Master's thesis

2021

Emilie Kristin Haugstulen

Master's thesis

NTNU
Norwegian University of
Science and Technology
Faculty of Economics and Management
Department of International Business

Emilie Kristin Haugstulen

# "You might also like"
# - The technological consumers understanding of transparent AI

June 2021

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of
Science and Technology

Norwegian University of
Science and Technology

# "You might also like"
# - The technological consumers understanding of transparent AI

**Emilie Kristin Haugstulen**

AE511816 1 Master Thesis International Business - discipline oriented
Submission date: June 2021
Supervisor: Mark Pasquine
Co-supervisor: Lena Vatne Bjørlo

Norwegian University of Science and Technology
Department of International Business

## Preface

This thesis represents the end of my Master of Science in International Business and Marketing at NTNU Ålesund. The past two years has been challenging and difficult, however it has been an interesting and a great learning experience. The research topic is based on a great personal interest of AI and an awareness that this technology will be more present in the years to come.

First, I would like to thank Lena V. Bjørlo for sharing all her knowledge about AI in marketing and taking the time to make this experiment with me. It was inspiring to be introduced to how AI influences our autonomy and that we as consumers should strive for more transparent algorithms. Therefore, I am deeply grateful for this partnership with Lena. I am also grateful for all the respondents providing the necessary data, so we were able to carry out our research.

Most importantly, I would like to thank my supervisor Mark Pasquine. His guidance and support have been helpful through this process, giving constructive feedback and sharing his valuable knowledge and expertise.

Ålesund, June 2021

Emilie K. Haugstulen

## Summary

Artificial intelligence is becoming an important tool in digital marketing employed by marketers to influence consumer in an online decision-making process. Recommendation algorithms can provide consumers with better products to fulfill needs and preferences based on their online behavior. However, consumers are rarely given information or explanations on which data that are used, and this could be solved by transparency in the algorithm. This thesis addresses how consumers tech competence might affect how they understand transparency in algorithms, and their awareness towards information privacy risks with using new age technology. Hereunder, hypotheses are presented to investigate how tech competence influences consumers understanding of transparency and their privacy awareness. In addition, it is explored if transparency is more important for high identity-relevant products. An experimental design was conducted to explore this topic and attempt to give a better understanding of transparency in algorithms. It was used three experimental conditions where consumers were exposed to different levels of transparency through an online survey. A total of 227 respondents were collected.

From the statistical analysis conducted in SPSS, results indicated that high tech competence consumers understand recommendation algorithm despite the transparency. It was also found that tech competence was positively related to consumers privacy awareness. Contradictive to the assumptions, it was found that transparency is not more important for high identity-relevant products. Furthermore, it is suggested that regulations and guidelines creating more transparent algorithms is necessary, to protect private information and improve customer experience. To the end of the thesis, it is discussion about findings, implications, limitations and suggestions for further research on the topic presented.

**Keywords**: Artificial intelligence; Transparency; Tech competence; Online decision-making; Privacy awareness; Identity-relevance

**Author contributions**: Conceptualization: L.B.; Introduction: E.H.; Literature review: E.H.; Hypotheses: E.H., L.B; Experimental design: L.B., M.P.; Survey structure: E.H., L.B.; Survey execution: E.H., L.B.; Discussion and implications: E.H. Main author E.H. with contributions from L.B.

## Norsk sammendrag

Kunstig intelligens begynner å bli et viktig verktøy i digital markedsføring brukt av markedsførere for å påvirke forbrukerne i en online beslutningsprosess.
Anbefalingsalgoritmer kan gi forbrukerne bedre produkter for å oppfylle behov og preferanser basert på deres online atferd. Imidlertid får forbrukerne sjelden informasjon eller forklaringer på hvilke data som brukes, og dette kan løses ved åpenhet i algoritmen. Denne oppgaven tar for seg hvordan forbrukernes teknologiske kompetanse kan påvirke hvordan de forstår gjennomsiktighet i algoritmer, og deres bevissthet rundt informasjonssikkerhetsrisiko ved bruk av moderne teknologi. Nedenfor presenteres hypoteser for å undersøke hvordan teknologikompetanse påvirker forbrukernes forståelse av åpenhet og deres personvernbevissthet. I tillegg undersøkes det om gjennomsiktighet er viktigere for produkter med høy relevans for identitet. Et eksperiment design ble utført for å utforske dette emnet og forsøke å gi en bedre forståelse av gjennomsiktighet i algoritmer. Det ble brukt tre eksperimentelle forhold der forbrukere ble utsatt for forskjellige nivåer av gjennomsiktighet gjennom en online undersøkelse. Totalt 227 respondenter ble samlet inn.
Fra den statistiske analysen som ble utført i SPSS, indikerte resultatene at høyteknologisk kompetente forbrukere forstår anbefalingsalgoritme til tross for gjennomsiktighet. Det ble også funnet at teknisk kompetanse var positivt relatert til forbrukernes personvern. I strid med antagelsene ble det funnet at åpenhet ikke er viktigere for produkter med høy relevans for identitet. Videre foreslås det at forskrifter og retningslinjer for å skape mer gjennomsiktige algoritmer er nødvendig for å beskytte privat informasjon og forbedre kundeopplevelsen. I slutten av oppgaven er det diskusjon om funn, implikasjoner, begrensninger og forslag til videre forskning om temaet som presenteres.

Nøkkelord: Kunstig intelligens; Åpenhet; Teknisk kompetanse; Online beslutningstaking; Bevissthet om personvern; Identitetsrelevanse

# Table of contents

## List of figures and tables

# 1. Introduction

Imagine entering a website and finding a product that catch your interest. Further down you notice recommendations made for you and you read the following words: “you might also like”. This scenario is familiar for most users of online services. Despite this, a lot of people are not aware of the influence this might have on us as consumers. Artificial intelligence (AI) and algorithms were first introduced in the 1950s, but the recent development of new age technology including social media and online shopping, has made AI an important part of digital marketing. Marketing has for centuries attempted to influence consumers in a decision-making process, and with new age technology the marketers are given new possibilities to connect with potential consumers and improve the customer experiences. Recommendation algorithms detect patterns, so the marketers can offer better services and get the ability to understand their customers’ needs and preferences in advanced ways. Also, it can help consumers navigate through choice overload and reduce search costs (Bjørlo, Moen, & Pasquine, 2021). However, algorithms use private and historical data to predict which products that will benefit the consumer, but the consumers are rarely given explanations of how their online behavior affects the personalized recommendations (Turilli & Floridi, 2009). Transparency in recommendations can provide information to the consumer about which information that have been used by the algorithm. This is one of the most important research fields today attempting to understand how the use of AI in marketing influences consumers decision-making, and that transparency in algorithms might improve consumer autonomy.

Simultaneously as AI and algorithms are being developed and improved, consumers are comprehending and using technology like never before. Based on this, consumers are becoming more tech competent which indicates that most consumers might have the potential to understand what AI is and how this works. Tech competent consumers might also be more aware of the information privacy risks with using new age technology, since algorithms use private and historical data for its recommendations. There is limited research on how consumers tech competence affects their understanding of transparency in recommendations, additionally how tech competence can affect their awareness towards information privacy risks connected to online services.

In this study we seek to advance our understanding on how consumers tech competence can affect their understanding of transparency in algorithms, and if high tech competence makes them more aware of the information privacy risks. At the end, the study will attempt to see if transparent algorithms are more important for high identity-relevant products.

### *1.2 Structure*

The study has the following structure:

**Chapter two** contains the literature review where the theoretical framework and hypotheses is presented.
**Chapter three** provides a description of the methodology, data collection and data cleaning.
**Chapter four** presents the results and analysis from the experiment.
**Chapter five** discuss the findings from the previous chapter.
**Chapter six** present implications and limitations of the study.
**Chapter seven** includes the concluding remarks.

## 2. Literature review

This chapter reviews the literature on important topics within digital marketing, artificial intelligence and consumer behavior, forming the theoretical foundation of the thesis. Firstly, a brief introduction to digital marketing and how this has enabled new marketing tools is explained. Secondly, artificial intelligence and recommendation algorithms is explained, and further how this affects consumers autonomy and privacy awareness concerning recommendation algorithms. Three hypotheses will be presented.

### *2.1 Digital marketing*

Marketing has existed for centuries and has always had the same purpose to influence people into making a decision. It is all about to persuade a consumer to take that action we want them to and choose the product we advertise (Ryan, 2016). To be able to succeed with influencing people and distribute goods and services, marketers need to be good at planning, implementation and follow-up of the activities that are put to action (Vikøren, 2020). Only then a company will successfully satisfy a customer’s need with their products or services offered.

Marketers have since the very beginning found a way of influencing people with the tools available at the time, but it has changed since the origin (Ryan, 2016). A strong tool has been the word-of-mouth, then more tools have become available when new technology has emerged, such as flyers, radio advertisement, TV-advertisements, e-mails and now: social media (Chaffey & Ellis-Chadwick, 2019). Social media and digital platforms are changing how businesses can communicate and share their message in more efficient ways. Digital marketing has created more tools for marketing including paid search placement, search optimization, pay-per-click advertisement, rich media and social media advertisement (Chaffey & Ellis-Chadwick, 2019). These keywords are all a part of an online revolution that have led to a new way for businesses to connect with new consumers, those who use and integrate technology into their everyday lives in ways that we could never have conceived a few decades ago (Ryan, 2016).

Social media and digital platforms i.e., Facebook, LinkedIn, Twitter and Instagram are some of the communication tools which effectively used to get the consumers interest (Balakrishnan, 2018). These platforms help marketers to get a foothold in the market, find new ways to become popular and take market share. Most importantly it is all about the people, which means marketers connect with customers to build trustful relationships and drive sales (Ryan, 2016). This has been the main concept of marketing of all times, but new technology is giving new possibilities for marketers to connect with potential customers worldwide. From this it is possible to define digital marketing as “achieving marketing objectives through applying digital technologies and media” (Chaffey & Ellis-Chadwick, 2019). Internet and digital platforms have transformed marketing giving access to billions of online users who regularly use online platforms and social media to find products, entertainment and friends.

For companies to succeed in the future, they will need to adapt to the technological changes and apply this in their digital marketing plans (Chaffey & Ellis-Chadwick, 2019). A firm must make an effort to acquire an understanding of their customers’ needs and behaviors across digital platforms using the technological tools available (Kumar, Ramachandran, & Kumar, 2021). A technological tool which is commonly used in digital marketing, is artificial intelligence (AI). Artificial intelligence is believed to transform business practices, increasingly changing how administrative planning processes are executed in both marketing, sales and management (Gentsch, 2018). It is developed to be an extension of human intelligence and can help humans to make better decisions. It can provide marketers with greater information about consumers and be able to provide better solutions for them (Gentsch, 2018).

AI has transformed many fields, and in marketing the interactions between firms and consumers are increasingly more individualized and generate a lot of big data (Ma & Sun, 2020). This data that consumers leave behind have driven companies to invest in machine learning that can be used to enhance the marketing capabilities. For consumers AI often reveal it selves through i.e., recommendations on e-commerce websites and content platforms such as Amazon and Netflix, deep learning engines who analyze and tag the billions of images on social media sites, automated bidding algorithms who examine a web surfer's profile in millisecond timescale to determine the optimal bid for ad delivery, and chatbots in customer service (Ma & Sun, 2020).

## *2.2 Customer experience*

To succeed, a company must also meet the demand and satisfy customers' needs and preferences, therefore create the best customer experience with focus on creating loyalty, value and a good journey (Chaffey & Ellis-Chadwick, 2019). Digital marketing is evolving to become more of a conversation, where marketers interact with the targeted segment, listen to opinions and participate. This can for example be through user-generated content where the marketer can increase the engagement with customers to increase loyalty, and further increase sales (Chaffey & Ellis-Chadwick, 2019). User-generated content is when consumers can freely create, share and exchange information and ideas in a virtual community which enable communication between consumers and firms (Chaffey & Ellis-Chadwick, 2019). The online presence of brands has increased in the past decades to improve customer relationships (Balakrishnan, 2018). Increasing marketing activities to have more effective communication with consumers, due to their possibilities to investigate products and services and share feedback with other consumers (Balakrishnan, 2018). User-generated content is a powerful tool for the consumers, being able to share honest and open thoughts about how a product, service or firm performs. This can be used to improve products, business models and values which in turn can help improve marketing strategies and strengthen customer relationships (Balakrishnan, 2018).

A customer journey can be defined as touchpoints or different types of paid, owned and earned media which influence consumers as they access different types of website and content when selecting products and services (Chaffey & Ellis-Chadwick, 2019). To create the best customer journey, the marketer needs to focus on creating customer loyalty which is the desire the customer has to continue doing business with a supplier (Chaffey & Ellis-Chadwick, 2019). One of the ultimate goals of interacting and influencing its customers through digital platforms is to create customer loyalty and satisfaction. It is two main drivers to create loyalty, whereas the first one is emotional loyalty and the other is behavioral loyalty. Emotional loyalty occurs when the loyalty to the brand is demonstrated by favorable perceptions, opinions and recommendations (Chaffey & Ellis-Chadwick, 2019). This gives companies unique insight in customer preferences.

With new-age technology, the consumer expects experiences that are effortless, intuitive, and seamless across touchpoints (Kumar et al., 2021). Therefore, it is important for firms to apply these technologies to their strategies to make an effortless and great experience for the consumer, to meet and exceed the expectations from the customer (Kumar et al., 2021). This is important if the firm wants the customer to repurchase a product at a later stage or engage in user-generated content for others to see.

### 2.2.1 Online decision-making process

The Internet of Things (IoT) has opened up a world full of products and services for online customers to choose from, which have left the online decision-making process complex (Kumar et al., 2021). An increasing number of consumers are engaged in online shopping, as well as the number of product options have increased (Karimi, Papamichail, & Holland, 2015). Customer are becoming more prone to shop online, and are also more knowledgeable and demanding since the new age technology provide them with more information (Chaffey & Ellis-Chadwick, 2019). Therefore, the online decision-making process is becoming more complex for the marketers, since brand, websites, social media and user-generated content needs to align with the customer experience (Chaffey & Ellis-Chadwick, 2019).

## *2.3 Consumer autonomy*

Consumer autonomy can be defined as the right of consumers to make their own decision (N. C. Smith, Goldstein, & Johnson, 2013). As individuals, we have a need to feel that we are making decisions that will fulfill our needs based on our own preferences, and experience that we have the full freedom to make these choices without feeling constrained or coerced (Matthew, 2006). This is among the most central values and rights consumers have in today's democratic society, given their ability to make well-informed choices (Bjørlo et al., 2021). With the rapid advancements in new age technology, a marketer’s ability to track, monitor, recommend and predict consumer choices has become better (Wertenbroch et al., 2020). In addition, the internet has reduced search and transaction costs for consumers, leaving them with the ability to obtain more choice options with the same budget as before (Wertenbroch et al., 2020). Despite this, consumers free-will require them to choose between all the options without feeling constrained or being manipulated by the firm's marketing strategies. It is important for the consumers to be able to make decisions on their own, without any external influences which is often applied without consumers’ knowledge and awareness (Wertenbroch et al., 2020). The lack of awareness and knowledge consumers have about this external influence, might impact their involvement in a purchase and their ability to make a decision based on their own preferences.

Consumers will attempt to exercise autonomy whenever they are trying to make a decision, but might have some constraints as price, time and information (Wertenbroch et al., 2020). Regardless of their ability to control the outcome, they can choose to play the game based on a mutual exchange whereby businesses and consumers trade products and services for money (Anker, 2020). This exchange is valid as long as both parties are aware and understand the exchange, however consumers can feel that there is a lack of information and not being able to make the best decisions. Being able to be in control of one's own identity, ability to act independently and to some extent be able to control its environment choosing what will fulfill one´s needs (Oyedele & Simpson, 2007). This includes being able to choose identity-relevant products to express who they are and how they want to be perceived by others (Berger & Heath, 2007).

All consumers will feel the need for autonomy and be in control of its choices, but how individuals perceive autonomy varies. The term was defined by Hertz in 1996 as a state of sensing and recognizing the ability to freely choose behaviors and courses of action on one's own needs and goals (Hertz, 1996). Perceived autonomy is a subjective experience and may be nuanced and vary in salience and intensity (Wertenbroch et al., 2020). This indicates that all consumers experience their autonomy individually in a decision-making process and emphasize different aspects and perceive autonomy individual over their choices.

