Osisioma, Nzewi, & Mgbemena, 2016; Wu, 2010). The resource-based concept of stagnating marketplaces is opposed by dynamic capacities (Teece et al., 1997). Teece (2007) demonstrated that dynamic capacities can adjust to changing technology and clients.

Sense-making, real-time decision-making, and change management are also critical talents (Li & Liu, 2014). Critical talents to implement BD and BDAC are required for organisations. These critical talents include deep understanding of data, data analysis, data governance and solution architectures (Akerkar, 2019). As a whole, BD and BDAC implementations in organisations can create advantages, but according to Teece (2007), finding and shaping threats and opportunities through market and technology exploration, capturing opportunities through data integration and interpretation, and sustaining competitive advantage through the creation, renewal, or reconfiguration of capabilities must align with organisational strategy to make a difference in an organisation. Teece et al. (1993), further point out that a competitive advantage can be gained by discovering new possibilities and growing organisational resources. To build a sustained competitive advantage, an organisation must also be able to adapt to opportunities and create new value (Erevelles et al., 2016). Also, dynamic capabilities are a key factor in understanding performance and future projects.

South African organisations are influenced by changing customer demands and technical advancements. Organisations engage in a highly competitive and saturated market (Millar et al., 2018), with rising competition which may lead to reduced profit margins. As a result, the dynamic capability (as a source of competitive advantage) concept is particularly relevant to this study.

2.5 CHAPTER SUMMARY

The focus of chapter two was on achieving competitive advantage and OP through BDAC, as reflected in the literature. The chapter explored possible conceptualisations, the relevance and rationale, models, and measures as well as research studies pertaining to BD, BDAC and finally OP.

CHAPTER 3:
LITERATURE REVIEW

RESEARCH STUDIES ON VALIDATION AND THE EFFECT OF BIG DATA ANALYTICS CAPABILITIES (BDAC) ON ORGANISATIONAL PERFORMANCE (OP)

3.1 INTRODUCTION

The previous chapter examined and provided an understanding of big data (BD), big data analytics capability (BDAC) and organisational performance (OP), their use and associated organisational value. It covered in detail the key concepts and terminologies that are needed and/or commonly used to understand the research objectives.

The purpose of this chapter is to provide an understanding of previous research studies on validation and the effect of BDAC on OP.

3.2 RESEARCH STUDIES ON BIG DATA ANALAYTICS CAPABILITIES (BDAC) AND ORGANISATIONAL PERFORMANCE (OP)

Numerous studies have been conducted on BDAC and OP. BDAC infrastructures collect, process, store, and analyse BD (Mikalef et al., 2019). This feature ensures that all technologies can process data in a variety of forms in a variety of contexts (Provost & Fawcett, 2013). The managerial competencies are critical for selecting the optimal BD infrastructure and information (Ferraris et al., 2019). Managers must decide on the best technical solution for their organisation. Similarly, studies suggest that managers must be able to process data, derive at useful data analytics to be able to make correct data-driven decisions (McAfee et al., 2012). Personnel must have BDAC skills for several reasons. Skilled personnel can lessen the danger of an organisation rejecting BDAC or not using new information management technologies. Moreover, skilled personnel can help the BDAC infrastructure perform better. Employees studying big data also need expertise, this allows for appropriate analysis and conclusions. As a result, organisations might gain a competitive advantage (Jeble et al., 2020). Organisations that make substantial investments enjoy the BDAC benefits. However, small and medium-sized organisations are often at a disadvantage, because they lack this investing ability. These organisations cannot afford data lakes, parallel computing, or staff training (Amado et al., 2018).

BDAC is linked to the structural aspect, BDAC infrastructure, organisational dynamics, and HR management. The organisational routines are tied to BDAC, and the main theoretical approach of dynamic capacities is based on BDAC research (Wamba et al., 2017). Teece (2012) defines dynamic capabilities as an organisation's ability to adapt to an ever-changing environment through changes in resources, internal or external processes, and strategies. Some definitions of dynamic capacities are directly linked to organisational improvisation. These are based on recognized routines (Teece, 2018). Some organisational processes and routines can become excellent organisational practices. An organisation's routines can be broken down into smaller procedures, or "bricks". Personnel and managerial BDAC practices are the bricks that can be employed in many situations. Thus, an organisation may gain a competitive advantage (Rialti et al., 2019).

Studies also suggest that BDAC infrastructures are interoperable, flexible, scalable, and adaptable (Mikalef et al., 2019). Moreover, they ensure timely information flow in any situation. BDAC's impact on an OP is evident (Wamba et al., 2017). This outcome is in line with research on BDAC value creation. As previously said, ambidexterity “is necessary to simultaneously pursue both exploration and exploitation for its inventive redesign of operational processes” (Lee et al., 2017). Organisations that can re-arrange present processes and resources to face new challenges are also capable of identifying changes in the environment and exploiting possibilities (O'Reilly & Tushman, 2008). This capacity to exploit opportunities is supported by a study by Mikalef et al., (2019). They confirmed that BDAC could improve the ability of the organisation to identify new risks and opportunities. Further, BDAC allows organisations to identify new opportunities and capitalise on them (Rialti et al., 2019). Adaptable information management systems may also help organisations recognise and exploit new opportunities (Lu & Ramamurthy, 2011).

Similarly, organisations can identify and exploit opportunities by using flexible information management systems. Thus, ambidexterity may affect performance. A resource-based view (RBV) states that an organisation's resources must be scarce, precious, and imitable (Howson, 2013). BDAC’s examination of organisational performance and commercial value provided by technology resources also received significant attention by Kiron et al., (2012). This study recognised the RBV and confirmed organisational value and organisational performance as the main determinants after competencies and resources. A previous study revealed that resource complementarity, such as system quality, can be achieved through information quality. This can further boost organisational performance and value (Chen et al., 2014). Big data's potential impact on organisational performance and value has aroused interest in this idea. An organisation has a competitive edge if it outperforms its competitors. According to the RBV, high organisational performance and organisational value show competitive advantage in big data environments. Thus, distinguishing between organisational performance and organisational value is important. To grasp the nomological net, these must be differentiated from resources like information and system quality.

Various notions of organisational value are found in this literature. Parasuraman et al. (1985) defined three categories of organisational value: automation, transformation, and information. Teece (2012) defined organisational IT value as strategic, informational, and transactional advantages. The efficacy and efficiency of organisational processes are significant predictors of economic value. According to researchers, this impacts an organisation's overall performance. Previous research has indicated that technological resources provide organisational value, improving overall organisational performance (Garca-Morales et al., 2012). The organisational value of BDAC was characterised as the information, strategy and information gained, while transactional value improves efficiency and reduce costs (Garmaki et al., 2016). The informational value focuses on time-based decision-making, but the strategic benefit is gaining a competitive advantage.

Wang et al. (2018) claim that predictive analytics of BDAC have altered the performance of healthcare organisations. The healthcare sector improved in quality and reduced fraud and waste. Similarly, Howson (2013) proposes that BDAC can boost both tangible and intangible advantages. Intangible benefits include increased organisational reputation (Basheer et al., 2018). A recent study found that matching value, performance, and quality can enhance operating profits by up to 60%. For example, using a trustworthy information and quality system can enhance the creation of new products and services by 70%, expand new markets by 72%, and satisfy customer needs by 79% (Akter et al., 2016).

3.3 STUDIES ON THE THEORETICAL RELATIONSHIP BETWEEN BIG DATA ANALYTICS CAPABILITIES AND ORGANISATIONAL PERFORMANCE

There is limited direct correlational research on the theoretical relationship between big data analytics capabilities (BDAC) and organisational performance (OP). Related empirical studies seem to support a positive relationship between these two variables. For example, the existing knowledge has made a valuable contribution to the impact of BDAC on OP which align with organisations' technological skills and innovation.

Big data analytics capability is the next frontier for innovation, competition, and productivity (McAfee et al., 2012). Academics and practitioners have consequently placed a premium on the value that organisations can derive from using BDAC to accomplish goals. In a study by Mikalef (2017), which discusses enabling high-velocity data capture, discovery, and analysis, BDAC is characterised as a new generation of technologies and architectures capable of extracting value from very large volumes of diverse data. In a study by Chen (2012) of BDAC, organisations can identify new opportunities, gain critical insights, and adapt their operations to a changing competitive environment, which has a positive impact on organisational performance. According to Abbasi (2016), BDAC enables more informed decision-making, by providing a significant competitive edge which has a statistically positive relationship with organisational performance. Numerous studies suggest that BDAC has the potential to greatly improve a variety of industries, including mining, healthcare, service delivery, supply chain management and marketing (Wang et al., 2016). According to an *MIT Sloan Management Review* article, BDAC can encourage creativity, with CEOs who use BDAC strategies more likely to offer new products and services than their direct competitors (Ransbotham & Kiron, 2017).

While many assert that BDAC benefits organisations, little is known about the organisational consequences and challenges (Wamba et al., 2017). While there is evidence that BD and BDAC can generate commercial value, Sharma et al. (2014) contend that further research is necessary. According to studies, generating value through BD and BDAC needs an organisation to commit to a BDAC strategy (Gupta &

George, 2016). BDAC is described by Mikalef et al. (2017) as an organisation's ability to successfully employ technology and talent to capture, store, and analyse data to gain insight. Concerning the management and use of BDAC, Vidgen et al. (2017) argue that organisations face several challenges in deriving value from them. Numerous organisational reports point to the same underlying issue as a result of BDAC, namely, to derive economic value from BDAC, organisations must first determine the key organisational elements (Abbasi et al., 2016). There is a lack of understanding of the context-dependent effects of such factors on organisational performance in relation to value creation, decision-making, competitive advantage and organisational performance as a whole (Günther et al., 2017).

Some studies have found isolated elements that contribute to an organisation's success in implementing BDAC. Similarly, Wamba et al. (2017) demonstrated empirically how investment in infrastructure, management, and employee expertise may help organisations enhance overall performance. Numerous studies have identified critical factors that contribute to the development of BDAC (Mikalef et al., 2017). Thus, value extraction from massive data is largely uniform, with little emphasis on context, and the success of IT projects is contingent on the environment in which they are executed, and several contingency factors, according to past research in the information systems sector (Bechor et al., 2010). Certain resources are thought to be more or less significant in producing performance gains in these studies, depending on the scenario (Petter et al., 2014).

Additionally, most studies presume that all organisations have the same challenges and therefore spend time in the same areas of improvement.

3.4 CHAPTER SUMMARY

The focus of chapter three was on understanding previous literatures and approaches to BDAC and OP. The chapter commenced by exploring research studies on BDAC and OP and concluded with the research studies' theoretical relationship between BDAC and OP.

CHAPTER 4:
RESEARCH METHODOLOGY

4.1 INTRODUCTION

The previous chapter looked at past literature and approaches to BDAC and OP. This chapter provides an understanding of the research methodology of this study. This includes the research approach, the research method, the unit of analysis, sampling techniques and size, and the measuring instruments. The chapter concludes with the statistical techniques used to analyse and interpret the data, and a chapter summary.

4.2 RESEARCH APPROACH

A quantitative research approach was used in this study. A quantitative approach was deemed necessary for this study because it was appropriate to design and validate a new scale, and to address the research questions to gain additional insight into the relationship between the variables. This is achieved through measuring and testing of quantitative data in a structured manner (Zikmund et al., 2012). The research was designed in layers, with the research philosophy and approach discussed and utilised to determine the type and strategy of the investigation. Because respondents' data was collected at a single point in time, this study can be classified as cross-sectional (Saunders & Lewis, 2014).

A structured, quantitative methodology was used to facilitate replication, and the research philosophy can be defined as positivism (Saunders & Lewis, 2014). This approach was also deemed appropriate because the researcher desired generalisable findings so that the study could be replicated in the future.

Given the explosion of research on big data (BD) and, more recently, big data analytics capability (BDAC) and data-driven decision-making, the researcher sought to answer research questions based on existing literature. Hence, this qualifies as deductive research (Saunders & Lewis, 2014). Additionally, this study provides an accurate representation of the organisations studied, and thus qualifies as a descriptive study. An online survey was chosen as the research strategy, as it facilitated collection of data from respondents in a structured manner (Saunders & Lewis, 2014).

4.3 RESEARCH METHOD

This section provides an understanding of the research method adopted. It covers in detail the population and sampling techniques used, the biographical characteristics of respondents, measuring instruments, demographical measures, and research procedures.

4.3.1 Research Respondents

A population is defined as a complete set of individual members (Saunders et al., 2015), which includes not only individuals but also locations and organisations accessible to the researcher.

The population of interest for this study included key identified senior managers, executives, and data analysts who worked in the context of BD and BDAC with organisations or project implementations. The population included industry experts who were responsible for defining, driving, and influencing BD and BDAC strategies within organisations. The population size was 400.

4.3.2 Unit of Analysis

The investigated subject of the study is referred to as the unit of analysis (Babbie & Mouton, 2001). The unit of analysis for this study was individuals who occupy senior management, executive, or subject matter experts (SMEs) within an organisation who have experience in relevant fields of BD and BDAC. The individuals selected are accountable or responsible for big data and big data analytics capability, and are expected to understand how BD and BDAC is currently being used to improve organisational impact and organisational performance.

4.3.3 Sampling Technique

Measuring the entire population was deemed unfeasible since it would be costly, difficult, and extremely time consuming to measure every constituent of the population (Zikmund et al., 2012). Hence, a subset of the larger population was used, known as a sample (Saunders & Lewis, 2014).

In this study, probability sampling and snowball sampling were used (Saunders & Lewis, 2014). Probability sampling involves random selection, allowing one to make strong statistical inferences about the whole group (Zikmund et al., 2012). This method was used to ensure that the responses obtained were representative of the study's population. The second method was snowball sampling, a technique in which more respondents are identified from information acquired from original sample respondents (Zikmund et al., 2012). This was important given the researcher's limited network and reliance on respondents to suggest future potential respondents.

To collect the sample required for this study, the researcher contacted professional networks via email and asked them to complete and forward the survey to others in their network (Accenture South Africa). This sample method rendered it impossible to calculate the response rate for the questionnaire with precision. Thus, also a form of non-probability sampling was used (non-random selection) based on convenience (Salkind, 2018). The following criteria were used to select participants:

1. Experience working (from less than two years' right through to 20 years' experience) with big data (BD) and big data analytics capability (BDAC).
2. Experience and knowledge of BD and BDAC project implementations in organisations.
3. Willingness to participate in the study.

4.3.4 Sample Description

A higher sample size is related with more accuracy in quantitative research (Zikmund et al., 2012), and the power of a test is greatly dependent on the size of the sample that was gathered and used in the test (Pallant, 2007). In addition, due to the lack of power, there is a potential that the findings will not be statistically significant. When performing factor analysis, using a sample size that is too small can potentially lead to less trustworthy correlation coefficients among the variables (Pallant, 2007).

When running multiple regression tests, the size of the sample being tested is also a significant consideration. For the findings of a research project to contribute to scientific advancement, they must be generalisable to other samples (Pallant, 2007). However, while doing multiple regression, the generalisability of results may be compromised by samples that are too small (Pallant, 2007). According to Tabachnick and Fidell's (2007) recommendations, the ideal size of the sample should be N > 50 + 8m, where m is the total number of independent variables. Table 4.1 provides a summary of the needed responses based on the formula that was suggested. This study had a final sample size of 239 usable questionnaires, which was enough for the research questions.

Table 4.1

Minimum Sample Size Calculation

Research Question	Number of Independent Variables	Response Required
Research Question 1	10	(50 + 80) = 130
Research Question 2	1	(50 + 8) = 58

4.3.5 Measuring Instruments

The purpose of this section is to present and to discuss the measuring instruments in the form of a survey which was used in the study. It covers in detail the survey, its design and pilot testing of the revised survey.

4.3.5.1 Data Collection

The following measuring instruments were used in the study.

- **Biographical questionnaire**

  This questionnaire was used to obtain the personal, biographical information needed for the statistical analysis of the data. This information included gender, age, educational level and length of service. These were selected based on a theoretical review of the variables with possible impact on the empirical results.
- **Newly developed online survey**

  Data was also collected from respondents using an online survey. By providing respondents with anonymity via an online questionnaire, respondents were more inclined to submit sensitive information (Zikmund et al., 2012) and avoided the chance of respondents engaging in social bias (Podsakoff, 2003). Thus, this strategy resulted not only in a cost-effective and efficient survey process, but also produced more candid and potentially more accurate feedback. The preliminary, newly designed survey was administered to individuals in the form of subject matter experts (SMEs) in big data and big data analytics capability.

For this study, the relationship between big data analytics capabilities (BDAC) and organisational performance (OP) was considered. (See Appendix A.) The newly developed survey questionnaire used in the study consists of previously published multi-item constructs with favourable psychometric properties (reliability and validity information) as shown in Table 4.2. All the constructs in the model were measured using a five-point Likert scale (from "strongly disagree" to "strongly agree").

Table 4.2

Constructs and Definitions

Construct and Definition	Reference
Big data analytics capability (BDAC) is broadly defined as the competence to provide organisational insights using data management, infrastructure (technology) and talent (personnel) capability to transform organisations into a competitive force.	Kiron et al., (2012)
BDAC infrastructure capability refers to the ability of the BDAC infrastructure (e.g., applications, hardware, data, and networks) to enable the BDAC staff to develop, deploy, and support necessary system components for an organisation quickly.	Kim et al., (2012)
Big data management capability refers to the BDAC unit's ability to handle routines in a structured (rather than ad hoc) manner to manage IT resources in accordance with business needs and priorities.	Kim et al., (2012)
Big data analytics personnel capability refers to the BDAC staff's professional ability (e.g., skills or knowledge) to undertake assigned tasks.	Kim et al., (2012)
Dynamic capabilities refers to the extent to which an organisation can develop or acquire required competences to make its processes more robust way than its	Kim et al., (2012)

competitors' in terms of coordination, integration, cost reduction, and business intelligence and learning related to BDAC projects.	
Organisational performance (OP) refers to the organisation's ability to gain and retain customers, and to improve sales, profitability, and return on investment.	Mithas et al., (2011)

However, not all questions were taken from existing literature, because of the construct measure with this study. Additional question items were taken from multiple sources as discussed in Table 2.3 Existing Models and Frameworks (Serhani et al., 2016), which is from the existing literature and adapted to the BD and BDAC context (Appendix A). The questions were customised to the context of the study to ensure that they were applicable to the audience, who were individuals working or SMEs in BD or BDAC.

The preliminary survey was subsequently pilot tested (Saunders & Lewis, 2014). Pilot testing was important because this aided in the understanding of the survey, the clarity of the questions, and the ease with which respondents were able to complete the survey. Pilot-testing also assisted in identifying any questions or items that may be unclear or offensive to certain respondents (Pallant, 2007). The pilot study of the survey allowed for the survey to be tested for robustness, effectiveness and if the questions were understood in the context of the study before the final data collection. All questions were measured using a five-point Likert scale. Appendix D is the original pilot study questions which was used to base-line the main study.

The pilot study selected respondents that have 1-20 years working experience with BD or BDAC. Respondents identified were Accenture South Africa employees who are SMEs on BD and BDAC. During the survey pilot-testing, a total of 27 responses were collected from these experts and analysed. All pilot-test responses were omitted from the study's actual sample.

The pilot-test feedback indicated that the majority of the questions were understood with a few questions not being clear enough to the respondents. The average duration to complete the survey was six minutes, which was deemed acceptable.

Recommendations by the experts for the survey were acknowledged and implemented where necessary. Finally, some respondent feedback raised concerns about the use of Likert scales in assessing response accuracy. This feedback was acknowledged, but given the reliability demonstrated by previous researchers, the researcher ultimately chose to maintain the question design. Table 4.3 indicates the changes from the feedback received from the pilot study.

Table 4.3

Changes to the following questions in the survey

Pilot Survey Question	Main Survey Question	Reasons for change
We consider and project about how much these options will help end-users make quicker decisions.	Big data analytics help end-users make quicker decisions.	The survey question confused respondents, and so the researcher used a more direct questioning approach.
We think about and estimate the cost of training that end-users will need.	Big data analytics thinks about the training that end-users will need.	
We consider and estimate the time managers will need to spend overseeing the change.	Big data analytics investment considers change management.	
Please indicate your perceptions regarding the following statements: In my organisation, the responsibility for big data analytics development is clear.	Please indicate your response regarding the following statements: The responsibility for big data analytics development is clear.	

The data collection of the main survey (Appendix A) for this study was undertaken by the researcher who has access to a leading technology organisation (Accenture South

Africa). The data collection was conducted in July 2022. To be more precise, an invitation to participate in the study was sent on to a random sample of 400 people who were using data in some way or form at Accenture South Africa and who were members of the following groups: business analysts, big data analytics, and IT professionals. After a careful analysis of all responses, 239 valid surveys were considered to have been correctly filled out and were used for further analysis, thus giving a response rate of 59.75%.

Table 4.4 shows that of the respondents, 9.2% are aged 18-20 while 14.6% are aged between 21 and 25 years old, while respondents aged between 26 and 35 years old and between 36 and 45 years old represent 58.2% and 16.3%, respectively. It is clear that our sample is dominated by people between the ages of 26 and 35 years. With regard to gender, 58.2% of respondents are men while 41.8% are women. In terms of level of education, the data analysis shows that 1.7% of respondents hold a doctorate, followed by 7.5% with a master’s degree, 23.8% with an honours or postgraduate degree, 47.7% with a bachelor’s degree, 15.9% with a diploma qualification and 3.3% with a matric education. In terms of the number of years working with their organisation, a breakdown of respondents shows that 15.1% have spent from less than two years with their organisation, followed by 50.2% with from 2–5 years with their organisation, 31.4% who have spent 6-10 years, 1.7% who have spent 11-15 years and 1.7% who have spent 16-20 years. Overall, 61.5% of the respondents are in organisations with 2,000 employees or more.

