Gordon Institute of Business Science
University of Pretoria

The influence of moral intensity on decision-making in the context of an artificial intelligence-based online personalised pricing model

20803576

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

02 November 2021

Abstract

This research sought to describe the influence that moral intensity has on managers' decision-making in the context of artificial intelligence (AI)-based online personalised pricing models. Moral intensity is a construct from the issue-contingent model, a decision-making framework stating that the more morally intense an issue, the more moral the judgement applied in decision-making. Jones (1991) described moral intensity as the proportionality of moral responsibility. Of the six factors of moral intensity in the issue-contingent model, three were used in this research: social consensus, magnitude of consequences and likelihood of effect. A hypothetical scenario was described about an online AI-based personalised pricing model for groceries. Experimental vignette methodology was used, in which eight vignettes were described with varying levels of moral intensity and questions on moral judgement were posed on each vignette. Personal characteristics of the decision-maker were also captured to account for variation they may cause in decision-making. Univariate analyses of variance and covariance were conducted. Findings were that personal characteristics have no influence on decision-making in this context, but each of the factors of moral judgement do. Implications are that moral decision-making in the use of AI-based online personalised pricing models can be improved by increasing the awareness of probable consequences and of the social opinion on whether these types of models are considered fair.

Keywords

Decision-making; moral intensity; issue-contingent; artificial intelligence; online personalised pricing

Declaration

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

Abigail Britton

Contents

1. Introduction to research problem 1

1.1 Background 1
1.2 Use of artificial intelligence 1
1.3 Real-world ethics concerns 2
1.2.1 Data privacy 2
1.2.2 Discrimination 2
1.2.3 Personalisation 4
1.4 Online personalised pricing 5
1.3.1 Types of price discrimination 5
1.3.2 Enablement of price discrimination by AI 6
1.3.3 Surplus distribution under price discrimination 7
1.3.4 Price discrimination in the digital economy 8
1.4 Research problem 9
1.5 Research objectives and scope 10

2. Literature review 12

2.1 Ethics and decision-making 12
2.2.1 Business ethics 12
2.2.2 Ethical decision-making 13
2.2 Issue-contingent model 17
2.3.1 Moral intensity 18
2.2.2 Influence of moral intensity on decision-making 20
2.3 Ethical decision-making and AI 22
2.4.2 On the fairness of price discrimination 23
2.4 Conclusion 24

3. Hypotheses 26

4. Research methodology and design 28

4.1 Methodology 28
4.1.1 Approach to theory development 28
4.1.2 Experimental vignette methodology 29
4.2 Population and unit of analysis 32
4.3 Sampling method and size 32
4.3.1 Sampling methodology 33

4.3.2 Sample size 33
4.4 Measurement instrument 34
4.5 Data transformations 36
4.6 Analysis approach 38
4.7 Quality controls 41
4.8 Design limitations 43
4.9 Conclusion 44
5. Results 46
5.1 Sample analysis 46
5.1.1 Descriptive statistics 46
5.1.2 Population representativity 47
5.2 Quality of the measurement instrument 48
5.2.1 Scale reliability 48
5.2.2 Validity of the data 50
5.2.3 Exploratory factor analysis 51
5.3 Statistical results 53
5.3.1 Influence of personal characteristics on moral judgement 53
5.5.2 Influence of moral intensity on recognition of an issue as a moral issue 57
5.3.3 Factors of moral intensity positively influence moral judgement 57
5.4 Conclusion 59
6. Discussion of results 61
6.1 Influence of personal characteristics on moral judgement 61
6.1.1 Personal characteristics 61
6.2 Influence of moral intensity on recognition of an moral issues 63
6.2.1 Moral and non-moral problems 63
6.3 Influence of moral intensity on moral judgement 64
6.3.1 Social consensus 64
6.3.2 Magnitude of consequences 65
6.3.3 Likelihood of effect 66
6.4 Influence of interaction effects on moral judgement 66
6.5 Conclusion 67
7. Conclusions and recommendations 69
7.1 Principle conclusions 69
7.2 Theoretical contribution 70

7.3 Implications 70
7.4 Limitations of the research 71
7.4.1 Design limitations 71
7.4.2 Data and analysis limitations 72
7.5 Suggestions for further research 73
8. References 74
9. Appendices 81
9.1 Full survey 81
9.2 Exploratory factor analysis results 86

List of figures

Figure 1: First degree price discrimination 6
Figure 2: Generalised workflow for an online personalised pricing model 7
Figure 3: The interactionist model of ethical decision-making 14
Figure 4: The Rest (1986) four-step ethical decision-making model 15
Figure 5: Jones' (1991) description of the ethical decision-making process 18
Figure 6: The issue contingent model 20
Figure 7: Conceptual model 39
Figure 8: Descriptive statistics of the sample 46
Figure 9: Spree plot of eigenvalue per factor number 52
Figure 10: Distribution of moral judgement per personal characteristic 53
Figure 11: Distribution of moral judgement 54

List of tables

Table 1: Kohlberg and Hersh's theory of moral development 13
Table 2: Personal and situational characteristics in ethical decision-making 16
Table 3: Vignette descriptions for high and low moral intensity 30
Table 4: Survey structure 35
Table 5: Analysis approach for each hypothesis 40
Table 6: Interaction effects of three factors 41
Table 7: Sample gender distribution for managerial levels 48
Table 8: Covariates recoded for analysis 37
Table 9: Independent variables recode logic 38
Table 10: Reliability statistics 49
Table 11: Cronbach's alpha to test reliability 49
Table 12: Pearson correlation to test validity 50
Table 13: Total variances explained 51
Table 14: Shapiro-Wilk tests of normality 55
Table 15: Levene statistic for tests of variance homogeneity 56
Table 16: The influence of demographic characteristics on moral judgement 56
Table 17: The influence of moral intensity on the recognition moral issues 57
Table 18: The influence of moral intensity on moral judgement 58
Table 19: The influence of interaction effects on moral judgement 59

1. Introduction to research problem

1.1 Background

The use of machine learning (ML) and artificial intelligence (AI) systems has grown rapidly, making it easier and less costly for companies to collect and process various types of information about individuals (Steinberg, 2020). Some companies are using such information to offer consumers online personalised pricing for goods and services. Understanding and explaining ethical issues associated online personalised pricing models and managers' decision-making in the context of these issues can be beneficial both for academic researchers and managers in business. For researchers, it is important that they identify factors associated with the ethical decision-making process. For managers in business, it is important they consider ethical issues in their decision-making processes. The primary purpose of this research is to describe the influence that moral intensity has on managers' decision-making in the context of an artificial intelligence (AI)-based online personalised pricing model.

1.2 Use of artificial intelligence

AI can be broadly described as computer programs developed to "simulate the intelligent behaviour" of humans (Piano, 2020). Fundamental to AI are ML algorithms, which learn from data and adjust their behaviour without being explicitly programmed to do so, through a mechanism called reinforcement learning (Piano, 2020; Liao, 2020). The concept of ML is not new. Samuel (1959) explored the use of ML and AI in games of checkers and concluded that computers can learn to play better games of checkers than the humans who initially programmed the computers. Some kinds of AI are programmed to act autonomously, such as chatbots interacting with people in real time. Others rely on human intervention to take the final action, such as granting a credit facility. Generally, the inclusion of AI in processes is beneficial for two reasons: firstly, it becomes possible to automate processes to a large degree; and secondly, it is a means to reduce biases that are inherent in humans in decision-making processes (Martin, 2019).

According to (Capgemini, 2020), AI is a transformational technology. In the COVID-19 pandemic alone, AI has played a role in limiting the spread of the virus by using voice-activated interfaces and in delivering food and medication by delivery robots and drones (Capgemini, 2020). The potential for AI to cause harm, however, is equally indisputable as many countries are exploring the use of autonomous weapons systems, which use sensors and algorithms to engage weaponry without any human intervention at all (Pandya, 2019; Liao, 2020). Examples of this type of weaponry are physical arms deployed in geophysical space, such as missiles, or may be virtual weaponry deployed in cyberspace, such as computer viruses.

1.3 Real-world ethics concerns

As the use of AI becomes more widespread, current and near-future ethical concerns become more pressing (Laio, 2020). The scope for ethical considerations in AI spans the entire algorithm development and deployment lifecycle: from data collection, feature selection and modelling to testing, deployment and revision. Outlined below are key concerns for ethical AI practice from various researchers.

1.2.1 Data privacy

The predictive power of algorithms is limited to the data available to them and this results in incentives for firms to buy or harvest as much data as they can (Liao, 2020). Ethical concerns related to data privacy can be segmented into three categories: *collection* of personal information which individuals may or may not be aware of; *blending* of data from disparate sources to reveal new facts about individuals; and *security* of locations where data is stored and processed to avoid data leaks (Mehmood, Natgunanathan, Xiang, Hua & Guo 2016).

1.2.2 Discrimination

A major benefit of using AI in decision-making is the ability to reduce bias that is inherent in human decision-making (Martin, 2019). When biased data and assumptions form the input for algorithms, however, the ability of algorithms to make fair and unbiased recommendations is diminished (Ahsen et al., 2019).

Placement of digital adverts are the result of algorithms, programmed to optimise reach and impressions of the advert. Lambrecht and Tucker (2019) found that an algorithm that was intended to simply optimise the cost effectiveness of a digital job advert, ended up displaying the advert to far fewer women than it did men. The advert was for employment opportunities in the science, technology, engineering and mathematics fields. This gender-bias was the result of women being a “prized demographic” and a more expensive demographic to display adverts to, on digital platforms. Consequently, the algorithm, which was intended to be “gender-neutral”, discriminated against women unexpectedly. Datta, Carl and Datta (2015) concluded similarly from their research. They found that adverts related to high-paying jobs, such as adverts for career coaching services, were more likely to be displayed to men, than to women.

Discriminatory digital advertising is not isolated to gender-based discrimination. Sweeney (2013) found that digital adverts for criminal record checks and adverts with other negative sentiment were disproportionately displayed upon the Google search of “black-sounding” names, compared to “white-sounding” names. The implication of this is that potential employers of people with black-sounding names are faced with adverts suggestive of arrest and other negative connotations, upon doing a Google search of these people.

The use AI for some medical diagnoses has been proven to be as accurate as that of medical professionals (Heinrichs & Eickhoff, 2020) though it does raise some concerns when diagnoses can be skewed by patients clinical information. Beyond diagnosis, algorithms are also used in the provision of care. Obermeyer and Powers (2019) found that a commercially-used predictive algorithm to determine the healthcare needs of patients dramatically underestimated the needs of black patients. This racial bias was the result of the algorithm using historical healthcare costs as a proxy for healthcare needs. Less money was historically spent on care of black patients and, thus, the algorithm, incorrectly, deducted that black patients were healthier than white patients and required less care when sick.

Algorithms can perpetrate gender, racial and other types of discrimination, oftentimes unintendedly so. This discrimination can be the result of historic social, cultural and institutional biases present in the data (Obermeyer & Powers, 2019; McLennan, Lee, Fiske & Celi, 2020; Liao, 2020) and unchecked assumptions (Lambrecht & Tucker, 2019; Datta et al., 2015; Sweeny, 2013).

1.2.3 Personalisation

Zanker, Rook and Jannach (2019) describe personalisation as the process of creating the most relevant user experience, by tailoring content based on information and assumptions about users' preferences. AI makes personalisation possible through the ability to analyse vast amounts of disparate data to elicit individuals' preferences. Personalisation surrounds consumers in a digital age and the expectations that consumers have of personalisation are growing (Weber, 2018). The applications of personalisation are widespread. Recommender systems form a subset of personalisation systems, where the former specifically recommends products and services based on user preferences, while the latter includes, as an example, personalisation of website attributes, like font and colour, to user preferences (Li & Karahanna, 2015). While there are benefits to personalisation, such as tailored content and product offerings, there are ethical concerns around the collection and analysis of individual-level data and the influence on customer choices.

Concerns around personalisation are that it requires processing of individual-level information, such as individuals' social networks and situational context (Aksoy, Kabadayi, Yilmaz & Alan, 2021). Additionally, personalisation has the "tremendous potential" to influence choices made by individuals (Aksoy et al., 2021). Huang and Zhou (2018) studied the influence that shoppers' motivations and personalised shop interfaces have on online shopping behaviour. The shopper motivations studied included convenience, bargain hunting and idea seeking. Huang and Zhou (2018) concluded that, based on the shoppers' motivations at the time, different interfaces result in different spending behaviours. In practice, when online shops can determine shopper motivation, through collection and analysis of online behavioural data, and tailor the shop interface accordingly, they may be able to influence the spending behaviour of shoppers.

Data privacy, unintended discrimination and personalisation are examples of real-world ethical concerns in the use of AI. In addition to these concerns, online personalised pricing is another, and this is explored from both a theoretical and practical perspective below.

1.4 Online personalised pricing

According to the OECD (2021) online personalised pricing is defined as:

> a form of price discrimination that involves charging different prices to different consumers according to their willingness to pay, where this is estimated from a consumer's personal data (e.g. personal information, search history, or the location, device, or browser from which they access a retailer's website). (p. 7)

This definition makes it clear that online personalised pricing is a form of price discrimination since "different consumers are charged different prices for the same item" (OECD, 2021, p. 7). This price discrimination takes different forms and is influence by several factors as discussed below.

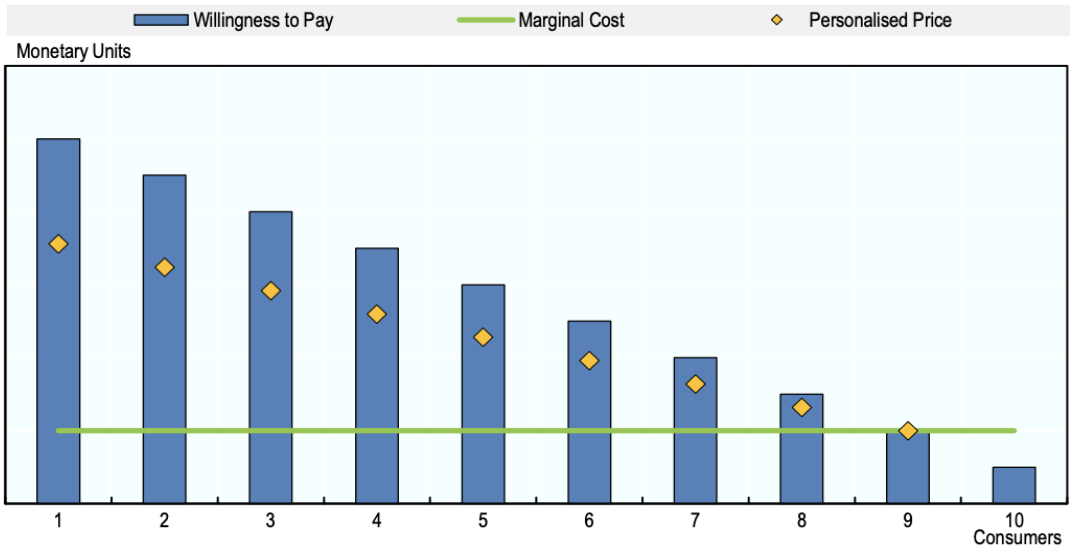
1.3.1 Types of price discrimination

The theory of marginal cost pricing states that prices are set where marginal cost meets marginal revenue, in competitive markets. Economists have identified three types of price discrimination:

1. first-degree price discrimination is when firms are able to charge customers at their individual willingness to pay, also known as perfect price discrimination;
2. second-degree price discrimination involves tiered pricing such that customers receive discounts once they buy certain thresholds, also known as non-linear pricing; and
3. third-degree price discrimination occurs when pricing is differentiated for different groups of customers, such as discounts for students (Botta & Wiedemann, 2020).

An illustration of first-degree price discrimination is represented in the following figure.

Figure 1: First degree price discrimination



Source: OECD (2018)

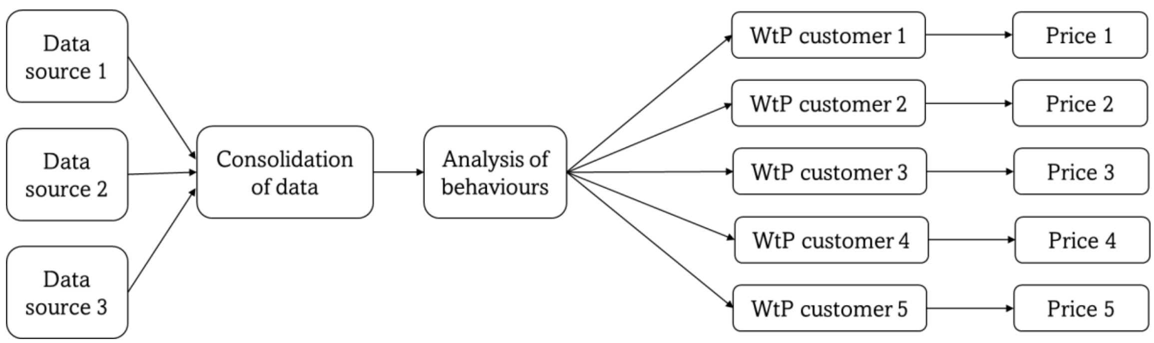
Online personalised pricing involves charging customers at their individual maximum willingness to pay for goods and services and is first-degree price discrimination. In practice, lower-end customers may be charged lower prices and higher-end customers may be charged higher prices than they would be charged in a standard pricing model (Cannataci, Falce & Pollicino, 2020).

Historically first-degree price discrimination has been accepted as impossible to implement due to the data and analyses required to compute willingness to pay at an individual customer level (Botta & Wiedemann, 2020). With the development of AI, however, this is changing quickly. In fact, the White House published a report in 2015, named "Big data and differential pricing", which stated that firms are moving closer to first-degree price discrimination as algorithmic prowess continues to grow (Executive Office of the President, 2015).