## *2.4 Artificial intelligence (AI)*

As we have attempted to introduce digital marketing and the use of tools to influence consumers, it is important to understand how AI works and can be applied. There are several ways to define artificial intelligence (AI) due to its complexity, but several scientists have attempted to create a definition that encloses this complexity. As early as in 1955, McCarthy defined AI as a problem that made a machine behave in ways that would be called intelligent if a human were so behaving (McCarthy, Minsky, Rochester, & Shannon, 2006). Another definition presented by Rust (2020) of AI is "the use of computerized machinery to emulate capabilities once unique to humans" (Rust, 2020). This indicates that developers of AI are attempting to make technology "think" like human beings, but most importantly AI is a set of technologies that works together to become "intelligent" (Bjørlo et al., 2021). The goal of developing AI is to achieve a level of automation of intelligent behavior (Zhang, Lu, & Jin, 2021).

AI is continuous learning and becoming more "intelligent", being able to self-learn and improve itself by updating and adding to its knowledge base (Kumar et al., 2021). This technology is able to take complex data, analyze it and find patterns and insights which the human mind would not be able to, having the capability to think and act like humans (Kumar et al., 2021).

### 2.4.1 Big data

To understand how AI works, it is important to first define big data. Big data is larger, more complex data sets that can be used to reveal patterns, trends and other human behavior (Gentsch, 2018). It refers to datasets whose size it's beyond the ability of typical database software tools to capture, store, manage, and analyze (Gentsch, 2018). This is gathered through the internet, social media, credit card sensors, mobile phones and so on (Ma & Sun, 2020). Big data has existed for a long time, but the amount has increased immensely with more people using the internet, mobile phones and social media. For marketers, big data is used to process and get a deeper understanding of consumers, but it can also be used for deep learning to exploit the data further (Gentsch, 2018).

### 2.4.2 Algorithms

Big data does not add value alone, it is first when it is put into algorithms that the value is created. Algorithm is a process or a set of rules to be followed in calculations or other problem-solving operations, especially performed by a computer (Gentsch, 2018). With the increasing amount of big data, it is important to use algorithms to analyze the data to get value and re-create operational functions (Gentsch, 2018). A perfect algorithm has been adjusted by human engineers to the factors of importance repeatedly until a desired outcome (Kumar et al., 2021). After this point, the algorithm is capable of adjusting the factors of importance, without human interaction (Kumar et al., 2021). Through the years the algorithms have been developed to solve more complex, unknown problems and will solve it through looking for similar, already solved problems in a known database (Gentsch, 2018).

### 2.4.3 Machine learning

An important part of an AI is machine learning and is an outcome of the algorithms combined. Mitchell (1997) defined machine learning as a computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E (T. M. Mitchell, 1997). An example to better understand this is if a chess computer program improves its performance in playing chess by experience, by playing as many games as possible and analyzing them (T. M. Mitchell, 1997). It has the possibility to collect, process and analyze huge amounts of data and use this to detect patterns and as a result become better at playing chess (Ma & Sun, 2020). Machine learning is a subset of AI that trains a machine on how to learn by using datasets to develop automated, self-training models and integrate multiple methods such that the machine is able to identify patterns and hidden insights without explicit instructions (Gentsch, 2018).

Machine learning is mostly done through three different ways: supervised learning, unsupervised learning and reinforcement learning (Gentsch, 2018). Supervised learning proceeds within clearly defined limits, using labeled datasets where the right possible answers are already known (Gentsch, 2018). Unsupervised learning, on the other hand, the system is not given target values labelled in advance. It is used to identify similarities in datasets and form clusters (Gentsch, 2018). Reinforcement learning is by using dynamic programming and supervised learning to solve problems (Gentsch, 2018). The complexity behind AI and machine learning can make it difficult to understand, but the main purpose is that the sets of technologies together will solve problems and give great predictions (Zhang et al., 2021).

### 2.4.4 Application of AI and recommendation algorithms

AI and machine learning is increasingly used in communication and interaction between businesses and consumers. Chatbots and messaging systems are highly relevant and have a big focus on making communication interfaces more efficient (Gentsch, 2018). A good example of this technology can be found in Amazon's Alexa, Google Home and Apple’s Siri, but is also increasing its popularity on web pages for self-help solutions (Gentsch, 2018). The purpose of this technology is to imitate human conversation and solve problems. In the beginning bots could only answer simple, repetitive questions, but with new and advanced AI and machine learning, bots can now solve more demanding tasks (Gentsch, 2018). In addition, there is increased use of algorithm-based recommendation systems which is a powerful tool to provide more personal and relevant content. With the high number of products and services offered online, it can be difficult for a consumer to navigate online to find products that will fit the need (Wertenbroch et al., 2020). A recommendation algorithm can easily navigate through the product overload to make the decision-making process easier where consumers have little knowledge or experience with a product group (Wertenbroch et al., 2020). This is why marketers are frequently using recommendation algorithms to help consumers find the information and products that will fit their needs and add value by offering personalized content and services (Wertenbroch et al., 2020).

The recommendation algorithms are based on consumers’ past experience, behaviors, preferences and interests, and gives firms opportunities to offer additional content to better satisfy demands and provide additional buying appeals (Gentsch, 2018). The intention of a recommendation algorithm is therefore to help the consumer in the online decision-making process and enhance user experience leading to higher customer satisfaction (Gentsch, 2018). In addition, it makes it easier for marketers to meet the right customer over the right channel at the right time (Gentsch, 2018). But as a result, consumers might not be exposed to options and content that does not correspond with their preferences and interests (Wertenbroch et al., 2020). Recommendation systems are able to provide consumers with personalized services and solutions by learning from previous behavior and from there be able to predict current and future preferences (Zhang et al., 2021).

The great amount of big data and information that consumers leave behind is used by algorithms and machine learning to make personalized content for consumers. The data available are used to become better at predicting which products will fulfill consumer preferences and become better at providing high quality recommendations as more data gets available for machine learning (Gentsch, 2018). It will recommend products or services that fit similar items which the consumer has shown interest for earlier. For example, Netflix can give recommendations to a consumer based on previous watched genres, actors, historical records and so on (Zhang et al., 2021). This way the consumer gets a narrower presentation of their content to better navigate and find something to enjoy in a choice overload. If the consumer then clicks on the recommendation, the system understands that it was able to provide a good recommendation for the consumer (Zhang et al., 2021). This will in turn make the AI learn more about the consumer to provide even higher quality recommendations (Zhang et al., 2021).

## *2.5 Tech competence*

As well as consumers have a need for autonomy when choosing a product to purchase, consumers do also have different demands to online companies than previous generations (Balakrishnan, 2018). More consumers are becoming tech competent and spend more time online research products and services before making an online purchase (Chaffey & Ellis-Chadwick, 2019). In addition, tech competent consumers expect companies to deliver and engage in customer experiences. The term tech savviness can be defined as knowing a lot about modern technology, especially computers ("Tech-savvy," 2021). This can be people educated in technology or those who have acquired technology knowledge through using it, but overall, those who consider themselves to be tech competent are more confident with technological and innovative solutions. Usually, these consumers have more than just a cursory understanding of technology but comprehend technology well (Swilley, 2019). Tech competent users are more prone to trying out new technology and studies show that this competence might influence how participants interact with recommendation systems (Y. Jin, Cai, Chen, Htun, & Verbert, 2019). For companies, it has therefore become important to develop new technologies, apps and platforms to appear more appealing to tech competent consumers(Y. Jin et al., 2019). These consumers are accessible through multiple digital touch points and there is an expectation that tech savvy consumers will engage more to user-generated content on companies' digital platforms. In addition, these consumers are more likely to supplement information they receive online with other sources to achieve greater benefits from online shopping (Balakrishnan, 2018).

## *2.6 Transparency*

As explained earlier, AI and machine learning are using already existing big data to make personalized recommendations. These are often based on consumers behaviors which can be measured such as ratings, clicks, purchases, and matching this with content attributes such as popularity, price and author (Harper et al., 2015). Despite having knowledge about which data that typically are used for recommendations, consumers are rarely given explanations of how their behavior online affects recommendations. It is important for consumers to be able to see what is going on, and understand why these recommendations can fulfill their needs, which is related to the term transparency. Transparency is the possibility for consumers to access information, intentions or behavior that have been intentionally revealed through a process of disclosure (Turilli & Floridi, 2009). The term is tightly linked to "openness" which is a concept framed with positive values such as open data, open source, open code and open access (Larsson & Heintz, 2020). This indicates that consumers' behavior online should be mapped in a way for human understanding, so the consumer is able to see why recommendations are made for them. The act of making a system knowable or visible can be referred to as algorithmic transparency (Rader, Cotter, & Cho, 2018). This term can be defined as the disclosure of information about algorithms to enable monitoring, checking, criticism, or intervention by interested parties (Diakopoulos & Koliska, 2016). It is believed that transparent algorithms can improve consumers' ability to make informed choices when being exposed to a recommendation system, and that the openness will allow more people to judge if the system works or not, and if it is appropriate for them as a consumer (Rader et al., 2018).

Businesses and consumers should strive to have meaningful transparency where recommendations that the consumers receive are based on transparent data and algorithms. A system where it is possible to explain and understand how AI creates recommendations for the individual consumer. Transparency in AI can develop more trustworthy systems where the consumers feel more taken care of (Larsson & Heintz, 2020). Despite this, the reality is opposite where the algorithms are closed and often referred to as “black box”. Black boxes occur when knowledge and processes get baked into the algorithm instead of the engineers and consumers being able to see how the AI learns and improves itself (Pedreschi et al., 2019). This is automated decision making where machine learning uses big data to categorize or group consumers without humans being able to understand how and on what grounds (Pedreschi et al., 2019) Those aware of this issue can ask if they as consumers are being treated fairly and able to make their own decisions and feeling autonomy or if they are exposed to external influence (Mittelstadt, Russell, & Wachter, 2019).

As explained above, more consumers spend time online researching products and services that will fulfill their needs and have new expectations for firm’s ability to deliver satisfying customer journeys (Chaffey & Ellis-Chadwick, 2019). These consumers that see themselves as tech competent have developed skills and confidence with using new age technology. By definition tech competent consumers are those who know a lot about modern technology, which includes artificial intelligence, machine learning and algorithms (Millecamp, Htun, Jin, & Verbert, 2018). Tech competent consumers might be more aware of the digital footprint they leave behind in social media and online services. This indicates that these consumers might have more knowledge about how recommendations algorithms works and what they are built on. Therefore, high tech competent consumers understand more about the algorithms and the transparency is not as important as for those with low tech competence. On the basis of this, the first hypothesis presented is:

$H_1$: Tech competence is positively related to transparency.

## *2.7 Privacy awareness*

AI and recommendations systems are creating new opportunities and ways for companies to connect with consumers and is continuing to innovate business practices. The growth of machine learning and recommendation algorithms has made it much more important to address the problem related to privacy (Mittelstadt et al., 2019). New technology is being developed and adopted by companies, creating more value for the customer and improving customer loyalty. In the meantime, AI and the use of personal and historical data are reshaping the risk connected to consumer privacy (G. Z. Jin, 2018). In other words, the same technological developments that have created internet as a marketplace with great potential have also increased the threats towards consumer privacy (Lwin, Wirtz, & Williams, 2007).

Personal information can be referred to as any information relating to an identified or identifiable natural person. The data can be directly linked to a person, such as a name, identification number or location data, but also indirectly linked data i.e., physical, physiological, genetic, mental, economic, cultural or social identity (DPA, 2018). Protection of personal information is becoming more important, and most people have a theoretical interest in keeping their privacy online and do not want everybody to know their personal information (Pötzsch, 2009). Privacy awareness can summarize to which extent consumers is informed about privacy practices and policies and about how disclosed information is used by marketers (Xu, Dinev, Smith, & Hart, 2008). Despite this, there are concerns related to people's awareness around information privacy issues. Studies shows that people tend to act differently online than what they intended to, creating what can be called a "privacy paradox" (Pötzsch, 2009). The privacy paradox refers to how people will have negative attitudes towards providing personal information to websites will despite this share a lot of information, even though there are no apparent benefit for them (Bjørlo et al., 2021).

A good way to describe consumer privacy is seeing it as a transaction; consumers might want to hide their willingness to pay, while the firm wants to hide their real costs, but they are interdependent (G. Z. Jin, 2018). The developments of new age technology have enabled companies to collect, store, process and use data in a larger scale with less costs. In return, consumers are matched with better products for their needs and demands creating more value for them (G. Z. Jin, 2018). Despite this, the consequences of providing personal information to a company can have a lot of negative side effects. Individuals do not have all resources available and can forget things, but computers do not forget personal and historical data.

Being aware of information privacy risks can include a consumers understanding about what data is collected by whom and for what purposes, with which third parties this data is shared with, and what corresponding risks and benefits may arise (Pötzsch, 2009). The lack of transparency in recommendation algorithms can prevent the consumers to understand which data that is used in algorithms, but they should have the right to select what personal information that is to be known to what people (Pötzsch, 2009). Having greater tech competence might indicate that consumers are more aware of the information privacy risks related to using online services, because they know that their online behavior might affect the algorithms and their personalized recommendations. On the basis of this, we purpose the second hypothesis to be:

$H_2$: Tech competence is positively related to consumer privacy awareness

## *2.8 Identity relevance*

As investigated earlier in the literature review, being able to be in control of one's own identity, and act independently when making a decision, is important for individuals (Oyedele & Simpson, 2007). Moreover, people all around the world have a drive to be different and make choices that diverge from others (Berger & Heath, 2007). Consumer will choose products that will help them signal their identity and make choices which depends on the set of people that share the taste (Berger & Heath, 2007).

Identity can be defined as "who a person is, or the qualities of a person or group that make them different from others" ("Identity," 2021). Making choices that diverge from others is an effectively way to communicate their identities and ensure others see this as well. Berger and Heath attempted to explain this as a social process of communication to express who you are as a person and signal an identity to others (Berger & Heath, 2007). Through all time, people have attempted to adopt tastes that distinguish from others to express who they are. For example, kids might feel a strong urge to separate themselves from their parents, or people want to identify themselves with a certain group by adopting a distinguish taste (Berger & Heath, 2007).

Identity-relevant products tend to diverge more in certain product ranges than others. If to many people adopt a taste, it can create a negative emotional reaction and it can be a fluctuation of tastes. Studies has shown that people want to feel unique and differentiated from others that do not belong to their group but feeling overly similar to others will make people attempt to behave in ways that make them feel different (Berger & Heath, 2007). Products we purchase, attitudes we profess, and the preferences we hold can act as signal of an identity and is an effectively way to communicate to the social world who we are and how we want to be perceived. Products can be purchased for what they symbolize, not only for their function and what they do (Berger & Heath, 2007). Studies shows that more consumers prefer personalized goods over mass production, which can be linked to peoples need for signaling their identity (Sheehan & Dommer, 2020). Also, a consumer's identity might be critical and an influential factor in purchase behavior, because consumers prefer products that are consistent with their identities (Sheehan & Dommer, 2020).

For this reason, it is believed that consumers want to understand why certain products are recommended for them and want control over which products we acquire to be able to show our identity to others. It would therefore be natural that consumers would appreciate transparency in recommendation algorithms before acquiring products which can be categorized as high identity relevant. Consumers will most likely want more control over those choices that affect their identity. Based on this, the third hypothesis presented is:

> $H_3$: Transparency in recommendation algorithms is more important for high identity-relevant products

## 3. Methodology

In this chapter the process from the beginning to a final master thesis is specified and explained. The chapter will start by outlining the quantitative research method that has been utilized in this thesis and discuss the choice of experimental design. After, the variables, survey structure and respondents will be introduced, lastly the statistical approach will be discussed together with reliability and validity.

## *3.1 Research design*

The purpose of the study indicates which research approach to choose, and the aim of the experiment was to understand how the consumer autonomy is influenced by increased transparency in the recommendation algorithms. A quantitative research approach is typically used to investigate a particular topic through the measurement of variables in quantifiable terms (Mertler, 2018). This means that it relies on collecting and analyzing numerical data to describe, explain, predict or control variables and phenomena of interest (Gay, Mills, & Airasian, 2009). It seeks to describe current situations, establish relationships between variables, and attempt to explain causal relationships between variables (M. L. Mitchell & Jolley, 2012). Based on this, quantitative research follows a well-established process in terms of flexibility and no aspect of the research should emerge during the process (Mertler, 2018). On the other side, qualitative research can provide more in-depth information about the topic (Johannessen, Christoffersen, & Tufte, 2016). To better understand the phenomena studied, a mixed method approach could have been utilized, but the experiment was conducted with a sample of the population so that the quantifiable insight may be produced (Wilson, 2011).