Table 4.4

Demographical Characteristics of Respondents

Variable	Frequency	Percentage
Gender		
Female	100	41.8%
Male	139	58.2%
Age		
18-20	22	9.2%
21-25	35	14.6%
26-35	139	58.2%
36-45	39	16.3%
46-55	4	1.7%
Education		
Matric	8	3.3%
Diploma	38	15.9%
Bachelor	114	47.7%
Honours / postgraduate	57	23.8%
Master's	18	7.5%
Doctorate	4	1.7%
Position in organisation		
BI consultant	19	7.9%
Technical business architect	24	10.0%
Project manager	31	13.0%
Product manager	18	7.5%
Data analytics expert	37	15.5%
Business analyst	16	6.7%
System analyst	15	6.3%
Executive	9	3.8%
Technology specialist	59	24.7%
Other	11	4.6%

How many years have you worked for this specific organisation?		
Less than 2 years	36	15.1%
2-5 years	120	50.2%
6-10 years	75	31.4%
11-15 year	4	1.7%
16-20 years	4	1.7%
Number of employees in your organisation? (Fulltime)		
101-250	19	7.9%
251-500	36	15.1%
501-1000	15	6.3%
1001-2000	22	9.2%
More than 2000	147	61.5%

4.3.5.2 Guidelines on Survey Design

When designing a survey, it is pertinent to include good practices to boost performance and prevent mistakes (Lohr, 2012). If survey questions are poorly constructed, even the greatest data collection, analysis, and display technologies cannot compensate, and the results are worthless. Table 4.5 discusses the six survey design steps taken to design the research questionnaire.

Table 4.5

Survey Design

Step	Description
Step 1 Define survey objectives, use of results and target population (Turcotte, 2010).	1. Define survey objectives and target group 2. Draft survey questions 3. Pilot and re-adjusting the questionnaire 4. Select respondents and the data collection method 5. Run the survey 6. Analyse the results

Step 2 Draft survey questions (Cosenza, 2008).	1. Do the answers to the questions help meet the objectives of the survey with the existing literature? 2. Do the questions address the most bothering issues of the target population? 3. Is the language simple and devoid of technical jargon? 4. Are key terms such as “regulation” clearly defined? 5. Do you avoid asking two questions in one, i.e., do all questions only ask one question at a time? 6. Do you avoid asking two questions in one, i.e., do all questions only ask one question at a time? 7. Are the formulation of questions and answer choices and their order as neutral as possible, i.e., do they avoid suggesting answers? 8. Are the answer choices and scales clearly defined and consistently understood across respondents? Have both been chosen carefully? 9. Does the target population have the capacity and knowledge to answer all questions? 10. Have screening questions been included, that is, has the same question been asked in different ways to identify consistent respondents and meaningful responses? 11. Have tricky questions been included towards the end of the survey when respondents feel more comfortable answering them?

	12. Is the questionnaire short enough to ensure that respondents will concentrate until the end?
Step 3 Pilot and re-adjust questionnaire (Fowler, 2009).	1. Questions are consistently understood across respondents 2. Answers accurately describe what respondents have to say 3. Answers provide valid measures of what the question is designed to measure 4. Respondents have the information needed to answer the questions
Step 4 Select respondents and the data collection method (Turcotte, 2010).	1. Online based surveys 2. Advantages a. Costs are low b. Potential for high-speed returns c. Respondents have time to give thoughtful answers 3. Disadvantages a. Challenge of getting people to reply (depending on people surveyed and topic) b. Respondents are limited to Internet users c. Correct set of email addresses is needed
Step 5 Running the survey (Fowler, 2009).	Running the actual survey is only one of the many steps in the process. Surveys that evaluate or measure awareness of regulatory reform should be timed to take into account the lag between reform implementation and diffusion. To maximise response rates in email surveys, at

	least three follow-up emails to non-respondents are appropriate, and sometimes more. Non-respondents should understand the importance of their answer.
Step 6 Analysing the results (Lohr, 2012).	• Interpret survey data not as facts, but as perceptions • Interpret results together with other data sources • Understand what is behind the results to draw policy conclusions • Take into account the number and the way respondents were selected in the interpretation of the results. Take into account the response rate in the interpretation of the results. If the response rate is too low, no generalisations about the views of the targeted population group can be drawn

Appendix A contains a detailed analysis of each question. To standardise and control responses, a five-point Likert scale was used. Appendix A's first section, titled Demographic Information, included a variety of questions about respondent demographics as well as information about their current employment. This enabled the researcher to provide descriptive information about the sample, ascertain the relevance of respondents (position), and establish a level of diversity. Questions 5 – 42 (see Appendix A) included questions aimed at determining the organisation's current state of BD, BDAC and OP.

The researcher used established measures from existing literature and thus chose to use a five-point Likert scale. Sections big data analytics planning, data analytics investment, big data analytics resources, connectivity, big data analytics capability, system design, technical knowledge, technology management knowledge, organisational knowledge, relation knowledge and organisational performance (see Appendix B). These questions are meant to quantify an organisation's various BD, BDAC and OP, by skillsets, toolsets, and datasets. These questions were adapted from previous research by various researchers as referenced in Appendix B. Additionally, while certain skills are applicable to contexts other than BDAC, such as "analytical applications, including trend analysis, 'what-if' scenarios," the online survey was clearly labelled as a BD, BDAC and OP survey, and respondents were informed via a consent form that the research involved BD, BDAC and OP, in order to ensure that the survey were answered appropriately.

The final sections, and ultimately the investment of BD and BDAC on OP (Appendix A), were used to assess an organisation's level of innovation, proactiveness, and risk-taking in order to determine the impact of BD, BDAC and OP in the South African context.

4.3.6 Research Procedures

Permission to conduct the research was obtained from Accenture South Africa and the University of South Africa (UNISA) CEMS/IOP Research Ethics Review Committee with the reference number, **2021/CEMS/IOP/025**. Each potential respondent received a link containing a cover page indicating the purpose of the study, the reference of ethical clearance and confirmation of the safekeeping and confidentiality of the responses, a consent form explaining that participation in this research was voluntary and the instructions for supplying the socio-demographic information and completing the survey. Each participant submitted the completed survey via the survey link, which was submitted to the researcher. The survey platform used was Lime Survey.

4.3.7 Statistical Analyses

In order to characterise the sample, descriptive statistics such as the mean, median, and standard deviation, and demographic and industry-related data, were analysed (Pallant, 2007). When attempting to examine relationships between variables and predict the result of a dependent variable based on one or more independent variables, correlational techniques are beneficial (Pallant, 2007). This study follows a similar methodology to Miller and Friesen (1982), and Linton and Kask (2017), involving relationships of association between variables (correlations) and multiple regression.

This study attempted to determine the links between several variables, such as BD, BDAC, and OP. Hence, correlation analysis was appropriate since it allowed the researcher to determine not only the strength of these relationships but also whether they were significant. This would answer the secondary study questions that sought to clarify the links between BDAC and OP.

The following statistical techniques were employed:

Step 1:

Item analysis for variable 1 (Cronbach's Alpha)

Step 2:

Exploratory factor analysis for variable 2

Determine factorial structure (Construct validity)

Step 3:

Measurements (AVE and CR – discriminant and convergent validity)

Step 4:

First and Second order (relationships structural equation modelling)

Step 5:

Repeat same process above for variable 2 (organisational performance)

Step 6:

Determine the effect of BDAC on OP

Step 7:

Summary of discriminant validity

Table 4.6 summarises the evaluation of the strength of the correlations based on the interpretation proposed by Pallant (2007).

Table 4.6

Correlation Classifications

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

FORMULATION OF HYPOTHESES

The following hypotheses are based on the theoretical integration and will subsequently be tested in the empirical study and reported on in chapter 5.

H1: There is an association between factors related to big data analytic capability and organisational performance.

H2: Factors related to big data analytic capability do influence organisational performance.

4.4 CHAPTER SUMMARY

The focus of this chapter was to report on and discuss the research methodology used in this study. This chapter discussed the quantitative research approach, followed by a description of the research strategy. The chapter included discussions of the research method pertaining to the research setting, the approach into the organisation, establishing researcher roles, sampling, data collection methods, recording of data, analysis, strategies employed to ensure quality data and ethical considerations that were applied in the study. In the next chapter, the results of the study are discussed.

CHAPTER 5:
RESULTS

5.1 INTRODUCTION

This chapter focuses on the results of the statistical analyses which were performed in order to answer to the hypotheses formulated for this study. The results of the empirical research are reported in form of tables. The results are interpreted and integrated with the literature review. The chapter starts with the item analysis, dimensionality analysis, descriptive statistics, correlations and inferential multivariate techniques.

5.2 ITEM ANALYSIS

5.2.1 Reliability Analysis Output for Big Data Analytic Planning Sub-Scale

As shown in Table 5.1 the three items of big data analytic planning (BDP) subscale depicted an acceptable value of Cronbach's Alpha coefficient of 0.718, which is above the permitted value of 0.70 (Pallant, 2016). The corrected item-total correlation value varies between 0.313 and 0.649, indicating a small to strong association among items (Hair et al., 2019). None of the items were regarded as problematic; therefore, all items were maintained for the next stage of analysis. The big data analytic subscale results are displayed below in Table 5.1.

Table 5.1

Big Data Analytic Planning

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
0.708	0.718	3

Inter-Item Correlation Matrix			
	BDP1	BDP2	BDP3
BDP1	1.000	0.786	0.292
BDP2	0.786	1.000	0.299
BDP3	0.292	0.299	1.000

Item-Total Statistics					
	Scale Mean if Item Deleted	Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
BDP1	8.51	0.604	0.649	0.621	0.458
BDP2	8.60	0.627	0.663	0.622	0.451
BDP3	8.42	0.765	0.313	0.098	0.879

5.2.2 Reliability Analysis Output for Data Analytic Investment Subscale

Table 5.2 below indicates that the four items of data analytic investment (DAI) subscale reflect a permitted value of Cronbach's Alpha coefficient of 0.695, which is above the acceptable value of 0.60 (Wiid & Diggines, 2015). The corrected item-total correlation value ranges from between 0.249 to 0.627, indicating a small to strong association among items (Hair et al., 2019). None of the items were regarded as problematic; therefore, all items were maintained for the next stage of analysis. The Data Analytic Investment subscale results are reflected below in Table 5.2.

Table 5.2

Data Analytic Investment subscale

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
0.693	0.695	4

Inter-Item Correlation Matrix				
	DAI1	**DAI2**	**DAI3**	**DAI4**
DAI1	1.000	0.574	0.515	0.277
DAI2	0.574	1.000	0.492	0.061
DAI3	0.515	0.492	1.000	0.256
DAI4	0.277	0.061	0.256	1.000

Item-Total Statistics					
	Scale Mean if Item Deleted	Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
DAI1	12.73	1.903	0.627	0.433	0.521
DAI2	12.69	2.450	0.504	0.401	0.618
DAI3	12.83	2.148	0.571	0.345	0.568
DAI4	12.96	2.687	0.249	0.121	0.765

5.2.3 Reliability Analysis Output for Data Analytic Resources Subscale

As indicated in Table 5.3 below, four items of data analytic resource (DAR) subscale indicates a permissible value of Cronbach's Alpha coefficient of 0.752, which is above the acceptable value of 0.70 (Pallant, 2016). The corrected item-total correlation value ranges from between 0.396 to 0.732, indicating a small to strong association among items (Hair et al., 2019). All items were acceptable; therefore, all of them were kept for the next stage of analysis. The data analytic resource subscale results are presented below in Table 5.3.

Table 5.3

Data Analytic Resources

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
0.747	0.752	4

Inter-Item Correlation Matrix				
	DAR1	DAR2	DAR3	DAR4
DAR1	1.000	0.770	0.334	0.267
DAR2	0.770	1.000	0.493	0.363
DAR3	0.334	0.493	1.000	0.358
DAR4	0.267	0.363	0.358	1.000

Item-Total Statistics					
	Scale Mean if Item Deleted	Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
DAR1	12.79	1.654	0.577	0.596	0.668
DAR2	12.80	1.587	0.732	0.665	0.584
DAR3	12.63	1.788	0.490	0.285	0.717
DAR4	12.54	1.855	0.396	0.174	0.770

5.2.4 Reliability Analysis Output of Model for Connectivity

As indicated in Table 5.4 below, four items for model of connectivity (CON) subscale have a value of Cronbach's Alpha coefficient of 0.718, which is above the cut-off of 0.60 (Wiid & Diggines 2015). The corrected item-total correlation value ranges between 0.426 and 0.536 indicating a medium to strong association among items (Hair et al., 2019). None of the items were regarded as problematic; therefore, all items were retained for the next stage of analysis. The model for connectivity subscale results are presented below in Table 5.4.

Table 5.4

Model for Connectivity

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
0.713	0.718	4

Inter-Item Correlation Matrix				
	CON1	CON2	CON3	CON4
CON1	1.000	0.722	0.141	0.355
CON2	0.722	1.000	0.276	0.287
CON3	0.141	0.276	1.000	0.554
CON4	0.355	0.287	0.554	1.000

Item-Total Statistics					
	Scale Mean if Item Deleted	Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
CON1	12.46	3.073	0.523	0.572	0.643
CON2	12.59	2.772	0.536	0.558	0.628
CON3	12.74	2.884	0.426	0.362	0.698
CON4	12.63	2.663	0.531	0.393	0.631

5.2.5 Reliability Analysis Output of Model for System Design

Table 5.4 below presents four items of model for system design (SYST) subscale having a value of Cronbach's Alpha coefficient of 0.771, which is above the cut-off of 0.70 (Pallant, 2016). The corrected item-total correlation value ranges between 0.369 and 0.681, indicating a medium to strong association among items (Hair et al., 2019). None of the items were viewed as problematic; therefore, all items were considered for the next stage of analysis. The model for system design subscale results are presented below in Table 5.4.

Table 5.5

Model for System Design

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
0.762	0.771	4

Inter-Item Correlation Matrix				
	SYST1	SYST2	SYST3	SYST4
SYST1	1.000	0.844	0.450	0.245
SYST2	0.844	1.000	0.484	0.262
SYST3	0.450	0.484	1.000	0.458
SYST4	0.245	0.262	0.458	1.000

Item-Total Statistics					
	Scale Mean if Item Deleted	Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
SYST1	12.34	2.973	0.653	0.715	0.655
SYST2	12.30	2.926	0.681	0.727	0.639
SYST3	12.40	3.342	0.587	0.355	0.697
SYST4	12.24	3.376	0.369	0.212	0.818

5.2.6 Reliability Analysis Output for Model for Technological Management of Knowledge

Seven items of model for technological management knowledge (TMK) subscale had a value of Cronbach's Alpha coefficient of 0.810, which is above the cut-off of 0.70 (Pallant, 2016) as shown in Table 5.5 below. The corrected item-total correlation value ranges from 0.404 and 0.740 indicating a medium to strong association among items (Hair et al., 2019). None of the items were regarded as problematic; therefore, all items were retained for the next stage of analysis. The model for technological management knowledge subscale results are presented below in Table 5.5.

Table 5.6

Technological Management Knowledge

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
0.813	0.810	7

Inter-Item Correlation Matrix							
	TECH1	TECH2	TECH3	TMK_1	TMK_2	TMK_3	TMK_4
TECH1	1,000	0,859	0,218	0,243	0,335	0,418	0,562
TECH2	0,859	1,000	0,358	0,248	0,362	0,317	0,667
TECH3	0,218	0,358	1,000	0,302	0,407	0,357	0,437
TMK_1	0,243	0,248	0,302	1,000	0,890	0,171	0,002
TMK_2	0,335	0,362	0,407	0,890	1,000	0,189	0,122
TMK_3	0,418	0,317	0,357	0,171	0,189	1,000	0,486
TMK_4	0,562	0,667	0,437	0,002	0,122	0,486	1,000

Item-Total Statistics					
	Scale Mean if Item Deleted	Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
TECH1	25,57	5,078	0,675	0,786	0,765
TECH2	25,51	5,343	0,740	0,823	0,751
TECH3	25,79	6,441	0,477	0,391	0,800
TMK_1	25,69	6,694	0,404	0,809	0,810
TMK_2	25,71	6,349	0,514	0,827	0,794
TMK_3	25,50	6,226	0,466	0,387	0,802
TMK_4	25,56	5,878	0.582	0,605	0,782

5.2.7 Reliability Analysis Output for Model for Relational Knowledge

Five items of model for relational knowledge (REK) subscale had a value of Cronbach's Alpha coefficient of 0.808, which is above the cut-off of 0.70 (Pallant, 2016) as shown in Table 5.7 below. The corrected item-total correlation value ranges from 0.434 to 0.682, indicating a medium to strong association among items (Hair et al., 2019). None of the items were regarded as problematic; therefore, all items were retained for the next stage of analysis. The model for technological management knowledge subscale results are presented below in Table 5.7.

Table 5.7

Model for Relational Knowledge

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
0.810	0.808	5

Inter-Item Correlation Matrix					
	ORK1	ORK2	REK1	REK2	REK3
ORK1	1,000	0,816	0,391	0,410	0,399
ORK2	0,816	1,000	0,322	0,350	0,337
REK1	0,391	0,322	1,000	0,891	0,271
REK2	0,410	0,350	0,891	1,000	0,390
REK3	0,399	0,337	0,271	0,390	1,000

Item-Total Statistics					
	Scale Mean if Item Deleted	Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
ORK1	17,29	2,468	0,665	0,695	0,753
ORK2	17,19	2,568	0,589	0,666	0,776
REK1	17,26	2,495	0,620	0,803	0,766
REK2	17,24	2,392	0,682	0,818	0,746
REK3	17,29	2,872	0,434	0,255	0,819

5.2.8 Reliability Analysis Output for Model for Organisational Performance

As indicated in Table 5.8 below, nine items of model for organisational performance (ORP) subscale present a value of Cronbach's Alpha coefficient of 0.921, which is above the cut-off of 0.70 (Pallant, 2016). The corrected item-total correlation value ranges between 0.536 and 0.795, indicating a strong association among items (Hair et al., 2019). None of the items were considered as problematic; therefore, all items were considered for the next stage of analysis. The model for organisational performance subscale results are presented below in Table 5.8.

Table 5.8

Model for Organisational Performance

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
0.921	0.921	9

Inter-Item Correlation Matrix									
	ORP1	ORP2	ORP3	ORP4	ORP5	ORP6	ORP7	ORP8	ORP9
ORP1	1.000	0.835	0.526	0.442	0.537	0.645	0.645	0.534	0.595
ORP2	0.835	1.000	0.517	0.483	0.575	0.596	0.589	0.487	0.543
ORP3	0.526	0.517	1.000	0.447	0.484	0.394	0.347	0.264	0.450
ORP4	0.442	0.483	0.447	1.000	0.737	0.571	0.555	0.601	0.606
ORP5	0.537	0.575	0.484	0.737	1.000	0.644	0.621	0.501	0.643
ORP6	0.645	0.596	0.394	0.571	0.644	1.000	0.966	0.504	0.618
ORP7	0.645	0.589	0.347	0.555	0.621	0.966	1.000	0.560	0.668

ORP8	0.534	0.487	0.264	0.601	0.501	0.504	0.560	1.000	0.543
ORP9	0.595	0.543	0.450	0.606	0.643	0.618	0.668	0.543	1.000

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
ORP1	33.05	18.897	0.766	0.766	0.909
ORP2	33.12	18.180	0.737	0.724	0.911
ORP3	33.34	20.318	0.536	0.407	0.923
ORP4	33.32	19.172	0.705	0.654	0.912
ORP5	33.36	18.290	0.758	0.660	0.909
ORP6	33.30	18.811	0.795	0.947	0.907
ORP7	33.34	18.587	0.794	0.951	0.906
ORP8	33.43	19.892	0.628	0.524	0.917
ORP9	33.38	19.029	0.744	0.619	0.910

5.3 DIMENSIONALITY ANALYSIS

5.3.1 Dimensionality Output for Big Data Analytics Planning

The big data analytic planning (BDP) scales depicted a KMO index value of 0.565 and Bartlett's test of sphericity value of 251.069 (df = 3, p < 0.000). This indicates that unidimensional factor analysis can be conducted. One factor with an eigenvalue greater than 1 was obtained. The factor BDP obtained an eigenvalue of 1.966, which accounted for 65.543% of the variance. The factor loadings were all above 0.50, as shown in Table 5.9.

Table 5.9

KMO and Bartlett's test for factor matrix of Big Data Analytic Planning

KMO and Bartlett's test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0,565
Bartlett's Test of Sphericity	Approx. Chi-Square	251,069
	Df	3
	Sig.	0.000
Factor matrix		
Items		**Factor**
BDP1		0,909
BDP2		0,911
BDP3		0,557
Eigenvalue		1,966
% of variance		65,543%
Cumulative %		65,543%

5.3.2 Dimensionality Output for Data Analytic Investment

The data analytic investments scales (DAI) depicted a KMO index value of 0.670 and Bartlett's test of sphericity value of 216.787 (df = 6, p < 0.000). This shows that factor analysis can be conducted. One factor (DAI) with an eigenvalue above the cut-off of 1 was obtained. The factor obtained an eigenvalue of 2.158, which accounted for 53.94.% of the variance. The factor loadings were all above 0.30, as shown in Table 5.10.