1.3.2 Enablement of price discrimination by AI

AI enables the creation of personalised pricing models through the collection and analysis of vast amounts of data at more granular levels than humans can. A general workflow for a personalised pricing model is: data on customers' personal characteristics and behavioural patterns is collected from various sources; this disparate data is consolidated; based on this data, estimations of willingness to pay (WtP) are calculated; and on these estimations, firms select prices charged to customers (OECD, 2018).

Figure 2: Generalised workflow for an online personalised pricing model



While perfect price discrimination may be a stretch even for firms using AI, what AI does allow is for populations of customers to be segmented into small groups based on their behaviours (Botta & Wiedemann, 2020). Many of these behaviours would be otherwise unknown to humans and the effort required to segment customers into infinitesimally small groups would be too substantial to be worthwhile.

1.3.3 Surplus distribution under price discrimination

Theoretically, when individuals are charged at their maximum WtP, there is an efficiency gain to be had from the surplus that is distributed between firms and customers. If the maximum WtP is greater than what the supplier had otherwise intended to charge, then the price difference contributes to the supplier's surplus. Conversely, if the maximum WtP is less than what the supplier would otherwise charge, the price difference contributes to the customers surplus.

It is not to say, however, that this surplus is distributed to firms and customers in equal parts (Cannataci et al., 2020). Ultimately, distribution of surplus between firms and customers depends on market structure and power balance: in a competitive market, customers have the power to shop elsewhere when adversely affected by personalised firm pricing practices; and, conversely, in uncompetitive or monopolistic markets, firms are able to price goods at customers' maximum willingness to pay and capture the entire surplus (Cannataci et al., 2020).

Steinberg (2020) found that customers' abilities to benefit from markets is undermined by personalised pricing, in favour of firm profits. Steinberg considered charging individuals at their personal reserve prices as a potentially equitable means to distribute goods based on individuals' welfare and resources. He stated that this is not so, however, when firms use personalised pricing purely for profit maximisation, rather than for a social end. Steinberg concluded that personalised pricing is an "unethical business practice" because it "exploits and exacerbates a market failure for the purpose of profit maximization".

1.3.4 Price discrimination in the digital economy

The benefit of the internet in creating virtual marketplaces for goods and services is indisputable, but the shift to the digital economy brings new concerns with price discrimination. By tracking online data such as website cookies, geoloactions and IP addresses, firms are able to elicit personal and private information about customers which they can use for price discrimination (Cahn, Alfeld, Barford & Muthukrishnan, 2016). In 2000, Amazon found itself in hot water after a shopper found that if he deleted his cookies, the price on a DVD was reduced (Wong, 2021). Seemingly, new shoppers on Amazon were charged higher prices than regular shoppers on the platform were charged for the same goods. Amazon attempted to explain this anomaly as stemming from its "random discounts experimentation" and defended itself by saying that shoppers who were overcharged were later refunded.

More recently, Hannak, Soeller, Lazer, Mislove and Wilson (2014) created fake shopper accounts to investigate the effect of different online behaviours on prices and they found evidence of price discrimination on a number of the top online stores. Pandey and Caliskan (2021) studied the fares charged for trips from ridehailing platforms, such as Uber. Using a sample of 100 million fares, Pandey and Caliskan (2021) found that "neighborhoods with larger non-white populations, higher poverty levels, younger residents and high education levels are significantly associated with higher fare prices". AI has made first-degree price discrimination possible through the ability to analyse vast amounts of data at customer-level granularity. Additionally, AI-based algorithms are used to collect data online about individuals' characteristics and behaviours and dynamically change the price of goods in online marketplaces. While this type of online personalised pricing may potentially benefit customers, Steinberg (2020) found that this is not the case when the use of such pricing algorithms is for firms' profit-maximisation only.

1.4 Research problem

Various guidelines for ethical AI practice have been published in popular media (Google, n.d; Microsoft Corporation, 2018; McKinsey Global Institute, 2019), but according to Hagendorff (2020) most often these guidelines are relatively vague and superficial, and have no actual impact on decision-making in the field of AI. Challenges such as poor explainability of AI systems (Shin, 2020), no clear accountability (Capgemini, 2020) and immature regulatory frameworks (Butcher & Beridze, 2019) result in poor operationalisation of these practices. Furthermore, these industry-published guidelines lack theoretical grounding and fail to explain the nature of decision-making in the AI field.

Beyond industry-published guidelines, authors have attempted to apply established ethics models to suggest what ethical behaviour comprises. Neubert and Montañez (2020) suggested a virtue framework based on prudence, temperance, justice, courage, faith, hope and love and they validated such a virtue framework against published AI ethics guidelines at Google. While virtue ethics provides a broad framework for ethical behaviour, the design of ethics codes and their relation to virtue dimensions is left to organisations to define for themselves.

Ferrell & Ferrell (2021) explored the idea of programmatically incorporating the normative values of the Hunt-Vitell model into AI algorithms. That is, explicitly coding rules that enable the algorithms itself, to make ethical decisions. The Hunt-Vitell model describes how humans make ethical decisions and because the aim of AI is to replicate, if not improve, human decision-making processes, it may be appropriate. Ferrell and Ferrell (2021), concluded that since the scope of assessment for algorithms is limited to the data they are fed at a particular point in time, the quality of decisions will be limited by such. This challenge applies to incorporating any ethics framework programmatically into algorithms. As such, the onus is on developers of algorithms and their managers to ensure that appropriate ethical considerations have been made before deploying these systems.

Failure to recognise and mitigate ethical consequences of AI solutions results in unintended harm (Sweeny, 2013; Datta, Carl & Datta; 2015; Huang & Zhou, 2018; Lambrecht & Tucker, 2019; Obermeyer & Powers, 2019). Ethical judgment cannot be explicitly programmed into algorithms (Ferrell and Ferrell, 2021) and, thus, a decision-making approach is better for incorporating ethics values into AI solutions. Though many industry-published guidelines exist for ethical AI practice, these are largely viewed as tick-box exercises and have been empirically proven to not impact the decision-making processes of developers (Hagendorff, 2020).

1.5 Research objectives and scope

The primary objective of this research is to describe the nature of managers' ethical decision-making around the use of personalised pricing models. Specifically, decisions of managers in the use AI-based online personalised pricing models will be investigated.

This research builds on the work done by decision-making researchers to find appropriate theoretical frameworks to describe ethical decision-making in the context of AI. The research will provide a framework through which managers can assess AI solutions as ethical or not. This is important because industry-published ethics guidelines have been shown to have little influence on decision-making in the implementation of AI technologies. Managers will be able to use this study's findings to influence their own decisions, as well as those of the AI developers they manage and the senior executives they report to. Regulators may use this research to identify and enforce decision criteria which are proven to influence more ethical decision-making around online personalised pricing. Firms will benefit from reduced reputational risk through increased capability of AI developers to recognise issues related to AI as ethical issues and, thus, make more ethical decisions around them. Consumers and people affected by AI solutions will also benefit if those developing the solutions make ethical considerations before deploying them.

2. Literature review

The field of study for this research is decision-making and, in particular, ethical decision-making in the context of AI. This chapter provides the literature review for the research that follows and is structured as follows: the concepts of business ethics and ethical decision-making are introduced; the issue-contingent model is described and validations of the model are presented; literature on ethical decision-making specifically in the field of AI is explored; and finally, the fairness of personalised pricing models is explored.

2.1 Ethics and decision-making

Ethical decision-making is a construct that has been well studied by academics. In this sub-chapter, foundational understanding of the construct as developed by Kohlberg and Hersch, Trevino and Rest is outlined.

2.2.1 Business ethics

Modern-day management philosopher, Peter Drucker (1981), defines ethics as the “rules of morality”. In normative ethics, these rules of right and wrong are universally applicable, whereas in a descriptive paradigm, rules differ for groups of populations (Crane, Mattan, Glozer & Spence, 2016). There exist many modernist and alternative theories for ethical decision making, with some of the most popular as follows: utilitarianism, the idea of doing what is best for most; egoism, advocating that people are morally obliged to behave in their own self-interest; virtue ethics, which posits that human should adopt virtuous characteristics; and postmodernism, based on listening to one’s instincts (Crane et al., 2016).

Business ethics, specifically, deals with questions of right and wrong in the greater economic environment and in the context of relationships with society (Hoffman & Moore, 1982). In some domains, there is intersection of ethics and the law since both deal with the idea of right and wrong. Crane et al. (2016) describe the law as an “institutionalisation or codification of ethics into specific rules, regulations and proscriptions”.

There is debate about the relationship between ethics and strategy. One explanation of this relationship is the “separation thesis”, which dictates that ethics and business have no relationship, nor should they, emphasising the fiduciary responsibilities that business has to stockholders (Minoja, 2012). Corporate social responsibility scholar, Freedman, is outspoken on his belief that the separation thesis must be rejected because “at its worst, the theory provides a destructive view of capitalism” (Freeman & Ramakrishna Velamuri, 2006). Additionally, business has social responsibilities beyond mere profit maximisation (Wicks, 1996) and multi-stakeholder perspectives create competitive advantages and reinforces better financial performance (Minoja, 2012). While there is general rejection of the separation thesis, the theory may have different interpretations and rejection based on trivial and semantic interpretations may result in valuable insights of the theory being missed, with a reason for varying interpretations being limited formalisation of the theory (Sandberg, 2008).

2.2.2 Ethical decision-making

Kohlberg and Hersh's (1977) theory of moral development is key early work in ethical decision-making. The authors defined moral development as “the transformations that occur in a person’s form or structure of thought”. They theorised that moral judgement develops in individuals over six stages as below.

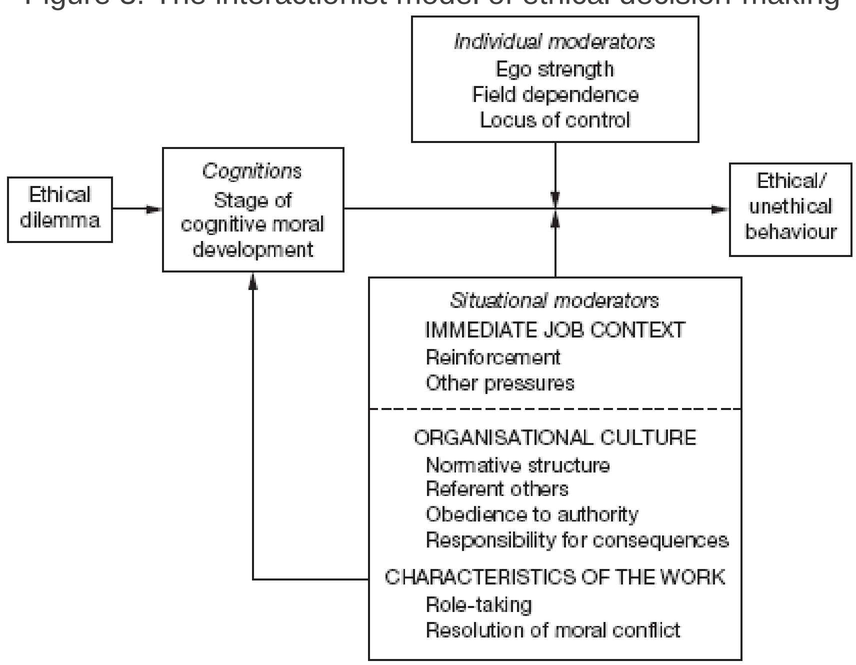
Table 1: Kohlberg and Hersh's theory of moral development

Level	Stage	Description
Preconventional	1 – Punishment and obedience	Difference between good and bad action is determined by what action is punished and what is not
Preconventional	2 – Instrumental relativist	The right actions are those which serve to satisfy one’s needs and, occasionally, the needs of others
Conventional	3 – Interpersonal concordance	Good behaviour is such which is approved and endorsed by others

Conventional	4 – Law and order	The right actions are those which are lawful and according to defined rules
Postconventional	5 – Social contract	Individual's rights and socially accepted standards determine the right behaviour
Postconventional	6 – Universal ethical principle	Individuals define what is right in accordance with their own self-chosen ethical principles

Expanding on the theory of moral development, Trevino (1986) proposed the interactionist model of ethical decision-making in organisations. The model builds on Kohlberg's earlier work by introducing situational factors to the earlier proposed personal factors in the decision-making process. Trevino argued that cognition from moral development stages is not enough to compel an individual to ethical action. Rather, there are additional individual and situation factors that are important. The individual factors are ego strength, field dependence and locus of control, while situational factors relate to the organisation's culture, job context and characteristics of the work.

Figure 3: The interactionist model of ethical decision-making

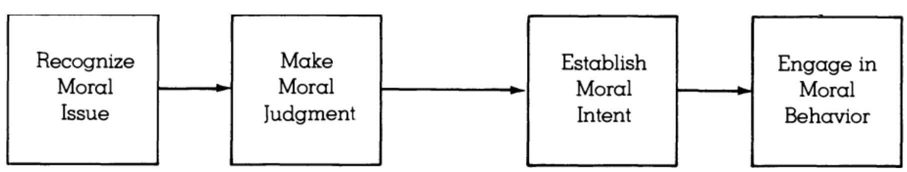


Source: Trevino (1986)

Rest (1986) developed a four-step ethical decision-making model, with the four steps as outlined below (Craft, 2013).

Recognise the moral issue:	the individual must have the agency to make the decision and must recognise that the decision will have an influence on others. If the individual does not recognise the decision as such, the remaining steps of the ethical decision-making process will not be mobilised.
Make a moral judgement:	based on the individual's level of moral development and his moral philosophy, moral reasoning is used to determine right from wrong.
Establish moral intent:	once a judgement has been made, intent to act on it must be established. This involves evaluating the practically and the consequences of the judgement in the context.
Moral behaviour:	the decision-making process culminates in action taken by the individual and a commitment to the moral outcomes.

Figure 4: The Rest (1986) four-step ethical decision-making model



Ford and Richardson (1994) reviewed the ethical decision-making literature to assess which characteristics are important in ethical decision-making processes. They segmented these characteristics into two groups: personal characteristics of the decision-maker; and characteristics of the situation. Their findings are summarized below.

Table 2: Personal and situational characteristics in ethical decision-making

Personal characteristics		**Situational characteristics**	
Characteristics	Finding	Characteristics	Finding
Religion	Of various factors considered, only strength of belief was significantly related to ethical decision-making	Top management	Some studies found a significant influence of top management on ethical decision-making, while others do not
Gender	Some studies find women to behave more ethically, while other studies find no difference between genders	Rewards and sanctions	Empirical evidence of significance of rewards and sanctions influencing ethical behaviour
Nationality	Mixed results on the significance of nationality in ethical decision-making	Organisation culture	Consensus that the more ethical the organisation, the more ethically individuals behave
Age	Some studies find younger managers to behave more ethically, while others find older managers to behave more ethically	Organisation size	Consensus that the larger the organization, the less ethical individuals
Type and years of education	Characteristics are significant in some studies, but not in others	Individual's level in the organisation	The higher the level in the organisation, the more ethical individuals are
Employment and years of employment	Characteristics are significant in some studies, but not in others	Industry ethical standards	Industry ethical standards are not

			related to individuals' ethical behaviour
Machiavellianism	Empirical evidence of significance of a negative relationship with ethical decision-making	Business competitiveness	Business competitiveness may or may not influence individuals' ethical behaviour

From their extensive meta-analysis, Ford and Richardson (1994) noted limitations in the extant literature at the time on characteristics influencing ethical decision-making: there are some clear gaps in the characteristics studied, such as marital status and whether or not the individual has children; and some potentially influential characteristics, such as individuals' personalities, are difficult to capture as part of a survey questionnaire.

Some of the gaps raised by Ford and Richardson were addressed post-1994. One of these gaps was the influence of personality on ethical decision-making and this has since been investigated by various researchers, as reported by Craft (2013) in her meta-analysis of ethical decision-making literature for the period 2004 to 2011. Craft (2013) found that mindfulness, pleasure-seeking, benevolence and subconscious and conscious attitudes are all personality traits that have a significant influence ethical decision-making.

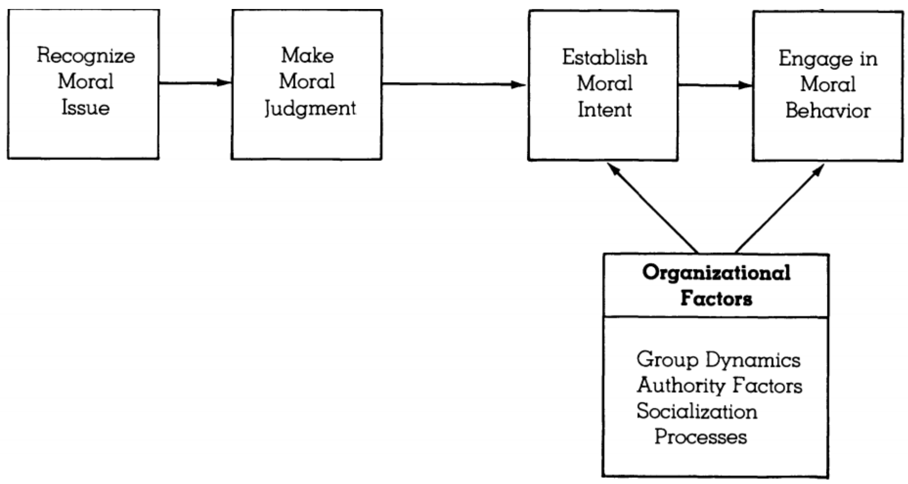
2.2 Issue-contingent model

Jones (1991) was the first to suggest that beyond the influence of personal and organisational factors, factors related to the ethical issue itself, affect the extent of ethical decision-making. In his seminal paper, Jones (1991) built on the model of Rest (1986) by introducing the issue-contingent model. This model describes how "moral intensity" influences the process of ethical decision-making by individuals in organisations.