### 3.1.1 Experimental design

One of methods in quantitative research approach is experimental design (Mertler, 2018). This is a particular type of study that allows researchers to make cause-effect statements to establish that the difference in behavior is probably not due to anything other than the manipulated variable (M. L. Mitchell & Jolley, 2012). It allows a researcher to establish different conditions and study if these conditions have an effect on the respondents (Mertler, 2018). In an experiment the cause-effect should be retrieved from the manipulated variable, but equally important is the randomization of the sample (M. L. Mitchell & Jolley, 2012). Random selection is the process of choosing random individuals for participation, such that every member of the population has an equal chance of being selected to be a member of the sample (Mertler, 2018). Randomly select individuals to participate in the study and assign, is important because it eliminate the risk of random error. Random assignment, in turn, means that every individual who has been randomly selected to participate in the experiment has an equal chance to be assigned to any of the groups (Mertler, 2018). The respondents were divided randomly into the three different conditions: low, medium and high. Random assignment and random selection are important in experimental design to avoid having other underlying variables explaining the cause-effect, as an elimination of random error. Random assignment to groups, the term means that every individual who has been randomly selected to participate in the experiment has an equal chance to be assigned to any of the groups. That the respondents were divided randomly into the three different conditions: low, medium and high. To avoid having other underlying variables explaining the cause-effect.

The first process of the master thesis was to review literature and research articles to map the knowledge of the use of AI in marketing. Research articles can provide guidance and create the theoretical framework when choosing variables to include in the experiment, but also with hypothesis testing. For this study, it was important to find research articles which mentioned consumer autonomy and transparency in recommendation algorithms, but also those describing tech competence. These were found through Google Scholar and Oria, and included keywords i.e., AI, digital marketing, consumer autonomy, tech competence, privacy awareness and identity-relevance. From this, a literature review was assembled providing a good theoretical framework for the thesis.

The research articles and the literature review provided a research base for the variables that we chose to include in our experiment. Reviewing related literature is important for the quality of the research by understanding how researchers have studied similar phenomena before (Mertler, 2018). From the research articles definitions and earlier reliable scales of how to measure the variables was found and organized in an own document. It was used scales that was previously tested in other studies to ensure that we measured the intended factors, and this would in turn increase the reliability of our experiment. For example, tech competence was defined by previous research articles and had been tested on a sample to ensure the reliability (Millecamp, Htun, Conati, & Verbert, 2019). This indicated that similar questions could be used in our experiment to attempt to correctly measure our respondent's tech competence.

## *3.2 Data collection*

### 3.2.1 Sawtooth software

When it comes to data collection, the software program Sawtooth was used to conduct the experiment. Sawtooth software enabled us to make a survey with all the variables and was used to ensure that the respondents were randomly assigned to the three conditions. A problem with online surveys is that it might gather personal data about the respondents which later on can be used to identify the respondents. In the beginning of the project, we decided to avoid gathering personal data since this would not add any value to the project. Therefore, to avoid the personal data issue, the survey was constructed in Sawtooth Software where it is possible to tick for “not gather IP-address” to keep it completely anonymous. It was also possible to program that the respondents would be randomly assigned to the three conditions, which is important to avoid random error.

### 3.2.2 Respondents

The target population was online consumers who are using online services, online shopping or social media. These were targeted because they will most likely have experience with recommendations algorithms and are most likely contributing with big data for the machine learning to use. Therefore, these will most likely be representative for the population (M. L. Mitchell & Jolley, 2012). To ensure that we would receive enough respondents, the survey sample was collected through Prolific. Prolific is a website to help researchers recruit high quality research participants to take part in the study. It is also possible to filter participants with demographic screeners, to have more relevant respondents. Therefore, we chose this website to improve the quality of our data and be sure that we fulfilled our requirements for the three conditions.

Our goal was to have 50 respondents in each condition, and since all respondents would be shown both identity-relevant and non-identity relevant products, it was a total of three conditions. On the basis of this, we required a minimum of 150 respondents randomly assigned to the three conditions but aspired to a larger sample size to reduce sampling error (M. L. Mitchell & Jolley, 2012). A larger sample will also help balance any random error (Mitchell & Jolley, 2012). After conducting the experiment, it was a total of 268 respondents, where 8 was rejected due to missing data and 33 were rejected due to failed attention check.

### 3.2.3 Pretest

A pretest was performed to classify identity-relevant and non-identity-relevant products, this was adapted from Berger and Heath (Berger & Heath, 2007). Respondents for the pretest were recruited through Reddit and personal contacts. On Reddit the pretest was shared at the pages r/Samplesize and r/SurveyExchange, to be able to reach those willing to answer surveys. A total of 26 respondents conducted the pretest, but three were removed due to not finishing the test. “Shoes” and “dress” were rated the most identity-relevant products, while “toothpaste” and “detergent” were rated non-identity-relevant products. Shoes and toothpaste were the two products that were chosen to include further in the study. The reason behind this was to make the study as gender neutral as possible to avoid having any underlying variables influence the result.

| Descriptive Statistics | | | | | |
|---|---|---|---|---|---|
| | N | Min | Max | Mean | Std. Deviation |
| Toothpaste | 28 | 1 | 5 | 2.23 | 1.36 |
| Detergent | 29 | 1 | 6 | 2.62 | 1.42 |
| Suitcase | 28 | 1 | 7 | 3.70 | 1.62 |
| Bike | 27 | 1 | 7 | 4.46 | 1.65 |
| Sofa | 28 | 1 | 7 | 4.71 | 1.40 |
| Jeans | 27 | 1 | 7 | 4.89 | 1.44 |
| Watch | 30 | 1 | 7 | 5.15 | 1.34 |
| Suits | 28 | 1 | 7 | 5.25 | 1.49 |
| Shoes | 27 | 1 | 7 | 5.52 | 1.32 |
| Dress | 28 | 1 | 7 | 5.64 | 1.41 |
| Valid N | 27 | | | | |

*Table 1: Descriptive statistics of pretest*

### 3.2.4 Pilot study

After constructing a survey in Sawtooth Software, a pilot study was shared on Prolific to a small sample group. The purpose of the pilot study was to test if the random assignment was working properly, and to detect any bugs or errors in the survey. It was also important to test if the Sawtooth software was working well when shared on Prolific. A total of ten respondents took the pilot study where one was rejected due to failing the attention check. The pilot study was also shared with friends and family to detect any misunderstandings with sentences or formulations in the survey. This enabled us to discuss our survey with others to revise and improve it further and make it more understandable before releasing the main study.

### 3.2.5 Main study

After revising and finalizing the survey, the main study was released on Prolific to a greater number of respondents. It was not used any filters, so all users of Prolific was able to enter the survey despite demographics etc. The final survey structure can be seen in appendix 1. In addition to the variables, an attention check was included to detect if the respondents had been paying attention to the survey or not. After conducting the main study, we had a total of 268 respondents where 33 were rejected due to failing the attention check and eight excluded cases due to missing data. This resulted in more respondents than we initially required, and therefore more data to use in the hypothesis testing.

## *3.3 Measure assessment and data validity*

### 3.3.1 Data cleaning

After collecting data, it is important to uncover any potential error in the data set and uncover any missing data. The process can be time consuming but is an important first step in any research. First, it was checked for errors in the data set. This was done through performing frequencies and descriptive analysis to detect any errors. The descriptive analysis is presented in chapter 3.5 in table 3. Outliers can be scores that falls outside the range of scores available (Pallant, 2016). It was not found any outliers in these analyses, since the minimum and maximum values were inside the established scales. But it was detected eight (8) missing data. This is based on the actions of the respondents, those missing are respondents who decided to not finish the survey. Missing data has a practical impact with making the sample size smaller (Hair, 2010), but since it was only eight (8) missing data, it is believed to have a rather small impact on the results. In addition, it was detected 33 cases where the respondents had failed the attention check. These cases were excluded from the analysis due to the uncertainty of the validity of these cases. Respondents who fail attention check use less time to complete experiments and can create noise in the data set, therefore eliminating these respondents might increase statistical power (Oppenheimer, Meyvis, & Davidenko, 2009).

### 3.3.2 Test of normality

Normality refers to the shape of the data distribution for an individual metric variable and its correspondence to the normal distribution (Hair, 2010). If the variation from the normal distribution is sufficiently large, all resulting statistical tests are invalid, because normality is required to use the f and t statistics (Hair, 2010). The statistical analysis performed in this thesis underline the assumptions of a normally distributed depend variable (Hair, 2010). Test of normality was conducted on all variables through nonparametric tests. In addition, Kolmogrov-Smirnov's test of normality was conducted to assess if the variables were normally distributed. A significant level of more than 0.05 indicates normality (Pallant, 2016). In the table below, the results shows that all variables used in this study has a sig. level that is less than 0.05, which indicates that the assumptions of normality are violated. Despite this, there can argued that non-normal distribution often occurs in social science due to the underlying nature of the construct being measured, not the scale (Pallant, 2016). Therefore, it was possible to continue with data analysis.

| Test of Normality | | | |
|---|---|---|---|
| | Kolmogorov-Smirnov | | |
| | Statistic | df | Sig. |
| Autonomy | 0.08 | 226 | 0.003 |
| Tech Competence | 0.22 | 226 | 0.000 |
| Privacy awareness | 0.08 | 226 | 0.002 |
| High Identity relevance | 0.14 | 226 | 0.000 |
| Low Identity relevance | 0.20 | 226 | 0.000 |
| Transparency Shoes | 0.21 | 226 | 0.000 |
| Transparency TP | 0.19 | 226 | 0.000 |
| Age | 0.49 | 226 | 0.000 |
| Gender | 0.35 | 226 | 0.000 |
| Education | 0.14 | 226 | 0.000 |

*Table 2: Test of Normality*

## 3.4 Description of variables

The study consists of a handful different variables which aim to explain how transparency in algorithms influences consumer autonomy. Despite this, only a few variables will be used to answer the three hypotheses presented in the previous chapter, because this thesis focuses on how tech competence influences the need for transparency. These variables are carefully picked as key variables to ensure that the study remains narrow and are able to answer the hypotheses (Mertler, 2018). All variables that were included in the experiment is presented in the final survey structure which can be seen appendix 1. To answer the hypotheses introduced in the previous chapter, the following variables will be processed further one in this thesis.

### 3.4.1 Transparency: experimental conditions

An experiment design has an independent variable and at least one dependent variable. The independent variable can be referred to as the manipulating variable, in our study transparency was treated as the independent variable. Cramer et al. (2008) found that explaining to consumers why recommendations were made increased acceptance of the recommendation (Cramer et al., 2008). Therefore, three experimental conditions were created with different levels of transparency and the respondents were randomly assigned to the conditions.

In low condition the respondents were given no information about what the recommendation algorithm based its recommendation on but was informed that products were presented to them. In the medium condition more information was provided for the respondents on what the algorithm based its recommendation on. This included activity on the website, i.e., purchase history, items in the shopping cart, recently viewed items and items in the wish list. The last condition with high transparency, the respondents were informed that the algorithm based its recommendation on activity on company websites as in medium, but also i.e., demographic data, geographical location, interest in social media, Google search history. How these manipulating conditions were presented to the respondents can be seen in appendix 2.

### 3.4.2 Transparency: manipulation check

After being exposed to the conditions, respondents were asked to answer the manipulation check. This was measured by three (3) items. The scales and the conditions are adapted from the study of Cramer et al. (2008), and the indicators 1 for “Strongly disagree”, 4 for “Neither agree or disagree", and 7 for “Strongly agree” (Cramer et al., 2008).

### 3.4.3 Perceived autonomy

A dependent variable can be defined as the factor that the experiment predicts is affected by the independent variable (M. L. Mitchell & Jolley, 2012). In other words, how consumer autonomy is affected by transparency. Smith, Goldstein and Johnson (2013) referred to consumer autonomy as the right of consumers to make their own decisions (N. C. Smith et al., 2013). In this experiment, how consumers perceive autonomy is used as a dependent variable and the scale was adapted from Chen et al. (Chen et al., 2015) and Michaelsen et al.

(Michaelsen, Johansson, & Hedesström, 2021) with a total of eight (8) items. This was measured by a likert scale 1-7 with the indicators 1 for "Strongly disagree", 4 for "Neither agree or disagree", and 7 for "Strongly agree".

### 3.4.4 Privacy awareness

Privacy can be referred to as the ability of the individual to personally control information about oneself (H. J. Smith, Milberg, & Burke, 1996). The purpose of this variable is to measure respondents' awareness of privacy and act as a control variable. The scale was adapted from Xu et al. (2008) to measure the overall awareness, with a total of three (3) items (Xu et al., 2008). A likert scale 1-7 was used with the indicators 1 for "Strongly disagree", 4 for "Neither agree or disagree", and 7 for "Strongly agree".

### 3.4.5 Identity-relevance

Berger and Heath (2007) proposed that consumers tend to make choices that diverge from others to ensure that they can communicate their desired identities (Berger & Heath, 2007). This was first tested in the pretest to distinguish between identity-relevant and non-identity-relevant products. The scale was adapted from Berger and Heath (2007) with a likert scale 1-7 with the indicators 1 for "Strongly disagree", 4 for "Neither agree or disagree", and 7 for "Strongly agree". Two items were asked for both two products, shoes and toothpaste, a total of four (4) items.

### 3.4.6 Tech competence

Tech competence was an important variable for Millecamp et al. when researching how music recommendations affected Spotify users (Millecamp et al., 2018). Their participants were asked to rate themselves on how confident they felt with using modern technology. This scale was therefore adapted from Millecamp et al. (2018) since it attempted to measure the competence. The respondents were asked to answer two (2) items with a likert scale 1-7 with the indicators 1 for "Not at all competent", 4 for "Neither incompetent or competent", and 7 for "Very competent".

### 3.4.7 Demographics

Demographics were included in the study as a control variable to map the characteristics of the sample group. First, the respondents were asked to indicate their age on a scale 1-5 with the indicators 1 for "Under 18", 2 for "18-34", 3 for "35-49", 4 for "50-65", and 5 for "Over 65". For gender the respondents were asked to indicate their gender choosing between 1 for "male", 2 for "female" and 3 for "other". At last, the respondents were asked to indicate their level of education. This was indicated through a scale 1-6 with 1 for "Some high school", 2 for "Completed high school", 3 for "Some college", 4 for "Completed college", 5 for "Some graduate studies", and 6 for "Completed advanced degree".

## *3.5 Descriptive statistics*

The first analysis that was conducted, was descriptive statistics. This provides more detailed information characteristics of the sample (Pallant, 2016)It provides information about total respondents and reports the central tendency i.e., mean scores, mode and median (Pallant, 2016). The total respondents after cleaning the data are 227, which consist of data from the pilot study and main study.

Table 4 shows the characteristics of the demographics of the sample, it shows the total respondents, distribution of age, gender and education level. The distribution of age extends from under 18 to over 65, which indicates that all age groups is represented, but the mean is 2.21, which means that the average respondent is between 18-34 years. For gender both male, female and those who define themselves as something else, are all represented in the sample. The distribution of education ranges from "Some high school" to "Completed advanced degree", with a mean score at 3.87. This indicates a slightly higher educated sample.

| **Descriptive Statistics** | | | | |
|---|---|---|---|---|
| | N | Min | Max | Mean |
| Age | 227 | 1 | 5 | 2.21 |
| Gender | 227 | 1 | 3 | 1.48 |
| Education | 227 | 1 | 6 | 3.87 |
| Valid N | 227 | | | |

*Table 3: Descriptive statistics of Demographics*

| Descriptive Statistics | | | | | | |
|---|---|---|---|---|---|---|
| | N | Min | Max | Mean | Skewness | Kurtosis |
| Autonomy | 227 | 2.0 | 7.0 | 5.33 | - 0.72 | 1.31 |
| Tech Competence | 227 | 2.5 | 7.0 | 5.81 | - 0.95 | 1.45 |
| Privacy awareness | 227 | 1.0 | 7.0 | 4.76 | - 0.31 | - 0.22 |
| High Identity relevance | 227 | 1.0 | 7.0 | 5.19 | - 1.07 | 1.41 |
| Low Identity relevance | 227 | 1.0 | 7.0 | 2.50 | 0.71 | - 0.20 |
| Transparency Shoes | 226 | 1.5 | 7.0 | 5.62 | - 1.21 | 1.87 |
| Transparency TP | 227 | 1.5 | 7.0 | 5.76 | - 1.07 | 2.64 |
| Valid N | 226 | | | | | |

*Table 4: Descriptive statistics of main variables*

The mean scores represent the average of the data, and is found by summating all the values and divide on the number of cases (Johannessen et al., 2016). The mode is that value, which is most frequently used in the data set, and can be used by all types of data i.e., nominal, ordinal and interval (Johannessen et al., 2016). Standard deviation an index for the extent to which individual scores differ from the mean, a measure of the degree of scatter in the scores (M. L. Mitchell & Jolley, 2012). This is also the most common used method for measuring validity, by taking square root of the summated squared deviations from the mean, divide them by the number of observations minus 1 (Hair, 2010). For the variables it shows that all of them has a high mean, except from low identity relevance which has 2.50, which is as expected.