Table 5.10

KMO and Bartlett's test for factor of Data Analytic Investment

KMO and Bartlett's test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0,670
Bartlett's Test of Sphericity	Approx. Chi-Square	216,787
	Df	6
	Sig.	0.000
Factor matrix		
Items		**Factor**
DAI2		0,784
DAI1		0,848
DAI3		0,804
DAI4		0,422
Eigenvalue		2,158
% of variance		53,944%
Cumulative %		53,944%

5.3.3 Dimensionality Output for Data Analytic Resources

The data analytic investments scales (DAR) obtained a KMO index value of 0.637 and Bartlett's test of sphericity value of 324.583 (df = 6, p < 0.000). This depicts that factor analysis can be conducted. The DAR scale was found to be unidimensional. One factor with an eigenvalue greater than 1 was obtained, with an eigenvalue of 2.328, which accounted for 58.20% of the variance. The factor loadings were all above 0.50, as shown in Table 5.11.

Table 5.11

KMO and Bartlett's test for factor of Data Analytic Resources

KMO and Bartlett's test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.637
Bartlett's Test of Sphericity	Approx. Chi-Square	324.583
	Df	6
	Sig.	0.000
Factor matrix		
Items		**Factor**
DAR1		0,818
DAR2		0,899
DAR3		0,701
DAR4		0,599
Eigenvalue		2,328
% of variance		58.200%
Cumulative %		58.200%

5.3.4 Dimensionality Output of Model for Connectivity

The model for connectivity scales (CON) depicted a KMO index value of 0.517 and Bartlett's test of sphericity value of 311.739 (df = 6, p < 0.000). This shows that factor analysis can be conducted. The scale was found to be unidimensional. One factor with an eigenvalue greater than 1 was obtained. The factor (CON) obtained an eigenvalue of 1.827, which accounted for 54.45% of the variances. The factor loadings were all above 0.50, as shown in Table 5.12.

Table 5.12

KMO and Bartlett's test for factor of model for Connectivity

KMO and Bartlett's test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.517
Bartlett's Test of Sphericity	Approx. Chi-Square	311.739
	Df	6
	Sig.	0.000
Factor matrix		
Items		**Factor**
CON2		0,804
CON1		0,786
CON4		0,725
CON3		0,623
Eigenvalue		2,178
% of variance		54,450%
Cumulative %		54,450%

5.3.5 Dimensionality Output of Model for System Design

The model for system design scales (SYS) depicted a KMO index value of 0.640 and Bartlett's test of sphericity value of 451.339 (df = 6, p < 0.000). This reveals that unidimensional factor analysis can be conducted. One factor with an eigenvalue greater than 1 was obtained, with an eigenvalue of 2.412. The factor accounted for 60.31% of the variance. The factor loadings were all above 0.50, as shown in Table 5.13.

Table 5.13

KMO and Bartlett's test for factor of model for System Design

KMO and Bartlett's test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.640
Bartlett's Test of Sphericity	Approx. Chi-Square	415.339
	Df	6
	Sig.	0.000
Factor matrix		
Items		**Factor**
SYS2		0,881
SYS1		0,865
SYS3		0,758
SYS4		0,650
Eigenvalue		2,412
% of variance		60,312%
Cumulative %		60, 312%

5.3.6 Dimensionality Output of Model for Technology Management Knowledge

The model for technology management knowledge scales (TMK) attained a KMO index value of 0.630 and Bartlett's test of sphericity value of 1093.826 (df = 21, p < 0.000). The findings indicate that factor analysis can be conducted. The TMK scale was found to be unidimensional. One factor with an eigenvalue greater than 1 was obtained. The factor (TMK) indicated an eigenvalue of 3.318, which accounted for 47.404% of the variance. The factor loadings were all above 0.30, as shown in Table 5.14.

Table 5.14

KMO and Bartlett's test for factor of model for Technology Management Knowledge

KMO and Bartlett's test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0,630
Bartlett's Test of Sphericity	Approx. Chi-Square	1093,826
	Df	21
	Sig.	0.000
Factor matrix		
Items		**Factor**
TECH2		0,830
TECH1		0.801
TMK4		0.712
TMK2		0.652
TECH3		0.618
TMK3		0,597
TMK1		0,549
Eigenvalue		3,318
% of variance		47,404%
Cumulative %		47,404%

5.3.7 Dimensionality Output of Model for Relational Knowledge

The relational knowledge scales (REK) depicted a KMO index value of 0.565 and Bartlett's test of sphericity value of 442.842 (*df* = 6, *p* < 0.000). This shows that factor analysis can be conducted. The REK scale was found to be unidimensional. One factor with an eigenvalue greater than 1 was obtained. The factor (REK) reflected an eigenvalue of 2.252. The factor accounted for 56.30% of the variance. The factor loadings were all above 0.50, as shown in Table 5.15.

Table 5.15

KMO and Bartlett's test for factor of model for Relational Knowledge

KMO and Bartlett's test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0,565
Bartlett's Test of Sphericity	Approx. Chi-Square	442,842
	Df	6
	Sig.	0.000
Factor matrix		
Items		**Factor**
REK2		0,931
REK1		0,896
REK3		0,570
REK4		0,506
Eigenvalue		2,252
% of variance		56,300%
Cumulative %		56,300%

5.3.8 Dimensionality Output for Organisational Performance

The organisational performance scales (ORP) depicted a KMO index value of 0.825 and Bartlett's test of sphericity value of 1891.637 (df = 36, p < 0.000). This shows that factor analysis can be conducted. The ORP scale was found to be unidimensional. One factor (ORP) with an eigenvalue greater than 1 was obtained. The factor, ORP, obtained an eigenvalue of 5.50. The factor accounted for 61.671% of the variance. The factor loadings were all above 0.50, as shown in Table 5.16.

Table 5.16

KMO and Bartlett's test for factor of Organisational Performance

KMO and Bartlett's test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0,825
Bartlett's Test of Sphericity	Approx. Chi-Square	1891,637
	Df	36
	Sig.	0.000
Factor matrix		
Items		**Factor**
ORP7		0,858
ORP6		0,855
ORP1		0,819
ORP5		0,816
ORP9		0,807
ORP2		0,799
ORP4		0,769
ORP8		0,705
ORP3		0,608
Eigenvalue		5,550
% of variance		61,671%
Cumulative %		61,671%

5.3.9 Summary of the Dimensionality Output for the BDP, DAI, DAR, CON, SYS, TECH, REK and ORP

To assess the correctness of data for conducting a confirmatory factor analysis, the exploratory factor analysis (EFA) using principal component Analysis was performed. KMO and Bartlett's test for sphericity were employed to attest the sample's adequacy. All KMO values for the eight constructs were in permissible level, and their corresponding Bartlett's values were statistically significant at $p < 0.001$. The EFA findings indicated that a sample size of 279 as used in the present study was suitable to conduct the EFA. In the same vein, confirmatory factor analysis (CFA) was performed on the eight constructs to find components prior to the validation of the measurement. The CFA will be presented in the next section.

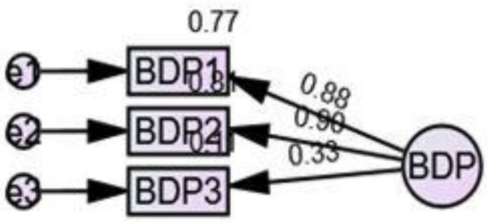
5.4 MEASUREMENT MODEL FOR THE BIG DATA CAPABILITY

5.4.1 Measurement Model for Big Data Analytic Planning

The four-item-model of the big data analytic planning construct suggested a good fit in the initial estimate model as shown in Figure 9 below. The indices (CMIN, DF, P, CMIN/DF, NFI, RFI, IFI, TLI, CFI, RMSEA, AVE and CR) were within the permissible levels (Hair et al., 2019). Since all the essential ratios were over 1.96. In this view, the most endorsed and supported indices, which are relative Chi square, CFI and RMSEA, showed that the model had a good fit (see Figure 5.1) (Kline, 2016).

Figure 5.1

Big Data Analytic Planning



Initial Model:

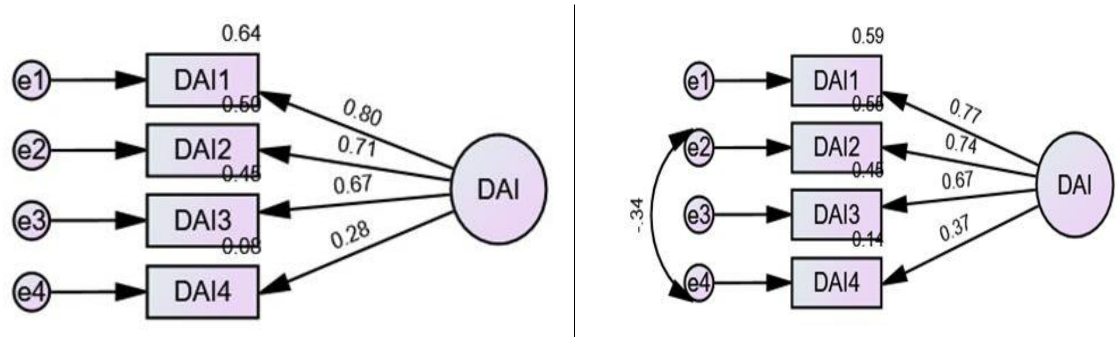
CMIN = 0	DF = 0
CMIN/DF = 0	p = 0.01
NFI = 1	RFI = 0.97
IFI = 0.98	TLI = 1
CFI = 1	RMSEA = 0.07
AVE = 0.655	CR = 0.85

5.4.2 Measurement Model for Data Analytic Investment

All four items of data analytic investment's component suggested a poor fit in the initial estimate model. The indices of CMIN, DF, P, CMIN/DF, NFI, RFI, IFI, TLI, CFI and RMSEA, AVE and CR below showed an unacceptable level (Hair et al., 2019) (see Figure 5.2 below). After the modification, all four items of the data analytic investment's component reflected a good fit model in the initial estimate model, as all the estimated paths in the modified model were acceptable with critical ratios above 1.96. In this view, the most recommended indices, which are relative Chi square, CFI and RMSEA, showed that the model had a good fit (Hair et al., 2019).

Figure 5.2

Data Analytic Investments



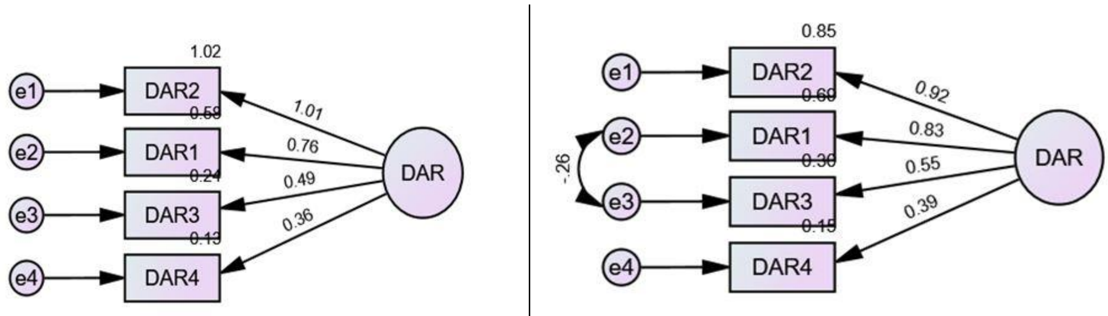
Initial Model:		**Modified Model:**	
CMIN = 16.147	DF = 2	CMIN = 0.079	DF = 1
CMIN/DF = 8.073	P = 0.00	CMIN/DF = 0.08	P = 0.78
NFI = 0.93	RFI = 0.78	NFI = 1	RFI = 1
IFI = 0.94	TLI = 0.80	IFI = 1	TLI = 1
CFI = 0.93	RMSEA = 0.17	CFI = 1	RMSEA = 0
AVE = 0.540	CR = 0.816	AVE = 0.540	CR = 0.816

5.4.3 Measurement Model for Big Data Analytics Resources

The items of data analytic resources' component displayed a poor fit in the initial estimate model. The indices of CMIN, DF, P, CMIN/DF, NFI, RFI, IFI, TLI, CFI and RMSEA, AVE and CR below showed an unacceptable level (Hair et al., 2019) (see Figure 5.3 below). After the modification, all four items of the data analytic resources' component indicated a good fit model in the initial estimate model as all the estimated paths in the modified model were acceptable with critical ratios above 1.96. In this view, the most recommended indices, which are relative Chi square, CFI and RMSEA, showed that the model had a good fit (Hair et al., 2019).

Figure 5.3

Big Data Analytic Resources



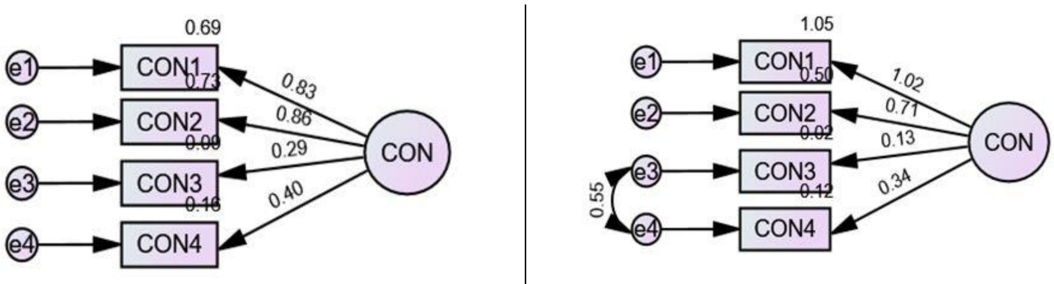
Initial Model:		Modified Model:	
CMIN =13.377	DF = 2	CMIN =9.846	DF = 1
CMIN/DF =6.688	p = 0.001	CMIN/DF =9.846	p = 0.002
NFI = 0.96	RFI=0.88	NFI = 0.97	RFI=0.82
IFI = 0.97	TLI = 0.89	IFI = 0.97	TLI = 0.84
CFI =0.97	RMSEA =0.16	CFI =0.97	RMSEA =0.19
AVE =0.542	CR=0.845	AVE =0.542	CR=0.845

5.4.4 Measurement Model for Connectivity

The four items of the model for connectivity's component indicated a poor fit model in the initial estimate model (see Figure 5.4 below). The indices of CMIN, DF, P, CMIN/DF, NFI, RFI, IFI, TLI, CFI, RMSEA, AVE and CR below showed an unacceptable level (Kline, 2016). After the modification of the initial model, the indices of the second model showed improvement (see Figure 5.4 below), as all the estimated paths displayed in the modified model were significant, with critical ratios above 1.96. In this view, the most recommended indices, which are relative Chi square, CFI and RMSEA, indicated a good fit model (Hair et al., 2019).

Figure 5.4

Connectivity



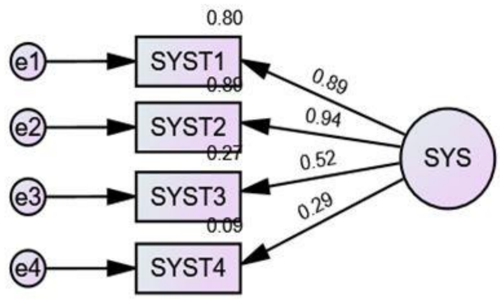
Initial Model:		Modified Model:	
CMIN =92.054	DF = 2	CMIN =18.896	DF = 1
CMIN/DF =46.027	p = 0.00	CMIN/DF =18.896	p = 0.00
NFI =0.71	RFI=0.12	NFI = 0.94	RFI=0.64
IFI =0.71	TLI =0.13	IFI = 0.94	TLI = 0.65
CFI =0.71	RMSEA =0.44	CFI =0.94	RMSEA =0.27
AVE =0.582	CR=0.826	AVE =0.582	CR=0.826

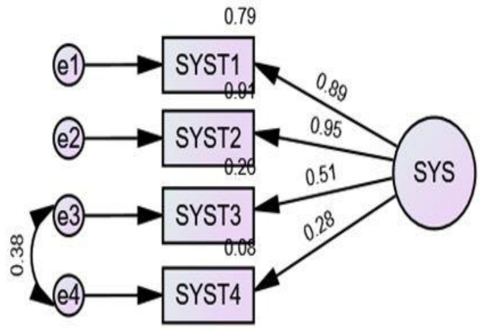
5.4.5 Measurement Model for System Design

The four items of the model for system design's component indicated a poor fit in the initial estimate model. The indices of CMIN, DF, P, CMIN/DF, NFI, RFI, IFI, TLI, CFI and RMSEA, AVE and CR below suggested an unacceptable level (Hair et al., 2019) (see Figure 5.5 below). After the modification of the initial model, the indices of the second model showed improvement (see Figure 5.5 below), as all the estimated paths in the modified model were significant with critical ratios above 1.96. In this view, the most recommended indices, which are relative Chi square, CFI and RMSEA, indicated a good fit model (Hair et al., 2019).

Figure 5.5

Model for System Design





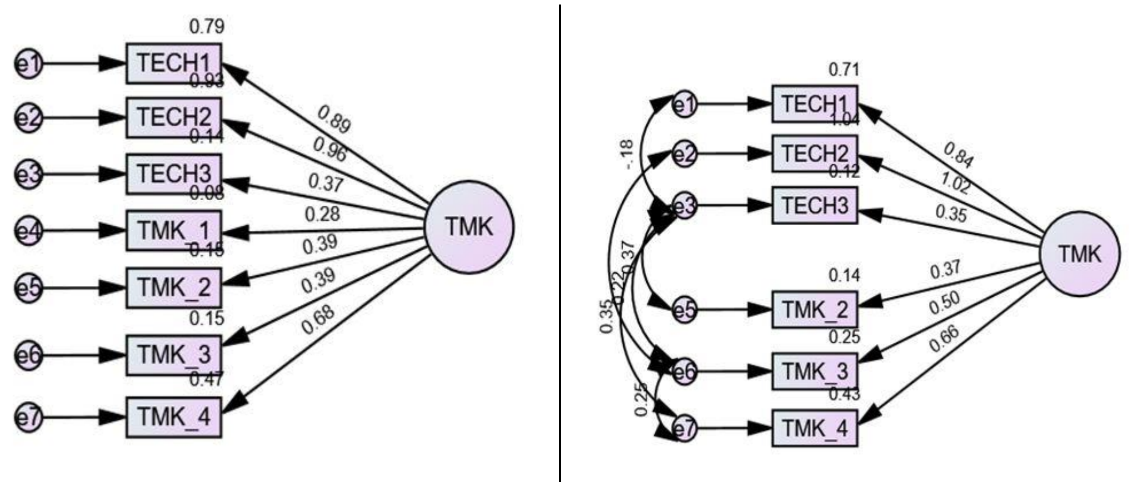
Initial Model:		**Modified Model:**	
CMIN =36.913	DF = 2	CMIN =0.003	DF = 1
CMIN/DF =18.457	p = 0.00	CMIN/DF =0.003	p = 0.95
NFI =0.91	RFI=0.74	NFI = 1	RFI=1
IFI =0.92	TLI =0.75	IFI = 1	TLI = 1
CFI =0.92	RMSEA =0.27	CFI =1	RMSEA =0
AVE =0.631	CR=0.871	AVE =0.631	CR=0.871

5.4.6 Measurement Model for Technology Management Knowledge

The items of model for technology management knowledge's component suggested a poor fit in the initial estimate model. The indices of CMIN, DF, P, CMIN/DF, NFI, RFI, IFI, TLI, CFI and RMSEA, AVE and CR below showed an unacceptable level (Kline, 2016) (see Figure 5.6 below). After the modification of the initial model, the indices of the second model showed improvement (Figure 5.6 below) as all the estimated paths in the modified model were significant with critical ratios above 1.96. In this view, the most recommended and supported indices, which are relative Chi square, CFI and RMSEA showed that the model had a good fit (Hair et al., 2019).

Figure 5.6

Technology Management Knowledge



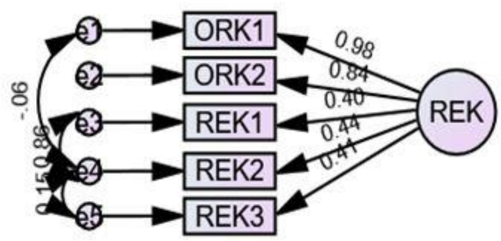
Initial Model:		Modified Model:	
CMIN =536.909	DF = 14	CMIN =8.507	DF = 3
CMIN/DF =38.357	p = 0.00	CMIN/DF =2.836	p = 0.037
NFI =0.52	RFI=0.27	NFI =0.99	RFI=0.94
IFI =0.52	TLI =0.28	IFI =0.99	TLI = 0.96
CFI =0.52	RMSEA =0.40	CFI =0.99	RMSEA =0.09
AVE =0.472	CR=0.860	AVE =0.472	CR=0.860

5.4.7 Measurement Model for Relational Knowledge

The five items of the relational knowledge's component indicated a poor fit in the initial estimate model. The indices of CMIN, DF, P, CMIN/DF, NFI, RFI, IFI, TLI, CFI and RMSEA, AVE and CR below suggested an unacceptable level (Kline, 2016) (see Figure 5.7 below). After the modification of the initial model, the indices of the second model showed improvement (see Figure 5.7 below), as all the estimated paths in the modified model were significant with critical ratios above 1.96. In this view, the most recommended and supported indices, which are relative Chi square, CFI and RMSEA, showed that the model had a good fit (Hair et al., 2019).