2.3.1 Moral intensity

Organisational factors influence moral intent and behaviour. While some of these factors may put pressure on individuals to behaviour ethically thus establishing intent, others may make it difficult or easy for individuals to actually behave ethically (Jones, 1991).

Figure 5: Jones' (1991) description of the ethical decision-making process



Source: Jones (1991)

Influencing ethical decision-making and behaviour is the moral intensity of the issue itself. Hence the name, "issue-contingent". Jones (1991) describes moral intensity as the "proportionality" of the moral responsibility. Moral intensity is the characteristics of the issue which drive the moral imperative and is made up of six components (Jones, 1991).

Magnitude of consequences:

Individuals considers the sum of the effect of their decisions on others, both in terms of harm and benefit. Practically, action causing harm to ten people would be taken over action causing harm to 100 people, all else being equal. Jones (1991) argues that since many of the issues faced by individuals in organisations have moral components and since individuals are not permanently grappling with moral dilemmas, individuals must have a threshold for the magnitude of the impact of their decisions and issues which fall under this threshold are not grappled with,

Social consensus:

The "social agreement" of whether certain behaviours are good or bad impact the ethical decision-making process (Jones, 1991). Social consensus reduces ambiguity and provides a benchmark to individuals for what is good and bad in situations. The greater the social consensus, the greater the moral intensity and the greater the effect on decision-making.

Likelihood of effect:

The likelihood of individuals' decisions resulting in benefit or harm have weighting on ethical decision-making (Jones, 1991). Practically, faced with two possibilities for action, individuals will choose the one with lesser chance of resulting in harm. Jones (1991) acknowledges that individuals may be poor at making accurate probability estimates but argues that even rough estimates will have a bearing on decision-making. The more probable the effect, the greater the moral intensity.

Temporal immediacy:

The sooner the impact of the decision will be, the higher the moral intensity and impact on decision making. Individuals discount future events and consider there to be time to intervene before future consequences arise (Jones, 1991).

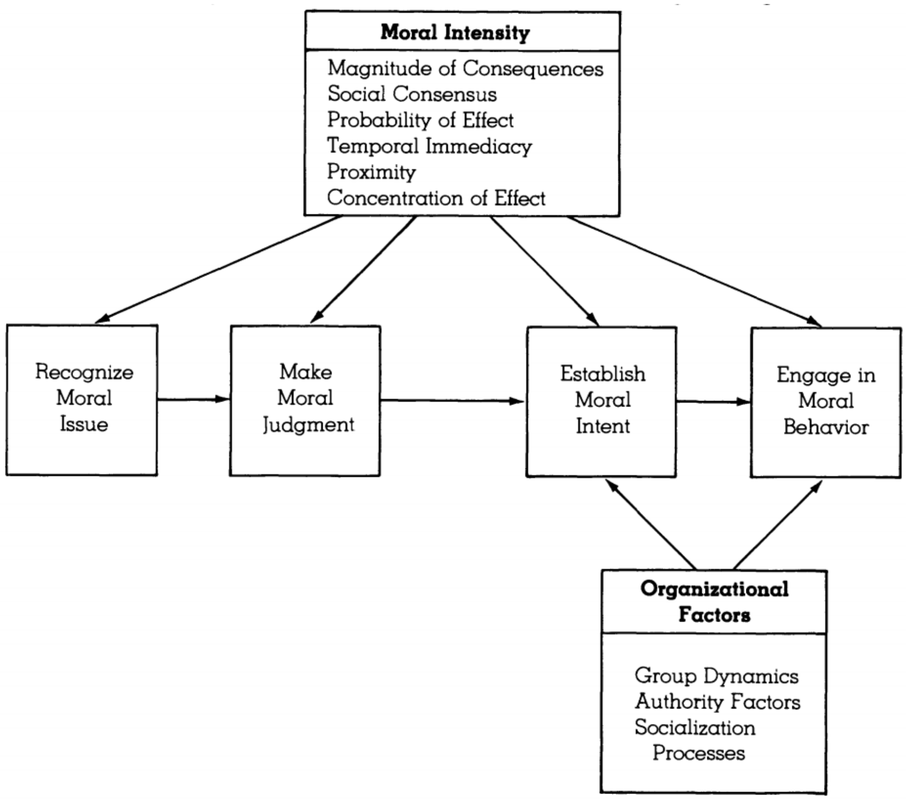
Proximity:

The closer an individual is to people impacted by their decision, the greater the moral intensity and impact on the ethical decision-making process (Jones, 1991). Closeness is comprised of physical, psychological, cultural and social closeness.

Concentration of effect:

The concentration of consequences is the number of people affected, given the magnitude of the consequences (Jones, 1991). Few people largely affected, is more morally intense than many people minimally affected.

Figure 6: The issue contingent model



Source: Jones (1991)

2.2.2 Influence of moral intensity on decision-making

The influence of moral intensity on decision-making has been validated in many research settings. Jones (1991) posited that distinction of a problem as a moral one, influences more ethical behaviour. Dukerich, Walker, George and Huber (2000) found moral intensity and managers' classifications of problems, as either moral problems or non-moral problems, are positively significantly related. Dukerich et al. (2000) found four of the six components of moral intensity to be significant, with temporal immediacy found to be unrelated to classification of problems and probability of effect not measured. This finding is a validation of the first step of decision-making in the issue-contingent model, since the way in which managers perceive and classify problems, influences the subsequent problem-solving processes and outcomes (Dukerich et al., 2000).

Morris and McDonald (1995, p. 724) found that social consensus is a “major determinant of what should be done in the judgement of the decision-maker”. The implication of this is that managers who strongly communicate what is accepted as good behaviour by the organisation may influence others to adopt this perspective in their own decision-making. Regarding the framing of magnitude of consequences, benefits to the organisation influence managers to make less morally-justifiable decisions and costs to the organisation influence managers to make more morally-justifiable decisions (Morris & McDonald, 1995). Morris and McDonald (1995) suggest that if this relationship is confirmed by other researchers, training on the negative consequences and spillover effects of decisions may result in more ethical decision-making.

Singer (1996) was interested in comparing whether the moral judgements of business managers differed from those of the general public. He found that they did not. He did find, however, that the processes employed to arrive at the moral judgement for each group were different. Amongst business managers, Singer (1996) found that social consensus was the greatest determinant of moral judgement of an issue, whilst for the general public, magnitude of consequences was the greatest determinant.

In his 2000 study on New Zealand business managers, Frey considered how the six components of moral intensity interact, using a factorial research design. As an example: how does high probability of effect combined with high magnitude of consequences affect moral judgement, compared to how either high probability of effect or high magnitude of consequence alone affect moral judgement. He found no interaction effects of “appreciable magnitude” and no difference in interaction effects across the various scenarios in his study. Additionally, Frey (2000) found that three (social consensus, magnitude of consequences and probability of effect) of the six moral intensity components were of particular importance in explaining the variation in moral judgement.

McMahon and Harvey (2007) studied the influence of moral intensity on moral judgement across two experiments: one between-subjects, that is across participants; and the other within-subjects. Their findings were that in the between-subjects experiment, actions taken in a high moral intensity environment were more harshly judged as unethical by participants, than the same actions in a low moral intensity environment. In the within-subjects experiment, McMahon and Harvey (2007) found that moral judgement was significantly influenced by manipulated moral intensity.

A criticism of Jones' model (1991) is how the organisational context affects the ethical decision-making process. While Jones (1991) posits that the organisational context affects the decision-maker himself, Kelley and Elm (2003) provide evidence for organisational context rather affecting moral intensity, in certain environments. Social services is the environment in which Kelley and Elm (2003) find this relationship, which, by nature, deals with problems of higher moral intensity than typical business problems.

2.3 Ethical decision-making and AI

The uses of AI can be segmented into two board categories: automation, whereby machines are programmed to act independently of humans; and augmentation, where humans collaborate with machines. Raisch and Krakowski (2021) highlight the automation-augmentation paradox given the interdependence between managers and machines. Managers influence how machines operate, through the parameters they set, data used to train models, feedback given and refinement of algorithms. Machines also have an influence on manager behaviour in that they provide formal reasoning that reinforces normative, expected behaviour. The paradox is such that there is no clear distinction between automation and augmentation. Rather, there are iterative interactions between managers and machines.

AI ethics is a field of ethics concerned with "moral problems" relating to data itself, algorithms and practices (Floridi & Taddeo, 2016). Floridi and Sanders (2004) attempted to formalise characteristics of entities which can be held morally responsible for action. They expanded the discourse on moral agents beyond adults, to artificial agents. "Agenthood" is defined by the ability to change with stimulus, change without stimulus and adapt to new rules in the environment, and is a function of the "level of abstraction" at which agents operate. Floridi and Sanders (2004) posited that when morality is modelled such that a threshold is defined under which behaviour is considered moral, and when artificial agents have agency at a given level of abstraction, artificial agents can be held morally responsible. Until then, though, it is humans alone that are responsible for moral behaviour.

An interesting question in the literature is whether humans have a responsibility towards treating robots ethically. (Dubber, Pasquale and Das, 2020) explored whether, and to what extent, humans do. They posited that beings must have the following two characteristics for humans have moral obligations towards them: sentience, the ability of beings to experience emotions and sensations, such as happiness or pain; and consciousness, the ability of sentient beings to be self-aware. Since robots do not embody these characteristics, humans do not have direct moral obligations in their treatment of robots. Dubber, Pasquale and Das (2020) add that there are, however, indirect obligations in the treatment of robots due to their relationships with other humans. As an example, one must respect another's robot because the other person values it.

2.4.2 On the fairness of price discrimination

Given the distribution of welfare under price discrimination, the fairness of the practice has been questioned a great deal in the literature. The impact of price discrimination is ambiguous, since lower-end customers may be able to afford goods they otherwise wouldn't, whilst higher-end customer pay more for goods they would otherwise pay less for (Borgesius & Poort, 2017; Botta & Wiedemann, 2020). They suggested that competition law that traditionally protects "industrial consumers" be extended to cover final consumers too, to protect individuals from firm dominance, and that the fairness of price discrimination be considered on a case-by-case basis.

Vulkan and Shem-Tov (2015) suggested that a fair pricing mechanism is one where customers pay a fixed percentage of their maximum willingness to pay. They arrived at this fixed percentage, of approximately 64%, in an experiment in which they positioned regular "everyday" customers as sellers of books who had access to willingness to pay and these sellers were asked to set prices for the books. By designing the experiment with everyday customers as sellers, they believed that their results represent what is fair in the minds of customers. Vulkan and Shem-Tov (2015) concluded that this pricing model leads to a “fair, whilst uneven, distribution of prices”.

To the contrary, Wong (2021) argued that it is within the rights of firms to price discriminate. He recommended that firms be transparent about their pricing methodologies and that regulations, such as Article 22 of the General Data Protection Regulation, should not be used to prohibit personalised pricing practices. Wong (2021) stated that according to European Union law, firms have the freedom to employ pricing methods as they choose to, including methods that allow firms to offer different prices to different customers. The debate is still raging on whether discriminatory pricing practices are acceptable and there is still much work to be done in ensuring that these practices are fair.

2.4 Conclusion

Ethical decision-making theory has developed from Kohlberg and Hersh's (1977) theory of moral development which proposes that ethical decision-making is a result of cognitive development of individuals. Trevino (1986) built on this by introducing the influence of situational factors in propelling individuals to behaviour ethically, in the interactionist model. From here, Jones (1991) developed the issue-contingent model, which suggests that there are factors inherent to the issue itself which influence ethical decision-making and behaviour.

Jones’ (1991) issue-contingent model may provide a framework for understanding how ethical decisions are made in the AI field. The model proposes that individuals make more ethical decisions as the moral intensity of the issue at hand increases. Moral intensity of the issue consists of the magnitude of consequences, social consensus, probability of effect, temporal immediacy, proximity and concentration of effect. As these measures increase, the more likely the individual is to recognise the issue as a moral one, judge the issue from a moral perspective and, ultimately, behave morally. The issue-contingent model has been validated in various contexts (Singer, 1996; Dukerich et al., 2000; Morris and McDonald, 1995; Frey, 2000; McMahon and Harvey, 2007), although never-before in the AI field.

Online personalised pricing is a practice enabled by AI and one that raises ethical concerns about fairness. While it may be within the rights of firms to charge different prices to different customers for the same goods, charging customers at their individual maximum willingness to pay is considered unethical by some because the welfare surplus is directed away from customers and towards firms. In this study, Jones' (1991) issue-contingent model will be used to examine managers' decision-making process in deployment of an AI-based online personalised pricing algorithm for grocery shopping.

3. Hypotheses

The overall purpose of this research is to determine whether moral intensity is a predictor of moral judgement regarding the deployment of an (AI)-based online personalised pricing model. Literature suggests that, aside from moral intensity, demographic characteristics of the decision-maker determine a baseline moral judgment. This notion was first proposed by (Kohlberg, 1976) in the theory of moral development and was supported by Jones (1991), who recognised that demographic characteristics of the decision-maker influence moral judgement, but that this is separate to moral intensity. Demographic characteristics have since been included by researchers (Morris & McDonald, 1995; Singer, 1996; Frey, 2000; McMahon & Harvey, 2007) when testing the influence of moral intensity on moral judgement. The first hypothesis for this research is as below:

H1: Age, gender, managerial level and involvement in AI projects influence moral judgement.

Recognition of an issue as a moral issue, influences the decision made. This relationship was defined by (Velasquez & Rostankowski, 1985) and Jones (1991) built on it by adding that moral intensity influences the recognition of an issue as a moral one, through the impact on "the individual's recognition of the consequences of decision". That is, the greater the factors of moral intensity are, the more "sophisticated moral reasoning elicited" will be and the more moral the decision made is (Jones, 1991). Factors of moral intensity for this research were limited to those found to be most influential in decision-making by previous researchers: social consensus (Morris & McDonald, 1995; Dukerich et al., 2000; Frey, 2000; McMahon & Harvey, 2007); magnitude of consequences (Morris & McDonald, 1995; Singer, 1996; Dukerich et al., 2000; Frey, 2000; McMahon & Harvey, 2007); and likelihood of effect (Singer, 1996; Frey, 2000; McMahon & Harvey, 2007). Based on the issue-contingent model (Jones, 1991), the following hypothesis will be tested:

H2: Factors of moral intensity (social consensus, magnitude of consequences, likelihood of effect) positively influence recognition of an issue as a moral one.

The fundamental contribution of the issue-contingent model is that factors of moral intensity positively influence moral judgement (Jones, 1991). That is, the more morally intense an issue, the greater the morality in the decision made. Several researchers (e.g., Singer, 1996; Dukerich et al., 2000; Morris & McDonald, 1995; Frey, 2000; McMahon & Harvey, 2007) have supported this proposition. Based on Jones' (1991) model and these findings, the following hypothesis will be tested:

H3: Factors of moral intensity (social consensus, magnitude of consequences, likelihood of effect) positively influence moral judgement.

Frey (2000) studied the influence of the factors of moral intensity on moral judgement when they are combined, although the issue-contingent model does not make explicit allowance for this. He sought to answer this question: do components of moral intensity interact to influence moral judgement? For example, how does moral judgement change when magnitude of consequences and likelihood of effects are both high together, compared to when each of them is high separately. He found that combined effects only have a marginal influence on moral judgement. In other words, components of moral intensity do not explain moral judgement any better when they are combined compared to when they are applied separately. Based on Jones' (1991) model and Frey's (2000) findings, the following hypothesis will be tested.

H4: Factors of moral intensity (social consensus, magnitude of consequences, likelihood of effect) interact to predict moral judgement.

4. Research methodology and design

The sub-sections below describe the research methodology and design in detail. The importance of such detailed description is to protect the credibility of authors and their findings, by providing all the information necessary for other researchers to reproduce the findings (Aguinis, Hill & Bailey, 2019). From a practice perspective, methodological transparency is important to ensure that when findings are acted on, they produce results as expected (Aguinis & Solarino, 2019).

4.1 Methodology

The use of experimental vignetted is described and defended in this sub-chapter. The scenario and vignette structure of the measurement instrument is also detailed.

4.1.1 Approach to theory development

The research philosophy underpinning this investigation is positivist. This research philosophy requires highly structured and meticulous approach to the data collection and analysis process, so that results may scrutinised statistically and replicated (Saunders & Lewis, 2018). A deductive approach to theory development was employed, whereby data was collected for the purpose of analysing the relationship between factors of moral intensity and moral judgement. This approach to theory development is most appropriate since the research question can be decomposed into testable propositions (Saunders & Lewis, 2018).

The purpose of the research design was to describe the relationship between factors of moral intensity, namely magnitude of consequences, social consensus and likelihood of effect, and moral judgement. Different factors comprising moral intensity were to be tested to understand whether certain components have a greater impact on moral judgement than others do. Given the positivist philosophy and deductive approach to theory development, it follows that a quantitative research design was most appropriate.

The research strategy was to collect data from the population through surveys. Surveys provide a simple and quick means to collect structured and standardised data from large numbers of participants. The time horizon was cross-sectional, meaning that data was be collected at one point in time. Longitudinal studies are used to describe changes over time and since there will be no intervention in the duration of this study, there is no reason to suspect that there would be changes in participants' experiences in the short time in which the research is conducted.

4.1.2 Experimental vignette methodology

The research methodology was experimental vignette methodology (EVM). EVM is based on presenting participants with realistic scenarios comprising the independent variables and capturing their responses to the scenarios as the dependent variables (Aguinis & Bradley, 2014). There are two types of EVM studies: paper people studies where the aim is to capture explicit outcomes known by respondents; and policy capturing and conjoint analyses where outcomes are implicit (Aguinis & Bradley, 2014). For the purpose of this research, a paper people study was conducted, in which participants were presented with a hypothetical scenario related to the use of AI and they were explicitly asked about their moral judgement of the scenario. EVM methodology was also used by Frey (2000) in his testing of the influence of moral intensity on pollution and vacation time-sharing.