In addition, the descriptive statistics shows the skewness and kurtosis of the data set. The skewness values measure the symmetry of a distribution, in most instances this is made to be normal distributed (Pallant, 2016). A positively skewed distribution has relatively few large values and tails off to the right, and a negatively skewed distribution has relatively few small values and tails off to the left. In addition, kurtosis offers information about the "peakedness" of the distribution, this means that it intends to measure the peakedness or flatness of a distribution when compared with a normal distribution (Pallant, 2016). A perfect normal distributed score would these two be equal to 0, but a positive value indicates a relatively peaked distribution, and a negative value indicates a relatively flat distribution (Hair, 2010). Table 5 shows that all the variables are skewed to the right, except from low identity relevance. Figure 1 shows a graphical examination of the tech competence variable and show the shape of the distribution. This provides a better understanding of the characteristics of a variable, in this case shown through a histogram. Other graphical techniques can be boxplots, scatterplots and so forth (Hair, 2010)

**TechComp_MEAN**

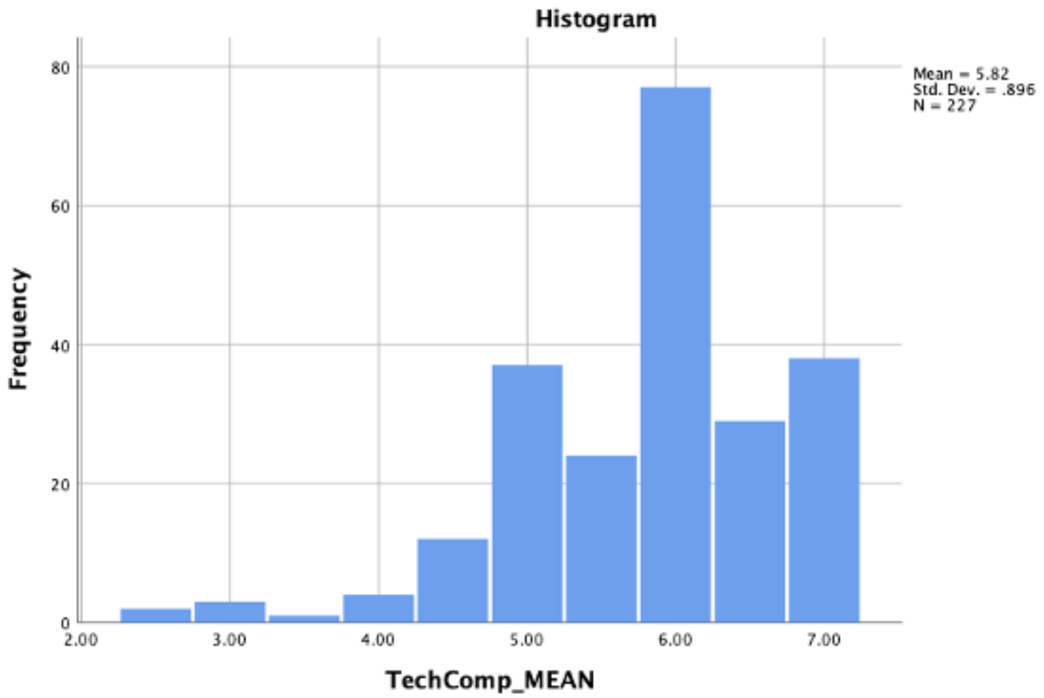



*Figure 1: Histogram of Tech competence*

## 3.6 Reliability

For a study it is important to measure the reliability which indicates how free it is from random error (M. L. Mitchell & Jolley, 2012) p.161). Reliability is important because this means that the scores are consistent and not influenced by random error (M. L. Mitchell & Jolley, 2012). To simplify this term, the interpretation of reliability is the correlation of the test with itself (Tavakol & Dennick, 2011). One way to measure this is through internal consistency. Internal consistency can be referred to as the degree of how the items in a scale "hang together", which means that the items successfully measure the same concept or construct (Tavakol & Dennick, 2011). The most commonly used indicator of internal consistency is the Cronbach's alpha coefficient (Pallant, 2016). To ensure validity to a scale, it is important that the internal consistency is determined before the study is used for research. Cronbach's alpha expresses the internal consistency with a number between 0 and 1. If the items are correlated to each other, the value of the alpha increases. A good Cronbach's alpha should have a value above 0.7, which indicates internal consistency and reliability of the scale (Pallant, 2016). For transparency the Cronbach's alpha was low, and it would be better if item 3 was deleted. Therefore, the following analysis will be using only transparency item 1 and 2.

The table below shows Cronbach's alpha for the variables used in the experiment. This indicates that the scales have good internal consistency with a reported Cronbach's alpha above 0.7 in all scales.

| **Reliability Statistics** | | |
|---|---|---|
| **Variables** | **Cronbach's alpha** | **N of items** |
| Autonomy | 0.845 | 8 |
| Transparency | 0.813 | 4 |
| Tech competence | 0.791 | 2 |
| Privacy awareness | 0.797 | 3 |
| Identity relevance | 0.704 | 4 |

*Table 5: Reliability statistics*

### *3.7 Validity*

Validity can be referred to as to which extent a measure or set of measures correctly represents the concept of study (Hair, 2010). In other words, validity is concerned with the extent to which an instrument measures what it is intended to measure (Tavakol & Dennick, 2011). It is closely associated with reliability, because a measurement cannot be valid unless it is reliable (Pallant, 2016). Validity operates with three common types which are internal, external and construct validity which is all important to ensure validity in a study (M. L. Mitchell & Jolley, 2012). Internal validity is the degree to which a study establishes that a factor causes a difference in behavior. If a study lacks internal validity, the researcher may falsely believe that a factor causes an effect when it really does not (M. L. Mitchell & Jolley, 2012).

To ensure internal validity, it is very important with independent random assignment because this is the only way to prove that the differences between the three conditions can be due to change and treatment (M. L. Mitchell & Jolley, 2012). In this study, it was important to ensure that all respondents were randomly assigned to the three conditions and avoid that any underlying variables caused the differences in the manipulated variable. This was ensured through test and retest the survey multiple times to check that the randomization worked correctly. The respondents were assigned to the three conditions randomly, which indicates a good internal validity. In addition, the respondents were not informed about the aim of the study. This to avoid the respondents to give inaccurate answer or guess what we want them to answer.

External validity can be referred to as the degree to which the results of a study can be generalized to other participants, settings, and times (M. L. Mitchell & Jolley, 2012). This means that to sustain external validity, the sample size should be of a certain size to be able to adopt the findings to a larger population. The sample size for this study ended on 268 respondents where 33 were rejected due to failed attention check and 8 due to missing data which can be determined as an appropriate sample size to say something about a larger population (M. L. Mitchell & Jolley, 2012). Furthermore, it is also important to consider the characteristics of the sample. In the descriptive statistics presented in chapter 4.1, table 4, it was found that age, gender and education level is all represented in the sample. This indicates that it might be possible to relate the findings to the population.

Construct validity can be referred to as the degree to which a study, test, or manipulation measures and/or manipulates what the researcher claims it does (M. L. Mitchell & Jolley, 2012). It is the extent to which indicators of a specific construct converge or share a high proportion of variance in common (Hair, 2010). There is not easy to measure transparency, consumer autonomy, tech competence and the other variables directly, due to how individuals perceive this differently. But all variables were measured with theoretical founded items and used already established reliable scales.

Content validity is the extent to which a measure represents a balanced and adequate sampling of relevant dimensions, knowledge, and skills (Hair, 2010). To ensure that the survey measured what we intended to, pretesting of the survey was important. The pretesting exposed several mistakes in the survey which we were able to improve before releasing the main study.

## 4. Results

In this chapter, the results from the survey will be presented through one-way ANOVA and univariate analysis. The presented results will then be used in hypothesis testing to see if the hypotheses presented in chapter 2 can be supported or not.

### *4.1 One-way ANOVA of Tech competence*

A one-way analysis of variance was conducted to explore if there was any difference in competence between the three conditions. This is applied when the researcher wants to measure the similarity or dissimilarity between the mean scores on certain variables or between groups of respondents (Hair, 2010; Pallant, 2016). The analysis of variance (ANOVA) is a statistical technique used when the researcher wants to compare the mean scores of many groups. The One-way analysis of variance will tell whether there are significant differences in the mean scores on the dependent variables across the three groups (Pallant, 2016). For tech competence this means that the one-way ANOVA is applied to see if the scores for tech competence differs between the three conditions.

**Test of homogeneity of variances** aims to test if the variance in scores is the same for each of the conditions (Pallant, 2016). A number greater than 0.05 means that the assumptions of homogeneity of variance are not violated. For tech competence the significance value is 0.949, which is greater than 0.05 and the assumptions are therefore not violated.

**ANOVA**: The ANOVA table tests if there is a significant difference somewhere among the mean scores on your dependent variable for the three groups (Pallant, 2016). A significant result is present if the value is less than or equal to 0.05, for tech competence the p value is 0.48 which is more than 0.05, and therefore not significant. This means that there is no difference in tech competence between the three conditions, which indicates that the transparency manipulation has not affected respondent's belief in their own tech competence. Despite the transparency condition that the respondents have been exposed to, respondents have approximately equal mean for the three conditions. Which means that tech competence has not influenced the outcome of the manipulation check.

| Test of Homogeneity of Variances | | | | | |
|---|---|---|---|---|---|
| | | Levene Statistics | df1 | df2 | Sig. |
| Tech Competence | Based on Mean | 0.052 | 2 | 224 | 0.949 |
| | Based on Median | 0.058 | 2 | 224 | 0.944 |
| | | | | | |

| Tech Competence | ANOVA | | | | |
|---|---|---|---|---|---|
| | Sum of squares | df | Mean square | F | Sig. |
| Between groups | 1.20 | 2 | 0.599 | 0.744 | 0.476 |
| Within groups | 180.40 | 224 | 0.805 | | |
| Total | 181.60 | 226 | | | |

*Table 6: Test of Homogeneity of Variances and ANOVA of Tech competence*

### 4.2 One-way ANOVA of Transparency

A One-way ANOVA was also conducted on transparency to explore if there was a difference between how the transparency was perceived in the three different conditions. Due to the respondents being divided into three conditions where they were exposed to different transparent recommendation algorithms, it is interesting to see if there are any significant results between the groups.

For shoes the test of homogeneity showed a score at 0.53, which is above the sig. level of 0.05. The assumptions of homogeneity are therefore not violated, which means that the variance of scores is the same for each condition. This was also the case for transparency for toothpaste where the sig. level was at 0.44. The ANOVA table, on the other hand, shows that the transparency for shoes is not significant on a level of 0.21, which means there is no difference in transparency for shoes between the three groups. This indicates that the respondents are not affected by the information they receive in the conditions but are confident that they understand on what grounds and how the recommendation algorithms are made.

On the other side, the ANOVA table for transparency for toothpaste is at sig. level equal to 0.04. This is a significant value and indicates that there are some differences between the three groups. The multiple comparisons table shows that there is a difference between low condition and high condition for low identity relevant products, indicating that the manipulation was successful in manipulating a difference when the respondent was given no information versus a lot of information.

| Test of Homogeneity of Variances | | | | | |
|---|---|---|---|---|---|
| | | Levene Statistics | df1 | df2 | Sig. |
| Transparency Shoes | Based on Mean | 0.636 | 2 | 223 | 0.530 |
| | Based on Median | 0.683 | 2 | 223 | 0.506 |
| | | | | | |

| Transparency Shoes | ANOVA | | | | |
|---|---|---|---|---|---|
| | Sum of squares | df | Mean square | F | Sig. |
| Between groups | 3.65 | 2 | 1.827 | 1.594 | 0.205 |
| Within groups | 255.62 | 223 | 1.146 | | |
| Total | 259.27 | 225 | | | |

*Table 7: Test of Homogeneity of Variances and ANOVA of Transparency shoes*

| Test of Homogeneity of Variances | | | | | |
|---|---|---|---|---|---|
| | | Levene Statistics | df1 | df2 | Sig. |
| Transparency TP | Based on Mean | 0.832 | 2 | 224 | 0.437 |
| | Based on Median | 1.251 | 2 | 224 | 0.288 |
| | | | | | |

| Transparency TP | ANOVA | | | | |
|---|---|---|---|---|---|
| | Sum of squares | df | Mean square | F | Sig. |
| Between groups | 5.687 | 2 | 2.843 | 3.242 | 0.041 |
| Within groups | 196.45 | 224 | 0.877 | | |
| Total | 202.14 | 226 | | | |

*Table 8: Test of Homogeneity of Variances and ANOVA of Transparency toothpaste*

### *4.3 One-way ANOVA of Privacy Awareness*

In addition to the other analysis, a one-way ANOVA was also conducted to explore privacy awareness and if there is any difference between respondent's privacy awareness and the three conditions. The test of homogeneity is at a sig. level of 0.67. This means that the assumptions were not violated, and we can move on to the ANOVA. The score for ANOVA between the groups is at 0.487 which is more than the sig. level of 0.05. This indicates that there is no difference between the three conditions concerned to privacy awareness, and that the experimental condition did not influence the outcome of the respondents privacy awareness.

| Test of Homogeneity of Variances | | | | | |
|---|---|---|---|---|---|
| | | Levene Statistics | df1 | df2 | Sig. |
| Privacy awareness | Based on Mean | 0.397 | 2 | 224 | 0.673 |
| | Based on Median | 0.346 | 2 | 224 | 0.708 |
| | | | | | |

| Privacy awareness | ANOVA | | | | |
|---|---|---|---|---|---|
| | Sum of squares | df | Mean square | F | Sig. |
| Between groups | 2.148 | 2 | 1.074 | 0.722 | 0.487 |
| Within groups | 333.48 | 224 | 1.489 | | |
| Total | 335.62 | 226 | | | |

*Table 9: Test of Homogeneity of Variances and ANOVA of Privacy awareness*

## 4.4 One-way ANOVA of Identity-relevance

An analysis was also conducted to see if there was any difference between the groups and the identity-relevant product the respondents were shown. For identity-relevance it is expected differences between the groups regarding which experimental condition the respondents were exposed to.

**Test of homogeneity of variances** aims to test if the variance in scores is the same for each of the conditions. A number greater than 0.05 means that assumptions of homogeneity of variance are not violated. For high identity-relevance (shoes) the significance value is 0.73 and for low identity-relevance (toothpaste) it is 0.311. Both of these are greater than 0.05, which means that the assumptions are not violated.

**ANOVA:** The ANOVA table tests if there is a significant difference somewhere among the mean scores between the three conditions. The result is significant if the value is less than or equal to 0.05 (Pallant, 2016). For high identity-relevance the p-value is 0.09 which is more than 0.05 and the result is not significant. This means that there are not differences between the three conditions and the manipulation of transparency. For low identity-relevance has a p-value on 0.54 which is more or less equal to 0.05 and can therefore be accepted as significant. This indicates that there might be a small difference between the three conditions and low identity-relevance, indicating that the experimental condition was able to provoke a difference between the conditions.

| Test of Homogeneity of Variances | | | | | |
|---|---|---|---|---|---|
| | | Levene Statistics | df1 | df2 | Sig. |
| High Identity relevance | Based on Mean | 0.309 | 2 | 224 | 0.734 |
| | Based on Median | 0.176 | 2 | 224 | 0.838 |
| | | | | | |

| High Identity relevance | ANOVA | | | | |
|---|---|---|---|---|---|
| | Sum of squares | df | Mean square | F | Sig. |
| Between groups | 7.95 | 2 | 3.973 | 2.361 | 0.097 |
| Within groups | 377.02 | 224 | 1.683 | | |
| Total | 384.97 | 226 | | | |

*Table 10: Test of Homogeneity of Variances and ANOVA of High Identity relevance*

| Test of Homogeneity of Variances | | | | | |
|---|---|---|---|---|---|
| | | Levene Statistics | df1 | df2 | Sig. |
| Low Identity relevance | Based on Mean | 1.17 | 2 | 224 | 0.311 |
| | Based on Median | 0.45 | 2 | 224 | 0.639 |
| | | | | | |

| Low Identity relevance | ANOVA | | | | |
|---|---|---|---|---|---|
| | Sum of squares | df | Mean square | F | Sig. |
| Between groups | 2.07 | 2 | 1.033 | 0.613 | 0.543 |
| Within groups | 377.43 | 224 | 1.685 | | |
| Total | 379.50 | 226 | | | |

*Table 11: Test of Homogeneity of Variances and ANOVA of Low Identity relevance*

## 4.5 One-way ANOVA of Consumer autonomy

Even though consumer autonomy, is not part of the hypotheses, it was performed an ANOVA of the variable to detect whether or not the experimental condition was able to make a difference in the perceived autonomy of the consumer.