Figure 5.7

Relational Knowledge



Initial Model:

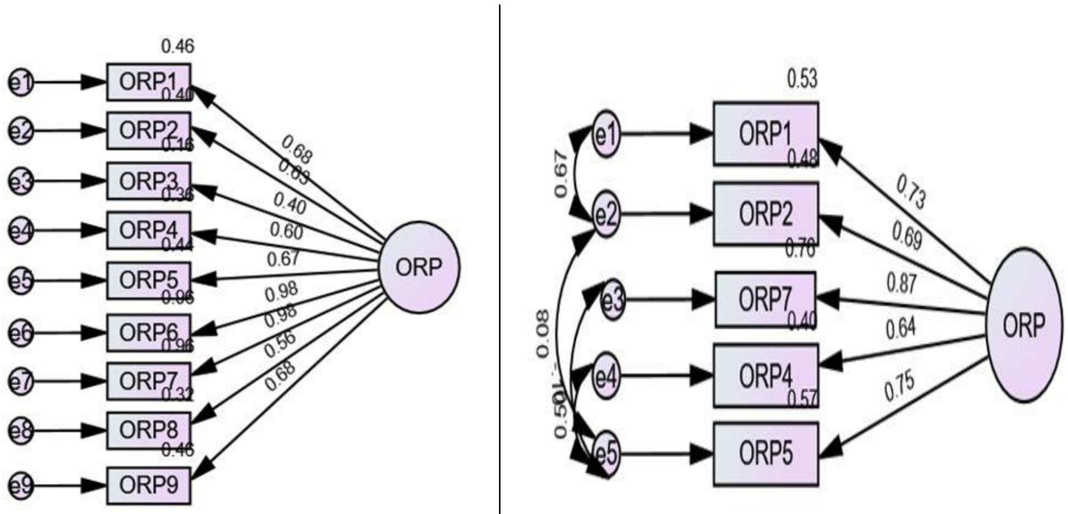
CMIN =9.246	DF = 2
CMIN/DF =4.623	p = 0.01
NFI =0.98	RFI=0.94
IFI =0.98	TLI =0.95
CFI =0.98	RMSEA =0.12
AVE =0.563	CR=0.828

5.4.8 Measurement Model for Organisational Performance

The items of the model for the component of organisational performance indicated a poor fit model in the initial estimate model (see Figure 5.8 below). The indices of CMIN, DF, P, CMIN/DF, NFI, RFI, IFI, TLI, CFI, RMSEA, AVE and CR below showed an unacceptable level (Kline, 2016). After the modification of the initial model, the indices of all the estimated paths in the final model were significant, with critical ratios above 1.96. In this view, the most recommended and supported indices, which are relative Chi square, CFI and RMSEA, showed that the model had a good fit (Hair et al., 2019) (see Figure 5.8 below).

Figure 5.8

Organisational Performance



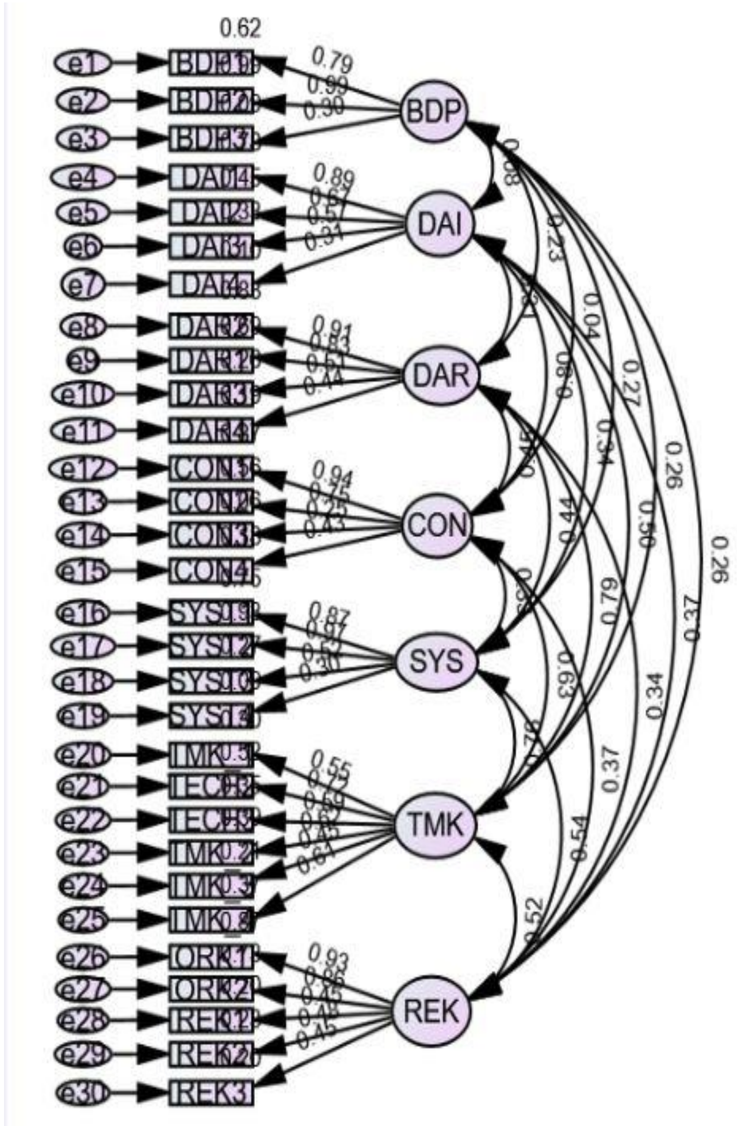
Initial Model:		Modified Model:	
CMIN =555.504	DF = 27	CMIN =6.884	DF = 1
CMIN/DF =20.574	p = 0.00	CMIN/DF =6.884	p = 0.01
NFI =0.71	RFI=0.52	NFI =0.99	RFI=0.86
IFI =0.72	TLI =0.53	IFI =0.99	TLI = 0.88
CFI =0.72	RMSEA =0.29	CFI =0.99	RMSEA =0.126
AVE =0.617	CR=0.935	AVE =0.617	CR=0.935

5.5 INITIAL AND FINAL MEASUREMENT MODEL FOR BIG DATA CAPABILITY

The seven essential constructs of the big data capability were linked in order to validate a single measurement model of big data capability comprising all seven constructs of big data capability. In the initial estimate model, the seven constructs of big data capability indicated an unfitting model (see Figure 5.9 below) with the indices of CMIN, DF, P, CMIN/DF, NFI, RFI, IFI, TLI, CFI, RMSEA, AVE and CR below an acceptable level. After the modification of the initial model, the indices of the second model suggested an improvement of the model (see Figure 5.9 below), following suggestions related to the acceptance of model fit made by Hair et al. (2019), as all the estimated paths displayed in the modified model were significant, with critical ratios above 1.96.

Figure 5.9

Big Data Capability



Initial Model:

CMIN =4390.396

DF = 384

CMIN/DF =11.433

p = 0.00

NFI =0.38

RFI=0.29

IFI =0.40

TLI =0.31

CFI =0.39

GFI= 0.51

RMSEA =0.21

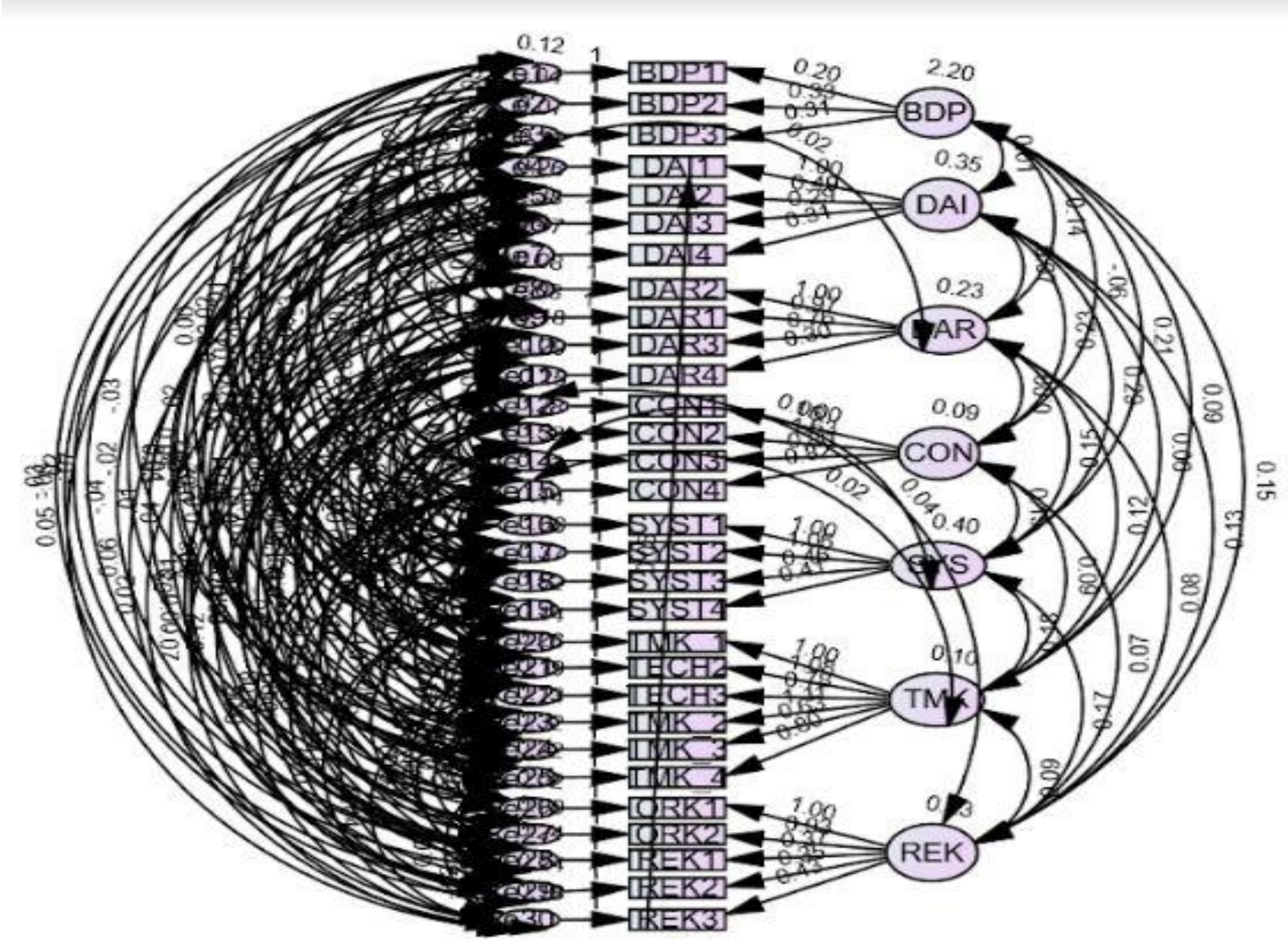
5.6 FINAL MEASUREMENT MODEL FOR BIG DATA ANALYTIC CAPABILITY

The seven essential constructs of the big data analytic capability (big data planning, data analytic investment, data analytic resources, connectivity, system design, technology knowledge management and relational knowledge) were combined in order to validate a single measurement model of big data analytic capability comprising all seven constructs of big data analytic capability. In the initial estimate model, the seven constructs of big data capability indicated an unfitting model (see Figure 5.9 above), with the indices of CMIN, DF, P, CMIN/DF, NFI, RFI, IFI, TLI, CFI and RMSEA below an acceptable level (Hair et al., 2019).

Figure 5.10 below displays the modified model of big data analytic capability, with a permissible level of good fit indices model.

Figure 5.10

Big Data Analytic Capability fitted model



Initial Model:

CMIN =851.323

DF = 199

CMIN/DF =4.28

p = 0.00

NFI =0.88

RFI=0.74

IFI =0.91

TLI =0.78

CFI =0.90

GFI= 0.83

RMSEA =0.18

As displayed in Figure 5.9 and Figure 5.10, confirmatory factor analysis (CFA) was used in the validation process of big data analytic capability scale, based on seven measurements comprising constructs such as big data planning, data analytic investment, data analytic resources, connectivity, system design, technology knowledge management and relational knowledge. The first model (see Figure 5.9) showed poor-fit indices of CMIN, DF, CMIN/DF, NFI, RFI, IFI, TLI, CFI, GFI and RMSEA that were below the acceptable level (Hair et al., 2019). However, after the modification of the first model, the second modified model which incorporates the aforementioned seven constructs (see Figure 5.10) presents an acceptable level of good-fit indices that validate big data analytic capability scale (measurement) (Hair et al., 2019). Table 5.17 illustrates the acceptable level of indices of goodness of fit that validated the big data analytic capability scale.

Table 5.17

Accepted value of good fit for Big Data Analytic Capability Scale

Indices of Good-Fit Model	Threshold of Good Indices	Accepted Value	Decisions
Chi-square degrees of freedom (d) (CMIN/DF)	≤ 5.0 to 2.0 Depending on the sample size and the number of available parameters, (Hair et al. 2019).	4.278	Accepted
P-value	< 0.05. (Hair et al., 2019)	0.000	Accepted
Bentler-Bonett Normed Fit Index NFI	Hu and Bentler (1999) propose a limit of NFI between 0 and 1	0.88	Accepted
Goodness-of-Fit Index (GFI)	From 0 to 1. But 0.90 is also acceptable (Kline, 2016).	0.83	Accepted
Tucker-Lewis Index TLI	> 0.9. (Hair et al., 2019)	0.78	Partially accepted
Comparative Fit Index CFI	≥ 0.9 to 1	0.91	Accepted
Root Mean Square Error of Approximation (RMSEA)	0.05 ≤ 0.08. (Kline, 2016)	0.18	Accepted
Root Mean Square Residual RMR	≤ 0.08. (Kline, 2016)	0.05	Accepted

Note: Author’s own data about acceptable model fit.

5.7 SUMMARY OF FINDINGS RELATED TO THE VALIDATION OF BIG DATA ANALYTIC CAPABILITY

Given the main objective of the study, that is to validate a big data analytic capability scale, the findings of the current study suggest the good-fit-indices that are featured the second model confirm the big data analytic capability scale that integrates constructs such as big data planning, data analytic investment, data analytic resources, connectivity, system design, technology knowledge management and relational knowledge in the context of South Africa (see Figure 5.10). This denotes that a well-designed system of technology knowledge management, with adequate data analytic planning, data analytic investment and data analytic resources, contributes to the performance of a given organisation (Kim et al., 2012; Chen et al., 2015). Table 5.17 above provides the accepted good-fit-indices validating the big data analytic capability scale (Olivier & Martins, 2018).

5.8 CORRELATIONAL ANALYSIS

As observed in Table 5.18 below, the Pearson product-moment correlation was performed to establish the relationship between big data analytic capability variables and organisational performance. In Table 5.18 a significant correlations were observed between big data analytic capability factors and organisational performance, with $p < 0.05$ determining a statistical significance and a Pearson's correlation coefficient ranging from small to large effect ($r = -0.10$; small effect; $p \leq 0.05$ to $r = 0.73$; large effect; $p \leq 0.05$). Table 29 below indicates the relationship between factors related to big data analytic capability and organisational performance.

Table 5.18

Correlation between Big Data Analytic Capability and Organisational Performance

Variables	Big data planning	Data analytic investment	Data analytic resources	Connectivity	System Design	Technology knowledge management	Relational knowledge	Organisational performance
Big data planning	1	0.18^{*}	0.29^{*}	0.03	0.28^{*}	0.20^{*}	0.15^{*}	-0.10^{*}
Data analytic investment		1	0.38^{**}	0.55^{***}	0.49^{**}	0.40^{**}	0.24^{*}	0.16^{*}
Data analytic resources			1	0.61^{***}	0.60^{***}	0.73^{***}	0.39^{**}	0.42^{**}
Connectivity				1	0.65^{***}	0.66^{***}	0.35^{**}	0.55^{***}
System design					1	0.73^{***}	0.47^{**}	0.58^{***}
Technology knowledge management						1	0.52^{***}	0.50^{***}
Relational knowledge							1	0.19^{*}
Organisational performance								1

Note. N = 239. ***p ≤ .001 **p ≤. 01 * p ≤.05. + r ≥ .29 (small effect); ++ r ≥ .30 ≥ r ≤ .49 (medium effect); +++ r ≥ .50 (large effect).

**. Correlation is significant at the 0.01 level. Source: Author's own data.

Table 5.18 shows that big data planning positively correlated with data analytic investment (*r* = 0.18; small effect; *p* ≤ 0.05), data analytic resources (*r* = 0.29; small effect; p ≤ 0.05), connectivity (r = 0.03; small effect; p ≤ 0.05), system design (r = 0.28; small effect; p ≤ 0.05), technology knowledge management (*r* = 0.20; small effect; *p* ≤ 0.05), and relational knowledge (r=0.15; small effect; p ≤ 0.05). A negative correlation was observed between big data planning and organisational performance (*r* = -0.10; small effect; *p* ≤ 0.05).

The findings in Table 5.18 indicates a positive correlation between data analytic investment and data analytic resources (r=0.38; medium effect; *p* ≤ 0.05), connectivity (*r* = 0.55; large effect; *p* ≤ 0.05), system design (r = 0.49; medium effect; p ≤ 0.05), technology management knowledge (*r* = 0.40; medium effect; *p* ≤ 0.05), relational knowledge (*r* = 0.24; small effect; *p* ≤ 0.05), and organisational performance (*r* = 0.16; small effect; *p* ≤ 0.05).

As Table 5.18 above depicts, a relationship is observed between data analytic resources and connectivity (*r* = 0.605; large effect; *p* ≤ 0.05), system design (*r* = 602; large effect; *p* ≤ 0.05), technology management knowledge (*r* = 0.733; large effect; *p* ≤ 0.05), relational knowledge (*r* = 0.393; large effect; *p* ≤ 0.05), and organisational performance (*r* = 0.420; medium effect; *p* ≤ 0.05).

The set of findings in Table 5.18 display a positive correlation between connectivity and system design (*r* = 0.65; large effect; *p* ≤ 0.05), technology knowledge management (r = 0.66; large effect; p ≤ 0.05), relational knowledge (r = 0.35; medium effect; *p* ≤ 0.05), and organisational performance (r = 0.55; large effect; *p* ≤ 0.05).

As observed in Table 5.18, there is an association between system design and technology knowledge management (*r* = 0.73; large effect; *p* ≤ 0.05), relational knowledge (*r* = 0.47; medium effect; *p* ≤ 0.05), organisational performance (r= 0.58; large effect; *p* ≤ 0.05). in addition, a significant correlation was found between technology knowledge management and relational knowledge (*r* = 0.52; large effect; p ≤ 0.05), and organisational performance (*r* =0.50; large effect; *p* ≤ 0.05). A correlation was also noticed between relational knowledge and organisational performance (r = 0.187; small effect; *p* ≤ 0.05).

The findings in the Table 5.18 confirm that there is an association between factors related to big data analytic capability and organisational performance, therefore hypothesis H1 was supported:

“There is an association between factors related to big data analytic capability (big data planning, data analytic resources, data analytic investment, system design, connectivity, relational knowledge, technology knowledge management and relational knowledge) and organisational performance.’’

5.9 MULTIPLE REGRESSION ANALYSIS

Multiple regression analysis is a multivariate method for investigating the mutual combined effects of the explanatory (independent) variables on the variance of the explained (dependent) variables. According to Cohen et al. (2011), following the items analysis, dimensionality and correlation, the stepwise multiple regression analysis was used to determine the influence of big data analytical capability on organisational performance. In the present study, a stepwise multiple regression was performed to yield only variables which have a large influence on the variance of the explained variable (organisational performance) as described in Table 29 below.

After performing stepwise multiple regression, five models were produced. At this stage of analysis, it is cautious to stress that from model 1 to model 4, no model was found to be conclusive with the objective of the study. Only the fifth model was retained for further analysis because it displayed strong effect in the change of the dependent variable (organisational performance).

Moreover, before carrying on with stepwise multiple regression analysis, an assessment of multicollinearity was performed to determine the effect of the variance inflation factor (VIF) on regression analysis by ensuring that the level of the VIF should not exceed 10. According to Tabachnick and Fidell (2016), a VIF below the threshold of 10 and a tolerance above 0.2 indicate that no problem of multicollinearity was found among independent variables. The VIF ranging between 1.15 and 2.20, and a tolerance varying between 0.46 and 087, show that there was correlation among independent variables.

It is importance to notice that the value of betas as represented in Table 5.19 below indicate the extent to which the selected variables related to big data capability (system design, big data planning, connectivity, data analytic resources, relational knowledge) generated by means of stepwise multiple regression largely influence the variance of dependent variable (organisational performance).

Table 5.19 below presents the regression model generated by the stepwise multiple regression analysis. In the same vein, Table 29, summarising the findings related to the multiple regression, indicates that big data capability selected factors (system design, big data planning, connectivity, data analytic resources, relational knowledge) are predictors of organisational performance. The system design (β=0.60; t= 8.65), big data planning (β = -0.21; t = -4.24), connectivity (β = 0.35; t = 5.19), data analytic resources (β = -0.26; t = -4.42), and relational knowledge (β = 0.12; t = -2.34) acted as predictors to explain the variance of organisational performance. The findings in Table 5.19 illustrate that system design (β = 0.60) is the most significant predictor in influencing organisational performance.