The scenario, developed from vignettes used by other researchers (e.g., OECD, 2021), was as follows:

> Artificial Intelligence (AI) can be broadly described as computer programs developed to simulate intelligent behaviour, by analysing vast amounts of data.
>
> The AI development team at SupaFood, a nationwide grocer, has deployed a machine learning model that personalises pricing on food items. The model works by estimating an individual's willingness to pay for food items, based on various factors. Since deploying the model, SupaFood's profits margins have increased substantially.
>
> There have, however, been concerns around the way the model price discriminates. The model could potentially exploit people who have a necessity for certain food items and no or little access to substitutes. This

would mean that some people may potentially be charged higher prices than other people are charged for the same items.

The choice of research design was either between-person, within-person or mixed (Atzmüller & Steiner, 2010). In the between-person design, all participants are exposed to the same scenario and differences in responses are measured across participants (Atzmüller & Steiner, 2010). In the within-person design, participants are exposed to different scenarios and responses are compared across the scenarios and mixed design comprises a combination of the two (Atzmüller & Steiner, 2010). The purpose of this research is to measure how different levels of moral intensity affects ethical judgement, rather than how ethical judgement varies across people. Thus, the within-person research design was most appropriate.

To measure the relationship between moral intensity and moral judgement, it is required that moral intensity be varied so that moral judgement may vary in response. Frey (2000) achieved this by presenting participants with a base line scenario in which all moral intensity factors are low and, thereafter, presenting the same scenario for a second time with at least one of the factors change to a high level of intensity.

In this research, each of the three factors of moral intensity were varied. Each of the factors took on two levels, high or low, resulting in eight possible variations of the scenario (2 levels raised to the power three). These eight variations of the scenario are henceforth referred to as the vignettes. These vignettes were presented to participants as follows.

Table 3: Vignette descriptions for high and low moral intensity

Vignette number	**Vignette description**	**Social consensus**	**Magnitude of effect**	**Likelihood of effect**
1	There is *no agreement* amongst the public on whether using such models is acceptable or not. The potential price increase experienced by some people would be *only a few cents*.	Low	Low	Low

	The chance that the above-mentioned price difference materialises is *low*.			
2	There is *no agreement* amongst the public on whether using such models is acceptable or not. The potential price increase experienced by some people would be *only a few cents*. The chance that the above-mentioned price difference materialises is *high*.	Low	Low	High
3	There is *no agreement* amongst the public on whether using such models is acceptable or not. The potential price increase experienced by some people would be *substantial*. The chance that the above-mentioned price difference materialises is *low*.	Low	High	Low
4	There is *clear agreement* amongst the public that models such as this are unfairly discriminatory. The potential price increase experienced by some people would be *only a few cents*. The chance that the above-mentioned price difference materialises is *low*.	High	Low	Low
5	There is *no agreement* amongst the public on whether using such models is acceptable or not. The potential price increase experienced by some people would be *substantial*. The chance that the above-mentioned price difference materialises is *high*.	Low	High	High
6	There *is clear agreement* amongst the public that models such as this are unfairly discriminatory. The potential price increase experienced by some people would be *only a few cents*. The chance that the above-mentioned price difference materialises is *high*.	High	Low	High
7	There is *clear agreement* amongst the public that models such as this are unfairly discriminatory. The potential price increase experienced by some people would be *substantial*. The chance that the above-mentioned price difference materialises is *low*.	High	High	Low
8	There is *clear agreement* amongst the public that models such as this are unfairly discriminatory. The potential price increase experienced by some people would be *substantial*. The chance that the above-mentioned price difference materialises is *high*.	High	High	High

4.2 Population and unit of analysis

The population of individuals for this research are business managers. This population spans all industries, in both the private and public sectors. The managers in this population are at varying levels of seniority: junior, mid and senior/executive. To reduce the population to a manageable target population, managers were limited to those based in South Africa.

Similar to Frey (2000), the unit of analysis was the manager. Since the issue-contingent model is a decision-making framework, it follows that decision makers be the unit of analysis. In the context of this particular research, managers are the people making decisions about developing and deploying AI models into the population. In the sample, it may be the case that multiple managers from one organisation are present and this is okay. The research proposed is about managers, rather than organisations, so any number of managers from one organisation is acceptable.

There was consideration of whether to include only managers who have direct involvement in the development and deployment of AI solutions. The decision was to open the population to all managers, for two reasons. Firstly, AI is becoming more ubiquitous in business and managers who are currently not involved in AI projects, may soon be, and their current moral judgements around AI solutions are unlikely to change dramatically in the short term. Secondly, finding the required sample size of managers directly involved in AI projects would not have been an easy feat. Instead, there was a survey question posed as part of the demographic questions to determine whether or not the manager has involvement in AI projects.

4.3 Sampling method and size

The way in which the sample was drawn and the size of the sample are important in determining whether the results of this research are representative of the decision-making behaviour of all managers. These are discussed below.

4.3.1 Sampling methodology

The sample frame consists of all the organisations which meet the conditions of the target population. It was not feasible to compile a list of all possible organisations making up the sample frame and it was, therefore, not possible to use probability sampling from this group (Saunders & Lewis, 2018). Instead, non-probability, and, in particular, purposive sampling was used. Purposive sampling consists of the researcher using judgement to actively select a group of participants based on who he believes to have the most information to share (Saunders & Lewis, 2018).

4.3.2 Sample size

Israel (1992) describes how sample sizes for surveys are dependent on three factors. Firstly, the acceptable sample error. This is the range from the sample values in which true values of the population are expected to exist and this is measured as a percent of the sample value. The smaller the sample error is, the closer to the sample values the population values are and the greater the sample required. Secondly is the confidence level. This is the proportion of the time that randomly drawn samples will yield the same results, given the sample error. The higher the confidence interval, the larger the required sample. Thirdly, the degree of homogeneity of population attributes. The more variable population attributes are, the larger the sample required to capture the variation.

Israel (1992) goes further to suggest how researchers may determine their required sample sizes. For small populations, he recommends a census. This involves sampling the entire population and, since this is a costly exercise, it is only appropriate for very small populations. This is not an appropriate method to for sampling managers in this research. Israel (1992) goes on to suggest that researchers consider the sample sizes of others who have conducted similar studies. Frey (2000) had a sample of 406 managers when he studied the influence of moral intensity in a business context. The final method Israel (1992) suggests in determining sample size is to use formulas, or tables of illustrative figures based on these formulas, which account for the population size, sample error and confidence level.

The sample size calculation as proposed by Charan and Biswas (2013) for quantitative research, is as follows:

$$required\ sample = \frac{zscore^2 \text{ x StdDev x } (1 - \text{ StdDev})}{margin\ of\ error^2}$$

where:

$zscore$ represents the confidence level required

$StdDev$ is the standard deviation seen in the population

$margin\ of\ error$ is the acceptable confidence interval, also known as the effect size

The population size of all managers in South Africa was reported to be 1,406,000 between April and June 2021 (Statistics South Africa, 2021). Due to proportionality, however, required samples vary very little after a population of 100,000, for a given sample error and confidence (Israel, 1992). As such, the population size in this research was considered as 100,000 managers. For the standard 95% confidence level, sample size requirements vary from 100 participants for 10% sample error, to 1,099 participants for 3% sample error. Ideally, a sample error of 5% should be targeted, which would require a sample size of 400 participants. Given limitations in time to collect the sample data, however, the sample size in this study was 120 participants.

The sample was drawn in two ways: firstly, by approaching managers in my personal network and secondly, through using LinkedIn. My personal network of managers consists of colleagues from my work and MBA classmates. I then used LinkedIn to extend the sample to managers with who I am connected, but do not necessarily have a relationship. Snowball sampling resulted from the purposive sampling. This happened because some participants shared the survey with relevant people in their own networks, resulting in the reach growing organically. This was a valuable way to extend the reach of the survey.

4.4 Measurement instrument

The measurement instrument was surveys, emailed to participants directly. The surveys comprised the following sections.

Table 4: Survey structure

Section	Description
Consent	Data privacy and consent clauses
Demographic information	Questions on gender, age group, management level and whether the participant is involved in AI projects
Survey instructions	Explanation of the scenario and vignette structure of the survey
Scenario description	Description of one hypothetical use of AI
Vignettes	Eight vignettes comprising variations on the scenario, with four questions on each

Participants were exposed to all eight vignettes, so that two-way and three-way interaction effects can be measured (Atzmüller & Steiner, 2010). These interaction effects capture the impact of two or more variables combined on the dependent variable.

After describing each vignette, participants were asked four questions relating to their moral judgement of each. These questions, comprising the dependent variables, were based on the questions asked by Frey (2000) to capture variation on moral judgment in his study. Frey's questions were altered slightly to suit the context, as follows:

Question 1: The decision made to deploy the model was morally correct
Question 2: I would oppose the decision to deploy this model
Question 3: This scenario involves a moral issue
Question 4: Should this model be used?

Questions 1 through 3 were answered on a five-point Likert scale and Questions 4 was answered with a binary *yes* or *no*. Please refer to the appendix for the full survey.

As a means to test the length and level of fatigue participants may experience in completing the survey, a test survey was done on one respondent. The time taken for him to complete the survey was eight minutes and minimal levels of fatigue were experienced by him. The survey was then tested on a pilot group of nine people before being sent to survey participants. The main feedback from the pilot participants was that vignettes were similar and, in some instances, pilot participants even believed vignettes to have been repeated, even though that was not the case. Based on this feedback, the vignettes were reworded and key words were highlighted to make differences more explicit. In terms of the response data from the pilot, two participants had the same responses for all vignettes. Their feedback was that the scenario is out and out immoral, regardless of the level of moral intensity variations. While this is a worthwhile finding, without variation in responses, there can be no model. The survey data was checked for other responses such as this. The survey was created using a tool called *Qualtrics*. The benefit of using this tool was that the order of the vignettes could be randomised. Data collected from *Qualtrics* was exported to a flat file and was stored in the cloud, where it was secure and available as needed.

The survey was sent to managers in my own network directly by me and further shared by some participants. Given this snowball sampling, it is not possible to determine how many managers in total the survey was shared with and compute the response rate. Having said that, I personally did not do any mass-sharing of the survey. I purely contact managers individually and asked for their completion. As such, the response rate was rather high amongst this subset of the sample. In total, the survey was opened 180 times, but how many unique managers opened the survey is indeterminable. Of those, the introductory demographic questions were completed 169 times and the survey was completed 120 times. These 120 complete responses make up the sample used in the analysis.

4.5 Data transformations

Data was exported from Qualtrics and imported into Google Sheets, where it was stored and the first iteration of transformation was done. This iteration consisted of: removing any unimportant fields that Qualmetrics had automatically included; assigning a numeric row identifier to each record for cross-referencing and transformation validation purposes later; transposing the data from having a row per participant, to having a row per participant and vignette; combining age groups: “36-45”, “46-55” and > “55”; and creating three new fields named “social consensus”, “magnitude of consequences” and “likelihood of effect” and assigning “low” or “high” to each, based on the vignette questioned in each row. From the initial dataset of 120 rows based on number participants, the dataset resulting from this iteration of transformation had 960 rows (120 x 8) at the participant-vignette level. Having duplications of participants in this way is not a concern for this study, since the design of this experiment is within-person, as opposed to between-person, and the aim is to measure the influence of social consensus, magnitude of consequences and likelihood of effect. After these transformations, spot checks were done to ensure that the transformations had worked as intended.

From here, the data was exported to a flat file and imported into SPSS, where the second iteration of transformation was done. This iteration consisted of recoding the data as required by the statistical analyses that were to be run. The data consists of three types of variables: independent, covariate and dependent. The independent variables make up moral intensity and these are set in the experiment. Each of these three variables have two levels: *low* and *high*. To incorporate these levels into the analyses, they needed to be transformed from string variables into numeric variables. As such, they were recoded as: *low* = 0 and *high* = 1.

The second set of variables to be recoded into numeric variables were the covariates. For age and management level, values were recoded such that the numerical recoded value increases as the value for the underlying data increases, as follows:

Table 5: Covariates recoded for analysis

Age

Original value	Recoded value
< 26	1
36 - 45	2
26 - 35	3

Management level

Original value	Recoded value
Junior	1
Middle	2
Senior/executive	3

Gender

Original value	Recoded value
Female	0
Male	1

Involvement in AI projects

Original value	Recoded value
No	0
Yes	1

The last set of variables to be recoded was the four dependent variables, of which three were answered on a five-point Likert scale of “strongly agree” to “strongly disagree”. These were recoded on a numerical scale of one to five. Like Frey (2000), these were recoded such that higher moral judgment was associated with a higher score. Based on the different ways in which the four questions were phrased, these were either recoded such that more agreement with the question was assigned a higher value, or that more disagreement with the question was assigned a higher value. Below the recode logic per question is described.

Table 6: Independent variables recode logic

Question	**Scale**	**Relationship with moral judgement**	**Recode logic**
The decision made to deploy the model was morally correct	Five-point Likert of strongly disagree of strongly agree	Inverse	Strongly agree = 1 to strongly disagree = 5
I would oppose the decision to deploy this model	Five-point Likert of strongly disagree of strongly agree	Direct	Strongly disagree = 1 to strongly agree = 5
This scenario involves a moral issue	Five-point Likert of strongly disagree of strongly agree	Direct	Strongly disagree = 1 to strongly agree = 5
Should this model be used?	Binary yes or no	No = greater moral judgement	Yes = 0 No = 1

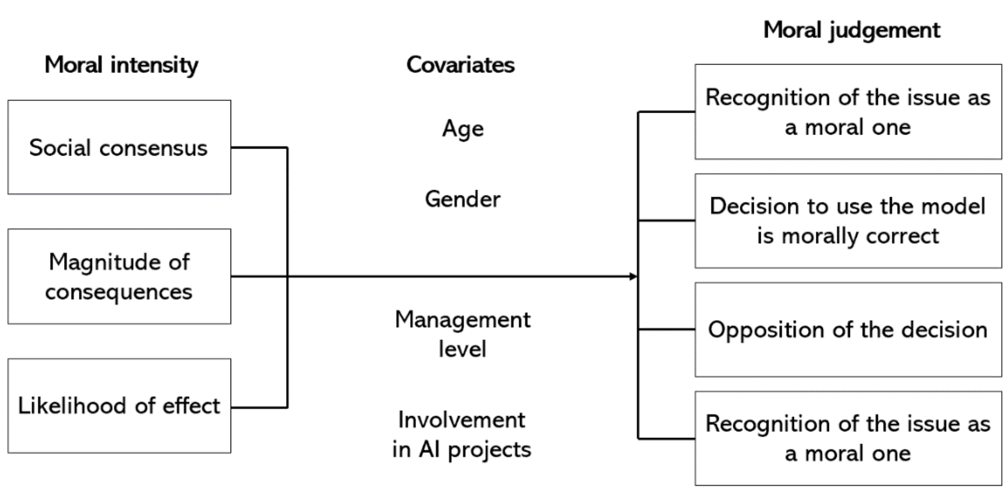
4.6 Analysis approach

As with Frey’s (2000) study, there are four dependent variables in this data. Given this complexity, I initially ended to do multivariate analysis, as Frey (2000) did. Multivariate models are a type of general linear model, where there are two or more dependent variables (Warne, 2014). Multivariate analyses are not uncommon in behavioural science studies, since most data resulting from real world problems is multivariate (Chatfield & Collings, 1980). In fact, Warne (2014) wrote that multivariate analysis of variance was used in 23% of all quantitative analyses in an educational psychology journal.

One of the assumptions for multivariate linear regression analysis is that the dependent variables are continuous. In this data, the dependent variables were measured on a five-point Likert scale. There are two problems with the scale used. Firstly, although Likert scales are considered to be continuous variables by many researchers, this is not theoretically true (Wu, 2007). Likert scales are in fact ordinal variables. That is, the values can be ordered, but no arithmetic processes can be run on them. Secondly, a scale with only five points is difficult to approximate to a normal distribution. Frey (2000) used a nine-point scale and this may have allowed him to assume continuous dependent variables. As such, the dependent variables in this data were not considered as continuous and therefore ordinary least squares regression was not utilized for statistical analyses.

Instead, analysis of variance (ANOVA) and analysis of covariance (ANCOVA) were the statistical techniques utilised to investigate the study's hypotheses (see Table 4). As an extension of the t-test, which compares means of two groups, ANOVA allows for comparisons of means across more than two groups, as is required in this analysis. A similar outcome may be achieved by conducting a series of t-tests, though this introduces a type-1 error since groups cannot be tested independently of each other (Armstrong, Slade & Eperjesi, 2000). Figure 7 shows the conceptual model for the study.

Figure 7: Conceptual model



Since hypotheses two to four two contain “covariates”, ANCOVA analyses were conducted. Covariates are an additional type of variable that may cause variances in the dependent variable, over and above variances caused they the independent variable (Şentürk & Müller, 2009). The benefit of including covariates into a model is that it provides a means to control for variation in the dependent variable which is not caused by the independent variable. In this context of this research, it has been confirmed by some previous researchers that personal characteristics influence moral judgement. As such, including demographic characteristics as covariates allows for the influence of moral intensity over moral judgement to be isolated, as was done by Frey (2000). The table below outlines the statistical analyses run for each hypothesis.