The Test of Homogeneity of Variances shows a significance level at 0.93 which is above 0.05. This means that the assumptions of homogeneity of variance is not violated, and the ANOVA table can be assessed. In the ANOVA table the result is significant if the value is less than or equal to 0.05. Consumer autonomy has a sig. level at 0.05 and is therefore accepted as significant. This indicates that there is a small difference between the three conditions and their perceived autonomy. The biggest difference is between low condition and high condition, indicating that the experimental condition was able to make a significant difference between the perceived autonomy when exposed to the three conditions.

| Test of Homogeneity of Variances | | | | | |
|---|---|---|---|---|---|
| | | Levene Statistics | df1 | df2 | Sig. |
| Consumer autonomy | Based on Mean | 0.07 | 2 | 224 | 0.93 |
| | Based on Median | 0.06 | 2 | 224 | 0.94 |
| | | | | | |

| Consumer autonomy | ANOVA | | | | |
|---|---|---|---|---|---|
| | Sum of squares | df | Mean square | F | Sig. |
| Between groups | 4.61 | 2 | 2.304 | 2.992 | 0.052 |
| Within groups | 172.49 | 224 | 0.770 | | |
| Total | 177.10 | 226 | | | |

*Table 12: Test of Homogeneity of Variances and ANOVA of Consumer autonomy*

## *4.6 Two-way ANOVA of Tech competence and Transparency*

The first hypothesis suggests that there is an effect between consumers technology competence and their understanding of transparency. In the one-way ANOVA table, we did not find a significant result, indicating that there was no difference in the scores between the three groups exposed to different transparency conditions. The analysis was conducted to check for the possibility that tech competence might explain transparency. A univariate analysis allows us to test the effect on an independent variable on the dependent variable and can identify if there is any interaction effect (Pallant, 2016). This can possibly explain why the mean scores for transparency for shoes was not significant, but perhaps the tech competence is the underlying factor.

Descriptive giving us the mean scores for the three conditions.

| Descriptive Statistics | | | |
|---|---|---|---|
| | Low | Medium | High |
| Transparency Shoes | 5.52 | 5.74 | 5.48 |
| Transparency TP | 5.84 | 5.85 | 5.51 |

*Table 13: Descriptive Statistics for Transparency conditions*

The Levene's Test of Equality of Error Variances is a test of one of the assumptions is the underlying analysis of variance (Pallant, 2016). In this test the significant level should be greater than 0.05, because a significant result would suggest that the variance of the dependent variable is not equal. For the transparency for shoes the significant level was greater than 0.05 at 0.45, while for toothpaste was 0.49. This means that the assumptions of homogeneity of variances were not violated, and we can therefore check for an interaction effect.

To check for the interaction effect means checking how the three conditions and tech competence influences transparency. If this has a value less than or equal to 0.05, there is a significant interaction effect. But here the value is 0.77, which indicates that there is no significant difference in the effect of tech competence on transparency for high identity-relevance for the three conditions. Tech Competence is 0.05, which is a significant value. This means that there might be a difference in understanding transparency based on respondents' tech competence.

For toothpaste there was also checked for an interaction effect, but the sig. value was at 0.99 which is above 0.05. There is no sig. interaction effect between the experimental conditions and tech competence for transparency for low identity-relevance. Also, the tech competence is sig. at 0.00, which again indicates that tech competence might explain that the respondents are not influenced by the experimental condition.

| Levene's Test of Equality of Error Variances | | | |
|---|---|---|---|
| F | df1 | df2 | Sig. |
| 0.814 | 2 | 223 | 0.445 |

| Test of Between-Subjects Effects | | | | | |
|---|---|---|---|---|---|
| | df | Mean square | F | Sig. | Partial Eta Squared |
| Intercept | 1 | 111.70 | 748.74 | 0.001 | 0.997 |
| ConditionTransp | 2 | 0.14 | 0.13 | 0.88 | 0.001 |
| Transparency | 1 | 4.64 | 4.08 | 0.05 | 0.018 |
| Transparency *ConditionTransp | 2 | 0.30 | 0.26 | 0.77 | 0.002 |

*Table 14:Two-way ANOVA of Tech competence and Transparency Shoes*

| Levene's Test of Equality of Error Variances | | | |
|---|---|---|---|
| F | df1 | df2 | Sig. |
| 0.814 | 2 | 223 | 0.445 |

| Test of Between-Subjects Effects | | | | | |
|---|---|---|---|---|---|
| | df | Mean square | F | Sig. | Partial Eta Squared |
| Intercept | 1 | 88.78 | 1633.60 | 0.00 | 0.999 |
| ConditionTransp | 2 | 0.05 | 0.06 | 0.941 | 0.001 |
| Tech competence | 1 | 13.55 | 16.39 | 0.00 | 0.069 |
| Tech competence *ConditionTransp | 2 | 0.006 | 0.008 | 0.99 | 0.00 |

*Table 15: Two-way ANOVA of Tech competence and Transparency Toothpaste*

## *4.7 Two-way ANOVA for Tech competence and Privacy awareness*

The second hypothesis suggest that there is a positive relation between consumers privacy awareness and tech competence. To check if tech competence has a relation on privacy awareness, a univariate analysis was conducted. The Levene's Test of Equality gives a sig. value above 0.6 which is not significant. Therefore, we can check for an interaction effect to see how tech competence influences privacy awareness within the three conditions. There is a significant interaction effect if the sig. value is less than or equal to 0.05, but with a value of 0.74 there is no interaction effect between the three conditions and tech competence. On the other side, tech competence has a sig. level of 0.005 which is significant. This might indicate that it is a positive relation between tech competence and respondents privacy awareness.

| Levene's Test of Equality of Error Variances | | | |
|---|---|---|---|
| F | df1 | df2 | Sig. |
| 0.510 | 2 | 224 | 0.601 |

| Test of Between-Subjects Effects | | | | | |
|---|---|---|---|---|---|
| | df | Mean square | F | Sig. | Partial Eta Squared |
| Intercept | 1 | 54.78 | 88.23 | 0.01 | 0.977 |
| ConditionTransp | 2 | 0.62 | 0.42 | 0.66 | 0.004 |
| Tech competence | 1 | 11.74 | 8.08 | 0.005 | 0.035 |
| Tech competence *ConditionTransp | 2 | 0.44 | 0.30 | 0.739 | 0.003 |

*Table 16: Two-way ANOVA of Tech competence and Privacy awareness*

## 4.8 Two-way ANOVA of Identity-relevance

The third hypothesis suggest that transparency is more important for high identity-relevant products. A univariate analysis was conducted to examine if there is more important with transparency for high identity-relevant products than for low identity-relevant.

The Levene's Test of Equality for high identity-relevance is 0.71 which is not significant because it is above 0.05. Therefore, we can check for interaction effect to see if there is an effect between transparency and identity-relevance for high identity-relevance. There is a significant interaction effect if the sig. value is less than or equal to 0.05, but with a value of 0.4 there is no interaction effect between the transparency and the three experimental conditions. In addition, transparency alone has a sig. level at 0.25, which is above the sig. level of 0.05. This indicates that the transparency is not as important for high identity-relevant products such as shoes, as first anticipated.

The test was also conducted for low identity-relevance to see if it could be that transparency was more important for low identity-relevance, as contradictive to the suggested hypothesis. The Levene's Test of Equality for low identity-relevance is 0.3, which is not significant, and we can check for the interaction effect between transparency and experimental conditions, to see if they have an effect on low identity-relevance. This was at 0.93, which means that there is no interaction effect as well as the transparency alone had sig. value at 0.88. The results from this analysis shows that transparency in recommendation algorithms is not more important for high identity-relevance than for low identity-relevance.

| **Levene's Test of Equality of Error Variances** | | | |
|---|---|---|---|
| F | df1 | df2 | Sig. |
| 0.344 | 2 | 223 | 0.709 |

| **Test of Between-Subjects Effects** | | | | | |
|---|---|---|---|---|---|
| | df | Mean square | F | Sig. | Partial Eta Squared |
| Intercept | 1 | 164.2 | 128.65 | 0.006 | 0.984 |
| ConditionTransp | 2 | 1.27 | 0.75 | 0.473 | 0.007 |
| Transparency | 1 | 2.23 | 1.32 | 0.252 | 0.006 |
| Transparency *ConditionTransp | 2 | 1.58 | 0.93 | 0.395 | 0.008 |

*Table 17: Two-way ANOVA of Transparency and High Identity-relevance*

| **Levene's Test of Equality of Error Variances** | | | |
|---|---|---|---|
| F | df1 | df2 | Sig. |
| 1.207 | 2 | 224 | 0.301 |

| **Test of Between-Subjects Effects** | | | | | |
|---|---|---|---|---|---|
| | df | Mean square | F | Sig. | Partial Eta Squared |
| Intercept | 1 | 37.45 | 1047.43 | 0.000 | 0.974 |
| ConditionTransp | 2 | 0.009 | 0.006 | 0.994 | 0.00 |
| Transparency | 1 | 0.042 | 0.025 | 0.876 | 0.00 |
| Transparency *ConditionTransp | 2 | 0.073 | 0.043 | 0.958 | 0.00 |

*Table 18: Two-way ANOVA of Transparency and Low Identity-relevance*

## *4.9 Summary of hypotheses*

$H_1$: Tech competence is positively related to transparency.

As suggested in the first hypothesis, there is a significant positive association between Tech competence and Transparency. The statistical findings suggest that there are reasons to support $H_1$, and that there are indications that tech competence has an effect on consumers need for transparency.

$H_2$: Tech comp is positively related with privacy awareness.
As suggested, there are a significant positive relation between Tech competence and Privacy Awareness. The statistical findings from the analysis indicate a relation between tech competence and consumers privacy awareness.

$H_3$: Transparency is more important for identity-relevant products.
Indicates that there is no relation between transparency and identity-relevance. The statistical findings from the analysis indicate that there is no effect of transparency in recommendation algorithms on how identity-relevant the product is.

| **Hypotheses** | **Relationship between variables** | **Findings** |
|---|---|---|
| $H_1$ | Tech competence is positively related to transparency | Supported |
| $H_2$ | Tech competence is positively related to privacy awareness | Supported |
| $H_3$ | Transparency is positively related to identity-relevance | Not supported |

*Table 19: Summary of hypotheses*

# 5. Discussion

In this chapter the empirical findings from the previous chapter will be discussed and reflected on.

## *5.1 The effects of tech competence on transparency*

Prior to carrying out the experiment, we believed that higher transparency in recommendation algorithms would increase consumers autonomy when making an online purchase. It was found in the experiment that the consumer autonomy had a significant difference between the three experimental conditions, implicating that more transparent algorithms might improve autonomy. Despite this, the experimental conditions were not able to create the big differences between the groups and the manipulation check had more or less the same result despite condition. This might be possible to explain due to consumers tech competence which proved to be important for transparency.

The first hypothesis is built on the assumption that consumers various degree of tech competence would affect their meeting with recommendation algorithms. Tech competence would improve their understanding of the concept and the technology behind AI. For those educated within the field or those who know a lot about modern technology after using it for long time, it can be assumed that they have a better understanding of the underlying principles. Artificial intelligence and machine learning would therefore be familiar terms for those consumers, and it can be assumed they understand better how their personal data can be used to create patterns and provide better solutions. On the other side, those who consider themselves to have less knowledge about technology would have a poorer understanding on why the personalized recommendations were targeted directly to them and on what basis. Which means that transparency in recommendation algorithms would be appropriate for those consumers.

The results showed that tech competence might be able to explain why consumers understand recommendation algorithms and therefore transparency in the experimental conditions were not as important. Tech competence can be an important factor when considering how transparency in algorithms can influence consumer autonomy. Increased knowledge about technology can make the consumer understand more about the recommendations and be able to make better choices that will fit their needs and preferences. The findings indicated that the hypothesis was to be true, due to significant result between tech competence and transparency. Therefore, consumers understand better AI and the concept of recommendations algorithms than those who do not consider themselves as tech competent.

Despite the findings, it can be discussed whether or not the tech competent consumers do really understand AI. As mentioned in the literature review, AI is a concept which is hard to define even for them working with and developing the technology. This is due to the complexity and the rapid development of AI making it harder to understand the whole concept. Therefore, transparency in recommendation algorithms will provide consumers with more information and explanations of how their online behavior affects personalized recommendations, despite their confidence with new age technology. Openness will give them insight in the data used for the algorithm, and therefore have better control over their data. Tech competent consumers might understand AI, but most will not be aware of all the factors and historical data which can be used by AI to create the recommendations. Therefore, transparency in recommendation algorithms is important.

## *5.2 The effects of tech competence on privacy awareness*

In addition to being important for transparency, it was also argued that tech competence would have an impact on privacy awareness. Privacy awareness includes being aware and stay updated of the information privacy risks and mechanism when using a website and follow news and developments within the field. The findings showed that regardless the transparency condition, there was no difference between the three experimental groups. This indicates that most of the respondents, regardless of the manipulation they have been exposed to, keep themselves updated on information privacy risks and possible solutions to ensure that their privacy data remains private.

The hypothesis suggested that tech competence could have an effect on privacy awareness. As explained and discussed previously, tech competence means that the consumer has a lot of knowledge about modern technology and therefore might understand better the concept of AI. Modern technology is being developed in a rapid speed giving marketers and consumers new way to connect and create trustworthy relationship. The consumers need to contribute with private data so the marketers and the technology can provide them with better predictions and products. Despite this interdependence, the use of privacy data can harm the consumers having privacy data going astray.

Consumers with high technology competence will be more aware of the privacy risks and data breaches where anyone can obtain personal data which can identify people. The findings also shows that the respondents are aware of the privacy risks connected to using online services. Due to the respondents overall high scores of tech competence, the findings do not show if those with low scores in tech competence is just as aware of privacy risks. But do to reports performed by the Norwegian Data Protection Authority (DPA, 2018), there is concerns around people's privacy awareness. A lot of consumers report that they are aware of the privacy risks when using online services, but their online behavior with accepting cookies and leaving data behind in every click, indicates that most consumers provide a lot of private information regardless of their negative attitude towards it. Therefore, it is important with transparency in algorithms to make people aware of how their private data and online behavior can be used by AI.

## *5.3 The effects of transparency on identity-relevant products*

The third hypothesis suggested that transparency would be more important for high identity-relevant products such as shoes, than for low identity-relevant products such as toothpaste. The reason for this would be that which shoe a person choose to buy and wear says more about their identity than which toothpaste they decide to use. A recommendation for shoes will therefore require more personal data and information to provide shoes that will best fulfill the consumer's needs and preferences. Therefore, transparency in this recommendation algorithm would be more necessary and important for the consumer so they will be able to understand why this particular shoe would be the perfect choice.

Due to the great access to big data and personal information, recommendation algorithms are becoming better at predicting consumers needs and preferences. Machine learning can detect patterns and recognize tastes, and a positive response from a consumer will affect the algorithm to find similar products. Consumer's online behavior can therefore make the AI recognize the same behavior in a group, and therefore suggest certain products that other in the same group responded to. Because of this, an open and transparent algorithm would give consumers more information about why the historical and private data has matched them with a certain pair of shoes.