The findings in Table 5.19 confirm that factors related to big data analytic capability acted as predictors of organisational performance; therefore, the hypothesis H2 was supported:

“Factors related to big data analytic capability do influence organisational performance.’’

Table 5.19

Factors related to Big Data Analytic Capability as predictors of Organisational Performance

Coefficients [a]													
Model		**Unstandardized Coefficients**		**Standardized Coefficients**	***t***	**Sig.**	**95.0% Confidence Interval for B**		**Correlations**			**Collinearity Statistics**	
		B	**Std. Error**	**Beta (*β*)**			**Lower Bound**	**Upper Bound**	**Zero-order**	**Partial**	**Part**	**Tolerance**	**VIF**
5	**(Constant)**	21,73	2,55		8,54	0,000	16,72	26,75					
	SysDes	0,89	0,10	0,60	8,65	0,000	0,68	1,09	0,58	0,49	0,40	0,46	2,20
	BIGD	-0,63	0,15	-0,21	-4,24	0,000	-0,92	-0,33	-0,10	-0,27	-0,20	0,87	1,15
	Con	0,55	0,11	0,35	5,18	0,000	0,34	0,76	0,55	0,32	0,24	0,48	2,09
	DATAANL	-0,45	0,10	-0,26	-4,42	0,000	-0,65	-0,25	0,16	-0,28	-0,21	0,65	1,53
	RelKnow	-0,22	0,09	-0,12	-2,34	0,020	-0,340	-0,03	0,19	-0,15	-0,11	0,77	1,29

Note. Source: Author's own data

5.10 SUMMARY OF CORRELATION AND MULTIPLE REGRESSION FINDINGS

In summary, Table 5.19 above reveals that findings from correlation confirm the occurrence relationships between big data analytic capability variables and organisational performance. The findings of the study suggest that big data analytic capability is associated with the organisational performance. This informs that the amount of change that occurs in the determinants of big data analytics capability will likely also affect the performance of the organisation. The findings of the present study corroborate the study by Otchere et al. (2022) who confirms the association between factors of big data capability and firm's performance. The study's findings are in line with Shabbir and Gardezi (2020) who investigate the nexus between application of big data analytics and organisational performance of small and medium enterprises. The results of their study indicate a significant connection between the application of big data analytics and performance of organisations. The results from the correlation suggest that when an organisation is well equipped in information and communication technology, the higher is its likelihood to adapt and improve its performance, and to resist change that occurs in the market (Ghasemaghaei, 2018). In the same vein, it is important to notice that an effective use of big data capability can direct the organisation to attain its goals.

Similarly, the above relationship (see Table 5.19) yielded by the stepwise multiple regression well illustrated factors related to big data capability as predictors of organisational performance. Previous studies had looked into the influence on big data analytic capability on organisational performance (Abassi, 2016; Chen, 201; Mikalef et al., 2019). The current study's findings agreed with those of Wamba et al. (2017), confirming that there is a link between big data analytics and organisational performance. Multiple stepwise regression results reflect that system design, connectivity, relational knowledge, big data planning and data analytics resource influence organisational performance positively and significantly. In the same vein, Walls and Barnard (2020) confirm the findings of the study by affirming that big data analytics capability can enhance the performance of the organisation. This implies that having access to big data capability can improve organisational performance (Walls & Barnard, 2020).

5.11 CHAPTER SUMMARY

Chapter five focused on presentation and interpretation of findings related to the validation process of a big data analytic capability scale (big data planning, data analytic resources, data analytic investment, system design, connectivity, relational knowledge, technology knowledge management) in the South African context. The confirmatory factor analysis confirmed a good-fit-model of big data analytic capability scale. The goodness of fit model from the confirmatory factor analysis validated the big data analytics capability scale; therefore, the main objective of the study was achieved. Additionally, the reliability and validity of big data analytic capability and organisational performance measurement instrument were determined by means of internal consistency (Cronbach's Alpha coefficient) and dimensional items analysis (exploratory factor analysis) (Wiid & Diggines, 2015). Moreover, interrelationships were found between factors related to big data analytic capability and organisational performance. The next chapter presents the conclusions, limitations, recommendations and contributions of the study.

CHAPTER 6:
CONCLUSIONS, LIMITATIONS AND RECOMMENDATIONS

6.1 INTRODUCTION

Chapter 6 discusses and summarises the findings and conclusions of this study and answers the research questions presented in Chapter 1, based on the analysis of the results presented in Chapter 5. Guided by the research objectives and questions, the researcher made use of the questionnaire survey to obtain quantitative data from respondents (BD specialists), which was used to validate a new scale and to establish if BDAC (big data analytics capability) can be used to improve OP (organisational performance). The researcher was able to determine the prerequisite conditions for BDAC and OP with the help of the survey.

6.2 CONCLUSIONS

Conclusions are made in the following sections in respect of the specific literature aims and the specific empirical aims of the study.

6.2.1 Specific Literature Aims

The general aim of this research was to develop and validated a BDAC scale for the South African context and to determine the relationship between BD, BDAC and OP. The general aim of this study was realised through the achievement of the specific aims, as set out in the subsections below. Conclusions were drawn about each of the specific aims regarding the relationship between BD, BDAC and OP.

6.2.1.1 Specific Literature Aim 1

To conceptualise the variables of BD, BDAC and OP from the literature.

Big Data (BD)

BD is a collection of data that is massive in volume and continues to increase exponentially in size over time. It is data that is so massive and complex that none of the usual data management methods can efficiently store or process it. BD is also data, but it is enormous in size (Ronda-Pupo et al., 2012).

Big Data Analytics Capability (BDAC)

Big data analytics capability (BDAC) is defined as "a new generation of technologies and architectures, designed to economically extract values from very large volumes of a wide range of data, by enabling high-velocity capture, discovery, and analysis" (Mikalef et al., 2020, p. 242). In addition to this definition, Wamba et al. (2017) described BDAC as a holistic strategy to managing, processing, and analysing volume, variety, velocity, and value in order to provide actionable ideas for generating long-term value, monitoring performance, and establishing competitive advantages. Chen et al. (2015), Mikalef et al. (2018) and Sheng et al. (2017) all agree that BDAC is a complicated technique that is used to uncover insightful information through the use of structured and unstructured data by revealing hidden patterns (Gandomi & Haider, 2015; Lee, 2017; Najafabadi et al., 2015). Hence, organisations are increasingly adopting BDAC for the sole goal of making operations and data-driven decision-making processes simpler and faster.

Organisational Performance (OP)

Organisational performance (OP) refers to an organisation's ability to meet its objectives and meet the expectations of its stakeholders, and to stay afloat in the economy (Griffin et al., 2003). It may also be defined as the process of examining and measuring an organisation's performance in relation to its objectives and goals, which includes a comparison of actual and planned outcomes (Richard et al., 2009). The actual productivity or outcomes of the organisation are contrasted to the desired outcome or objectives in terms of OP. Higher performance, according to Teece (2019), depends on the organisation's ability to deal with innovation, safeguard, and employ intangible knowledge assets beneficially. Further, OP can be defined as the process of ensuring that organisational resources are properly used, and it encompasses all actions or activities undertaken by managers at various levels of the organisational hierarchy in order to assess the extent to which an organisation has met its goals (Teece, 2000).

This aim was realised in Chapter 2. That chapter focused on a conceptual understanding of and the relationship between these concepts, big data analytics capability (BDAC), big data (BD) and organisational performance (OP).

This study contributes to the emerging literature on the importance of effective BDAC utilisation in the context of organisational performance. First, it shows how an organisation's efforts towards the implementation of BDAC and data analysis, the training of resources and staff, and the promotion of a data-driven culture all favour the implementation of BDAC processes. Second, it discovers the existence of a positive relationship between BDAC and OP, thus suggesting that the availability of adequate BDAC resources and capabilities encourages the adoption of a strategic propensity towards decision-making characterised by high degrees of innovativeness, proactivity and risk-taking, which, in turn, facilitates the identification and implementation of effective changes regarding organisational performance.

6.2.1.2 Specific Literature Aim 2

To report on research studies regarding BDAC and OP from the literature.

This aim was realised in Chapter 2. That chapter provides an understanding of previous studies of validation and the effect of BDAC on OP. Although prior research has highlighted the influence of BDAC on other organisational strategic (e.g., market orientation, learning orientation (Gnizy, 2019)), BDAC and OP are the ideal mediators because they reflect a propensity for organisations to seek opportunities and competitive advantages in the present (Zhong et al., 2016). Indeed, organisations with robust BDAC are able to effectively collect and analyse data from the external environment, through which opportunities can be sensed and shaped (Garmaki et al., 2016).

This translates into the development of a BDAC that, capitalising on the valuable insights extracted, can enable organisations to overcome the flaws in their BDAC by promoting innovative and consistent new product and process development efforts, typically involving a high level of investment (Usai et al., 2021).

The association between BDAC and OP is hypothesised to be positively mediated by the organisation's ability to use BDAC. First, it is fair to suppose that organisations capable of adopting good BDAC practices acquire a tendency for innovation, creativity, and future thinking (Lumpkin & Dess, 1996), and pursue data-driven strategies that have the potential to disrupt their competitors (Wang et al., 2018). Secondly, organisations with robust BDAC skills are likely to be highly receptive to market signals and latent needs of both current and potential customers (Hughes & Morgan, 2007), allowing them to anticipate and even cause changes in the external environment through radical modifications to their logistical process. Lastly, as BDAC enhances an organisation's intelligence and data analysis systems, they encourage the pursuit of innovation opportunities outside the methods and thought patterns in which the organisation usually operates and competes, thereby encouraging managers to assume greater risks and be more receptive to adopting profound changes to the organisation's value mechanisms (Roberts et al., 2016).

6.2.1.3 Specific Literature Aim 3

To determine the effect of BDAC on OP from the literature.

This aim was realised in Chapter 2, the purpose was to extend BDAC's well-established impact on organisational performance (OP) (Yasmin et al., 2020). Despite the contributions of prior research to the advancement of organisational performance (OP) knowledge, there remain gaps in the literature that hinder our understanding of how and when BDAC becomes an enabler of organisational performance. Thus, the present study's findings contribute to the BDAC and OP literature in two ways. As Xu et al. (2016) and Chaudhary et al. (2016) suggest, a higher level of BDAC may not contribute to competitive advantage. Rather, from an OP perspective, sustained competitive advantage may emerge when organisations leverage their BDACs to develop data-driven knowledge and insights to proactively develop disruptive technologies and innovations that their competitors find impossible to replicate.

This study extends the BDAC to new contexts of analysis by assessing the role of BD as a facilitating mechanism in the interaction between the BDAC and OP. A second weakness in the BDAC literature is the absence of circumstances under which investments in BDAC are profitable. It appears from the research in this field that spending more in BDAC is a good idea, as big data provides organisations with valuable economic benefits (Mills, 2019). Nevertheless, Côrte-Real et al. (2017) and Ross et al. (2013) have cast doubt on whether big data are always an effective predictor of an organisation's success.

For instance, Ross et al. (2013) contend that investments in BD may not be profitable because organisations already struggle to manage existing data. Existing literature claims that the efficacy of BDAC as a predictor of OP is enhanced when the competition to identify and satisfy customer wants is intense and when operating data to inform customer value creation decisions are of crucial importance (McAfee et al., 2012). In addition, the economic gain for BDAC is contingent on finding a new competitive base in the market; therefore, the degree to which BDAC effects OP may depend on the competitive intensity. The literature supports that in times of intense competition, BDAC provides an organisation with a distinct competitive advantage over market rivals, which informs the development of innovative business models to promote OP.

6.2.2 Research Aims

Conclusions in terms of the specific research aims of the study.

6.2.2.1 Empirical Aim 1

To develop and validate the identified BDAC scale for the South African context.

This aim was realised in Chapter 5. Chapter 5 focused on presentation and interpretation of findings related to the validation process of big data analytic capability (BDAC) scale (big data planning, data analytic resources, data analytic investment, system design, connectivity, relational knowledge, technology knowledge management) in the South African context. The confirmatory factor analysis confirmed a good fit model of big data analytic capability scale. The goodness-of-fit model from the confirmatory factor analysis validated the big data analytics capability scale. Therefore, the main objective of the study was achieved.

Additionally, the reliability and validity of big data analytic capability and organisational performance measurement instrument were determined by means of internal consistency (Cronbach's Alpha coefficient) and dimensional items analysis (exploratory factor analysis) (Wiid & Diggines, 2015). Moreover, positive interrelationships were found between factors related to big data analytic capability and organisational performance.

6.2.2.2 Empirical Aim 2

To make recommendations to the participating organisation, industrial and organisational psychology, and future research, based on the results of the study.

Recommendations are made in section 6.4 below.

6.3 LIMITATIONS

Despite its obvious strengths, this study also has limitations.

The research was mostly limited to South African respondents. The study was done in South Africa, and the majority of respondents were from Accenture South Africa, hence, the results are specific to BDAC in the South African environment and may not be applicable or generalisable elsewhere.

Another limitation concerns the research setting. South African is not a technologically advanced country, and the adoption of BDAC is not yet well established here, which this study can assist in future research.

Overall, the survey form was used so that respondents could answer some questions from their practical knowledge of BDAC management. However, it cannot be ruled out that some individuals responded using their theoretical understanding of BDAC. Despite these limitations, the study's results remain legitimate and significant for BDAC and OP's use, as well as for academics' future research.

6.4 RECOMMENDATIONS

The following recommendations are made based on the findings of this study.

Recommendations to the participating organisation

It is recommended that BD individuals increase the rate of adopting and using BDAC in OP. If organisations adopt and use BDAC, this will improve organisational performance and assist organisations in alignment with their goals. Organisations should play a more active role in encouraging the adoption and use of BDAC, as the study revealed that there are not many organisations investing in BDAC resources.

Organisations should incentivise individuals who adopt and use BDAC, which would help to improve OP.

Recommendations for IOP

Industrial and Organisational Psychology (IOP) is a long-standing subject of study that has never been as popular as it is today. This study may have implications for IOP and the training of IO professionals. This study confirms the increasing importance of big data and big data analytics capability in general and its potential impact on organisational performance. Thus, it is recommended that new IO psychologists should be empowered with cutting-edge expertise in data science, statistics, and computing (King et al., 2016).

There has been a definite movement from traditional IOP research methods to those that use more data analytics. However, the lack of adequate training in contemporary analytic methods may disadvantage IOP professionals (Putka & Oswald, 2016), especially as modern organisations now speak the language of data analytics. Increasing numbers of organisations and IOP departments see the benefits of BD and BDAC and are transitioning to using it to make data-driven strategic choices, hence affecting organisational performance. To compete in this digital era, organisations require an expanding number of employees with computer science and data analysis skills. IOP should be setting the tone or run the risk of being left behind.

IO professionals are uniquely positioned to benefit from BD and BDAC. One reason for this is that they frequently perform research or give advisory services for organisations with enormous quantities of data. These professionals are also uniquely qualified to interpret statistical analyses, as their training encompasses the human side of statistics and the importance of interpreting results in a way that makes sense to people outside the field. This enables them to communicate findings more effectively to technically proficient individuals and executives in organisations.

However, IO professionals may be at a disadvantage if they lack knowledge of developing BD and BDAC (i.e., data science, and data visualisation tools and software, including as R, Python, SQL, and data visualisation tools (Oswald et al., 2020)). These tools have gained great popularity and are routinely employed in organisations. Some of the world's largest organisations use R for data science and research. IO professionals with these abilities may assist organisations in maximising their talent potential through the use of predictive analytics and other data-driven decision-making techniques.

The lack of analytical abilities among IO psychology professionals is a significant issue, due to an uneven emphasis on theory as opposed to practical application. This places the field in a position where new researchers and practitioners in IO psychology may have outdated statistics and methods training, limiting the subject's progress, value, and multidisciplinary potential (Putka & Oswald, 2016).

The issue with this disparity is that it does not appear to be closing, especially as it becomes increasingly difficult for IOP professionals to obtain employment in their sector. People working in the subject of Industrial Psychology will need to reconcile or at least recognise the conflict between theoretical instruction and practical application.

IOP master's programmes provide adequate training in traditional statistical techniques (e.g., regression, correlations) using SPSS, the most popular statistical software in graduate school. Nevertheless, many universities fail to teach classic statistical processes using the analytical software and tools that are routinely used in organisational contexts. The relevance of having a computer or machine learning background is gaining support within the IOP community, despite the fact that many programmes continue to lack the use of modern analytical techniques.

Despite a growing emphasis on the need of machine learning and more robust data analytics approaches, there has been little change in the way statistics is taught to IOP professionals. It is easier said than done to revamp IOP professional skill sets to incorporate R and Python programming. Programmes face a substantial number of obstacles and difficulties. However, the first crucial step is to acknowledge the widening skills gap, followed by efforts to close it.

IOP professionals can introduce programming to their core skills in a variety of ways:

- Certify in the latest trends in data analytics software, such as Tableau and Power BI, and encourage them to seek out opportunities to expand their proficiency with these tools.

- University professors adopting SPSS in their statistical and research classes could offer SPSS homework and then challenge their students to acquire the same findings using R.

- University could offer electives including data science, data analytics, and computer programming courses.

These examples illustrate only a few of the numerous ways IOP professionals can develop crucial data analysis skills.

Because IOP professionals provide organisations with unique and significant expertise, they have a stronger organisational influence if they increase their knowledge of BDAC and the applied methodology to OP. IOP professionals benefit from programmes that teach statistics using R and/or Python instead of or in conjunction with SPSS. Instead of knowing how to use a single analytical tool, IOP professionals will be able to conduct statistical analyses using various tools. Having a broader exposure with a variety of analytical tools can provide organisations with additional alternatives for analysing and displaying data, enhance their adaptability to organisational needs, and ultimately increase their worth as IOP professionals.

Recommendations for future research

In light of the findings of this study, the following recommendations are made for future research:

This study provides researchers with important information for further studies to be conducted to establish how BD and BDAC can also be used to establish OP methodologies. The feedback from participants who have used both the traditional and agile methodologies of BDAC provide a valuable basis for further research. Future studies could be done to assess the impact of BDAC use on individual dimensions per study to devote more time and resources and have an in-depth understanding of the individual dimensions.

The study revealed that big data analytics planning, data analytics investment, big data analytics resources, connectivity, big data analytics capability, system design, technical knowledge, technology management knowledge, organisational knowledge, relational knowledge, and organisational performance improved by the use of BDAC, but further research should be conducted for in-depth analysis of how these dimensions are individually improved by BDAC use, and how the rest of the dimensions are also impacted with the inclusion of a qualitative methodology,

This study was conducted in the South African context. A comparative study could be conducted in more technologically advanced countries to establish if the study conducted in better-resourced environments would reach different findings and conclusions, especially in relation to the adoptions and use of BDAC. It could be that organisations in other countries that are technologically ahead of South Africa may have embraced BDAC better, leading to a different outcome of the study.

6.5 CHAPTER SUMMARY

This research builds a theory of BDAC strategy that shows how to leverage the BDAC dimensions and sub-dimensions to build an overall BDAC strategy. Although several studies highlight the importance of big data analytics planning, data analytics investment, big data analytics resources, connectivity, big data analytics capability, system design, technical knowledge, technology management knowledge, organisational knowledge, relational knowledge and organisational performance, this research illuminates the role of these dimensions and entanglement view in proposing an integrated BDAC model and its overall impact on organisational performance.

With the growing interests in organisational analytics across various industries, the current study advances BDAC conceptualisation and the role of analytics in enhancing organisational performance. A notable strength of the current study is that data were collected from multiple industries to test the model empirically.

Overall, the study leads to a better understanding of big data analytics capability in the data economy, and is likely to open new avenues of research into academic and organisational processes and practices in efforts to improve organisational performance.

REFERENCES

Abbasi, A., Sarkar, S., & Chiang, Roger H. L. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the Association for Information Systems*, *17*(2), 1-32.

Ackoff, R. L. (1981). *Creating the corporate future: Plan or be planned for*. John Wiley & Sons.

Adams, N. M. (2010). Perspectives on data mining. *International Journal of Market Research*, *52*(1), 11-19.

Addo-Tenkorang, R., Helo, P. T., Shamsuzzoha, A., Ehrs, M., & Phuong, D. (2016). *Logistics & supply chain management tracking networks: Data-management system integration / interfacing issues*. PICMET '12: Technology Management for Emerging Technologies. Vancouver, Canada 29 Jul - 02 Aug 2012 IEEE.

Agrawal, D., Das, S., & El Abbadi, A. (2011). *Big data and cloud computing: Current state and future opportunities*. Proceedings of the 14th International Conference on Extending Database Technology, ACM, 2011, 530-533.

Akbay, S. (2015). How big data applications are revolutionizing decision-making. *Business Intelligence Journal*, *2*0(1), 25-29.

Akerkar, R. (2019). *Introduction to Artificial Intelligence*. PHI Learning.10.1007/978-3-319-97436-1_1.

Akter, S., Wamba, F. S., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*, 182, December, 113-131.

Alexander, B. (2017). Does loyalty span domains? Examining the relationship between consumer loyalty, other loyalties and happiness. *Journal of Business Research, 68*(12), 2464-2476.

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

Al-Sai, Z. A., & Abualigah, L. M. (2017). *Big data and E-government: A review*. 2017 8th International Conference Information Technology, pp. 580–587.