Table 7: Analysis approach for each hypothesis

Hypothesis	Independent variables	Covariates	Dependent variables	Statistical analysis
H1: Age, gender, managerial level and involvement in AI projects affect moral judgement	Age; gender; management level; involvement in AI projects	n/a	*Decision was morally correct;* *oppose the decision;* *scenario involves a moral issue;* *should this pricing model be used*	ANOVA (univariate analysis of variance)
H2: Factors of moral intensity (social consensus, magnitude of consequences, likelihood of effect) positively influence recognition of an issue as a moral one	Social consensus; magnitude of consequences; likelihood of effect	Age; gender; management level; involvement in AI projects	*Scenario involves a moral issue*	ANCOVA (univariate analysis of covariance)
H3: Factors of moral intensity (social consensus, magnitude of consequences, likelihood of effect) positively influence moral judgement	Social consensus; magnitude of consequences; likelihood of effect	Age; gender; management level; involvement in AI projects	*Decision was morally correct;* *oppose the decision;* *scenario involves a moral issue;* *should this pricing model be used*	ANCOVA (univariate analysis of covariance)

H4: Factors of moral intensity (social consensus, magnitude of consequences, likelihood of effect) interact to influence moral judgement	Social consensus; magnitude of consequences; likelihood of effect; and interaction effects	Age; gender; management level; involvement in AI projects	*Decision was morally correct; oppose the decision; scenario involves a moral issue; should this pricing model be used*	ANCOVA (univariate analysis of covariance)

Hypothesis four includes interaction effects among the independent variables. That is, levels magnitude of consequences and likelihood of effects interact to influence moral judgement. Since all eight vignettes are judged by each respondent, it was possible to measure these interaction effects, in addition to the main effects of each independent variable on its own. As such, in addition to the three main effects independent variables, there are four interaction variables considered in hypothesis four:

Table 8: Interaction effects of three factors

Social consensus	**Magnitude of consequences**	**Likelihood of effect**
Low	High	High
High	Low	High
High	High	Low
High	High	High

4.7 Quality controls

There are various biases which may influence participants' responses in this research. In their study of the influence of moral intensity on moral judgement, McMahon and Harvey (2007) control for various biases: *response bias* by limiting the number of words presented in each of their scenarios; *social desirability* bias by not writing the scenario with the respondent as the decision-maker in the scenario; and *gender bias* by varying the gender of the person in the scenario across questionnaires. Similar controls were implemented in this research: to control for response bias, each vignette consisted of only three short sentences; the scenario in the survey described a decision-maker which is not the survey respondent to limit social desirability bias; and to reduce gender bias, the scenario was written such that the gender of the decision-maker is not identifiable.

Order effects is the impact that the order of questions can have on the responses of participants. This happens because the first questions in a survey can serve as an anchor for later questions, by establishing understandings in the participant's mind as to what kind of responses are expected (Perreault, 1975). Additionally, in an attempt to be consistent, participants may let their responses in early questions guide their later responses, even if they do not necessarily align with the participant's true beliefs (Perreault, 1975). To reduce the impact that order effects may have, the order of vignettes was randomised.

Since participation in the research is not compulsory, a response bias may exist such that people who do elect to respond to the study are different from those who elect not to. Morris and McDonald (1995) account for this by measuring the response rate they received on their survey and deemed this acceptable at 75%. In this research, it was difficult to account for everyone whom the survey was shared since there was some snowball sampling in effect. To limit the response bias to some extent, the researcher reached out to most of the sample personally to ask them to complete the survey. When participants are approached individually, it is more difficult for them to ignore the request.

Participants had to complete preceding questions before moving onto new questions. This assisted in ensuring that no questions were accidently missed by participants. Additionally, since only surveys where every question was completed were used as part of the sample, the data can be considered as complete with no missing values. All questions were set as multiple-choice questions, having answers from a pre-defined list. Having no free-text fields ensures that participants could not capture data that was ambiguous or nonsensical. Accuracy of the data captured by participants could not be tested. Testing accuracy would involve cross-referencing the demographic data with another data source. For example, the age group may have been checked for accuracy using the date of birth component of a South African identification number.

4.8 Design limitations

Vignettes may look similar to participants and this may distort their responses. This was initially raised as a concern during the survey pilot and although work was done to make vignettes more distinct before the survey went live, some participants may still not have identified all differences in all vignettes. One survey participant noted that perhaps the instructions should have been clearer in instructing participants to read the survey questions very carefully and complete the survey free of distraction. To address this concern, questions measuring the "perceived moral intensity" could have been included, as was the case in Frey's (2000) research. In the current research, moral intensity is objectively changed in each vignette, but there is no certainty over whether participants observed these changes. As an example, to measure whether the intended change in magnitude of consequences was observed, a question such as "this model is likely to impact a large number of people a great deal" could have been asked.

Additionally, the order of vignettes may impact the responses of participants. While it is recommended that the order of questions be randomised to avoid order effects, randomisation of questions can result in certain questions being more salient than others since the preceding and following questions cannot be controlled (McFarland, 1981). This results in the participant being primed to respond to the question in a different way than they would have had the order of questions been different. In this research this is a limitation because the vignettes are constructed around high and low levels. Respondents may have, unknowingly, compared the level of moral intensity in one question with its predecessor and based their moral judgment on the change in levels they identified. If the order of questions were different, however, they may have made different comparisons and responded differently.

The impact of social desirability reporting should not be underestimated in studies of moral behaviour. This bias describes the tendency of people to overreport their virtuous characteristics and underreport their undesirable characteristics (Morris & McDonald, 1995). Morris and McDonald (1995) account for this incorporating the Balanced Inventory of Desirable Responding scale into their survey. This allowed them to account for the impact of social desirability reporting on their results. In this research, there was no control for social desirability reporting and it may be the case that survey participants responded with greater moral judgement than they would respond with in reality.

Lastly, there may be limitations in the sample. The sample size may be too small to capture all the variation in the population. The population for the research is all business managers, whether they are involved in AI projects or not, across all industries, managerial levels and age groups. There is a great deal of variation in this population and populations such as this require large samples to capture the variation sufficiently (Israel, 1992). More so, the snowball sampling which arose may have resulted in non-managers completing the survey. While anyone who shared the survey was asked to only share with managers in their networks, there is no way to ensure this was the case. A qualifying question at the beginning of the survey asking whether the participant is a manager or not, would have been a good way to exclude non-managers before analysing the data.

In conclusion, there are potential limitations with the scenario used, differences in vignettes may have been unclear to some participants, participants' responses may have been guided by what is considered socially acceptable and the sample size may have been too small to sufficiently represent the population. Nevertheless, the methodological choices serve as an exploration into moral intensity and AI, and these choices can be improved upon by future researchers.

4.9 Conclusion

The aim of this research is to understand the influence that moral intensity has on decision-making about an AI-based online personalised pricing model. Since the onus for ethical decision-making in this context falls upon managers, managers were the population from which the sample of 120 participants were drawn, using purposive and snowball sampling methodologies.

The impact that moral intensity has on moral judgement in the use of AI can be explored using an experimental vignette methodology. In this research, participants were exposed to a scenario relating to the use of an AI model for personalised pricing. Based on this scenario, eight vignettes were presented to participants, with varying levels of moral intensity. Three components of moral intensity (social consensus, magnitude of consensus and likelihood of effect) were varied according to high and low levels. Participants were then asked four questions on each vignette to capture their moral judgment of the vignette. Answers to three of the four questions were captured on a five-point Likert scale and the final question was capture with a binary “yes” or “no” response. Additionally, four personal characteristics were captured on each participant: age, gender, management level and involvement in AI projects.

The personal characteristics data and the responses to the four questions were recoded into numerical values. The four questions were recoded such that a higher value represents a more moral decision. Univariate analysis of variance and covariance was the analysis approach used for testing each of the hypotheses.

5. Results

This chapter presents the results of the analyses described in chapter four. The sample analysis is described first, followed by the data preparation process, the statistical analyses and the results of the analyses for each hypothesis. Along with the statistical analyses, for the analyses are tested. Finally, limitations in the data and analysis are discussed.

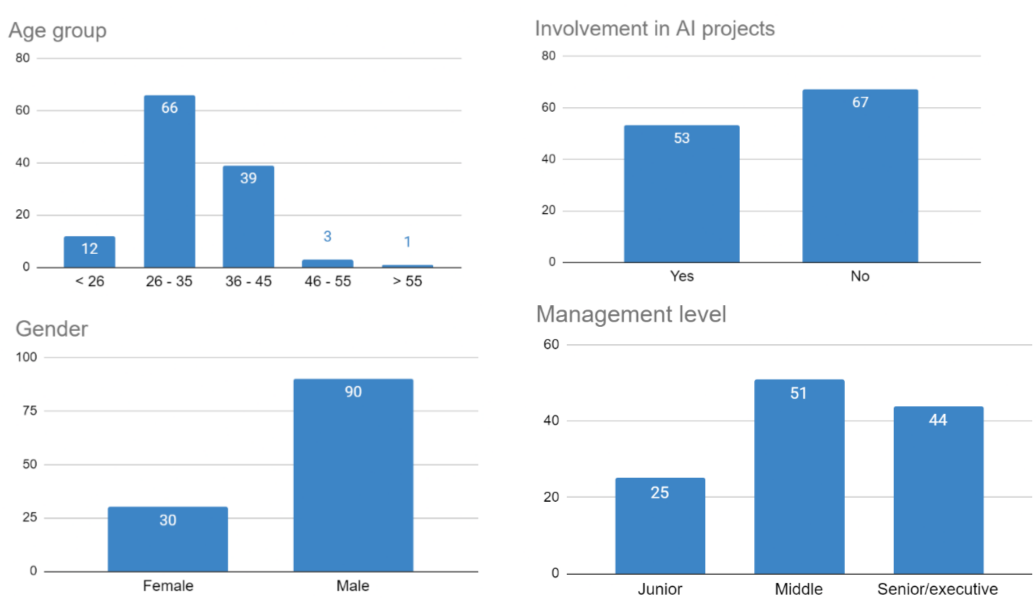
5.1 Sample analysis

It is necessary that the sample used is representative of the population of all managers if the findings are to be generalisable to all managers. Below a description of the sample is provided and the representativity of all manager is explored.

5.1.1 Descriptive statistics

Females represented only 25% of the sample and 98% of the sample was aged 35 and below. The management level spread was more equally at 21% junior managers, 42% middle managers and 37% senior managers and executives. 44% of the participants reported to be involved in AI projects.

Figure 8: Descriptive statistics of the sample



5.1.2 Population representativity

For the results of this experiment to be applicable to all managers in South Africa, the sample must represent the population of managers in South Africa. The age distribution sample in the is such that 10% of managers are below 26 years old, 55 % are between 26 and 35 years old and 32% are between 36 and 45 years old. Given that the number of participants in the “46-55” and “>55” age groups make up only 3% of the population and comprise three and one participants respectively, these groups are too small to derive valid analyses from. As such, these groups were combined into one group named “>35”.

While there is no publicly available data for managers specifically, Statistics South Africa (2021) reports that the age distribution of employed people in South Africa is as follows: 5% are between the ages of 15 and 24; 28% are between the ages of 25 and 34; and the remaining 67% are between 35 and 64 years old. Based on these population estimates, one may argue that 26 to 35 years olds in the sample may be over represented, while the above 36 age group is under-represented, although this cannot be concluded with certainty.

From the sample distribution, it appears that female managers are under-represented in the sample when compared to Statistics South Africa estimates that females make up 44% of employed people in South Africa (Statistics South Africa, 2021). In contrast, only 25% of the managers in the sample are female. To check whether the gender distribution of managers in the sample represents the population of managers, the distribution of genders across different managerial levels is explored below.

Table 9: Sample gender distribution for managerial levels

Managerial level	**Female**		**Male**	
	N	*% of level*	*N*	*% of level*
All	25.0%	30	75.0%	90
Junior	44.0%	11	56.0%	14
Middle	17.6%	9	82.4%	42
Senior/executive	22.7%	10	77.3%	34

Females are well-represented at the junior level, very poorly represented at the middle level, and somewhat better, but still poorly, represented at the senior and executive level. In their analysis of women leader in business, the University of Stellenbosch Business School found that 22% of executive managers at South Africa's top 40 companies are women (Mwagiru, 2020). This finding validates the gender distribution of senior managers and executives in the sample is representative of the population at this managerial level, although there is no validation of the gender distribution across other levels. Despite there being limited means to test the sample representability of the population, this data is deemed to be adequate for the purpose of this experiment.

5.2 Quality of the measurement instrument

To ensure that results from the hypothesis testing are reliable and valid, it is important that the input data is of sufficient quality and that the sample is representative of the greater population it seeks to represent. The following subsections provide comfort of these requirements.

5.2.1 Scale reliability

Reliability is about the internal consistency of a scale (Sijtsma, 2009). In this experiment, there were two scales used: the first three dependent variables were answered on a five-point Likert scale and the final question was answered with a binary yes or no. Cronbach's alpha is used to compare reliability of one scale, and so only the first three questions were tested for reliability. Cronbach's alpha is measured on a score of zero to one, where the closer to one the score is, the more consistency there is in the scale used. A Cronbach's alpha of above 0.7 is accepted as a good level of internal validity.

Cronbach's alpha was test on the three questions using the Likert scale in SPSS. The results were as follows:

Table 10: Reliability statistics

Cronbach's alpha	Cronbach's alpha based on standardised items	Number of items
0.826	0.827	3

Since Cronbach's alpha is greater than 0.7, the scale used to measure the three dependent variables can be considered consistent. Additionally, it is possible to test whether each of the three dependent variables contribute positively to internal consistency of the scale, or whether any of them should rather be excluded.

Table 11: Cronbach's alpha to test reliability

	Scale mean if item deleted	Scale variance if item deleted	Correct item-total correlation	Squared multiple correlation	Cronbach's alpha if item deleted
The decision made to deploy the model was morally correct	7.25	4.175	0.638	0.439	0.805
I would oppose the decision to deploy the model	7.36	3.511	0.774	0.601	0.664
This scenario involves a moral issue	7.08	4.554	0.652	0.471	0.794

From the above table it is clear that if any of these three dependent variables were excluded, Cronbach's alpha would perform worse. As such, all three of these dependent variables are retained.

5.2.2 Validity of the data

Validity of the data is the measure to which the measurement instrument is valid for the relationship that is intended to be captured. In this experiment, validity means that the vignette design and four questions on each were an appropriate means to measure moral judgement. To do this, a Pearson correlation was run in SPSS for each of the dependent variables and the construct of moral intensity. Results for this test are as follows.

Table 12: Pearson correlation to test validity

	The decision made to deploy the model was morally correct	I would oppose the decision to deploy this model	This scenario involves a moral issue	Should this model be used	Moral judgement
The decision made to deploy the model was morally correct	--				
I would oppose the decision to deploy this model	.659**	--			
This scenario involves a moral issue	.500**	.683**	--		
Should this model be used	.633**	.720**	.542**	--	
Moral judgement	.839**	.915**	.817**	.800**	--

*** Correlation is significant at the 0.01 level (2- tailed)*

The significance of each of the correlation coefficients indicates that the measurement instrument in valid for the purpose of measuring moral judgement.

5.2.3 Exploratory factor analysis

To combine the four dependent variables into an index measuring "moral judgement", it was first necessary to do an exploratory factor analysis (EFA). The purpose of EFA is to reduce the dimensionality of data by combining variables that measure the same behaviour (Yong & Pearce, 2013). In this context, it was necessary to test whether the four questions posed to participants on each vignette do measure the same behaviour, called "moral judgement". In chapter 5.2.3, a Pearson correlation matrix is presented and it is evident that the four dependent variables are correlated, with the smallest of the bivariate correlations being 0.5 and the largest being 0.72.

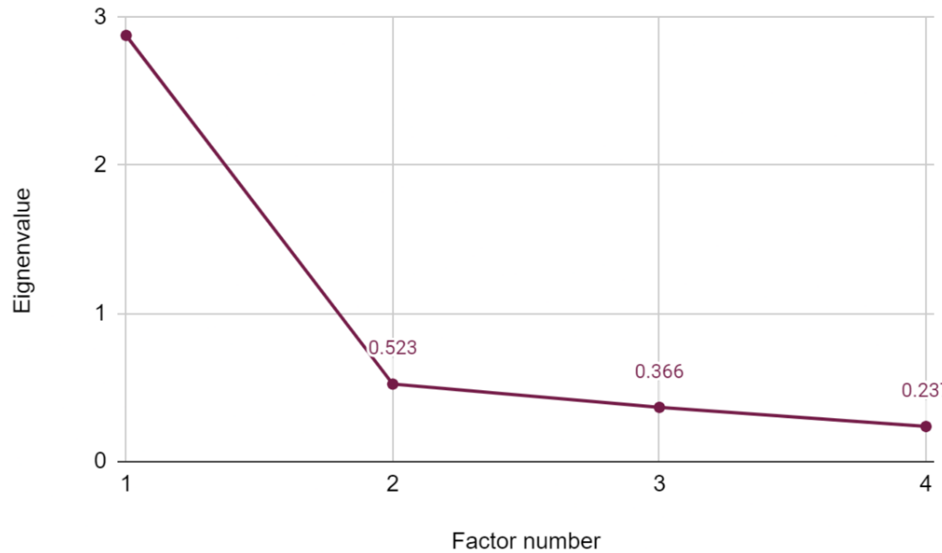
The EFA was conducted in SPSS. As the first step, the sample was checked to ensure that was appropriate for EFA. From the Kaiser-Meyer-Olkin measure of sampling adequacy being 0.80 and Bartlett's test of sphericity being significant at $p < 0.05$, I concluded that the sample was adequate for EFA (Yong & Pearce, 2013). There was only one factor with total initial eigenvalue greater than 1, explaining 71.9% of the variance.