Contradictive to the expected results, transparency has no significant importance on identity-relevance. The findings show that despite the identity-relevance of a product, a transparent recommendation algorithm is not more important for the consumer even though it is expected that consumers want to control their identity. Previously research indicates that identity-relevant products is important for consumers to feel unique and be able to express themselves, and make sure other perceive their identity the same way (Berger & Heath, 2007). The products consumers acquire and wear, is an effectively way to communicate who they are and how they want to be perceived by others. Therefore, it was expected that identity-relevant products would require more transparency in the recommendation algorithms, because it is nearby people's identities and how they choose to express themselves. Despite this, the findings might indicate that as long as the consumer feel that he or she belongs to the group they identify themselves with, it is not important to understand which data the algorithm bases its recommendation on. In turn, this might indicate that AI already detected this pattern and found a shoe that will match the preferences a consumer of a certain group holds. Openness in algorithms would provide the consumers with information about why a certain shoe was recommended to them, and then perhaps become more aware of the data that is used. Therefore, even though the findings shows that transparency is not important for identity-relevant products, it is likely to believe that transparency in algorithms would give people more autonomy over the products they acquire.

# 6. Implications and limitations

In this chapter implications of the study will be presented, both theoretical, managerial and policy implications. Further, research limitations will be discussed, and in the end suggestion for further research will be presented.

## *6.1 Theoretical implications*

From our findings, it is uncovered that an important factor for the understanding of algorithms is tech competence. Those with higher technological competence understand AI better than those who are not considered tech competent and are not equally dependent on transparency in recommendation algorithms. In addition, the tech competence is also important for consumers privacy awareness. The findings indicates that those with high tech competence is more aware of information privacy risks and keep themselves more updated on how to protect their personal data.

The findings also suggest that despite the importance of identity-relevant products, transparency in algorithms is not as important as we first indicated. There was no distinguish difference for high identity-relevance and low identity-relevance and the scores for transparency. Overall, the study has revealed that consumers tech competence can be an important factor when understanding how transparency in algorithms might influence consumers autonomy.

## 6.2 Managerial implications

Prior to carrying out this study, it was believed that consumers who considered themselves as tech competent, understood more of the recommendation algorithms. With the rapid development of AI and the frequently use of this in marketing, it is important to understand how consumers is affected by it. Our findings shows that even though high transparency will give consumers more information on what the recommendation bases its output on, the less autonomy they feel when making the decision. This might emphasize that the more information the consumer receives about a personalized recommendation, the more fear consumers feel. Which in turn can be used to question if tech competence really enhances their understanding of AI and algorithms.

In addition, findings indicate that AI and algorithms might influence us in a greater way than we are aware of. Even though most consumers states that they are aware of the information privacy risks and consider themselves as tech competent, the findings indicate that consumers do not know enough about how AI is applied in digital marketing and the potential it has to influence consumers in an online decision-making process.

For businesses and marketers, AI has become a great marketing tool to reach out and connect with consumers in new and improved ways. Tech competent consumers spend more time searching for online services and products that can fulfill their needs and preferences. Based on this, these consumers also require more from the marketers and what they offer. The findings shows that tech competent consumers understand more about the recommendation algorithms, which means that the marketers need to be transparent in their marketing strategies. Marketers can benefit from using AI to detect patterns and get segment analysis faster and more in-depth than before and can therefore use this to improve customer experience and satisfaction. Therefore, it can be assumed that transparency in algorithms can be beneficial for both marketers and consumers in creating more trustworthy relationships.

## 6.3 Policy implications

In addition to have theoretically and managerial implications, a few policy implications are to be suggested. AI is solving a lot of problems, but at the same time creating others that is important to address especially considering privacy. This study has shown that most consumers are aware of the informational privacy risks when using online services, despite this, algorithms are often closed, and consumers are given little or no information about which data the algorithm uses. Findings also indicated that high tech competence also made people more aware of privacy risks, but despite this, people share a lot of private information online with or without their awareness. Based on this, there should be more guidelines and regulations for use of AI in digital marketing for both businesses and consumers to follow. This might create more trustworthy relationships between marketers and consumers, but also to protect private information and give consumers a better control over their data. Guidelines will be beneficial for both marketers and consumers, because this will help with not crossing a line where it can get awkward or where the consumers feel intruded.

## *6.4 Research limitations*

The findings have several implications; however, some limitations should be discussed and highlighted. From our findings, it is clear that the sample size had a lot of knowledge about new age technology and can therefore be considered to be a high tech competent sample, an internet panel. The sample size indicated that they already had a great understanding of how algorithms and new age technology works, and therefore it could be important to retest the survey on another sample to see if there were any differences in the findings if the sample had a lower tech competence.

Another important limitation of the study was that the manipulation of transparency was not successful. Despite the three experimental conditions respondents were assigned to, the differences between the results from the three groups was almost insignificant. This means that the manipulation check did not work as we intended to, and we are not able say that there is a cause-effect between transparency manipulation and consumer autonomy. This might be because the respondents had high tech competence and already knew a lot about recommendations, therefore retesting the survey on another sample could augment the cause-effect.

### *6.5 Further research*

The study has revealed that tech competent consumers do understand more about recommendation algorithms. Therefore, a purpose for further research could be to include more qualitative data gathered through in-depth interviews. In-depth interviews could be used to cover the phenomena more in-depth with richer information about how a consumer experiences recommendation algorithm and if they would be able to explain how it works. Consumers will most likely have both positive and negative experiences with recommendations, since their previous online behavior not always reflects who they are today. In-depth interviews can uncover how consumers actually understanding of recommendations, what the algorithm it is built on and the historical and private data it uses. This would extent the literature within AI in digital marketing further, because it is important to understand how AI influences consumers and have guidelines for marketers and businesses to follow.

## 7. Conclusions

With the rapid development of new age technology and the use of this in digital marketing, it has enabled marketers to connect with their consumers in new and improved ways. AI is used to detect patterns for better segmentation, make personalized recommendations and provides marketers with opportunities to create better customer experiences. Historical and private data is used by machine learning and algorithms to create recommendations that will better satisfy a consumer needs and preferences and help consumers in a choice overload. On the other side, it can prevent the consumer from seeing all the opportunities and products that can be found online. Despite this, algorithm's ability to detect patterns can create recommendations which enables consumers to express their identity. The increased use of AI technology can be difficult to understand for most consumers, and that is why transparency in algorithms can be beneficial for both consumers and marketers. Furthermore, the findings indicates that high tech competent consumers understand more of AI and how recommendations are made for them. Based on this, we can argue that it is important and highly relevant to keep investigating this topic since it influences consumers and might restrict their autonomy when making decisions.

Even though our findings indicates that tech competent consumers understand AI better, the transparency is therefore not as important. It is reason to believe that consumers still are not aware of all the aspects connected to AI in marketing, and transparency can make consumers more aware of how their online behavior affects recommendations. This will, in turn, make consumers more aware of privacy risks connected to online services, and guidelines and regulations can be executed to be protect individual's privacy information. Transparency in algorithms will also provide consumers with information about why particular products is recommended for them, which can help the consumer make well-informed choices when deciding which products that will enhance their identity the most.

It is well known that machine learning and algorithms uses private and historical data to create personalized recommendations and detect patterns based on our online behavior. Because of this, transparency in algorithms can help consumers understand how their private data is used and enable them to protect themselves when using online services. Even though this might indicate that about the fact that businesses and corporations know so much about us, it is important to remember that ignorance is bliss. Therefore, it is important to keep researching how AI in marketing influences us as consumers so we can take better choices for ourselves. So, the next time you read the words "you might also like", you understand why this was made for you.

## References


Anker, T. (2020). Autonomy as license to operate: Establishing the internal and external conditions of informed choice in marketing. *Marketing theory, 20*(4), 527-545. doi:10.1177/1470593120926255

Balakrishnan, M. (2018). Influence of Customer Tech Savviness and Proneness to Social Media, Facilitating Online Purchase Intention, Supplemented Through Firms Content Generation in Facebook. *Drishtikon : a management journal, 9*(1), 10.

Berger, J., & Heath, C. (2007). Where Consumers Diverge from Others: Identity Signaling and Product Domains. *The Journal of consumer research, 34*(2), 121-134. doi:10.1086/519142

Bjørlo, L., Moen, Ø., & Pasquine, M. (2021). The Role of Consumer Autonomy in Developing Sustainable AI: A Conceptual Framework. *Sustainability (Basel, Switzerland), 13*(4), 2332. doi:10.3390/su13042332

Chaffey, D., & Ellis-Chadwick, F. (2019). *Digital marketing : strategy, implementation and practice* (7th edition. ed.). Harlow, England: Pearson.

Chen, B., Vansteenkiste, M., Beyers, W., Boone, L., Deci, E. L., Van der Kaap-Deeder, J., . . . Verstuyf, J. (2015). Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motivation and Emotion, 39*(2), 216-236. doi:10.1007/s11031-014-9450-1

Cramer, H., Evers, V., Ramlal, S., Someren, M., Rutledge, L., Stash, N., . . . Wielinga, B. (2008). The Effects of Transparency on Trust and Acceptance in Interaction with a Content-Based Art Recommender. *User Model. User-Adapt. Interact., 18*, 455-496. doi:10.1007/s11257-008-9051-3

Diakopoulos, N., & Koliska, M. (2016). Algorithmic Transparency in the News Media. *Digital Journalism*, 1-20. doi:10.1080/21670811.2016.1208053

DPA. (2018). *Artificial intelligence and privacy*. Retrieved from https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf

Gay, L. R., Mills, G. E., & Airasian, P. W. (2009). *Educational research : competencies for analysis and applications* (9th ed. ed.). Upper Saddle River, N.J: Pearson Education International.

Gentsch, P. (2018). *AI in Marketing, Sales and Service: How Marketers Without a Data Science Degree Can Use AI, Big Data and Bots*. Cham: Cham: Springer International Publishing AG.

Hair, J. F. (2010). *Multivariate data analysis : a global perspective* (7th ed., Global ed. ed.). Upper Saddle River, N.J: Pearson.

Harper, F., Xu, F., Kaur, H., Condiff, K., Chang, S., & Terveen, L. (2015). Putting Users in Control of their Recommendations. In *ACM Conference On Recommender Systems* (pp. 3-10): ACM.

Hertz, J. E. (1996). Conceptualization of Perceived Enactment of Autonomy in the Elderly. *Issues in Mental Health Nursing, 17*(3), 261-273. doi:10.3109/01612849609049919

Identity. (2021). In *Cambridge Dictionary*: Cambridge Univeristy Press.

Jin, G. Z. (2018). Artificial Intelligence and Consumer Privacy. doi:10.3386/w24253

Jin, Y., Cai, W., Chen, L., Htun, N. N., & Verbert, K. (2019). *MusicBot: Evaluating Critiquing-Based Music Recommenders with Conversational Interaction*. Paper presented at the Proceedings of the 28th ACM International Conference on Information and Knowledge Management, Beijing, China. https://doi.org/10.1145/3357384.3357923

Johannessen, A., Christoffersen, L., & Tufte, P. A. (2016). *Introduksjon til samfunnsvitenskapelig metode* (5. utg. ed.). Oslo: Abstrakt.

Karimi, S., Papamichail, K. N., & Holland, C. P. (2015). The effect of prior knowledge and decision-making style on the online purchase decision-making process: A typology of consumer shopping behaviour. *Decision Support Systems, 77*, 137-147. doi:10.1016/j.dss.2015.06.004

Kumar, V., Ramachandran, D., & Kumar, B. (2021). Influence of new-age technologies on marketing: A research agenda. *Journal of business research, 125*, 864-877. doi:10.1016/j.jbusres.2020.01.007

Larsson, S., & Heintz, F. (2020). Transparency in artificial intelligence. *Internet policy review, 9*(2), 1. doi:10.14763/2020.2.1469

Lwin, M., Wirtz, J., & Williams, J. D. (2007). Consumer online privacy concerns and responses: a power–responsibility equilibrium perspective. *Journal of the Academy of Marketing Science, 35*(4), 572-585. doi:10.1007/s11747-006-0003-3

Ma, L., & Sun, B. (2020). Machine learning and AI in marketing – Connecting computing power to human insights. *International journal of research in marketing, 37*(3), 481-504. doi:10.1016/j.ijresmar.2020.04.005

Matthew, T. (2006). Human Brands: Investigating Antecedents to Consumers' Strong Attachments to Celebrities. *Journal of marketing, 70*(3), 104-119. doi:10.1509/jmkg.70.3.104

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the Dartmouth summer research project on artificial intelligence: August 31, 1955. *The AI magazine, 27*(4), 12.

Mertler, C. A. (2018). *Introduction to Educational Research* (Vol. 2): SAGE Publishing.

Michaelsen, P., Johansson, L.-O., & Hedesström, M. (2021). Experiencing default nudges: autonomy, manipulation, and choice-satisfaction as judged by people themselves. *Behavioural Public Policy*, 1-22. doi:10.1017/bpp.2021.5

Millecamp, M., Htun, N. N., Conati, C., & Verbert, K. (2019). To explain or not to explain: the effects of personal characteristics when explaining music recommendations. In: Association for Computing Machinery.

Millecamp, M., Htun, N. N., Jin, Y., & Verbert, K. (2018). Controlling Spotify Recommendations: Effects of Personal Characteristics on Music Recommender User Interfaces. In: ASSOC COMPUTING MACHINERY.

Mitchell, M. L., & Jolley, J. M. (2012). *Research design explained* (8th ed. ed.). Belmont, Calif: Wadsworth Cengage Learning.

Mitchell, T. M. (1997). *Machine learning*. New York: McGraw-Hill.

Mittelstadt, B., Russell, C., & Wachter, S. (2019). Explaining Explanations in AI. In *Conference on Fairness, Accountability, and Transparency* (pp. 279-288): ACM.

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of experimental social psychology, 45*(4), 867-872. doi:10.1016/j.jesp.2009.03.009

Oyedele, A., & Simpson, P. M. (2007). An empirical investigation of consumer control factors on intention to use selected self-service technologies. *International journal of service industry management, 18*(3), 287-306. doi:10.1108/09564230710751497

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

Pedreschi, D., Giannotti, F., Guidotti, R., Monreale, A., Ruggieri, S., & Turini, F. (2019). Meaningful Explanations of Black Box AI Decision Systems. *Proceedings of the AAAI Conference on Artificial Intelligence, 33*, 9780-9784. doi:10.1609/aaai.v33i01.33019780

Pötzsch, S. (2009). Privacy Awareness: A Means to Solve the Privacy Paradox? In (pp. 226-236). Berlin, Heidelberg: Berlin, Heidelberg: Springer Berlin Heidelberg.

Rader, E., Cotter, K., & Cho, J. (2018). Explanations as Mechanisms for Supporting Algorithmic Transparency. In *Conference on Human Factors in Computing Systems* (pp. 1-13): ACM.

Rust, R. T. (2020). The future of marketing. *International journal of research in marketing, 37*(1), 15-26. doi:10.1016/j.ijresmar.2019.08.002

Ryan, D. (2016). *Understanding Digital Marketing: Marketing Strategies for Engaging the Digital Generation*. London: London: Kogan Page, Limited.

Sheehan, D., & Dommer, S. L. (2020). Saving Your Self: How Identity Relevance Influences Product Usage. *The Journal of consumer research, 46*(6), 1076-1092. doi:10.1093/jcr/ucz037

Smith, H. J., Milberg, S. J., & Burke, S. J. (1996). Information Privacy: Measuring Individuals' Concerns about Organizational Practices. *MIS quarterly, 20*(2), 167-196. doi:10.2307/249477

Smith, N. C., Goldstein, D. G., & Johnson, E. J. (2013). Choice without Awareness: Ethical and Policy Implications of Defaults. *Journal of Public Policy & Marketing, 32*(2), 159-172. doi:10.1509/jppm.10.114

Swilley, E. (2019). *Investigating the Tech-Savvy Consumer: An Abstract*, Cham.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *Int J Med Educ, 2*, 53-55. doi:10.5116/ijme.4dfb.8dfd

Tech-savvy. (2021). In *Cambridge Dictionary*: Cambridge University Press.

Turilli, M., & Floridi, L. (2009). The ethics of information transparency. *Ethics and information technology, 11*(2), 105-112. doi:10.1007/s10676-009-9187-9

Vikøren, B. M. P., Roger. (2020, 30.11.2020). Markedsføring i Store norske leksikon. Retrieved from https://snl.no/markedsf%C3%B8ring

Wertenbroch, K., Schrift, R. Y., Alba, J. W., Barasch, A., Bhattacharjee, A., Giesler, M., . . . Zwebner, Y. (2020). Autonomy in consumer choice. *Marketing letters, 31*(4), 429-439. doi:10.1007/s11002-020-09521-z

Wilson, A. M. (2011). *Marketing research : an integrated approach* (3rd ed. ed.). Harlow: Financial Times Prentice Hall.

Xu, H., Dinev, T., Smith, H., & Hart, P. (2008). *Examining the Formation of Individual's Privacy Concerns: Toward an Integrative View*.