Altman, D. G., Deeks, J. J., & Higgins, J. (2008). *Analysing data and undertaking meta-analyses. Handbook for systematic reviews of interventions*. Cochrane Book Series.

Aluri, A. J., Slevitch, L., & Larzelere, R. (2015). The effectiveness of embedded social media on hotel websites and the importance of social interactions and return on engagement. *International Journal of Contemporary Hospitality Management*, *27*(4), 671-689 10.1108/IJCHM-09-2013-0415.

Alvesson, M., & Skoldberg, K. (2009). *Reflexive methodology: New vistas for qualitative research* (2nd ed.). Sage.

Amado, A., Cortez, P., & Rita, P., & Moro, S. (2017). Research trends on big data in marketing: A text mining and topic modeling based literature analysis. *European Research on Management and Business Economics*, *24*(1), 1-7. 10.1016/j.iedeen.2017.06.002.

Andale, A. (2015). *Probability sampling: Definition, types, advantages and disadvantages.* Statistics How To. https://www.statisticshowto.com/probability-and-statistics/sampling-in-statistics/probability-sampling/

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

Banerjee, A. (2013). *Big data & advanced analytics in telecom: a multi-billion-dollar revenue opportunity* [White paper]. Heavy Reading. https://carrier.huawei.com/~/media/cnbg/downloads/technical%20topics/heavyreading%20huawei%20big%20data%20customized%20report/heavyreading%20%20huawei%20big%20data%20customized%20reportfinal1a.pdf

Barry, A., Chaney, B., Stellefson, M., & Chaney, D. (2011). So, you want to develop a survey: Practical recommendations for scale development. *American Journal of Health Studies, 26*(1), 97-105.

Barton, D., & Davenport, T.H, & Harris J. G. (2012). Making advanced analytics work for you. *Harvard Business Review*, *90*(10), 78-83.

Basheer, M., Wheeler, K., Ribbe, L., Majdalawi, M., Abdo, G., & Zagona, E. (2018). Quantifying and evaluating the impacts of cooperation in transboundary river basins on the water-energy-food nexus: The Blue Nile Basin. *Science of the Total Environment*, 630, 1309-1323. 10.1016/j.scitotenv.2018.02.249.

Battaglia, M., Sampling, N., & Lavrakas, P. J. (2008). *Encyclopedia of survey research methods.* Sage Publications.

Bean, R. (2018). How big data and AI are driving business innovation. *MIT Sloan Management Review*, February 5, 2018. Website: https://sloanreview.mit.edu/article/how-big-data-and-ai-are-driving-business-innovation-in-2018/

Bechor, T., Neumann, S., Zviran, M., & Glezer, C. (2010). A contingency model for estimating success of strategic information systems planning. *Information & Management*, *47*, 17-29. 10.1016/j.im.2009.09.004.

Becker, J., Knackstedt, R., & Poeppelbuss, J. (2009). Developing maturity models for IT management. *Business & Information Systems Engineering*, *1*(3), 213-222. 10.1007/s12599-009-0044-5.

Bersin, J., & Ferrar A. (2014). Big data in human resources: A world of haves and have-nots. *Forbes*. http://www.forbes.com/sites/joshbersin/2013/ 10/07/big-data-in-human-resources-a-world-of-haves-and-have-nots/

Bigelow S. J. (2020). *Big data*. TechTarget. https://www.techtarget.com/searchdatamanagement/definition/big-data

Bitici, U. (2005). Performance measurement systems in SMEs: A review for a research agenda. *International Journal of Management Reviews*, *7*(1), 25-47. 10.1111/j.1468-2370.2005.00105.x.

Bitici, U. (2017). Interplay between performance measurement and management, employee engagement and performance. *International Journal of Operations & Production Management*, *37*(9), 1207-1228. 10.1108/IJOPM-06-2015-0313.

Blaikie, N. (2007). *Approaches to social enquiry* (2nd ed.). Polity Press.

Bowling , A & Ebrahim, S. (2006). *Handbook of health research methods. Investigation, measurement and analysis*. Open University Press.

Brackertz, N., & Kenley, R. (2002). Evaluating community facilities in local government: Managing for service enablement. *Journal of Facilities Management*, *1*(3), 283-299. 10.1108/14725960310807971.

Bradley, S., Cámara, P., Javier, C., & Bernardino, J. (2016). *Big data in cloud computing: Features and issues.* Conference paper. https://www.researchgate.net/publication/302973843_Big_Data_in_Cloud_Computing_Features_and_Issues

Brown, B., Manyika, J., & Chui, M. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute.

Brownell, P., & Abernethy, M. A. (1997). Management control systems in research and development organizations: the role of accounting, behavior and personnel controls. *Accounting, Organizations and Society*, *22*, 233–248.

Bryman, A., & Bell, E. (2011). *Business research methods* (3rd ed.). Oxford University Press.

Brynjolfsson, E., McAfee, A., Davenport, T. H., Patil, D. J., & Barton, D. (2011). Big data: The management revolution. *Harvard Business Review*, *90*(10), 60–68.

Burke, W. W., & Litwin, G. H. (1992). A causal model of organisational performance and change. *Journal of Management*, *18*(3), 523–546.

Chaudhary, A. (2016). *Forest management meta-analysis SI*. ResearchGate. https://www.researchgate.net/publication/299599534_Chaudhary_et_al_2016_Forest_management_meta-analysis_SI

Check J., & Schutt R. K. (2012). *Research methods in education*. Sage.

Checkland, P. (1999). *Systems thinking, systems practice*. Wiley.

Chen, C. (2012). Data-intensive applications, challenges, techniques and technologies: A survey on big data. *Information Sciences*, *275*, 314–347.

Chen, Y., Wang, Y., Nevo, S., Jin, J., Wang, L., & Chow, W. S. (2014). IT capability and organizational performance: the roles of business process agility and environmental factors. *European Journal of Information Systems*, *23*(3), 326-342.

Chen, H., Chiang, R. H., & Storey, V. C. (2015). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, *36*(4), 1165-1188.

Chesbrough, H., & Bogers, M. (2014). Explicating open innovation: Clarifying an emerging paradigm for understanding innovation, in *New frontiers in open innovation.* Oxford University Press, forthcoming.

Church, A. H., & Dutta, S. (2013). The promise of big data for OD: Old wine in new bottles or the next generation of data-driven methods for change. *OD Practitioner*, *45*, 23-31.

Cohn D. L., & Marshall A. (2014). *Driving innovation through data*. IBM Institute for Business Value.

Collingridge, D. (2014). The quality of qualitative research. *American Journal of Medical Quality*, *34*, 439-445. 10.1177/1062860619873187.

Columbus, L. (2014). 84% of enterprises see big data analytics changing their industries' competitive landscape in the next year. *Forbes*. Website: https://www.forbes.com/sites/louiscolumbus/2014/10/19/84-of-enterprises-see-big-data-analytics-changing-their-industries-competitive-landscapes-in-the-next-year/?sh=1b2ea7f517de

Comuzzi, M., & Patel, A. (2016). How organisations leverage big data: A maturity model. *Industrial Management & Data Systems*, *116*, 1468-1492. 10.1108/IMDS-12-2015-0495.

Cooper, D. R., & Schindler, P. S. (2006). *Business research methods* (9th ed.). McGraw-Hill.

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.

Cosenza, C. (2008). Writing effective survey questions. In de Leeuw, E. Hox, J., & Dillman, D., *The international handbook of survey methodology*, pp. 136-160. Routledge.

Courtney M. (2013). *Engineering & technology: Puzzling out big data*. ResearchGate. https://www.researchgate.net/publication/260619931_Puzzling_out_big_data_Information_Technology_Analytics

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Sage Publications.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Sage Publications Ltd.

Cross, K. F., & Lynch, R. L. (1988). The SMART way to define and sustain success. *National Productivity Review*, *8*(1), 23-33.

Davenport, T. H. (2017). *Keeping up with the quants: Your guide to understanding and using analytics*. Harvard Business Review Press.

Dent, E. B. (1997). *The design, development, and evaluation of measures to survey worldview in organisations* [George Washington University doctoral dissertation].

Diaz-Aviles, E., Fisichella, M., Kawase, R., Nejdl, W., & Stewart, A. (2015). *Unsupervised auto-tagging for learning object enrichment*. ResearchGate. https://www.researchgate.net/publication/221549630_Unsupervised_Auto-tagging_for_Learning_Object_Enrichment

Du Plessis, A., & Subramanien, B. (2014). Voices of despair: Challenges of multi-grade teachers in a rural district in South Africa. *Educational Research for Social Change*, *3*(1), 20-36.

Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. (2013). *Fundamentals of business process management.* Springer.

Dutta, P., & Bose, S. (2015). *Gender diversity in the boardroom and financial performance of commercial banks: Evidence from Bangladesh*. ResearchGate. https://www.researchgate.net/publication/24115700_Gender_Diversity_in_the_Boardroom_and_Financial_Performance_of_Commercial_Banks_Evidence_from_Bangladesh

Edu, A. S. (2022). Positioning big data analytics capabilities towards financial service agility. *Aslib. J. Inf. Manag*., *74*, 569–588. doi: 10.1108/AJIM-08- 2021-0240

Eisenhardt, K., & Martin, J. (2000). Dynamic capabilities: What are they? *Strategic Management Journal*, *21*, 1105-1121. 10.1002/1097-0266(200010/11)21:10/113.0.CO;2-E

Elgendy, N., & Elragal, A. (2016). Big data analytics in support of the decision making process. *Procedia Computer Science*. *100*, 1071-1084. 10.1016/j.procs.2016.09.251

Erevelles, S., Fukawa, N., & Swayne, L. (2015). Big data consumer analytics and the transformation of marketing. *Journal of Business Research*, *69*(2), 897-904. 10.1016/j.jbusres.2015.07.001.

Espinosa, J. and Armour, F. (2016). *The big data analytics gold rush: a research framework for coordination and governance*. 49th Hawaii International Conference on System Sciences (HICSS). 1112-1121. 10.1109/HICSS.2016.141.

Ferraris, A., Mazzoleni, A., Devalle, A., & Couturier, J. (2019). Big data analytics capabilities and knowledge management: Impact on firm performance. *Management Decision*, *57*(8), 1923-1936. https://doi.org/10.1108/MD-07-2018-0825

Fidel, P., Schlesinger, W., & Cervera, A. (2015). Collaborating to innovate: Effects on customer knowledge management and performance. *Journal of Business Research*, *68*(7), 1426-1428.

Fitzgerald, J. T., White, C., & Gruppen, L. (2001). A longitudinal study of self-assessment accuracy. *Medical Education. 37*, 645-9. 10.1046/j.1365-2923.2003.01567.x.

Fowler, Floyd. (2008). *Survey research methods* (5th ed.). Sage.

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

Gandomi, A., & Haider, M. (2015). Beyond the hype: big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137-144.

García-Morales, V., Jiménez-Barrionuevo, M., & Gutierrez, L. (2012). *Transformational leadership influence on organizational performance through organizational learning and innovation*. ResearchGate. https://www.researchgate.net/publication/238500512_Transformational_Leadership_Influence_on_Organizational_Performance_through_Organizational_Learning_and_Innovation

Garmaki, M., Boughzala, I., & Wamba, S. F. (2016). *The effect of big data analytics capability on firm performance*. Pacific Asia Conference on Information Systems (PACIS).

George, G., Lavie, D., Osinga, E. C., & Scott, B. A. (2014). Big data and data science methods for management research. *Academy of Management Journal*, *59*(5), 1493–1507.

Ghasemaghaei, M., Ebrahimi, S., & Hassanein, K. (2018). Data analytics competency for improving firm decision-making performance. *The Journal of Strategic Information Systems*, *27*(1), 101-113.

Gnizy, I. (2018). Big data and its strategic path to value in international firms. *International Marketing Review*, *36*(3), 318-341.

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

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

Griffin, B., & Hesketh, B. (2003). Adaptable Behaviours for Successful Work and Career Adjustment. *Australian Journal of Psychology*, *55*, 65-73. 10.1080/00049530412331312914.

Günther, W., Rezazade M., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *The Journal of Strategic Information Systems*, *26*(3), 191-209. 10.1016/j.jsis.2017.07.003.

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management*, *53*(8), 1049-1064.

Hair J. F., Black W. C., Babin B. J., Anderson R. E., & Tatham R. L. (2019). *Multivariate data analysis* (6th ed.). Pearson.

Harris, A. (2012). Distributed leadership and school improvement: leading or misleading? *Educational Management, Administration and Leadership*, *32*, 11-24.

Hasan, H., & Tibbits, H. (2000). Strategic management of electronic commerce: An adaptation of the balanced scorecard. *Internet Research*, *10*, 439-450. 10.1108/10662240010349453

Helfat, C., & Peteraf, M. (2014). Managerial cognitive capabilities and the microfoundations of dynamic capabilities. *Strategic Management Journal*, *36(*3) 831-850. 10.1002/smj.2247.

Hevner, A., Alan, R., March, S., Salvatore, T., Park, J., & Ram, S. (2004). Design science in information systems research. *Management Information Systems Quarterly*, *28*(1), 75-105.

Hirschheim, R., & Klein, H. (1989). *Four paradigms of information systems development*. ResearchGate. https://www.researchgate.net/publication/213802262_Four_Paradigms_of_Information_Systems_Development

Horga, G. (2012). *Leadership on organizational performance*. Expert Publishing.

Howson, J. (2013). Does firm performance reveal its own causes? The role of Bayesian inference. *Strategic Management Journal*, *31*, 39–57. 10.1002/smj.799.

Hughes, M., & Morgan, R. (2007). Deconstructing the relationship between entrepreneurial orientation and business performance at the embryonic stage of firm growth. *Industrial Marketing Management*, *36*, 651-661. 10.1016/j.indmarman.2006.04.003

Hung, S., Chang, C., & Yu, T. (2006). Determinants of user acceptance of the e-Government services: The case of online tax filing and payment system. *Government Information Quarterly*, *23*, 97-122. 10.1016/j.giq.2005.11.005

Inamdar, S., Benias, P., Liu, Y., Sejpal, D., Satapathy, S., & Trindade, A. (2020). Prevalence, risk factors, and outcomes of hospitalized patients with COVID-19 presenting as acute pancreatitis. *Gastroenterology*, *159*(6), 2226-2228. 10.1053/j.gastro.2020.08.044.

Jeble, S., Dubey, R., Childe, S., Papadopoulos, T., Roubaud, D., & Prakash, A. (2020). Impact of big data and predictive analytics capability on supply chain sustainability. *The International Journal of Logistics Management*, *29*(2), 513-538.

Ji, C., Li, Y., Qiu, D., Awada, U., & Li, K. (2012). *Big data processing in cloud computing environments*. Proceedings of the 2012 International Symposium on Pervasive Systems, Algorithms, and Networks, I-SPAN 2012. 10.1109/I-SPAN.2012.9.

Jiang, J. J., Klein, G., Slyke, C. V., & Cheney, P. (2003). A note on interpersonal and communication skills for IS professionals: Evidence of positive influence. *Decision Sciences*, *34*, 799-812.

Johnson, J. E. (2012). Big data + big analytics = big opportunity. *Financial Executive*, *28*, 50-53.

Jones, B. B., & Brazzel, M. (2014). *The NTL handbook of organization development and change: Principles, practices, and perspectives* (2nd ed.). Wiley.

Joubert, G., & Ehrlich, R. (2009). *Epidemiology: a research manual for South Africa* (2nd ed.). Oxford University Press.

Kambatla, K., Kollias, G., Kumar, V., & Grama, A. (2014). Trends in big data analytics. *Journal of Parallel and Distributed Computing*, *74*, 2561-2573. 10.1016/j.jpdc.2014.01.003.

Kaplan, R. S., & Norton, D. P. (1996). Using the balanced scorecard as a strategic management system. *Harvard Business Review*, January–February, 75–85.

Kaplan, R., & Norton, D. (2001). *The strategy-focused organization: How balanced scorecard companies thrive in the new business environment*. Harvard Business School Press.

Kaplan, R., & Norton, D. (2004). The strategy map: Guide to aligning intangible assets. *Strategy & Leadership*, *32*, 10-17. 10.1108/10878570410699825.

Karimi, J., Somers, T. M., & Gupta, Y. P. (2001). Impact of information technology management practices on customer service. *Journal of Management Information Systems, 17*, 125-158.

Katz, M. H. (2006). *Study design and statistical analysis*. Cambridge University Press.

Kim, G., Shin, B., & Kwon, O. (2012). Investigating the value of sociomaterialism in conceptualizing IT capability of a firm. *Journal of Management Information Systems*, *29*, 327-362.

Kiron D., Prentice P. K., & Ferguson R. B. (2012). Innovating with analytics. *MIT Sloan Management Review*, *54*(1), 47–52.

Kitchens, B., Dobolyi, D., Li, J., & Abbasi, A. (2018). Advanced customer analytics: strategic value through integration of relationship-oriented big data. *Journal of Management Information System*s, *35*, 540-574. 10.1080/07421222.2018.1451957.

Kline R. B. (2016). *Principles and Practice of Structural Equation Modelin*g (4th ed.). Guilford Press.

Kothari, C. R. (2004). *Research methodology: An introduction*. New Age.

Kubick, W. R. (2012), Big data, information and meaning. *Applied Clinical Trials*, *21*(2), 26–28.

Kubina, M., Varmus, M., & Kubinova, I. (2015). Use of big data for competitive advantage of company. *Procedia Economics and Finance*, *26*, 561-565. 10.1016/S2212-5671(15)00955-7.

Kueng, P. (2000). Process performance measurement system: A tool to support process-based organizations. *Total Quality Management*, *11*, 67-85. 10.1080/0954412007035.

Kueng, P., & Wettstein, T. (2001). Performance measurement systems must be engineered. *Communications of the Association for Information Systems*, 7. 10.17705/1CAIS.00703.

Kumari, V., & Kaur, H. (2018). Predictive modelling and analytics for diabetes using a machine learning approach. *Applied Computing and Informatics*. ahead-of-print. 10.1016/j.aci.2018.12.004.

Kumari, S., Patil, Y., & Jeble, S. (2018). Role of big data in decision making. *Operations and Supply Chain Management: An International Journal*, *11*(1), 36-44. 10.31387/oscm0300198

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT Sloan Management Review*, *52*(2), 21-32.

Lee, H. S. (2017). Big data facilitation, utilization, and monetization: Exploring the 3Vs in a new product development process. *Journal of Product Innovation Management*, *34*(5), 640–658. https://doi.org/10.1111/jpim.12397.

Levy, P. E. (2006). *Industrial/organisational psychology: Understanding the workplace* (2nd ed.). Houghton Mifflin.

Li, D., & Liu, J. (2014). Dynamic capabilities, environmental dynamism, and competitive advantage: Evidence from China. *Journal of Business Research*, 67, 2793–2799. 10.1016/j.jbusres.2012.08.007.

Lin, Y., & Wu, L. (2014). Exploring the role of dynamic capabilities in firm performance under the resource-based view framework. *Journal of Business Research*, *67*, 407–413. 10.1016/j.jbusres.2012.12.019.

Lind, D. A., Marchal, W. C., & Wathen, S. A. (2008). *Statistical techniques in business and economics with global data sets* (15th ed.). McGraw-Hill Irwin.

Linton, G., & Kask, J. (2017). Configurations of entrepreneurial orientation and competitive strategy for high performance. *Journal of Business Research*, *70*, 168–176. 10.1016/j.jbusres.2016.08.022.

LoBiondo-Wood, G., & Haber, J. (2002). *Methods, critical appraisal, and utilization* (5th ed.). Mosby.

Lohr, Steve. (2012). The age of big data. *New York Times*, 11 February.

Lu, Ying & Ramamurthy, K. (2011). Understanding the link between information technology capability and organizational agility: An empirical examination. *MIS Quarterly*, *35*, 931-954. 10.2307/41409967.

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

Lunde, A., Okaty, B., Dymecki, S., & Glover, J. (2019). Molecular profiling defines evolutionarily conserved transcription factor signatures of major vestibulospinal neuron groups. *eNeuro*, 6(1). https://www.eneuro.org/content/6/1/ENEURO.0475-18.2019.long

Maguire, J. (2018). *Using data analytics for competitive advantage: Expert advice*. Datamation. https://www.datamation.com/big-data/using-data-analytics-expert-advice.htm/.

Marr, B. (2021). *How much data is there in the world?* Bernard Marr & Co. Website: https://bernardmarr.com/how-much-data-is-there-in-the-world/

Martins, N., & Coetzee, M. (2009). Applying the Burke-Litwin model as a diagnostic framework for assessing organisational effectiveness. *South African Journal of Human Resource Management*, *7*(1), 144-156. 10.4102/sajhrm.v7i1.177.

Masa'deh, R., Obeidat, B., & Tarhini, A. (2016). A Jordanian empirical study of the associations among transformational leadership, transactional leadership, knowledge sharing, job performance, and firm performance: A structural equation modelling approach. *Journal of Management Development*, *35*(5), 681-705. 10.1108/JMD-09-2015-0134.

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

McAbee, S., Landis, R., & Burke, M. (2017). Inductive reasoning: The promise of big data. *Human Resource Management Review*, *27*, 277-290. 10.1016/j.hrmr.2016.08.005.

McAfee, A., Brynjolfsson, E., Davenport, T., Patil, D. J., & Barton, D. (2012). Big data: The management revolution. *Harvard Business Review*, *90*, 61-67.