Table 13: Total variances explained

	Initial eigenvalues			Extraction sums of squared loadings		
Factor	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	2.875	71.867	71.867	2.526	63.151	63.151
2	0.523	13.063	84.390			
3	0.366	9.151	94.080			
4	0.237	5.920	100.00			

The scree plot following is a visual representation of the eigenvalue per factor number. It is clear that the eigenvalues after factor one drop to below one, indicating that there is only one factor and one underlying behaviour described by the four dependent variables.

Figure 9: Spree plot of eigenvalue per factor number



After confirming that there was one factor in the data, factor analysis was done to check the correlation that each of the variables had with the factor. Each of the variables had a correlation of greater than 0.5 and so all variables were retained in creating the moral judgement construct. Based on this, I concluded that the four dependent variables do describe the same underlying behaviour and that these can be combined into one factor called "moral judgement". Please refer to the appendix for the full set of results from EFA.

Based on the results of the EFA, the four dependent variables were then combined to create an index that measured “moral judgement”. After recoding, three of the four dependent variables have values from one to five, while the fourth variable has a value of zero or one. Given the difference in these two scales, the variables scoring a potential of five will have a greater impact in the construct score than the variable scoring a maximum of one (Lakshmanan, 2019) . In reality, this would mean that whether the participant believes the model should be deployed or not has less of an impact on moral judgment than the other variables. This is not the case. As such, all dependent variables were normalised to be within the range of zero and one. This was done by taking the difference between each data point and the minimum possible score in the scale, and dividing that by the range of the scale. The sum of the four normalized score were the taken and divided by four.

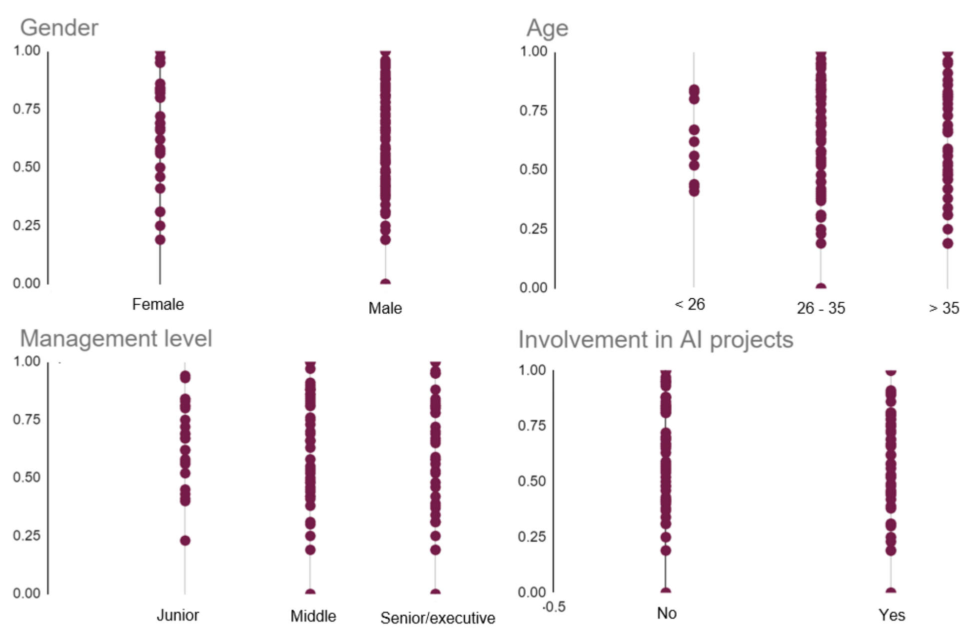
5.3 Statistical results

The following subsections explore the results of each hypothesis tested. Along with the results of each statistical test are the assumptions which give the results validity.

5.3.1 Influence of personal characteristics on moral judgement

The influence of personal characteristics (age, gender, managerial level and involvement in AI projects) on moral judgement was the first of the hypotheses tested. Since the data was at participant-vignette level, it was summarised to participant level by taking the mean moral judgement per participant across vignettes. The distribution of moral judgement for each of the personal characteristics is as follows.

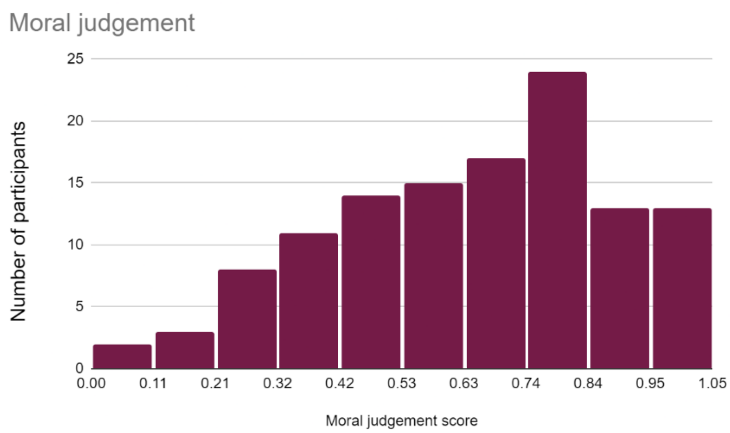
Figure 10: Distribution of moral judgement per personal characteristic



From a graphical perspective, it is difficult to determine whether there is a relationship between moral judgement and personal characteristics. The influence of each of the four personal characteristics was tested with a univariate analysis of variance (ANOVA). ANOVA compares the effect that each independent variable has on the dependent variable (Armstrong et al., 2000). Specifically in this research hypothesis, ANOVA was used to test the influence of personal characteristics on moral judgement. Before performing the ANOVA, assumptions of underlying data were tested (Laerd Statistics, n.d.).

The first assumption for the ANOVA is that the dependent variable is continuous. Continuous variables are variables for which there is an infinite number of values between the lowest and highest possible values. The four dependent variables were measured on a discrete scale and were not continuous in their raw form. There were transformations done on these four dependent variables to create the “moral judgement” construct and this was ultimately a score between zero and one which allowed for a very large, although not unlimited, number of values between zero and one. Based on this transformation, moral judgment resembles a continuous variable and is assumed to be so for the remainder of the analysis.

Figure 11: Distribution of moral judgement



The second assumption for ANOVA is that two or more categorical groups of independent variables exist. In this data, all independent variables have two or more level and this assumption holds. The third assumption of independence of observations also holds. There were no significant outliers in the data, as seen in figures 9 and 10, satisfying assumption four. The fifth assumption is that the dependent variable is normally distributed for each level of the independent variables and this was tested using the Shapiro-Wilks test for normality. This is a test of “departure from normality”, with the null hypothesis that the distributions are not normal (Royston, 1992). This was tested at the 95% confidence level and the results are reported below.

Table 14: Shapiro-Wilk tests of normality

		Statistic	df	Sig.
Gender	*Female*	0.933	30	0.058
	Male	0.964	90	0.013**
Age	*< 26*	0.916	11	0.287
	26 - 35	0.943	66	0.004**
	> 35	0.951	43	0.062
Management level	*Junior*	0.965	25	0.533
	Middle	0.948	51	0.027**
	Senior/executive	0.956	44	0.091
Involvement in AI projects	*No*	942	67	0.004**
	Yes	0.974	53	0.305

*** Significance at the 95% level*

The *Sig.* value in this test represents the probability of finding a deviation from normality in the sample, if the population was normally distributed. As such, a significant result means that the population is likely not normal and neither is the sample. The significant results on some of the categories presented in the output above shows that there are some categories which are normally distributed and others which are not. We reject the null hypothesis of normality for categories: male; 26 - 35 years of age; middle management; and no involvement in AI projects.

The sixth and final assumption is homogeneity of variances. That is, variances are equal for all groups of the independent variables. This was tested used Levene's test at the 95% confidence interval. The null hypothesis for this test is that the variances are equal and the results are reported below.

Table 15: Levene statistic for tests of variance homogeneity

	Levene Statistic	Sig.
Age	1.553	0.216
Managerial level	2.789	0.066
Involvement in AI projects	0.131	0.718
Gender	0.598	0.441

Since no *sig.* values in the output are less than 0.05, we can conclude there is no significance at the 95% confidence level, and, thus, we reject the null hypothesis of equal variances. From here the ANOVA was run and results are as below.

Table 16: The influence of demographic characteristics on moral judgement

	F	Sig.
Gender	0.697	0.406
Age	0.832	0.438
Management level	0.087	0.916
Involvement in AI projects	0.642	0.425

There is no significance to the role of demographic details in describing moral judgement at the 95% confidence level. The adjusted R-squared is 1.9%. Similarly, Frey (2000) found that gender no influence on moral judgement, while McMahon and Harvey (2007) found that decisions made by females were more moral. Regarding age, Morris and McDonald (1995) found that age was the only significant characteristic in one of their two scenarios studied. Their explanation of this was that older people are more likely to remember consequences of immoral behaviour of past lived experiences, while younger people are less aware of social norms. Another of the characteristics found to have no influence on moral judgement is management level. Singer (1996), too, studied the difference between managers and the general population in moral judgements and found no differences between the two groups. As such, the first hypothesis that demographic characteristics influence moral decision-making in an AI-related issue is rejected.

5.5.2 Influence of moral intensity on recognition of an issue as a moral issue

The second of the hypotheses specially focuses on one of the dependent variables and includes the demographic details as covariates. For this analysis, univariate analysis of covariance (ANCOVA) was used. The dependent variable was “this scenario involves a moral issue”. Results of the ANCOVA are as follows.

Table 17: The influence of moral intensity on the recognition moral issues

	F	Sig.
Demographics		
Gender	0.891	0.346
Age	3.634	0.057
Management level	1.642	0.2
Involvement in AI projects	2.577	0.109
Moral intensity		
Social consensus	2.743	0.098
Magnitude of consequences	12.726	<.001**
Likelihood of effect	3.181	0.075

*** Significant at the 95% level*

Of the moral intensity components, magnitude of consequences is significant in recognising an issue as a moral one, at the 95% confidence level.

5.3.3 Factors of moral intensity positively influence moral judgement

Hypothesis three states that social consensus, magnitude of consequences and likelihood of effect. Similarly to the previous hypothesis, ANCOVA is used for this analysis, except here the dependent variable is the moral judgement construct. Results for this hypothesis are as below.

Table 18: The influence of moral intensity on moral judgement

	F	Sig.
Demographics		
Gender	0.53	0.467
Age	1.305	0.254
Management level	1.985	0.159
Involvement in AI projects	11.805	<.001**
Moral intensity		
Social consensus	6.394	0.012**
Magnitude of consequences	36.486	<.001**
Likelihood of effect	10.621	0.001**

*** Significant at the 95% level*

Considering the demographic details as covariates first, there is no change in the influence on age in moral judgement, though this relationship has been found to be otherwise by previous researchers. Morris and McDonald (1995) and Frey (2000) found that change in moral judgement, from increasing moral intensity, was inversely related to age, although the coefficient on age was found to be small, such that the more morally intense the issue became, the less moral their judgement of the issue became. Their explanation of this relationship was that younger people are more likely to go along with morally questionable decisions, than older people are. The influence of gender and management level remain unchanged, while there is now a significant influence of involvement in AI projects. Each of the factors of moral intensity have a significant influence on moral judgement at the 95% confidence level.

5.3.4 Factors of moral intensity interact to influence moral judgement

Beyond the main effects of each component of moral intensity, it is hypothesised that when these components are applied together, they interact to cause additional effects. These are called the “interaction effects” and the significance of these is as reported below.

Table 19: The influence of interaction effects on moral judgement

	F	Sig.
Demographics		
Gender	0.53	0.467
Age	1.305	0.254
Management level	1.985	0.159
Involvement in AI projects	11.805	<.001**
Moral intensity main effects		
Social consensus	6.394	0.012**
Magnitude of consequences	36.486	<.001**
Likelihood of effect	10.621	0.001**
Moral intensity interaction effects		
Social consensus * Magnitude of consequences	0.967	0.326
Social consensus * Likelihood of effect	0.028	0.866
Magnitude of consequences * Likelihood of effect	0.096	0.757
Social consensus * Magnitude of consequences * Likelihood of effect	0.348	0.555

*** Significant at the 95% level*

Components of moral intensity do not interact to cause any significant effect at the 95% confidence level. The adjusted R-squared is 7.1%. Similarly, Frey (2000) found interaction effects of change in *magnitude of consequences, social consensus* and *likelihood of effects* to be unimportant in his study, with these only improving the adjusted R squared by between 1.6% and 2.2% above the main effects.

5.4 Conclusion

The sample was interrogated to check whether it may be representative of the population. It appears the females were under presented for some management levels, although this cannot be concluded with certainty. Reliability of the scale used was tested with Cronbach’s alpha and this was found to be reliable. Validity of the data was tested with Pearson correlation between each of the dependent variables and the moral judgment construct. The significance of the correlation between these items was proof that the data was valid. EFA was used to check that there was only one underlying construct measured by the four dependent variables, moral judgement and this was found to be true given eigen values for less than one after the first factor.

The first hypothesis tested whether personal characteristics (gender, age, management level and involvement in AI projects) influenced decision-making in the context of an AI-based online personalised pricing model and it was concluded that they do not. The second hypothesis tested whether factors of moral intensity (social consensus, magnitude of consequences and likelihood of effect) affected recognition of the AI-based online personalised pricing model as presenting a moral issue and it was found only magnitude of consequences affected the recognition. The third hypothesis tested the influence of the factors of moral intensity on moral judgement of the AI-based online personalised pricing model and each three of the factors were found to be significant. The final hypothesis tested whether there were any interaction effects between the factors of moral intensity and there were found to be no interaction effects.

6. Discussion of results

In this chapter, the findings described in the previous chapter will be discussed. Findings are compared and contrasted to existing research findings and the implications of this is discussed for the context of personalised pricing AI models in particular.

6.1 Influence of personal characteristics on moral judgement

The first of the hypotheses tested was whether personal characteristics influence moral judgment. It is common in the extant literature that personal characteristics are included in studies of moral judgement. Consideration is given to age, gender, involvement in AI projects and management level in this research. A discussion of these is presented below.

6.1.1 Personal characteristics

In their meta-analysis of the moral decision-making literature, Ford and Richardson (1994) found that age had significant influence on moral judgement in some studies and not in others. Of the studies where age was significant, some found that younger people made more moral decisions, while others found that older people make more moral decisions.

In this research, age was found to not have a significant effect on moral judgement. This finding is in line with McMahon and Harvey (2007) who too found age to be insignificant in their study, but contradicts Frey (2000) and Morris and McDonald (1995), who found age does influence moral judgement. Frey (2000) found that as moral intensity is increased, younger people are less likely to change their initial judgment of the scenario. Frey (2000) explained this by younger managers being less willing to oppose decisions than older managers are. Morris and MacDonald (1995), who found the age significance in a scenario related the bribery, explained the significant positive relationship they found by older managers remembering past bribery scandals and being more aware of the legal implications of bribery.

Similar to age, Ford and Richardson (1994) find mixed significance of gender on moral judgement. They reported that in some studies women were reported to behave more ethically, while in other studies, men were reported to behave more ethically. Nevertheless, gender was the second of the demographic characteristics considered in this research, with the finding that gender did not influence moral judgement. Frey (2000) and Morris and McDonald (1995) too found that gender did not have a significant effect on moral judgement.

Perhaps the most interesting of the personal characteristics tested is the involvement in AI projects. There is no literature support that a decision-maker's experience in a given field may influence his or her moral judgement in the field. The issue contingent model, however, does make allowance for the impact of organisational setting on moral judgement. As such, whether a manager is involved in AI projects or not is considered as an organisational setting in this research and is tested along with the personal characteristics of the decision-maker.

A question was posed to managers of whether they are involved in AI projects or not. 44% of managers said they are involved in AI projects. This question was asked to test whether this influences moral judgement. Findings were that involvement in AI projects does not have a significant influence on the moral judgement of this AI-based online personalised pricing model. That is, managers who are involved in AI projects have no different judgement of the solution than those who do not. This finding is noteworthy because one would expect that managers who are developing AI solutions would be sensitive to the moral consequences of them, more so than other managers. These findings illustrate the need for awareness and training on AI ethics among managers. Such training could be at both company and industry level.

In conclusion, personal characteristics were found not to have a significant influence on moral judgement. These findings are largely supported by previous research, though there are some contradictions. It is disappointing to find that managers who are involved in AI projects do not have different moral judgements from managers who do not work in this field. Limitations such as accountability, understanding of societal consequences and lack of professional bodies need to be addressed if organisations and industries which prioritise ethical practice are to be established.

6.2 Influence of moral intensity on recognition of an moral issues

The issue-contingent model builds on the decision-making process first developed by Rest, in which he described the first step in the ethical decision-making process as recognising an issue as a moral one. Rest’s model says that unless decision-makers recognise issues as moral issues, the remaining steps in the moral decision-making process will not be mobilised. Jones (1991) added that the greater the moral intensity, the more recognition of issues as moral issues there will be. In the second hypothesis, this relationship is proven for magnitude of consequences only. That is, social consensus and likelihood of effect are found to have no influence on recognition of an issue as a moral one, but the greater the consequences are, the more recognition of the moral issue there is.

6.2.1 Moral and non-moral problems

This significant influence of magnitude of consequences on recognition of moral issues is supported by the research of Dukerich et.al. (2000) who found that managers were able to distinguish between moral and non-moral problems using four components of moral intensity. From their qualitative research, Dukerich et.al. noted that the phrases managers used in describing moral and non-moral problems differed. Specifically, managers recounted moral problems in terms of how they were affected personally and in terms of loss, that is, the magnitude of personal consequences suffered. Their conclusion from this was that managers may be more cognisent of moral issues when the he or she has personal experience and there is a cost involved.