Zhang, Q., Lu, J., & Jin, Y. (2021). Artificial intelligence in recommender systems. *Complex & intelligent systems, 7*(1), 439-457. doi:10.1007/s40747-020-00212-w

# Appendix

## *Appendix 1: Survey structure*

### Survey – Pilot study

1. Transparency manipulation (random assignment to condition): Stimuli
2. Transparency (manipulation check): 3 items
3. Self-attribution: 2 items
4. Privacy concern: 5 items
5. Privacy awareness: 3 items + 1 item attention check
6. Perceived autonomy: 8 items
7. Purchase involvement: 2 items x2
8. Familiarity/knowledge of product (category): 2 items x2
9. Tech competence/savviness: 2 items
10. Identity-relevance: 2 items x2
11. Attitude towards recommendations: 4 items
12. Demographics (age, gender, education): 3 items

   Total: 40 items

To what extent would you agree with these statements?
Please rate yourself on the following scales (1-7)

| Variable | Measure | Reference |
|---|---|---|
| Transparency + Identity-relevance | Manipulation: 3(Low(control)/Medium/High Transparency) x 2(Identity-relevant/Non-identity relevant) | Cramer et al. 2008; Dogruel 2019 |
| Transparency | 1. I understand why the recommendation algorithm recommended the products it did<br>2. I understand what the recommendation algorithm bases its recommendations on<br>3. I would be able to explain how the recommendation algorithm works to a friend | Cramer et al. 2008 |
| Self-attribution | 1. My choice was, above all, attributable to my specific actions<br>2. My choice was due to my skills for searching and evaluating products/options | Hoffman & Post 2014; Dorn and Huberman 2005 |
| Privacy concern | 1. How concerned would you be that your personal data may be used for purposes other than the reason you provided the information for?<br>2. How concerned would you be about your online personal privacy?<br>3. How concerned would you be about the fact that sites you visited might be known/tracked?<br>4. How concerned would you be about your personal information being shared with other parties?<br>5. How concerned are you about disclosing your financial information? | Wirtz et al. 2007 |

| | | |
|---|---|---|
| Privacy awareness | 1. I am aware of the information privacy risks and mechanisms related to online shopping<br>2. I follow the news and developments about the information privacy risks related to online shopping<br>3. I keep myself updated about information privacy risks and possible solutions to ensure my information privacy<br>4. It is important to pay attention to this study. Please tick "Strongly disagree" (Attention check) | Adapted from Xu et al. 2011 |
| Perceived autonomy | 1. I feel a sense of choice and freedom in the choice I made (SC)<br>2. I feel that my decision reflected what I really want (SE)<br>3. I feel my choices expresses who I really am (SE)<br>4. I felt pressured to make the choices (SC)<br>5. I felt in control of my choices<br>6. I felt that my choices belonged to me<br>7. My choices reflected my preferences<br>12. The choices I made were free from external influence | Adapted from Chen et al. 2015; Michaelsen et al. 2017 |
| Purchase involvement | 1. Purchasing toothpaste is important to me<br>2. For me, purchasing toothpaste does not matter (R)<br>3. Purchasing shoes is important to me<br>4. For me, purchasing shoes does not matter (R) | Laurent and Kapferer 1985 |
| Familiarity/ knowledge of product (category) | 1. How experienced do you consider yourself with using shoes?<br>2. How experienced do you consider yourself with shopping for shoes?<br>3. How experienced do you consider yourself with using toothpaste?<br>4. How experienced do you consider yourself with shopping for toothpaste? | Arnthorsson et al.<br>1991 |
| Intent to use the recommendation algorithm | 1. I would rather choose products by hand than use the recommendation if I would have to perform this task again<br>2. I would like to use the recommendation again for similar tasks<br>3. The next time I am looking for a product I would like to use this recommendation | Cramer et al. 2008 |
| Likelihood to click | 1. I would likely click on one of the recommendations provided<br>2. I would likely purchase one of the recommendations provided<br>3. I would likely purchase the item I selected<br>(One for shoes and one for toothpaste) | |
| Attitude towards recommendation/ algorithm | 1. I found the recommendations helpful<br>2. I found the recommendations annoying<br>3. I found the recommendations intrusive<br>4. I found the recommendations convenient<br>(One for shoes and one for toothpaste) | |
| Tech competence | 1. How competent do you consider yourself with regards to using technology?<br>2. How competent do you think others would consider you regards to using technology? | Millecamp et al. 2018 |

| | | |
|---|---|---|
| | (Level of Competence – from Importance scale) | |
| Identity-relevance | 1. You can say a lot about someone based on their choice of toothpaste<br>2. I believe that you can express yourself by using toothpaste<br>3. You can say a lot about someone based on their choice of shoes<br>4. I believe that you can express yourself by using shoes | Berger & Heath 2007 |

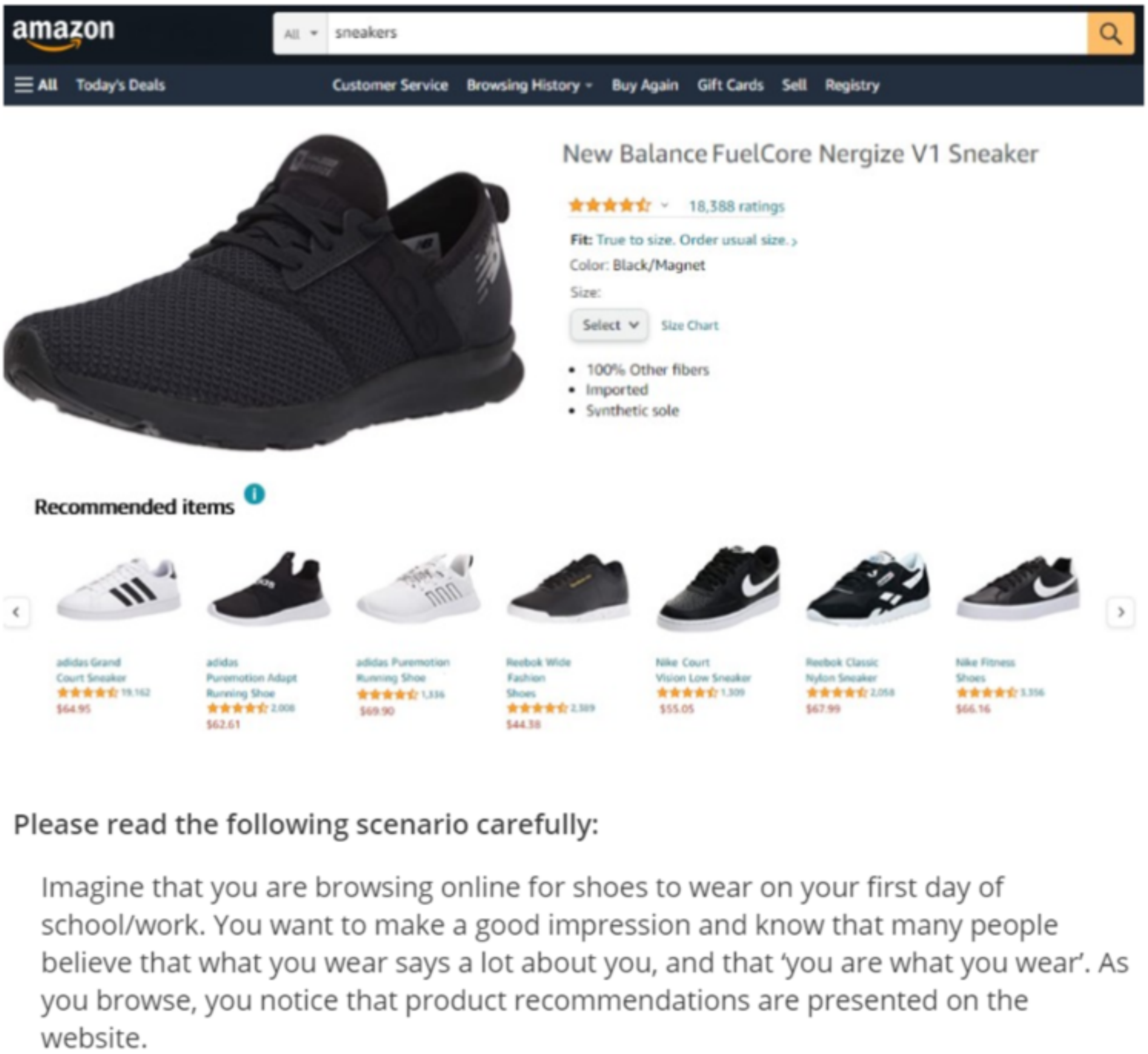
| **Variable** | **Definition** | **Reference** |
|---|---|---|
| Transparency | 1. I understand why the system recommended the products it did<br>2. I understand what the system bases its recommendations on<br>3. Would you be able to explain shortly how the system works? | Cramer et al. 2008; Sinha 2002 |
| Self-attribution | 1. My choices were, above all, attributable to my specific skills<br>2. The recent performance of my choice accurately reflects my skills for searching and evaluating products/options | |
| Privacy concern | 1. How concerned are you that your personal data may be used for purposes other than the reason you provided the information for?<br>2. How concerned are you about your online personal privacy on this web site?<br>3. How concerned are you about the fact that this web site might know/track the sites you visited?<br>4. How concerned are you about this web site sharing your personal information with other parties?<br>5. How concerned are you about receiving e-mails from this online company?<br>6. How concerned are you about disclosing your financial information to this website? | Wirtz et al. 2007 |
| Privacy awareness | 1. I am aware of the information privacy risks and mechanisms related to using this website<br>2. I follow the news and developments about the information privacy risks and preserving mechanisms.<br>3. I keep myself updated about information privacy risks and possible solutions to ensure my information privacy | Adapted from Xu et al. 2011 |
| Perceived autonomy | 1. I feel a sense of choice and freedom in the choice I made (SC)<br>2. I feel that my decision reflected what I really want (SE)<br>3. I feel my choice expresses who I really am (SE)<br>4. I feel I chose what really interests me (SE)<br>5. Choosing made me feel like ‘‘I had to’’ (SG) | Adapted from Chen et al. 2015; Michaelsen et al. 2017 |

| | | |
|---|---|---|
| | 6. I felt forced to make a choice which I normally wouldn't do (SC)<br>7. I felt pressured to make the choice (SC)<br>8. Making a choice felt like an obligation (SG)<br>9. I felt in control of my choice<br>10. I felt that my choices belonged to me<br>11. My choice reflected my preferences<br>12. The choice I made were free from external influence | |
| Purchase involvement | 1. Prescription filling/Purchasing books is important to me<br>2. For me, prescription filling/purchasing books does not matter(R) | Laurent and Kapferer 1985 |
| Familiarity/ knowledge of product (category) | 1. Familiar with clothing/batteries VCRs vs. Unfamiliar with clothing/batteries<br>2. Experienced in using clothing/batteries vs. Inexperienced in using clothing/batteries<br>3. Experienced in shopping for clothing/batteries vs. Inexperienced in shopping for clothing/batteries | Arnthorsson et al. 1991 |
| Intent to use the system | 1. I would rather choose the 6 artworks by hand from the collection of artworks than use the system if I would have to perform this task again. (inverted for analysis)<br>2. I would like to use the system again for similar tasks<br>3. The next time I am looking for a recommendation for an artwork I would like to use this system. | Cramer et al. 2008 |

*Appendix 1: Survey structure*

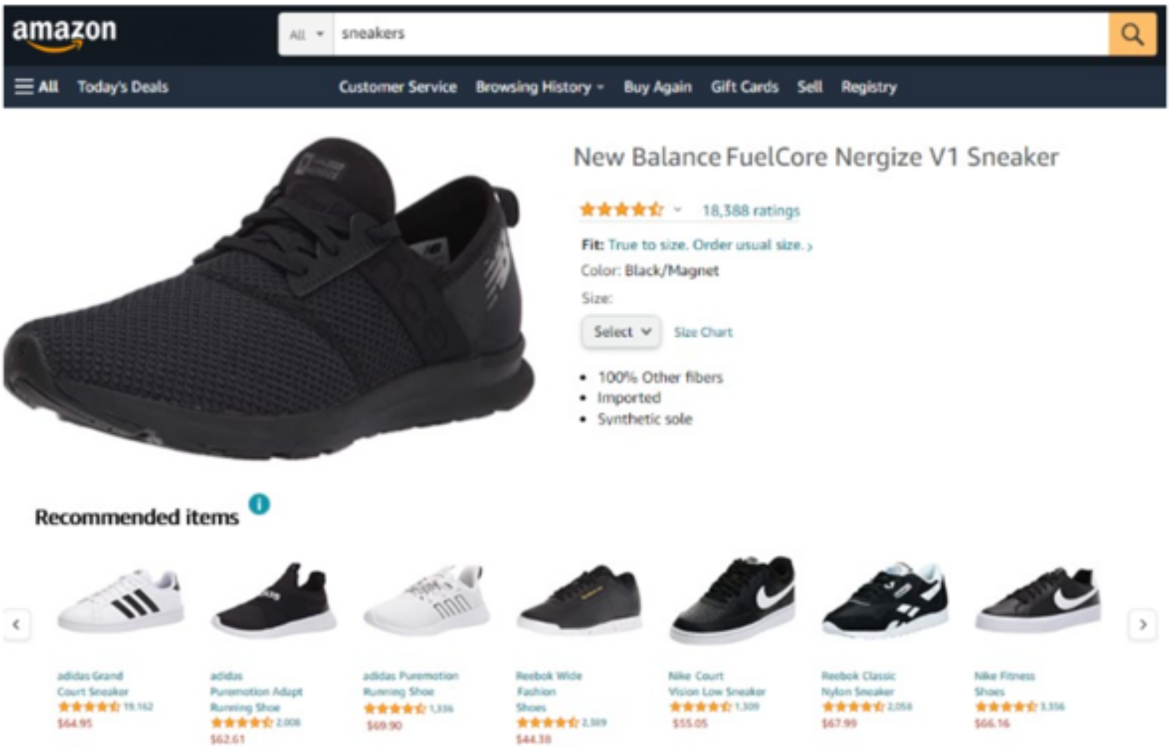
## *Appendix 2: Experimental conditions*

### *Low condition – shoes*




*Appendix 2: Shoes - low condition*

***Medium condition – shoes***




**Please read the following scenario carefully:**

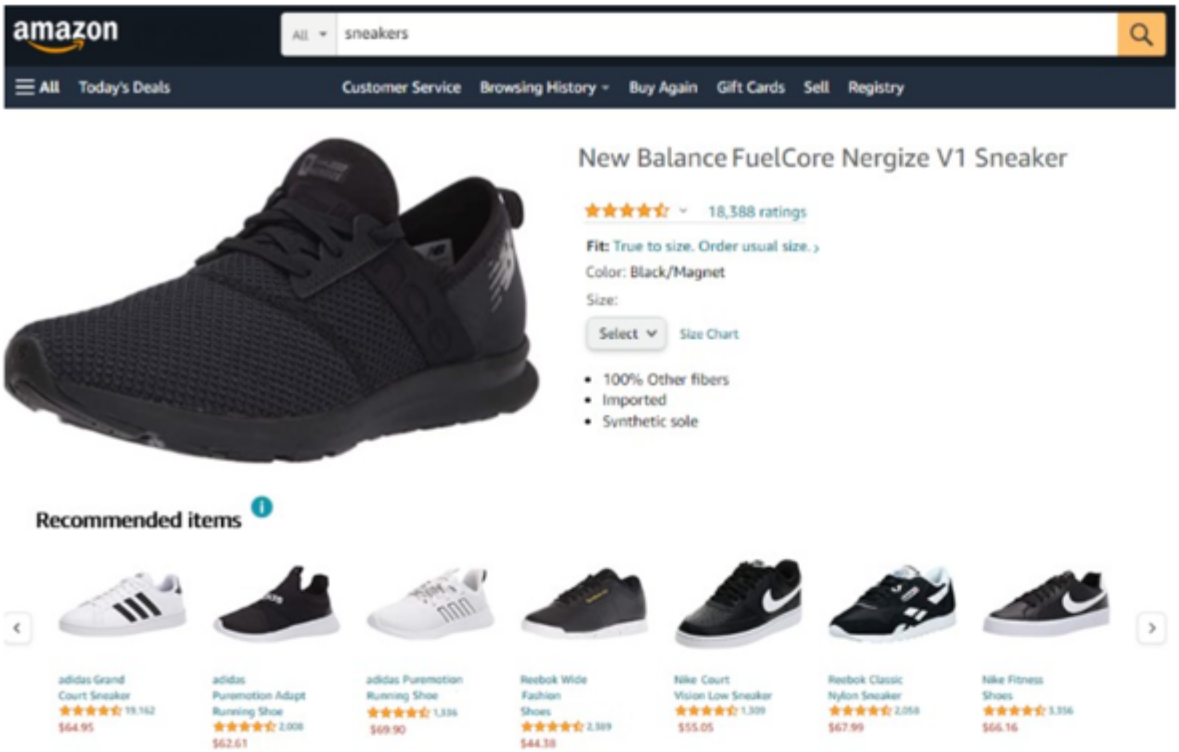
Imagine that you are browsing online for shoes to wear on your first day of school/work. You want to make a good impression and know that many people believe that what you wear say a lot about you, and that 'you are what you wear'. As you browse, you notice that product recommendations are presented on the website.