McNair, H., Brzezinski, M., & Krause, J. (2015). Quantifying diatom silicification with the fluorescent dye, PDMPO. *Limnology and Oceanography: Methods*, *13*(10), 511-599. 10.1002/lom3.10049

Mdluli, S., & Makhupe, O. (2017). *Defining leadership competencies needed for the fourth industrial revolution: leadership competencies 4.*0. Bank Seta South Africa.

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*, 1-16. 10.1016/j.jbusres.2016.09.004.

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

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics capabilities and innovation: The mediating role of dynamic capabilities and moderating effect of the environment. *British Journal of Management*, *30*(2), 272-298. 10.1111/1467-8551.12343.

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2020). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research*, *98*(2019), 261–276. https://doi.org/10.1016/j.jbusres.2019.01.044

Mikalef, P., & Krogstie, J. (2018). *Big data governance and dynamic capabilities: The moderating effect of environmental uncertainty*. Pacific Asia Conference on Information Systems (PACIS) At: Yokohama, Japan.

Millar, C., Groth, O., & Mahon, J. (2018). Management innovation in a VUCA world: Challenges and recommendations. *California Management Review*, *61*(1). 000812561880511. 10.1177/0008125618805111.

Miller, D., & Friesen, P.H (1982). Innovation in conservative and entrepreneurial firms: Two models of strategic momentum. *Strategic Management Journal*, *3*, 1-25. https://doi.org/10.1002/smj.4250030102

Mills, T. (2019). Five benefits of big data analytics and how companies can get started. *Forbes*. https://www.forbes.com/sites/forbestechcouncil/2019/11/06/five-benefits-of-big-data-analyticsand-how-companies-can-get-started/?sh=118f81b117e4

Minges, M. (2003). *Is the internet mobile? Measurements from Asia-Pacific*. Presented at The International Telecommunications Society Asia-Australasian Regional Conference, M-business, E-commerce and the Impact of Broadband on Regional Development and Business Prospects. Perth. 22-24 June.

Mithas, S., Ramasubbu, N., & Sambamurthy, V. (2011). How information management capability influences firm performance. *MIS Quarterly*, *35*, 237-256. 10.2307/23043496.

Moe, J., & Kallin, H. (2011). *U.S. Patent No. 7,957,743. Self-configuring and optimization of cell neighbors in wireless telecommunications networks*. U.S. Patent and Trademark Office.

Monino, J. (2016). The big data revolution, in *Big Data, Open Data and Data Development*. 10.1002/9781119285199.ch1.

Najafabadi, M., Villanustre, F., Khoshgoftaar, T., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. *Journal of Big Data*, *2*(1). 10.1186/s40537-014-0007-7.

Nannetti, P. (2012). *The deciding factor: Big data & decision making*. Capgemini.

Neely, A., Mills, J., Platts, K., Richards, H., Gregory, M., Bourne, M., & Kennerley, M. (2000). Performance measurement system design: Developing and testing a process-based approach. *International Journal of Operations & Production Management*, *20*(10), 1119-1145. 10.1108/01443570010343708.

Olivier, M., & Pieterse, H. (2018). Classifying the authenticity of evaluated smartphone data. *Advances in Digital Forensics*, *15*, 39-57. 10.1007/978-3-030-28752-8_3.

Opresnik, D., & Taisch, M. (2015). The conceptualization of sustainability in operations management. *Procedia CIRP*, *29*, 532-537. 10.1016/j.procir.2015.01.038.

O'Reilly, C., & Tushman, M. (2008). Ambidexterity as a dynamic capability: Resolving the innovator's dilemma. *Research in Organizational Behavior*, *28*, 185-206. 10.1016/j.riob.2008.06.002.

Osisioma, H., Nzewi, H., & Mgbemena-Nsofor, I. (2016). Dynamic capabilities and performance of selected commercial banks in Awka, Anambra State, Nigeria. *European Journal of Business and Social Sciences*, *4*, 98-110.

Oswald, T., Rumbold, A., Kedzior, S., & Moore, V. (2020). Psychological impacts of "screen time" and "green time" for children and adolescents: A systematic scoping review. *PloS one*. 15. e0237725. 10.1371/journal.pone.0237725.

Otchere, I., Afum, T., Morgan, P., Musah, A., Ohene-Aboagye, S., Osei-Wusu, S., Yirenkye, S., Tetteh-Ocloo, G., Ansa, G., Laryea, C., Forson, A., Asante-Poku, A., & Yeboah-Manu, D. (2022). Evaluation of a loop-mediated amplification test for rapid diagnosis of tuberculosis in Ghana. *Health Sciences Investigations Journal*, *3*(1), 335-340. 10.46829/hsijournal.2022.6.3.1.335-340.

Özköse, H., Arı, E. S., & Gencer, C. (2015). Yesterday, today and tomorrow of big data. *Procedia – Social and Behavioral Sciences*, *195*, 1042–1050. https://doi.org/10.1016/j.sbspro.2015.06.147

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

Pallant, J. (2016). *SPSS survival manual: A step-by-step guide to data analysis using SPSS program* (6th ed.). McGraw-Hill Education.

Parasuraman, A., Zeithaml, V., & Berry, L. (1985). A conceptual model of service quality and its implication for future research (SERVQUAL). *The Journal of Marketing*, *49*, 41-50. 10.2307/1251430.

Perrons, R., & Jensen, J. (2015). Data as an asset: What the oil and gas sector can learn from other industries about “Big Data”. *Energy Policy*, *81*, June, 117-121. 10.1016/j.enpol.2015.02.020.

Petter, S., McLean, E., & Delone, W. (2014). Information systems success: The quest for the independent variables. *Journal of Management Information Systems*, *29*(4), 7-62. 10.2753/MIS0742-1222290401.

Pigni, F., & Watson, R. (2016). Alea iacta est: Now is the time to extract value from digital data streams. *MIS Quarterly Executive*, *15*, i-ix.

Pisano, G. (2017). Toward a prescriptive theory of dynamic capabilities: Connecting strategic choice, learning, and competition. *Industrial and Corporate Change*, *26*, 747-762. 10.1093/icc/dtx026.

Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*, 879–903. doi: 10.1037/0021-9010.88.5.879

Polit, D. F., & Beck, C. T. (2008). *Nursing research: Generating and assessing evidence for nursing practice* (8th ed.). Wolters Kluwer Health/Lippincott Williams & Wilkins.

Popovič, A., Hackney, R., Coelho, P. and Jaklič, J. (2012). Towards business intelligence systems success: Effects of maturity and culture on analytical decision-making. *Decision Support Systems*, *54*, 729-739. 10.1016/j.dss.2012.08.017.

Pospiech, G. (2019). Framework of mathematization in physics from a teaching perspective, in *Mathematics in physics education*. 10.1007/978-3-030-04627-9_1.

Prigogine, I., & Stengers, I. (1984). *Order out of chaos: Man's new dialogue with nature*. Flamingo Editions.

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data*, *1*(1). 10.1089/big.2013.1508.

Putka, D., & Oswald, F. (2015). Implications of the big data movement for the advancement of I-O science and practice, in *Big data at work*. Routledge.

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. *Health Information Science and Systems*, *2*(3). https://doi.org/10.1186/2047-2501-2-3

Rahman, N., & Aldhaban, F. (2015). Assessing the effectiveness of big data initiatives. *PDXScholar*. https://pdxscholar.library.pdx.edu/cgi/viewcontent.cgi?article=1076&context=etm_fac

Ransbotham, S., & Kiron, D. (2017). Analytics as a source of business innovation. *MIT Sloan Management Review*, February 28.

Remenyi, D., Williams, B., Money, A., & Swartz, E. (1998). *Doing Research in Business and Management*. Sage.

Rialti, Ri., Zollo, L., Ferraris, A., & Alon, I. (2019). Big data analytics capabilities and performance: Evidence from a moderated multi-mediation model. *Technological Forecasting and Social Change*, *149*(C). 10.1016/j.techfore.2019.119781.

Richard, P., Devinney, T., Yip, G., & Johnson, G. (2009). Measuring organizational performance: Towards methodological best practice. *Journal of Management*, *35(*3). 10.1177/0149206308330560.

Roberts, J., & Simpson, K. (2016). Stakeholders perspectives on inclusion of students with autism in mainstream schools. *International Journal of Inclusive Education*, *20*(16), 1084-1096. doi: 10.1080/13603116.2016.1145267.

Ronda-Pupo, G & Guerras-Martín, L. (2012). Dynamics of the evolution of the strategy concept 1962–2008: A co-word analysis. *Strategic Management Journal*, *33*, 162-188. 10.1002/smj.948.

Ross, J. W., Beath, C., & Quaadgras, A. (2013). You may not need big data after all. *Harvard Business Review*, December.

Russom, P. (2013). *Big data analytics*. TDWI Best Practices Report, Fourth Quarter.

Ryan, S. D., Harrison, D. A., & Schkade, L. L. (2002). Information-technology investment decisions: when do costs and benefits in the social subsystem matter? *Journal of Management Information Systems*, *19*, 85-128.

Sagiroglu, S., & Sinanc, D. (2013). *Big data: A review*. https://ieeexplore.ieee.org/document/6567202

Sarstedt, M., & Mooi, E. (2019). Regression analysis, in *A concise guide to market research*. Springer.

Sathi, A. (2012). *Big data analytics*. MC Press Online LLC.

Saunders, M., Lewis, P., & Thornhill, A. (2007). Research methods for business students (5th ed.). Prentice Hall.

Saunders, M. N., & Lewis, P. (2012). *Doing research in business and management: An essential guide to planning your project*. Pearson Higher Ed.

Saunders, M. N., & Lewis, P. (2015). *Doing research in business and management: An essential guide to planning your project*. Pearson Higher Ed.

Schwab, K. (2019). *The global competitiveness report 2019*. Website: https://www3.weforum.org/docs/WEF_TheGlobalCompetitivenessReport2019.pdf

Seang, T. (2003). *Best practices in KPI*. Paper presented at the National Conference of Key Performance Indicators, Pan Pacific Hotel, Kuala Lumpur, 21-23.

Serhani, M., Dssouli, R., Bouhaddioui, C., Taleb, I., & El Kassabi, H. (2016). *Big data quality: A quality dimensions evaluation*. 2016 Second IEEE International Conference on Cloud and Big Data Computing (CBDCom 2016, 18-21 July, 2016, Toulouse, France). 10.1109/UIC-ATC-ScalCom-CBDCom-IoP-SmartWorld.2016.0122.

Setia, P., & Patel, P. C. (2013). How information systems help create OM capabilities: Consequents and antecedents of operational absorptive capacity. *Journal of Operations Management*, *31*, 409-431.

Shabbir, Q. M., & Gardezi, S. (2020). Application of big data analytics and organizational performance: the mediating role of knowledge management practices. *Journal of Big Data*, *7*(10), 1684-1692. 1186/s40537-020-00317-6.

Sharma, J., Weston, M., Batterham, A., & Spears, I. (2014). Gait retraining and incidence of medial tibial stress syndrome in army recruits. *Medicine and Science in Sports and Exercise*, *46*(9).

Sheng, J., Amankwah-Amoah, J., & Wang, X. (2017). A multidisciplinary perspective of big data in management research. *International Journal of Production Economics*, *191*(June), 97–112.

Song, Z. (2009). Optimising product configurations with a data-mining approach. *International Journal of Production Research*, *47*, 1733-1751. 10.1080/00207540701644235.

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

Strong, D.M., & Volkoff, O. (2010). Management information systems, *MIS Quarterly*, *34*(4), 731-756.

Tabachnick, B., & Fidell, L. L. S. (2007). Using Multivariat Statistics. Allyn and Bacon.

Tabachnick, B., & Fidell, L. L. S. (2016). Using Multivariat Statistics. Allyn and Bacon.

Teddlie, C., & Tashakkori, A. (2003). Major issues and controversies in the use of mixed methods in the social and behavioral sciences, in *Handbook of mixed methods in social and behavioural sciences*. Sage.

Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (long-run) enterprise performance. *Strategic Management Journal*, *28*(13), 1319-1350.

Teece, D. J. (2012). Business models, business strategy and innovation. *Long Range Planning*, *43*(2-3), 172-194.

Teece, D. J., (2014). The foundations of enterprise performance: Dynamic and ordinary capabilities in an (economic) theory of firms. *The Academy of Management Perspectives*, *28*, 328-352.

Teece, D. J., Pisano, G., & Shuen, A. M. Y. (1997). *Dynamic capabilities and strategic management*. Wiley Stable. https://www.jstor.org/stable/3088148

Teece, D. (2000). Strategies for managing knowledge assets: The role of firm structure and industrial context. *Long Range Planning*, *33*(1), 35-54. 1016/S0024-6301(99)00117-X.

Teece, D. (2018). Business models and dynamic capabilities. *Long Range Planning*, *51*(1), 40-49. 10.1016/j.lrp.2017.06.007.

Terry, A. B. (2000). Measuring the flexibility of information technology infrastructure: Exploratory analysis of a construct. *Journal of Management Information Systems*, *17*, 167-208.

Tippins, M. J., & Sohi, R. S. (2003). IT competency and firm performance: is organizational learning a missing link? *Strategic Management Journal*, *24*, 745-761.

Tsai, C. W., Lai, C. F., Chao, H. C., & Vasilakos, A. V. (2015). Big data analytics: a survey. Journal of Big Data, *2*(1), 1–32. https://doi.org/10.1186/s40537-015-0030-3

Tseng, S. (2014). The effect of knowledge management capability and dynamic capability on organizational performance. *Journal of Enterprise Information Management*, 27(2). 10.1108/JEIM-05-2012-0025.

Turcotte S. (2010). Interventions for improving the adoption of shared decision-making by healthcare professionals. *Cochrane Database System Review*. https://pubmed.ncbi.nlm.nih.gov/20464744/

Usai, A., Fiano, F., Petruzzelli, A., Paoloni, P., Briamonte, M., & Orlando, B. (2021). Unveiling the impact of the adoption of digital technologies on firms' innovation performance. *Journal of Business Research*, *133*, 327-336. 10.1016/j.jbusres.2021.04.035.

Vaill, P. B. (1996). *Learning as a way of being: Strategies for survival in a world of permanent white water*. Jossey-Bass.

Vargo, S., & Lusch, R. (2004). Evolving to a new dominant Logic: The service-dominant logic of marketing. *Dialog, debate, and directions*, *68*, 1-17.

Verhoef, P., & Lemon, K. (2013). Successful customer value management: Key lessons and emerging trends. *European Management Journal*, *31*, 1–15. 10.1016/j.emj.2012.08.001.

Verhoef, P., Neslin, S., & Vroomen, B. (2007). Multichannel customer management: Understanding the research-shopper phenomenon. *International Journal of Research in Marketing*, *24*(2), 129-148. 10.1016/j.ijresmar.2006.11.002.

Verhoef, P., Kannan, P. K., & Inman, J. (2015). From multi-channel retailing to omni-channel retailing. *Journal of Retailing*, 91(2), 174-181. 10.1016/j.jretai.2015.02.005.

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

Villalpando, B. (2014). Performance analysis model for big data applications in cloud computing. *Journal of Cloud Computing*, *3*(19). 10.1186/s13677-014-0019-z.

Wah, B., Cheng, X., Wang, Y., & Jin, X. (2015). Significance and challenges of big data research. *Big Data Research*, *2*(2), 59-64. 10.1016/j.bdr.2015.01.006.

Walker, R. S., & Brown, I. T. J. (2019). Big data analytics adoption: A case study in a large South African telecommunications organisation. *SA Journal of Information Management*, *21*(1), 1-10. https://www.researchgate.net/publication/336932690_Big_data_analytics_adoption_A_case_study_in_a_large_South_African_telecommunications_organisation

Walliman, N. (2011). *Research methods: The basics*. Routledge.

Wallis, S. E. (2013). How to choose between policy proposals: A simple tool based on systems thinking and complexity theory. *Emergence: Complexity & Organization*, *15*(3), 94-120, ISSN: 1521-3250.

Wallis, S. E. (2014). Existing and emerging methods for integrating theories within and between disciplines. *Organisational Transformation and Social Change*, *11*(1), 3-24. ISSN: 1477-9633

Walls, C., & Barnard, B. (2020). Success factors of big data to achieve organisational performance: Theoretical perspectives. *Expert Journal of Business and Management*, *8*(1),1-16.

Wamba, F. S., Akter, S., & Bourmont, M. (2017). Quality dominant logic in big data analytics and firm performance. *Business Process Management Journal*, *25*(3). 10.1108/BPMJ-08-2017-0218.

Wamba, F. S., Angappa, G., Papadopoulos, T., & Ngai, E. (2018). Big data analytics in logistics and supply chain management. *The International Journal of Logistics Management*, *29*(2), 478–484.

Wang, C., & Ahmed, P. (2007). Dynamic capabilities: A review and research agenda. *International Journal of Management Reviews*, 9(1), 31-51. 10.1111/j.1468-2370.2007.00201.x.

Wang, G., Gunasekaran, A., Ngai, E.W & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics*, *176*, 98-110.

Wang, Y., Kung, L., Wang, D. R., William Y. C., & Cegielski, C. (2018). An integrated big data analytics-enabled transformation model: Application to health care. *Information & Management*, *55*, 64-79. 10.1016/j.im.2017.04.001.

Wiid, J. & Diggines, C. 2015. *Marketing research* (3rd ed.). Juta.

Wu, L. (2010). Applicability of the resource-based and dynamic-capability views under environmental volatility. *Journal of Business Research*, *63*, 27-31. 10.1016/j.jbusres.2009.01.007.

Xu, Z., Frankwick, G. L., & Ramirez, E. (2016). Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective. *Journal of Business Research*, *69*(5), 1562-1566.

Yasmin, R., Paracha, A., Malik, S., & Azeem, M. (2020). *Impact of organizational justice on employee performance: Mediating role of emotional intelligence: An analysis of public sector organizations of Pakistan.* ResearchGate. https://www.researchgate.net/publication/338409136_Impact_of_Organizational_Justice_on_Employee_Performance_Mediating_Role_of_Emotional_Intelligence_An_Analysis_of_Public_Sector_Organizations_of_Pakistan

Zheng, L. J., Zhang, J. Z., Wang, H., & Hong, J. F. (2022). Exploring the impact of big data analytics capabilities on the dual nature of innovative activities in MSMEs: a data-agility-innovation perspective. *Annals of Operations Research*, *1*(29). doi: 10.1007/ s10479-022-04800-6

Zhong, R. Y., Newman, S. T., Huang, G. Q., & Lan, S. (2016). Big Data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and future perspectives. *Computers & Industrial Engineering*, *101*, 572–591. https://doi.org/10.1016/j.cie.2016.07.013

Zikmund, W. G., Babin, B. J., Carr, J. C., & Griffin, M. (2012). *Business research methods* (8th ed.). Cengage Learning.

Zikopoulos, P., Parasuraman, K., Deutsch, T., Giles, J., & Corrigan, D. (2013). *Harness the power of big data. The IBM big data platform*. McGraw Hill Professional.

APPENDICES

APPENDIX A. SURVEY

Dear Participant,

I would hereby like to invite you to complete a survey provided below as part of the survey I am doing for my Master of Commerce in Industrial & Organisational Psychology at the University of South Africa (UNISA). The purpose of this survey is to assess, The relationship between Big Data (BD), big data analytics capabilities (BDAC) and organisational performance (OP). The sole purpose of this study is to obtain information from employees in the South African technology industry, such as yourself to determine the nature of your everyday experience related to the research topic.

Please note that your participation is entirely voluntary, and you are free to decline to participate in this survey. This research study by UNISA Student, Renee Naicker, has been approved by University of South Africa (UNISA), ethical clearance certificate reference number 2021/CEMS/IOP/025 and will be conducted according to the accepted and applicable UNISA CEMS/IOP ethics review committee with applicable ethics guidelines and principles. The survey is anonymous and response data will only be analysed at aggregate level.

SECTION ONE: DEMOGRAPHIC DATA

Listed below is a series of different personal information that is required for the research. The demographical information requested for this study is not for individual identification but for statistics analysis purposes and for validating data.

Please highlight or tick the appropriate box in relation to the details requested.

1.1 Indicate your gender.

Female	Male	Other	Do not want to disclose
1	2	3	4

1.2 Your age category.

18-20	21-25	26-35	36-45	46-55	56-60	60+
1	2	3	4	5	6	7

1.3 Education

No formal education	Matric	Diploma	Bachelor	Honours / Postgraduate	Masters	Doctorate
1	2	3	4	5	6	7

1.4 Position in the organisation.

BI Consultant	1
Technical Business Architect	2
Project Manager	3
Product Manager	4
Data Analytics Expert	5
Business Analyst	6
System Analyst	7

Executive	8
Operational User	9
Technology Specialist	10
Other, and please specify (Insert textbox for answers on the online form)	11

1.5 How many years have you worked for this specific organisation?

Less than 2 years	2-5 years	6-10 years	11-15 years	16-20 years	20 years +
1	2	3	4	5	6

1.6 Number of employees in your organisation? (Fulltime)

50 or fewer	1
51-100	2
101-250	3
251-500	4
501-1000	5
1001-2000	6
More than 2000	7

SECTION TWO: BIG DATA ANALYTICS PLANNING

This section is based on your experience towards Big Data Analytics Planning in the organisation. Please highlight or tick the appropriate box in relation to the details requested.