In the context of this research, there was no indication of whether the managers participating had suffered losses (economically or otherwise) or had experiences (personal or professional) with online personalised pricing models. If the finding by Dukerich et al. (2000) holds, it would present possible mechanisms for growing recognition amongst managers that their work may indeed involve moral issues and improving the moral judgement amongst these managers. Practically, managers (and developers) could be asked to recount times when they, or people they know, were affected by online personalised pricing and, specifically, times when this interaction resulted in a loss for the individuals. By doing so at project initiation, managers may recognise that they are dealing with moral issues early on in projects and their subsequent decisions may reflect their intentions to behave morally in the situation.

In addition to this, Dukerich et al. (2000) concluded that managers were more able to identify moral issues as ones where employees were directly affected, such as lay-offs, as opposed to operational decisions, such as purchasing new equipment, where employees may have been indirectly affected. The implication of Dukerich et al.'s (2000) finding for this research is that managers should understand that AI solutions may have a direct influence on people and their well-being. Education for managers around these influences may mean they recognise AI projects as having a moral consequence and ultimately make more moral decisions.

6.3 Influence of moral intensity on moral judgement

The central hypothesis in this study is the influence of moral intensity on moral judgement regarding an AI-based online personalised pricing model. In this experiment, three factors of moral intensity were considered: social consensus, magnitude of consequences and likelihood of effect. Results showed that all three of these factors had a significant influence on moral judgement. Each of these factors are discussed below.

6.3.1 Social consensus

In the issue-contingent model, Jones (1991) described social consensus as the "social agreement" around what behaviours are considered good and what behaviour are considered bad. This influences the moral decision-making processes by providing a benchmark for decision-makers to judge good and bad behaviour. In this experiment, the greater the social agreement on whether a personalised pricing model is accepted by the industry, the more moral the decision around deploying the model. High and low levels of social consensus were presented in the vignettes as follows: for the low moral intensity case – "there is no agreement amongst the public on whether using such models is acceptable or not"; for the high intensity case – "there is clear agreement amongst the public that models such as this are unfairly discriminatory".

Findings were that social consensus does indeed influence moral judgement. This finding is supported by previous researcher. Singer (1996) found that social consensus was in fact the greatest contributor to moral judgment, compared to the other factors of moral intensity. This contrasted with the general public, for which magnitude of consequences was the greatest influencer of moral judgement. Singer's reasoning for this phenomenon was that managers are more aware of the actions of their peers and the "prevailing practices" of the time. Singer (1996) even likened this behaviour to the "herd" mentality that is witnessed in stock markets.

6.3.2 Magnitude of consequences

Magnitude of consequences is the extent of the impact that decisions can have (Jones, 1991). The issue-contingent model proposes that the greater the potential effect a decision can have, the greater the moral judgement applied will be. This happens because decision-makers weigh the impact of their decisions and are more considerate in decisions which will have a greater impact. The magnitude of consequences was varied in this experiment, with the low intensity vignettes indicating a price difference of only a few cents, while the high intensity vignettes had a substantial price difference.

Magnitude of consequences was found to have a significant effect on moral judgement, such that the greater the consequences the more moral the decision. McMahon and Harvey (2007) concluded similarly from their research. In their discussion of their results, McMahon and Harvey (2007) raised an interesting point that magnitude of consequences can act as an informant or a deterrent of ethical behaviour. Literature has focused on the former, in that the greater the consequences from a set of actions, the more unethical it is to continue with a given set of actions. The example given by McMahon and Harvey (2007) was an individual running to catch a train. If he slightly bumps one person on his way, it seems acceptable. However, if he shoves the person to the extent that the person becomes seriously injured, the behaviour would be considered unethical. This is an example of magnitude of consequences informing ethical behaviour.

On the contrary, McMahon and Harvey (2007) provided another example of the influence of magnitude of consequences, where this time it acts as a deterrent of ethical behaviour. The example they provided is stealing a $100 compared to stealing $1,000. As the issue-contingent theory would suggest, the decision to not steal $1,000 is more moral than the decision to not steal $100, but is it really? In this way, magnitude of consequences actually serves as a deterrent of ethical behaviour.

In the context of an AI-based online price discriminating model, is an impact of only a few cents more morally justifiable than a substantial price difference? A similar sentiment was found during the pilot of this survey, in which two pilot participants stated that the pricing model is immoral for any level of magnitude of consequences, or any other of the factors of moral intensity. McMahon and Harvey (2007) suggested that researchers cannot conclude that magnitude of consequences has a positive or negative influence on moral decision-making, unless these nuances were planned for, which they say are usually not. Nevertheless, the significance of magnitude of effect is supported by the findings of Morris and McDonald (1991), who recommended that training on the perceived costs can positively influence moral judgement.

6.3.3 Likelihood of effect

The likelihood of effect is the probability that the potential consequences will be realised. The influence of this on moral intensity is such that the greater the probability of consequences realising, the greater the care taken in decision-making and the more mora the ultimate decision (Jones, 1991). In this research design, likelihood of effect was varied between low and high with the following statement: the chance that the price difference materialises is low/high. Findings were that the likelihood of effect is a significant influencer of moral judgement.

6.4 Influence of interaction effects on moral judgement

Beyond the main effects of each factor of moral intensity, the interactions effects of these factors were tested too. That is, how do social consensus, magnitude of consequences and likelihood of effect interact to influence decision-making. Similar to Frey (2000) who found no interaction effects of “appreciable magnitude”, there was no significance of the interaction effects in this research either. Ultimately, the influence of each of these factors combined do not influence moral judgment by any more than they do alone.

6.5 Conclusion

In earlier chapters, the potential harm from AI-based personalised pricing models was explored. The ethics of online personalised pricing have been questioned by many researchers and there have been proven incidences of personalised pricing models being unfairly discriminatory. In this research, the moral intensity construct from Jones (1991) issue-contingent model was put to the test. The objective was to understand whether this construct may influence the decisions of managers in an AI-related issue.

The influence of personal characteristics of managers on decision-making was tested first. The insignificance of three of the personal characteristics (age, gender, management level) are largely unsurprising and supported by literature. The fourth characteristic of involvement in AI projects, however, is concerning. One would expect managers who work with AI solutions to have some difference in moral judgement compared to managers who do not. In particular, managers in the AI field should identify this AI solution as presenting moral issues, more so than managers who do not work in this field.

The three components of moral intensity studied were found to have an impact on decision-making. The implication of this is that if business managers are trained to include considerations of moral intensity in their decision-making, they may be able to make more ethical decisions. The benefit of using the issue-contingent model as a decision-making framework is that it guides managers on the things they should focus on when making a moral decision (Frey, 2000). In the context of AI, this focus is even more crucial, given the technical complexity, unpredictable consequences and ethical grey areas which come with many AI solutions. As a straightforward and practical framework, managers are able to use the issue-contingent model to structure their decision-making around the size, immediacy, concentration and likelihood of potential consequences arising and the social agreeability of actions. Further to this, managers should train their development teams on the issue-contingent model so that AI developers themselves know how to identify moral issues relating to their work.

On the issue of salience, McMahon and Harvey (2007) raised the question of whether it is ever appropriate to study ethical issues using surveys as the measurement instruments. They noted that the salience of these issues in real life may not come across in equal magnitude to what they do on paper. Additionally, they suggested that some scenarios may be more salient to certain participants than to others, and this may result in the findings of studies not being generalisable to populations. In support of paper-based scenarios, such as are used in this study, McMahon and Harvey (2007) argued that oftentimes business managers are faced with scenarios on paper on which to base their decisions. The example they gave is of a manager deciding to lay off a proportion of his workforce, while in the boardroom with one to-be laid off employee. In the company of one employee, the salience is substantially lower than all employees who face potential lay off, but this is a typical moral issue that a business manager may be faced with. As such, the salience of the price discrimination model used in this study can be considered as a typical moral issue business managers may face in their use of AI.

7. Conclusions and recommendations

The final chapter in this paper consolidates the most important findings and highlights the theoretical contribution and impactions from these. Additionally, limitations of this research are discussed and recommendations are made for future research.

7.1 Principle conclusions

The issue of AI ethics is topical, with many consulting (McKinsey and Capgemini for example) and technology firms (Google and IBM for example) racing to publish their AI ethics “principles” and “guidelines”. Given so many real-world ethics violations still being reported from these solutions, however, it would appear that these are ineffective and the literature offers various reasons for this. AI-based online personalised pricing has come under scrutiny from scholars due to the re-distribution of welfare it causes away from customers and towards sellers. Since it is not possible to program ethical decision-making into AI solutions such an online personalised pricing, the onus remains on business managers to ensure that AI solutions they are deploying do not perpetrate harm.

Jones’ (1991) issue-contingent model of moral decision-making may be one framework that describes the nature of ethical decision-making in the context of AI. The model builds on work by previous ethical decision-making scholars, to include factors of the issue itself. These factors are: magnitude of consequences, social consensus, likelihood of effect, proximity, temporal immediacy and concentration of effect. The influence of the first three of these factors on moral judgement were explored in this study. Using experimental vignette methodology, a hypothetical scenario was presented to managers detailing the use of a personalised pricing AI solution. The three factors of moral intensity were then varied from low to high levels and the moral judgement of the managers was captured in four questions relating to their thoughts around whether the AI solution presents a moral issue and whether the solution should be used.

The first of the hypotheses tested whether personal characteristics of the decision maker (age, gender, management level and involvement in AI projects) influences the moral judgement. Moral judgement was a construct created from the four dependent variables, normalised to be a score out of one. Results indicated that these factors do not influence moral judgement and some previous research supports these results, though other researchers have found age to be significant in determining moral decision making. The second hypothesis was that the factors of moral intensity influence the recognition of an issue as a moral issue, such that the higher the level of intensity of the factors, the more recognition of the issue as a moral one. Findings were that only magnitude of consequences influences recognition of moral issues.

Hypothesis three was that the factors of moral intensity influence moral judgement of the personalised pricing AI model and this was found to be true. From the final hypothesis, which tested the significance of interaction effects, it was concluded that there are no significant interaction effects in this model and this is in line with Frey's (2000) findings.

7.2 Theoretical contribution

The relevance of personal characteristics in the issue-contingent model has been supported by some researchers and not supported by others. In this particular context, personal characteristics have not been found to have a significant influence on decision-making. Jones' (1991) proposition that moral intensity influences recognition of moral issues is somewhat true in the AI context. The only significant factor of moral intensity was found to be magnitude of consequences.

7.3 Implications

A crucial step in the moral decision-making process is recognition that one is dealing with issues of a moral nature. Without this recognition early in the decision-making process, the decision-maker will not recognise that the outcome of the decision has a substantial effect on the well-being of others and will not give the decision the thoughtful consideration it demands. Findings in this research indicated that the scale of consequences from the AI-based online personalised pricing model is the only factor significant to managers in the recognition of the solution as presenting moral issues.

Social agreement on the fairness of the model and the probably of the potential consequences realising, do not influence the recognition of moral issues in this context. The implication of this finding is that managers must be made aware of the scale of consequences of their AI solutions. In particular, there is research to suggest that if the consequences are posed as losses and if the consequences can be framed as affecting managers personally, there will be more recognition of the moral issues at hand.

Further to recognition of the moral issue, overall moral judgement is improved as the social consensus, magnitude of consequences and likelihood of effect rise. In practice, if managers are aware: of the social disagreement in the use of AI-based online personalised pricing models; the negative effects these may have on individuals; and the probability that these effects will realise, their decisions around deployment of such models may be more moral.

7.4 Limitations of the research

Design, data and analysis limitations are present in this research. These are discussed thought this paper and the most important of which are presented below.

7.4.1 Design limitations

Social desirability reporting is an issue in studies that attempt to measure ethical behaviours. The phenomenon of social desirability reporting is one in which participants overstate their propensity to behave ethically in their survey answers, compared to how they would behave in real life. Earlier researchers of moral intensity have included measures to account for the effects of social desirability reporting in their results. Morris and McDonald (1995), for example, used an amended version of the Blanchard Inventory of Desirable Reporting (BIDR–6) scale to measure participants' propensity for self-deception and impression management. Including a measure such as this would have indicated how far participants' survey answers were to their behaviors in the real world and this would influence the implications of the study. As an example, if participants report their ethical decision making to be very high but they also score highly on social desirability reporting, it would indicate that, in reality, these participants may not behave as ethically as they would have led the researcher to believe.

The measurement instrument used may play a role in determining how participants respond to the survey questions. As an example, the order of questions may result in some questions being more salient than others. The scenario used in this survey, of the personalised pricing AI model, and the vignettes based off this, representing low of high levels of moral intensity, may or may not have been appropriate to measure the construct of moral judgement. Additionally, participants may not have noticed the difference between vignettes. In future research, this measurement instrument should be reviewed by ethics scholars to ensure that it is appropriate for measuring moral judgement.

7.4.2 Data and analysis limitations

The sample size of 120 participants is smaller than one would have liked. For the population of all managers in South Africa, which can be assumed to be over 100,000, a much larger sample would have been needed to capture all the possible variation in the population (Israel, 1992). Additionally, the number of participants per each group of the personal traits varied greatly, with females generally being under-represented in the sample.

The univariate analyses of variance and covariance done require that the dependent variable is continuous. The four original dependent variables were not continuous since they were measured on a five-point Likert scale. Combining these variables and normalising them to score between zero and one, called “moral judgment”, allowed them to take on a more continuous form, but future researchers should use a scale which contains more points, such as Frey (2000), who used a nine-point Likert scale.

After combining these variables into one score, some of the assumptions of the underlying distributions were still not met. One of these assumptions was that the dependent variable is normally distributed for each level of the independent variables and this was tested using the Shapiro-Wilks test for normality. We rejected the null hypothesis of normality for categories: male; 26 - 35 years of age; middle management; and no involvement in AI projects. When the underlying data fails assumptions such as this, the reliance that can be placed on results is influenced.

7.5 Suggestions for further research

Several limitations have been discussed in the body of this research and summarised in this chapter. Future researchers should, firstly, ensure that these are addressed so that reliance can be placed on the results of the analysis. Secondly, in this research only three of the six factors of moral intensity were investigated. Factors not investigated were: temporal immediacy, concentration of effect and proximity. The reasons for excluding these were largely related to making the scenario-vignette design of the measurement instrument more manageable for participants to complete, however, these three factors may have a significant influence on decision-making in this context. Further researchers should attempt to include all six factors when measuring decision making in this context. This would mean that the scenario-vignette design used in this research needs to be reconsidered such that it is not too cumbersome for participants to complete, but that the within-subjects approach can still be maintained.

Given the significance of magnitude of consequences in the recognition of AI solutions as causing moral issues, this should be explored more deeply. In particular, the findings by Dukerich et al. (2000) should be tested in this context. That is, do managers recognise moral issues from AI solutions more when they have experienced personal losses due to such solutions.

8. References

Aguinis, H., & Bradley, K. J. (2014). Best Practice Recommendations for Designing and Implementing Experimental Vignette Methodology Studies. *Organizational Research Methods*, *17*(4), 351–371. https://doi.org/10.1177/1094428114547952

Aguinis, H., Hill, N. S., & Bailey, J. R. (2019). Best practices in data collection and preparation: Recommendations for reviewers, editors, and authors. *Organizational Research Methods*, 1–16. https://doi.org/10.1177/1094428119836485

Aguinis, H., & Solarino, A. M. (2019). Transparency and replicability in qualitative research: The case of interviews with elite informants. *Strategic Management Journal*, *40*(8), 1291–1315. https://doi.org/10.1002/smj.3015

Ahsen, M. E., Ayvaci, M. U. S., & Raghunathan, S. (2019). When algorithmic predictions use human-generated data: A bias-aware classification algorithm for breast cancer diagnosis. *Information Systems Research*, *30*(1), 97–116. https://doi.org/10.1287/isre.2018.0789

Aksoy, N. C., Kabadayi, E. T., Yilmaz, C., & Alan, A. K. (2021). A typology of personalisation practices in marketing in the digital age. *Journal of Marketing Management*, *37*(11–12), 1091–1122. https://doi.org/10.1080/0267257X.2020.1866647

Armstrong, R. A., Slade, S. V., & Eperjesi, F. (2000). An introduction to analysis of variance (ANOVA) with special reference to data from clinical experiments in optometry. *Ophthalmic and Physiological Optics*, *20*(3), 235–241. https://doi.org/10.1016/S0275-5408(99)00064-2

Atzmüller, C., & Steiner, P. M. (2010). Experimental vignette studies n survey research. *Methodology*, *6*(3), 128–138. https://doi.org/10.1027/1614-2241/a000014

Borgesius, Frederik, Z., & Poort, J. (2017). Online Price Discrimination and EU Data Privacy Law. *Journal of Consumer Policy*, *40*(3), 347–366. https://doi.org/10.1007/s10603-017-9354-z

Botta, M., & Wiedemann, K. (2020). To discriminate or not to discriminate? Personalised pricing in online markets as exploitative abuse of dominance. *European Journal of Law and Economics*, *50*(3), 381–404. https://doi.org/10.1007/s10657-019-09636-3

Butcher, J., & Beridze, I. (2019). What is the state of artificial intelligence governance globally? *RUSI Journal*, *164*(5–6), 88–96. https://doi.org/10.1080/03071847.2019.1694260

Cahn, A., Alfeld, S., Barford, P., & Muthukrishnan, S. (2016). An empirical study of web cookies. *25th International World Wide Web Conference, WWW 2016*, 891–901. https://doi.org/10.1145/2872427.2882991

Cannataci, J., Falce, V., & Pollicino, O. (2020). *Legal challenges of big data*. Edward Elgar Publishinh Limited.

Capgemini. (2020). *The AI-powered enterprise: Unlocking the potential of AI at scale*. https://www.capgemini.com/research/the-ai-powered-enterprise/

Charan, J., & Biswas, T. (2013). *How to Calculate Sample Size for Different Study Designs in Medical Research ? 35*(2), 24049221. https://doi.org/10.4103/0253-7176.116232

Chatfield, C., & Collings, A. J. (1980). *Introduction to multivariate analysis* (First). Routledge.

Craft, J. L. (2013). A Review of the empirical ethical decision-making literature: 2004-2011. *Journal of Business Ethics*, *117*(2), 221–259. https://doi.org/10.1007/s10551-012-1518-9

Crane, A., Matten, D., Glozer, S., & Spencer, L. J. (2016). Chapter 3: Evaluating business ethics. In *Business Ethics*. Oxford.

Datta, A Carl, M Datta, A. (2015). Automated experiments on ad privacy settings. *Proceedings on Privacy Enhancing Technologies*, 92–112.

Drucker, P. (1981). What is business ethics? *The Public Interest*, 133–150. https://doi.org/10.1007/978-94-015-8183-7_7

Dubber, A. D., Pasquale, F., & Das, S. (2020). The Oxford handbook of ethics of AI. In *Oxford University Oress*.

Dukerich, J. M., Waller, M. J., George, E., & Huber, G. P. (2000). Moral intensity and managerial problem solving. *Journal of Business Ethics*, *24*(1), 29–38. https://doi.org/10.1023/A:1006030109047

Executive Office of the President. (2015). Big data and differential Pricing. *Whitehouse Government*, *February*, 1–22. https://www.whitehouse.gov/sites/default/files/docs/Big_Data_Report_Nonembargo_v2.pdf

Ferrell, O. C., & Ferrell, L. (2021). Applying the Hunt Vitell ethics model to artificial intelligence ethics. *Journal of Global Scholars of Marketing Science*, *31*(2), 178–188. https://doi.org/10.1080/21639159.2020.1785918

Floridi, L., & Sanders, J. W. (2004). On the morality of artificial agents. *Minds and Machines*, *14*(3), 349–379. https://doi.org/10.1023/B:MIND.0000035461.63578.9d

Floridi, L., & Taddeo, M. (2016). What is data ethics? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 1–4.

Ford, R. C., & Richardson, W. D. (1994). Ethical decision making review. *Journal of Business Ethics*, *13*(3), 205–221.

Freeman, R. E., & Ramakrishna Velamuri, S. (2006). A new approach to CSR: Company stakeholder responsibility. *Corporate Social Responsibility: Reconciling Aspiration with Application*, 9–23. https://doi.org/10.1057/9780230599574

Frey, B. F. (2000). The impact of moral intensity on decision making in a business context.

Journal of Business Ethics, *26*(3), 181–195. https://doi.org/10.1023/A:1006139124110

Google. (n.d.). *Our principles*. https://ai.google/principles/

Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines*, *30*(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8

Hannak, A., Soeller, G., Lazer, D., Mislove, A., & Wilson, C. (2014). Measuring price discrimination and steering on E-commerce web sites. *Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC*, 305–318. https://doi.org/10.1145/2663716.2663744

Heinrichs, B., & Eickhoff, S. B. (2020). Your evidence? Machine learning algorithms for medical diagnosis and prediction. *Human Brain Mapping*, *41*(6), 1435–1444. https://doi.org/10.1002/hbm.24886

Hoffman, W. M., & Moore, J. M. (1982). What is business ethics? A reply to Peter Drucker. *Journal of Business Ethics*, *1*(4), 293–300. https://doi.org/10.1007/BF00382818

Huang, J., & Zhou, L. (2018). Timing of web personalization in mobile shopping: A perspective from uses and gratifications theory. *Computers in Human Behaviour*, *88*, 103–113.

Israel, G. D. (1992). Determing sample size. *Program Evaluation and Organizational Development. University of Florida*, *10*(2), 84–86.

Jones, T. M. (1991). Ethical decision making by individuals in organizations: An issue-contingent model. *Academy of Management Review*, *16*(2), 366–395. https://doi.org/10.1353/frf.2017.0009

Kelley, P. C., & Elm, D. R. (2003). The effect of context on moral intensity of ethical issues: Revising Jones's issue-contingent model. *Journal of Business Ethics*, *48*(2), 139–154. https://doi.org/10.1023/B:BUSI.0000004594.61954.73

Kohlberg, L. (1976). *Moral stages and moralisation: The cognitive development approach*. Holt, Rinehart and Winston.

Kohlberg, L., & Hersh, R. H. (1977). Moral development: A review of the theory. *Theory Into Practice*, *16*(2), 53–59. https://doi.org/10.1080/00405847709542675

Laerd Statistics. (n.d.). *Two-way ANOVA in SPSS statistics*. Retrieved October 14, 2021, from https://statistics.laerd.com/spss-tutorials/two-way-anova-using-spss-statistics.php

Lakshmanan, S. (2019). How, When, and Why Should You Normalize / Standardize / Rescale Your Data? *Towards AI — The Best of Tech, Science, and Engineering*, 1. https://towardsai.net/p/data-science/how-when-and-why-should-you-normalize-standardize-rescale-your-data-3f083def38ff

Lambrecht, A., & Tucker, C. (2019). Algorithmic bias? An empirical study of apparent gender-based discrimination in the display of stem career ads. *Management Science*, *65*(7), 2966–2981. https://doi.org/10.1287/mnsc.2018.3093

Li, S., & Karahanna, E. (2015). Online recommendation systems in a B2C e-eommerce contex: A review and future directions. *Journal of the Association for Information Systems*, *16*(2), 72–107.

Liao, S. M. (2020). The ethics of artificial intelligence. In *IFIP Advances in Information and Communication Technology* (Vol. 555). Oxford University Press. https://doi.org/10.1007/978-3-030-64246-4_5

Martin, K. (2019). Ethical implications and accountability of algorithms. *Journal of Business Ethics*, *160*(4), 835–850. https://doi.org/10.1007/s10551-018-3921-3

McFarland, S. G. (1981). Effects of question order on survey responses. *Public Opinion Quarterly*, *45*, 208–215.

McKinsey Global Institute. (2019). *Tackling bias in artificial intelligence (and in humans)* (pp. 1–8). https://www.mckinsey.com/featured-insights/artificial-intelligence/tackling-bias-in-artificial-intelligence-and-in-humans

McMahon, J. M., & Harvey, R. J. (2007). The effect of moral intensity on ethical judgment. *Journal of Business Ethics*, *72*(4), 335–357. https://doi.org/10.1007/s10551-006-9174-6

Mehmood, A., Natgunanathan, I., Xiang, Y., Hua, G., & Guo, S. (2016). Protection of big data privacy. *IEEE Access*, *4*, 1821–1834.

Microsoft Corporation. (2018). *Microsoft AI principles*. 1. https://www.microsoft.com/en-us/ai/our-approach-to-ai

Minoja, M. (2012). Stakeholder management theory, firm strategy, and ambidexterity. *Journal of Business Ethics*, *109*(1), 67–82. https://doi.org/10.1007/s10551-012-1380-9

Morris, S. A., & McDonald, R. A. (1995). The role of moral intensity in moral judgments: An empirical investigation. *Journal of Business Ethics*, 463–479. https://doi.org/10.1007/978-94-007-4126-3_23

Mwagiru, N. (2020). Women leaders in business and higher education in Africa: Are they reaching the top? *USB Management Review*, 2020.

Neubert, M. J., & Montañez, G. D. (2020). Virtue as a framework for the design and use of artificial intelligence. *Business Horizons*, *63*(2), 195–204. https://doi.org/10.1016/j.bushor.2019.11.001

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453. https://doi.org/10.1530/ey.17.12.7

OECD. (2018). *Personalised pricing in the digital era – Note by the United Kingdom*. *33*, 1–2. www.oecd.org/daf/competition/personalised-pricing-in-the-digital-era.htm%0Awww.oecd.org/daf/competition/market-concentration.htm.

OECD. (2021). The effects of online disclosure about personalised pricing on consumers: Results from a lab experiment in Ireland and Chile. *OECD Digital Economy Papers*, *303*. https://doi.org/10.1787/1ce1de63-en

Pandey, A., & Caliskan, A. (2021). Disparate Iimpact of artificial intelligence bias in ridehailing economy's price discrimination algorithms. *AIES 2021 - Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, 822–833. https://doi.org/10.1145/3461702.3462561

Pandya, J. (2019). *The weaponization of artificial intelligence* (p. 86283686).

Perreault, W. (1975). Controlling order-effect bias. *Public Opinion Quarterly*, *39*(4), 544–551. https://doi.org/https://doi.org/10.1086/268251

Piano, S. Lo. (2020). Ethical principles in machine learning and artificial intelligence: cases from the field and possible ways forward. *Humanities and Social Sciences Communications*, *7*(1), 1–7. https://doi.org/10.1057/s41599-020-0501-9

Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review*, *46*(1), 192–210. https://doi.org/10.5465/AMR.2018.0072

Royston, P. (1992). Approximating the Shapiro-Wilk W-test for non-normality. *Statistics and Computing*, *2*(3), 117–119. https://doi.org/10.1007/BF01891203

Samuel, A. L. (1959). Some studies in machine learning usng the game of checkers. *IBM Journal of Research and Development*, *3*(3), 210–229. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5392560

Sandberg, J. (2008). Understanding the separation thesis. *Journal of Business Ethics*, *18*(2), 213–232. https://www.jstor.org/stable/pdf/27673229.pdf?refreqid=excelsior%3A9b8cbf05dca6e74701044f96b228434b

Saunders, M., & Lewis, P. (2018). *Doing research in business and management* (Second). Pearson.

Şentürk, D., & Müller, H. G. (2009). Covariate-adjusted generalized linear models. *Biometrika*, *96*(2), 357–370. https://doi.org/10.1093/biomet/asp012

Shin, D. (2020). User perceptions of algorithmic decisions in the personalized AI system: Perceptual evaluation of fairness, accountability, transparency, and explainability. *Journal* *of Broadcasting and Electronic Media*, *64*(4), 541–565. https://doi.org/10.1080/08838151.2020.1843357

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of cronbach's alpha. *Psychometrika*, *74*(1), 107–120. https://doi.org/10.1007/s11336-008-9101-0

Singer, M. S. (1996). The role of moral intensity and fairness perception in judgments of ethicality: A comparison of managerial professionals and the general public. *Journal of Business Ethics*, *15*(4), 469–474. https://doi.org/10.1007/BF00380366

Statistics South Africa. (2021). Quarterly Labour Force Survey. *Quarterly Labour Force Survey*, *August*, 1–70. http://www.statssa.gov.za/publications/P0211/P02111stQuarter2019.pdf

Steinberg, E. (2020). Big data and personalized pricing. *Business Ethics Quarterly*, *30*(1), 97–117. https://doi.org/10.1017/beq.2019.19

Sweeney, L. (2013). Discrimination in online add delivery. *Communications of the ACM*, *56*(5), 44–54. https://doi.org/10.1145/2447976.2447990

Trevino, L. K. (1986). Ethical Decision Making in Organizations: A Person-Situation Interactionist Model. *Academy of Management Review*, *11*(3), 601–617. https://doi.org/10.5465/amr.1986.4306235

Velasquez, M., & Rostankowski, C. (1985). *Ethics: Theory and practice*. Prentcie Hall.

Vulkan, N., & Shem-Tov, Y. (2015). A note on fairness and personalised pricing. *Economics Letters*, *136*, 179–183. https://doi.org/10.1016/j.econlet.2015.09.012

Warne, R. T. (2014). A primer on multivariate analysis of variance (MANOVA) for behavioral scientists. *Practical Assessment, Research and Evaluation*, *19*(17), 1–10.

Weber, B. Y. E. (2018). Creating advanced learning transfer with AI coaching. *Training and Development*, *December*, 9–12.

Wicks, A. C. (1996). Overcoming the separation thesis: The need for a reconsideration of business and society research. *Business and Society*, *35*(1), 89–118. https://doi.org/10.1177/000765039603500108

Wong, B. (2021). Online personalised pricing as prohibited automated decision-making under Article 22 GDPR: a sceptical view. *Information and Communications Technology Law*, *30*(2), 193–207. https://doi.org/10.1080/13600834.2020.1860460

Wu, C.-H. (2007). An Empirical Study on the Transformation of Likert-scale Data to Numerical Scores. *Applied Mathematical Sciences*, *1*(58), 2851–2862.

Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*, *9*(2), 79–94. https://doi.org/10.20982/tqmp.09.2.p079

Zanker, M., Rook, L., & Jannach, D. (2019). Measuring the impact of online personalisation: Past, present and future. *International Journal of Human Computer Studies*, *131*(June), 160–168. https://doi.org/10.1016/j.ijhcs.2019.06.006

9. Appendices

9.1 Full survey

Consent

Firstly, thank you for your time to participate in my research. This data will be used in my research project, in partial fulfilment of my MBA from the University of Pretoria's Gordon Institute of Business Science.

I am studying issues related to artificial intelligence (AI). No personally identifiable data will be collected during the process and your response is anonymous. The survey should take no more than ten minutes to complete.

Your participation is voluntary and you may withdraw at any time without penalty. By proceeding to the survey, you indicate that you voluntarily participate in the research.

For any comments or questions, feel free to reach out to me or my supervisor.

Researcher:	Abigail Britton	208036576@mygibs.co.za	072 805 6249
Supervisor:	Dr Frank Magwegwe	magwegwef@gibs.co.za	

Demographic details

Age:	*<26*	*26 - 35*	*36 - 45*	*46 - 55*	*>55*
Gender:	*male*	*female*	*other*		
Management level:	*junior*	*middle*	*senior/executive*		
Are you involved in AI projects?		*yes*	*no*		

The scenario

Artificial Intelligence (AI) can be broadly described as computer programs developed to simulate intelligent behaviour, by analysing vast amounts of data.

The AI development team at SupaFood, a nationwide grocer, has deployed a machine learning model that personalises pricing on food items. The model works by estimating an individual's willingness to pay for food items, based on various factors. Since deploying the model, SupaFood's profits margins have increased substantially.

There have, however, been concerns around the way the model price discriminates. The model could potentially exploit people who have a necessity for certain food items and no or little access to substitutes. This would mean that some people may potentially be charged higher prices than other people are charged for the same items.

Vignette 1:

There is *no agreement* amongst the public on whether using such models is acceptable or not.
The potential price increase experienced by some people would be *only a few cents*.
The chance that the above-mentioned price difference materialises is *low*.

Q1a: The decision made to deploy the model was morally correct

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q1b: I would oppose the decision to deploy this model

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q1c: This scenario involves a moral issue

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q1d: Should this model be used?

1 - Yes 2 – No

Vignette 2:

There is *no agreement* amongst the public on whether using such models is acceptable or not.
The potential price increase experienced by some people would be *only a few cents*.
The chance that the above-mentioned price difference materialises is *high*.

Q2a: The decision made to deploy the model was morally correct

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q2b: I would oppose the decision to deploy this model

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q2c: This scenario involves a moral issue

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q2d: Should this model be used?

1 - Yes 2 – No

Vignette 3:

There is *no agreement* amongst the public on whether using such models is acceptable or not.
The potential price increase experienced by some people would be *substantial*.
The chance that the above-mentioned price difference materialises is *low*.

Q3a: The decision made to deploy the model was morally correct

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q3b: I would oppose the decision to deploy this model

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q3c: This scenario involves a moral issue

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q3d: Should this model be used?

1 - Yes 2 – No

Vignette 4:

There is *clear agreement* amongst the public that models such as this are unfairly discriminatory.
The potential price increase experienced by some people would be *only a few cents*.
The chance that the above-mentioned price difference materialises is *low*.

Q4a: The decision made to deploy the model was morally correct

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q4b: I would oppose the decision to deploy this model

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q4c: This scenario involves a moral issue

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q4d: Should this model be used?

1 - Yes 2 – No

Vignette 5:

There is *no agreement* amongst the public on whether using such models is acceptable or not.
The potential price increase experienced by some people would be *substantial*.
The chance that the above-mentioned price difference materialises is *high*.

Q5a: The decision made to deploy the model was morally correct

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q5b: I would oppose the decision to deploy this model

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q5c: This scenario involves a moral issue

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q5d: Should this model be used?

1 - Yes 2 – No

Vignette 6:

There *is clear agreement* amongst the public that models such as this are unfairly discriminatory.
The potential price increase experienced by some people would be *only a few cents*.
The chance that the above-mentioned price difference materialises is *high*.

Q6a: The decision made to deploy the model was morally correct

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q6b: I would oppose the decision to deploy this model

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q6c: This scenario involves a moral issue

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q6d: Should this model be used?

1 - Yes 2 – No

Vignette 7:

There is *clear agreement* amongst the public that models such as this are unfairly discriminatory.
The potential price increase experienced by some people would be *substantial*.
The chance that the above-mentioned price difference materialises is *low*.

Q7a: The decision made to deploy the model was morally correct

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q7b: I would oppose the decision to deploy this model

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q7c: This scenario involves a moral issue

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q7d: Should this model be used?

1 - Yes 2 – No

Vignette 8:

There is *clear agreement* amongst the public that models such as this are unfairly discriminatory.
The potential price increase experienced by some people would be *substantial*.
The chance that the above-mentioned price difference materialises is *high*.

Q8a: The decision made to deploy the model was morally correct

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q8b: I would oppose the decision to deploy this model

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q8c: This scenario involves a moral issue

1 - Strongly disagree 2 - Disagree 3 - Neural 4 - Agree 5 - Strongly agree

Q8d: Should this model be used?

1 - Yes 2 – No

9.2 Exploratory factor analysis results

KMO and Bartlett's Test

Kaiser-Meryer-Olkin measure of sampling adequacy		0.802
Barlett's test of sphericity	Approximate Chi-Square	1,950.801
	df	6
	Sig.	0.000

Factor matrix

Dependent variable	**Factor 1**
Should this model be used	0.802
The decision made to deploy the model was morally correct	0.740
I would oppose the decision to deploy the model	0.916
The scenario involves a moral issue	0.705