- Strongly disagree = 1
- Disagree = 2
- Neither agree nor disagree = 3
- Agree = 4
- Strongly agree = 5

Big Data Analytics Planning					
Please indicate your response regarding the following statements	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
1. We continuously examine the innovative opportunities for the strategic use of Big Data Analytics.					
2. We enforce adequate plans for the introduction and utilisation of Big Data Analytics.					
3. We perform Big Data Analytics planning processes in systematic and formalised ways.					
4. We frequently adjust Big Data Analytics plans to better adapt to changing conditions.					

Data Analytics Investment					
Please indicate your response the following statements.	**1**	**2**	**3**	**4**	**5**
In my organisation…………					
5. When we invest in big data analytics, we consider the impact on staff productivity.					
6. Big data analytics help end-users make quicker decisions.					
7. Big data analytics thinks about the training that end-users will need.					
8. Big data analytics investment considers change management.					

Big Data Analytics Resources					
Please indicate your response regarding the following statements.	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
9. The responsibility for Big Data Analytics development is clear.					
10. We are confident that big data analytics project proposals are properly appraised.					
11. We constantly monitor the performance of the Big Data Analytics function.					
12. Our analytics department is clear about its performance criteria.					

Connectivity					
Please indicate your response regarding the following statements.	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
13. It has the foremost available analytics systems.					
14. All remote, branch, and mobile offices are connected to the central office for analytics.					
15. It utilises open systems network mechanisms to boost analytics connectivity.					
16. There are no identifiable communications bottlenecks within our organisation when sharing analytics insights.					

Big Data Analytics Capability					
Please indicate your response regarding the following statements.	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
17. Our user interfaces provide transparent access to all platforms and applications.					
18. Analytics-driven information is shared seamlessly across our organisation, regardless of the location.					
19.It provides multiple analytics interfaces or entry points for external end-users.					

System Design					
Please indicate your response regarding the following statements.	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
20. Reusable software modules are widely used in new analytics model development.					
21. End-users utilize object-oriented tools to create their own analytics applications.					
22. Object-oriented technologies are utilized to minimize the development time for new analytics applications.					
23. Applications can be adapted to meet a variety of needs during analytics tasks.					

Technical Knowledge					
Please indicate your response regarding the following statements.	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
24. Our analytics resources(staff) are very capable in terms of managing project life cycles.					
25. Our analytics resources (staff) are very capable in the areas of data and network management and maintenance.					
26. Our analytics resources (staff) create very capable decision support systems driven by analytics.					

Technology Management Knowledge					
Please indicate your response regarding the following statements.	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
27. Our analytics resources (staff) show superior understanding of technological trends.					
28. Our analytics resources (staff) show superior ability to learn new technologies.					
29. Our analytics resources (staff) are very knowledgeable about the critical factors for the success of our organisation.					
30. Our analytics resources (staff) are very knowledgeable about the role of big data analytics as a means, not an end.					

Organisational Knowledge					
Please indicate your response regarding the following statements.	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
31. Our analytics resources (staff) understand our organisation's policies and plans at a very high level.					
32. Our analytics resources (staff) are very capable in interpreting organisational problems and developing appropriate technical solutions.					
33. Our analytics resources (staff) are very knowledgeable about organisational functions.					

34 Our analytics resources (staff) are very knowledgeable about the organisation's industry.					

Relational Knowledge					
Please indicate your response regarding the following statements.	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
35. Our analytics resources (staff) are very capable in terms of planning, organising, and leading projects.					
36. Our analytics resources (staff) are very capable in terms of planning and executing work in a collective environment.					
37. Our analytics resources (staff) are very capable in terms of teaching others.					
38. Our analytics resources (staff) work closely with customers and maintain productive user/client relationships.					

SECTION THREE: ORGANISATIONAL PERFORMANCE

This section is based on your feelings towards organisational performance in the organisation. Please highlight or tick the appropriate box in relation to the details requested.

- Strongly disagree = 1
- Disagree = 2
- Neither agree nor disagree = 3
- Agree = 4
- Strongly agree = 5

Organisational Performance					
Please indicate your response regarding the following statements.	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
1. The Big Data Analytics plan aligns with the organisation's mission, goals, objectives, and strategies.					
2. The Big Data Analytics plan contains quantified goals and objectives.					
3. The Big Data Analytics plan contains detailed action plans/strategies that support organisational direction.					
4. We prioritize major Big Data Analytics investments by the expected impact on organisational performance.					
5. Using Big Data Analytics improved customer retention during the last 3 years relative to competitors.					
6. Using Big Data Analytics improved sales growth during the last 3 years relative to competitors.					

7. Using Big Data Analytics improved profitability during the last 3 years relative to competitors.					
8. Using Big Data Analytics improved return on investment during the last 3 years relative to competitors.					

Thank you for taking the time to complete this survey. I truly value the information you have provided.

APPENDIX B. SURVEY MEASURES

Section	Questions	Sources
Big data analytics planning	1. We continuously examine the innovative opportunities for the strategic use of Big Data Analytics.	Karimi et al., 2001 Kim et al., 2012
	2. We enforce adequate plans for the introduction and utilisation of Big Data Analytics.	
	3. We perform Big Data Analytics planning processes in systematic and formalised ways.	
	4. We frequently adjust Big Data Analytics plans to better adapt to changing conditions.	
Data analytics investment	5. When we invest in big data analytics, we consider the impact on staff productivity.	Kim et al., 2012 Ryan et al., 2002
	6. Big data analytics help end-users make quicker decisions.	
	7. Big data analytics thinks about the training that end-users will need.	
	8. Big data analytics investment considers change management.	
Big data analytics resources	9. The responsibility for Big Data Analytics development is clear.	Karimi et al., 2001 Kim et al., 2012
	10. We are confident that big data analytics project proposals are properly appraised.	
	11. We constantly monitor the performance of the Big Data Analytics function.	
	12. Our analytics department is clear about its performance criteria.	
Connectivity	13. It has the foremost available analytics systems.	Kim et al., 2012 Terry, 2000
	14. All remote, branch, and mobile offices are connected to the central office for analytics.	
	15. It utilises open systems network mechanisms to boost analytics connectivity.	

	16. There are no identifiable communications bottlenecks within our organisation when sharing analytics insights.	
Big data analytics capability	17. Our user interfaces provide transparent access to all platforms and applications.	Kim et al., 2012 Terry, 2000
	18. Analytics-driven information is shared seamlessly across our organisation, regardless of the location.	
	19.It provides multiple analytics interfaces or entry points for external end-users.	
System design	20. Reusable software modules are widely used in new analytics model development.	Kim et al., 2012 Terry A. B., 2000
	21. End-users utilize object-oriented tools to create their own analytics applications.	
	22. Object-oriented technologies are utilized to minimize the development time for new analytics applications.	
	23. Applications can be adapted to meet a variety of needs during analytics tasks.	
Technical knowledge	24. Our analytics resources(staff) are very capable in terms of managing project life cycles.	Kim et al.,2012 Terry, 2000
	25. Our analytics resources (staff) are very capable in the areas of data and network management and maintenance.	
	26. Our analytics resources (staff) create very capable decision support systems driven by analytics.	
Technology management knowledge	27. Our analytics resources (staff) show superior understanding of technological trends.	Kim et al., 2012 Terry, 2000 Tippins & Sohi, 2003
	28. Our analytics resources (staff) show superior ability to learn new technologies.	
	29. Our analytics resources (staff) are very knowledgeable about the critical factors for the success of our organisation.	

	30. Our analytics resources (staff) are very knowledgeable about the role of big data analytics as a means, not an end.	
Relational knowledge	35. Our analytics resources (staff) are very capable in terms of planning, organising, and leading projects.	Jiang et al., 2003 Kim et al., 2012 Terry, 2000
	36. Our analytics resources (staff) are very capable in terms of planning and executing work in a collective environment.	
	37. Our analytics resources (staff) are very capable in terms of teaching others.	
	38. Our analytics resources (staff) work closely with customers and maintain productive user/client relationships.	
Organisational performance	39. The Big Data Analytics plan aligns with the organisation's mission, goals, objectives, and strategies.	Setia & Patel, 2013 Tippins & Sohi, 2003
	40. The Big Data Analytics plan contains quantified goals and objectives.	
	41. The Big Data Analytics plan contains detailed action plans/strategies that support organisational direction.	
	42. We prioritize major Big Data Analytics investments by the expected impact on organisational performance.	
	43. Using Big Data Analytics improved customer retention during the last 3 years relative to competitors.	
	44. Using Big Data Analytics improved sales growth during the last 3 years relative to competitors.	
	45. Using Big Data Analytics improved profitability during the last 3 years relative to competitors.	

	46. Using Big Data Analytics improved return on investment during the last 3 years relative to competitors.	

APPENDIX C. ETHICAL CLEARANCE

UNISA university of south africa

UNISA CEMS/IOP RESEARCH ETHICS REVIEW COMMITTEE

23 September 2021

Dear Mrs Renee Naicker,

Decision: Ethics approval from 23 September 2021 to 23 September 2024

NHREC Registration # : (if applicable)	
ERC Reference # :	***2021/CEMS/IOP/025***
Name :	Mrs Renee Naicker
Student #:	46492968
Staff #:	N/a

Researcher(s): Name: Mrs Renee Naicker
Address: Johannesburg, South Africa
E-mail address, telephone: 46492968@mylife.unisa.ac.za, 0835072860

Supervisor (s):) Name: Prof Aden-Paul Flotman
Address: Unisa, Muckleneuk Campus, Preller Street, Pretoria, 0003
E-mail address, telephone: flotma@unisa.ac.za, 0124294879

The development and validation of a Big Data and Big Data Analytics assessment instrument for the South African context.

Qualification: Masters (MCom)- Post graduate degree

Thank you for the application for research ethics clearance by the Unisa CEMS/IOP Research Ethics Review Committee for the above-mentioned research. Ethics approval is granted for **Three** years.

*The **low risk application** was **reviewed** by the CEMS/IOP Research Ethics Review Committee on the 31st August 2021 in compliance with the Unisa Policy on Research Ethics and the Standard Operating Procedure on Research Ethics Risk Assessment. The decision was approved on 23rd September 2021.*

The proposed research may only commence with the provision that:

1. ***The researcher will ensure that the research project adheres to the relevant guidelines set out in the Unisa COVID-19 Position Statement on research ethics dated 26 June 2020 which is attached.***

University of South Africa
Preller Street, Muckleneuk Ridge, City of Tshwane
PO Box 392 UNISA 0003 South Africa
Telephone: +27 12 429 3111 Facsimile: +27 12 429 4150
www.unisa.ac.za

2. The researcher(s) will ensure that the research project adheres to the values and principles expressed in the Unisa Policy on Research Ethics.
3. Any adverse circumstance arising in the undertaking of the research project that is relevant to the ethicality of the study should be communicated in writing to the Unisa CEMS/IOP Research Ethics Review Committee.
4. The researcher(s) will conduct the study according to the methods and procedures set out in the approved application.
5. Any changes that can affect the study-related risks for the research participants, particularly in terms of assurances made with regards to the protection of participants' privacy and the confidentiality of the data, should be reported to the Committee in writing, accompanied by a progress report.
6. The researcher will ensure that the research project adheres to any applicable national legislation, professional codes of conduct, institutional guidelines and scientific standards relevant to the specific field of study. Adherence to the following South African legislation is important, if applicable: Protection of Personal Information Act, no 4 of 2013; Children's act no 38 of 2005 and the National Health Act, no 61 of 2003.
7. No field work activities may continue after the expiry date **(23 September 2024)**
8. Submission of a complete research ethics progress report will constitute an application for the renewal of Ethics Research Committee approval.

Note:
*The reference number **2021_CEMS/IOP_025** should be clearly indicated on all forms of communication with the intended research participants, as well as with the Committee.*

Yours sincerely,

Vieker t

Signature
Chair of IOP ERC
E-mail : vnieka2@unisa.ac.za

Tel: (012) 429-8231

Signature
Acting Executive Dean : CEMS
E-mail: Mpofurt@unisa.ac.za

Tel: (012) 429-4808

APPENDIX D. PILOT SURVEY

SECTION ONE: DEMOGRAPHIC DATA

1.1 Indicate your gender.

Female	Male	Other	Do not want to disclose
1	2	3	4

1.2 Your age category.

18-20	21-25	26-35	36-45	46-55	56-60	60+
1	2	3	4	5	6	7

1.3 Education

No formal education	Matric	Diploma	Bachelor	Honours / Postgraduate	Masters	Doctorate
1	2	3	4	5	6	7

1.4 Position in the organisation.

BI Consultant	1
Technical Business Architect	2
Project Manager	3
Product Manager	4
Data Analytics Expert	5
Business Analyst	6
System Analyst	7
Executive	8
Operational User	9
Technology Specialist	10

Other, and please specify (Insert textbox for answers on the online form)	11

1.5 How many years have you worked for this specific organisation?

Less than 2 years	2-5 years	6-10 years	11-15 years	16-20 years	20 years +
1	2	3	4	5	6

1.6 Number of employees in your organisation? (Fulltime)

50 or fewer	1
51-100	2
101-250	3
251-500	4
501-1000	5
1001-2000	6
More than 2000	7

SECTION TWO: BIG DATA ANALYTICS PLANNING

Big Data Analytics Planning					
Please indicate your perceptions regarding the following statements	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
1. We continuously examine the innovative opportunities for the strategic use of Big Data Analytics.					
2. We enforce adequate plans for the introduction and utilisation of Big Data Analytics.					

3. We perform Big Data Analytics planning processes in systematic and formalized ways.					
4. We frequently adjust Big Data Analytics plans to better adapt to changing conditions.					
Data Analytics Investment					
Please indicate your perceptions regarding the following statements.	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
5. We think about and estimate the effect they will have on the productivity of the employees' work.					
6. We consider and project about how much these options will help end-users make quicker decisions.					
7. We think about and estimate the cost of training that end-users will need.					
8. We consider and estimate the time managers will need to spend overseeing the change.					

Big Data Analytics Resources					
Please indicate your perceptions regarding the following statements.	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
9. The responsibility for Big Data Analytics development is clear.					
10. We are confident that big data analytics project proposals are properly appraised.					

11. We constantly monitor the performance of the Big Data Analytics function.					
12. Our analytics department is clear about its performance criteria.					

Connectivity					
Please indicate your perceptions regarding the following statements.	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
13. It has the foremost available analytics systems.					
14. All remote, branch, and mobile offices are connected to the central office for analytics.					
15. It utilizes open systems network mechanisms to boost analytics connectivity.					
16. There are no identifiable communications bottlenecks within our organisation when sharing analytics insights.					

Big Data Analytics Compatibility					
Please indicate your perceptions regarding the following statements.	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
17. Our user interfaces provide transparent access to all platforms and applications.					

18. Analytics-driven information is shared seamlessly across our organisation, regardless of the location.					
19.It provides multiple analytics interfaces or entry points for external end-users.					

System Design					
Please indicate your perceptions regarding the following statements.	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
20. Reusable software modules are widely used in new analytics model development.					
21. End-users utilize object-oriented tools to create their own analytics applications.					
22. Object-oriented technologies are utilized to minimize the development time for new analytics applications.					
23. Applications can be adapted to meet a variety of needs during analytics tasks.					

Technical Knowledge					
Please indicate your perceptions regarding the following statements.	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
24. Our analytics resources(staff) are very capable in terms of managing project life cycles.					
25. Our analytics resources (staff) are very capable in the areas of data and network management and maintenance.					
26. Our analytics resources (staff) create very capable decision support systems driven by analytics.					

Technology Management Knowledge					
Please indicate your perceptions regarding the following statements.	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
27. Our analytics resources (staff) show superior understanding of technological trends.					
28 Our analytics resources (staff) show superior ability to learn new technologies.					
29. Our analytics resources (staff) are very knowledgeable about the critical factors for the success of our organisation.					
30. Our analytics resources (staff) are very knowledgeable about the role of big data analytics as a means, not an end.					

Organisational Knowledge					
Please indicate your perceptions regarding the following statements.	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
31. Our analytics resources (staff) understand our organisation's policies and plans at a very high level.					
32. Our analytics resources (staff) are very capable in interpreting organisational problems and developing appropriate technical solutions.					
33. Our analytics resources (staff) are very knowledgeable about organisational functions.					

34 Our analytics resources (staff) are very knowledgeable about the organisation's industry.					

Relational Knowledge					
Please indicate your perceptions regarding the following statements.	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
35. Our analytics resources (staff) are very capable in terms of planning, organising, and leading projects.					
36. Our analytics resources (staff) are very capable in terms of planning and executing work in a collective environment.					
37. Our analytics resources (staff) are very capable in terms of teaching others.					
38. Our analytics resources (staff) work closely with customers and maintain productive user/client relationships.					

SECTION THREE: ORGANISATIONAL PERFORMANCE

Organisational performance					
Please indicate your perceptions regarding the following statements.	**1**	**2**	**3**	**4**	**5**
In my organisation …………					
1. The Big Data Analytics plan aligns with the organisation's mission, goals, objectives, and strategies.					
2. The Big Data Analytics plan contains quantified goals and objectives.					

3. The Big Data Analytics plan contains detailed action plans/strategies that support organisational direction.					
4. We prioritize major Big Data Analytics investments by the expected impact on organisational performance.					
5. Using Big Data Analytics improved customer retention during the last 3 years relative to competitors.					
6. Using Big Data Analytics improved sales growth during the last 3 years relative to competitors.					
7. Using Big Data Analytics improved profitability during the last 3 years relative to competitors.					
8. Using Big Data Analytics improved return on investment during the last 3 years relative to competitors.					

Thank you for taking the time to complete this survey. I truly value the information you have provided.

APPENDIX E. LANGUAGE-EDITING CONFIRMATION

Certificate of Editing

R.J. Thompson

Editing and proofreading of theses and manuscripts

143 The Landmark, 31 Dover Street, Ferndale,
Johannesburg, 2194
rjthompson84@hotmail.com
richardt@regenesys.net
Cell: 082-890-5264

13 December 2022

This is to certify that:

I, Richard James Thompson, identity number 630722 5095 085, am employed by Regenesys Business School, 165 West Street, Sandton, as a copy editor. For the most part I edit study guides and other learning material for Regenesys.

In my spare time I undertake private editing work. In that capacity I have edited Ms Renee Naicker's Unisa M.Com. dissertation titled "The Validation of a Big Data Analytics Capability Scale for the South African Context."

The dissertation conforms to the requirements of the APA 7 style guide.

Should there be any queries, please contact me on the cell phone number or at any of the addresses indicated above.

Yours faithfully

R. J. Thompson

Richard Thompson

APPENDIX F. TURN-IT-IN CERTIFICATE/REPORT

Document Viewer

Turnitin Originality Report

Processed on: 09-Jan-2023 11:16 SAST
ID: 1990128977
Word Count: 41993
Submitted: 1

Revision 1 By Renee Naicker

Similarity Index	Similarity by Source	
18%	Internet Sources:	19%
	Publications:	11%
	Student Papers:	9%

include quoted | include bibliography | excluding matches < 1% | mode: quickview (classic) report | print | download

2% match (Internet from 08-Feb-2019)
https://repository.up.ac.za/bitstream/handle/2263/64902/Niland_Toward_2018.pdf?isAllowed=y&sequence=1

2% match (Internet from 30-Jun-2021)
https://repository.up.ac.za/bitstream/handle/2263/79645/Qwabe_Big_2020.pdf?sequence=

2% match (Internet from 08-Feb-2019)
https://repository.up.ac.za/bitstream/handle/2263/64890/Mourinho_Leveraging_2018.pdf?isAllowed=y&sequence=1

2% match (Internet from 14-Oct-2022)
https://ro.uow.edu.au/cgi/viewcontent.cgi?article=2021&context=buspapers

1% match (Internet from 15-Dec-2022)
https://repository.up.ac.za/bitstream/handle/2263/68899/Balkissoon_Big_2019.pdf?isAllowed=y&sequence=1

1% match (Internet from 05-Sep-2017)
http://uir.unisa.ac.za

1% match (Internet from 02-Nov-2020)
http://uir.unisa.ac.za

1% match (Internet from 10-Sep-2022)
https://uir.unisa.ac.za/bitstream/handle/10500/27848/dissertation_tlhagale_fk.pdf?sequence=

1% match (Internet from 22-Nov-2022)
https://uir.unisa.ac.za/bitstream/handle/10500/28311/dissertation_pikanegore_r.pdf?isAllowed=y&sequence=2

1% match (Internet from 15-Nov-2021)
https://www.oecd-ilibrary.org/docserver/9789264167179-6-en.pdf?accname=guest&checksum=A43D55CB99D5722561F3E2E1EADB75BA&expires=1637026246&id=id

1% match (student papers from 08-Jun-2020)
Submitted to University of the Western Cape on 2020-06-08

1% match (Internet from 26-Feb-2014)
http://www.gwu.edu

1% match (Internet from 19-Aug-2022)
http://eprints.lincoln.ac.uk

1% match (Internet from 12-Oct-2020)
https://repository.nwu.ac.za/bitstream/handle/10394/35923/Mokaba_MS.pdf?isAllowed=y&sequence=1

1% match (Eleonora Gabriela Conţu. "Organizational performance – theoretical and practical approaches; study on students' perceptions", Proceedings of the International Conference on Business Excellence, 2020)
Eleonora Gabriela Conţu. "Organizational performance – theoretical and practical approaches; study on students' perceptions", Proceedings of the International Conference on Business Excellence, 2020

1% match (student papers from 19-May-2022)
Submitted to Kwame Nkrumah University of Science and Technology on 2022-05-19