The website states that the recommendations are based on your activity on the website and interactions with the company, i.e.:

- Your purchase history (Items you have previously purchased)
- Items in your shopping cart
- Items your recently viewed
- Items in your wish list

*Appendix 3: Shoes - medium condition*

***High condition – Shoes***




**Please read the following scenario carefully:**

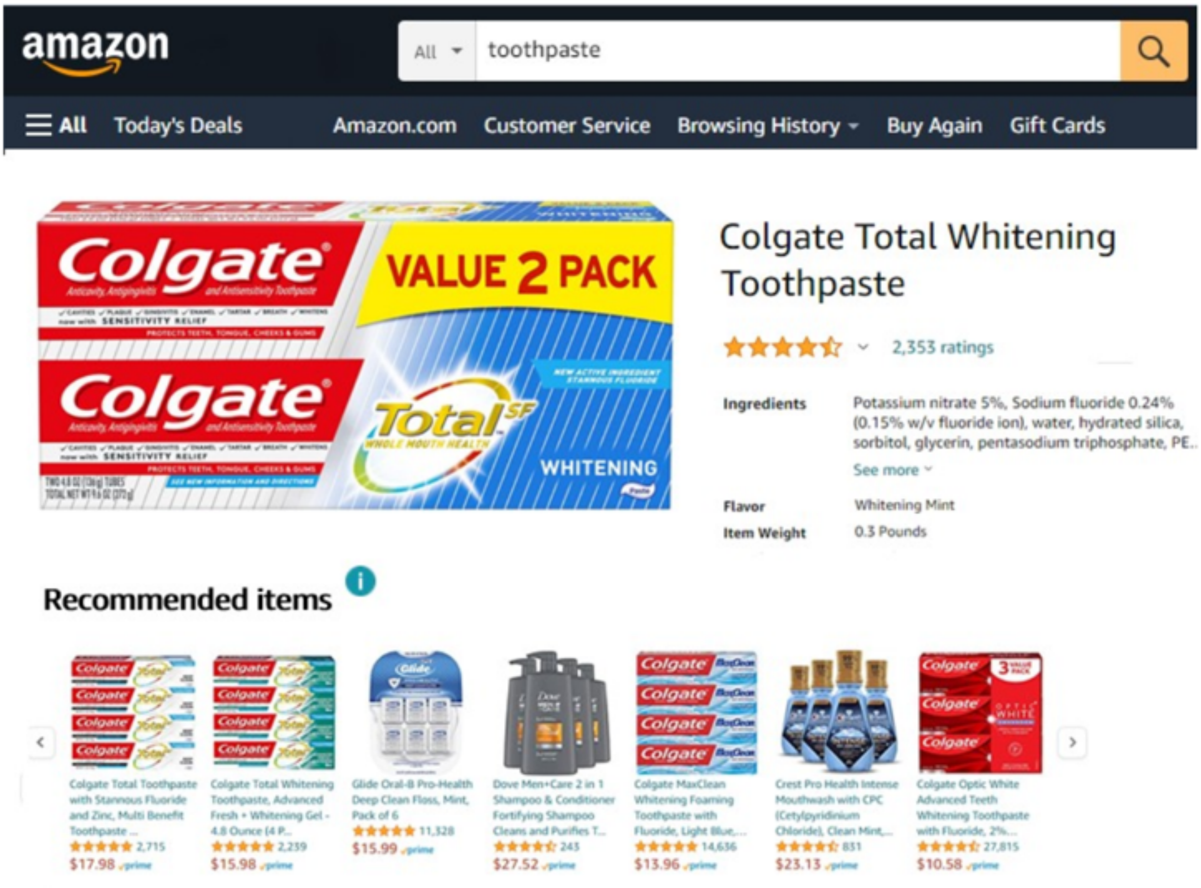
Imagine that you are browsing online for shoes to wear on your first day of school/work. You want to make a good impression and know that many people believe that what you wear say a lot about you, and that 'you are what you wear'. As you browse, you notice that product recommendations are presented on the website.

The website states that the recommendations are based on an algorithm which uses your personal data that is available online. Based on the data, the algorithm performs advanced calculations, makes inferences about you, and predicts the items which you are most likely to be interested to click on. The type of data that the algorithm uses as input could include i.e.:

- Your demographic data (i.e., age, gender, marital status, education level)
- Your geographical location
- Your interests on social media (i.e., Facebook "likes")
- Information you have entered on other websites or apps
- Your Google search history
- Contents from your email
- Activity on company website (purchase history, items in shopping cart or wish list, recently viewed items)

*Appendix 4: Shoes - high condition*

***Low condition – TP***

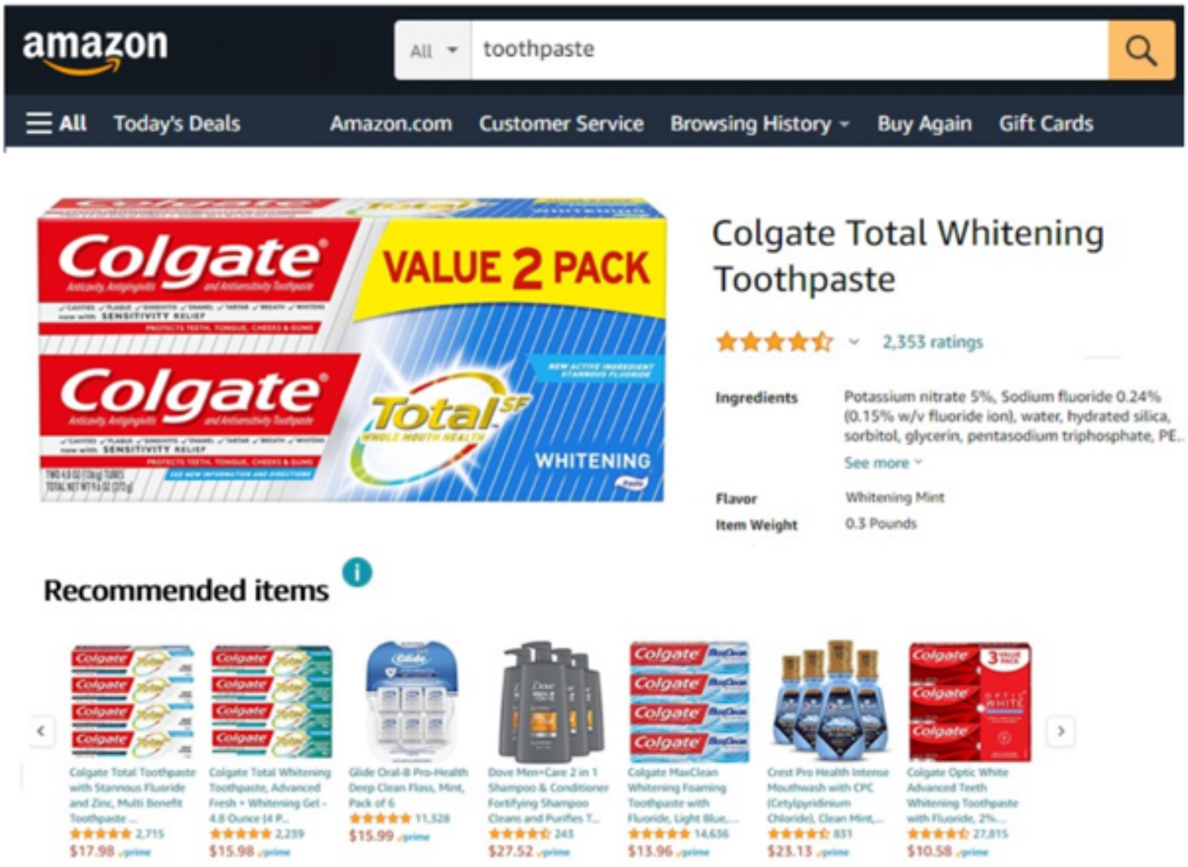



Please read the following scenario carefully:

Imagine that you are browsing online for toothpaste. As you browse, you notice that product recommendations are presented on the website.

*Appendix 5: TP - low condition*

***Medium condition - TP***



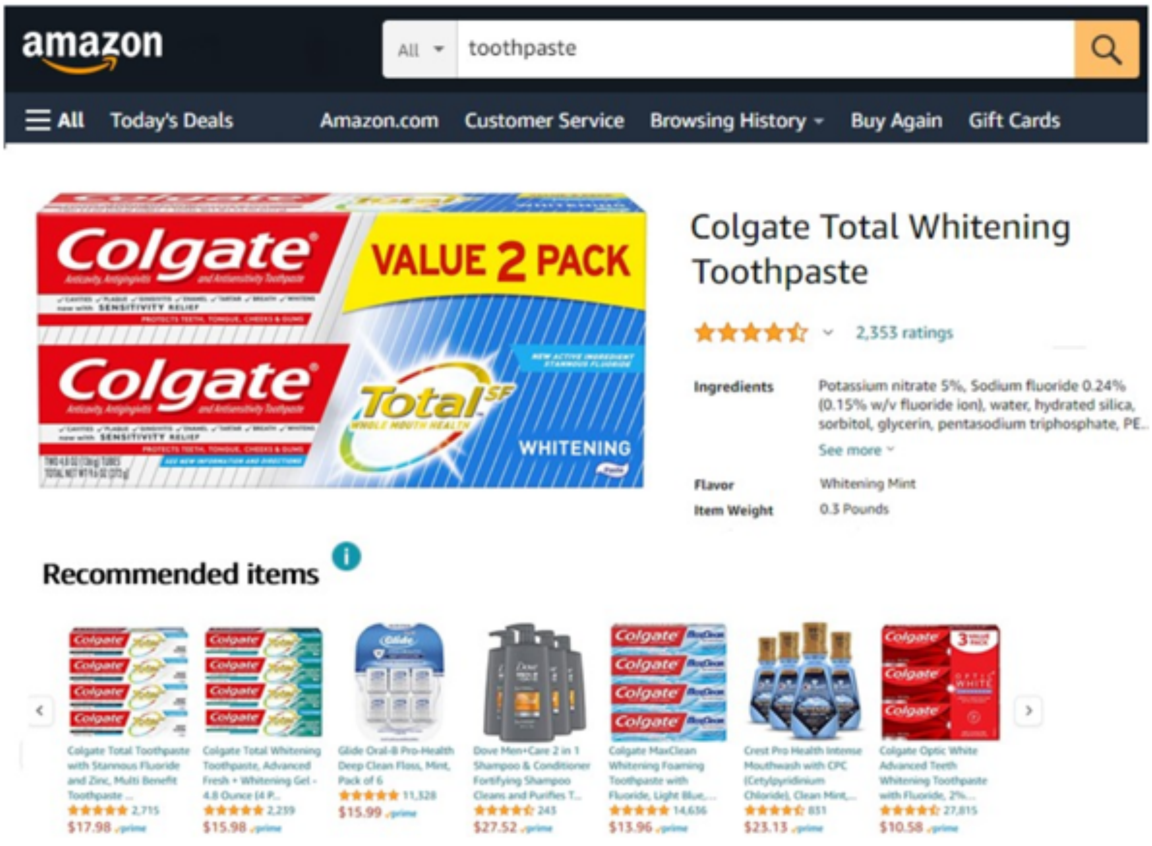

Please read the following scenario carefully:

Imagine that you are browsing online for toothpaste. As you browse, you notice that product recommendations are presented on the website. The website states that the recommendations are based on your activity on the website and interactions with the company, i.e.:

- Your purchase history (Items you have previously purchased)
- Items in your shopping cart
- Items your recently viewed
- Items in your wish list

*Appendix 6: TP - medium condition*

***High condition – TP***




Please read the following scenario carefully:

Imagine that you are browsing online for toothpaste. As you browse, you notice that product recommendations are presented on the website. The website states that the recommendations are based on an algorithm which uses your personal data that is available online.

Based on the data, the algorithm performs advanced calculations, makes inferences about you, and predicts the items which you are most likely to be interested to click on. The type of data that the algorithm uses as input could include i.e.:

- Your demographic data (i.e., age, gender, marital status, education level)
- Your geographical location
- Your interests on social media (i.e., Facebook "likes")
- Information you have entered on other websites or apps
- Your Google search history
- Contents from your email
- Activity on company website (purchase history, items in shopping cart or wish list, recently viewed items)

*Appendix 7: TP - high condition*

# *Appendix 3: SPSS output*

## SPSS-output

### Pretest

**Descriptive Statistics**

| | N | Minimum | Maximum | Mean | Std. Deviation |
|---|---|---|---|---|---|
| Toothpaste | 28 | 1,00 | 5,00 | 2,2321 | 1,36410 |
| Detergent | 29 | 1,00 | 6,00 | 2,6207 | 1,42463 |
| Suitcase | 28 | 1,00 | 7,00 | 3,6964 | 1,62355 |
| Bike | 27 | 1,00 | 7,00 | 4,4630 | 1,65207 |
| Sofa | 28 | 1,00 | 7,00 | 4,7143 | 1,39728 |
| Jeans | 27 | 1,00 | 7,00 | 4,8889 | 1,43670 |
| Watch | 30 | 1,00 | 7,00 | 5,1500 | 1,34003 |
| Suits | 28 | 1,00 | 7,00 | 5,2500 | 1,49381 |
| Shoes | 27 | 1,00 | 7,00 | 5,5185 | 1,31910 |
| Dress | 28 | 1,00 | 7,00 | 5,6429 | 1,40671 |
| Valid N (listwise) | 27 | | | | |

*Appendix 8: Pretest*

### Descriptive statistics

**Descriptive Statistics**

| | N | Minimum | Maximum | Mean | Std. Deviation | Skewness | | Kurtosis | |
|---|---|---|---|---|---|---|---|---|---|
| | Statistic | Statistic | Statistic | Statistic | Statistic | Statistic | Std. Error | Statistic | Std. Error |
| Age | 227 | 1 | 5 | 2.21 | .540 | 2.500 | .162 | 6.346 | .322 |
| Gender | 227 | 1 | 3 | 1.48 | .518 | .290 | .162 | -1.471 | .322 |
| Education | 227 | 1 | 6 | 3.87 | 1.527 | -.098 | .162 | -1.151 | .322 |
| Valid N (listwise) | 227 | | | | | | | | |

*Appendix 9: Descriptive statistics demographics*

**Descriptive Statistics**

| | N | Minimum | Maximum | Mean | Std. Deviation | Skewness | | Kurtosis | |
|---|---|---|---|---|---|---|---|---|---|
| | Statistic | Statistic | Statistic | Statistic | Statistic | Statistic | Std. Error | Statistic | Std. Error |
| Autonomy_MEAN | 227 | 2.00 | 7.00 | 5.3254 | .88522 | -.721 | .162 | 1.307 | .322 |
| TechComp_MEAN | 227 | 2.50 | 7.00 | 5.8194 | .89639 | -.951 | .162 | 1.454 | .322 |
| PrivAware_MEAN | 227 | 1.00 | 7.00 | 4.7562 | 1.21863 | -.314 | .162 | -.216 | .322 |
| Identrel_MEAN_Shoes | 227 | 1 | 7 | 5.19 | 1.305 | -1.065 | .162 | 1.411 | .322 |
| Identrel_MEAN_TP | 227 | 1 | 7 | 2.50 | 1.296 | .708 | .162 | -.199 | .322 |
| Transparency item 1 and 2 - Shoes | 226 | 1.50 | 7.00 | 5.6195 | 1.07347 | -1.208 | .162 | 1.871 | .322 |
| Transparency item 1 and 2 - TP | 227 | 1.00 | 7.00 | 5.7643 | .94574 | -1.074 | .162 | 2.638 | .322 |
| Valid N (listwise) | 226 | | | | | | | | |

*Appendix 10: Descriptive statistics variables*

**Explore**

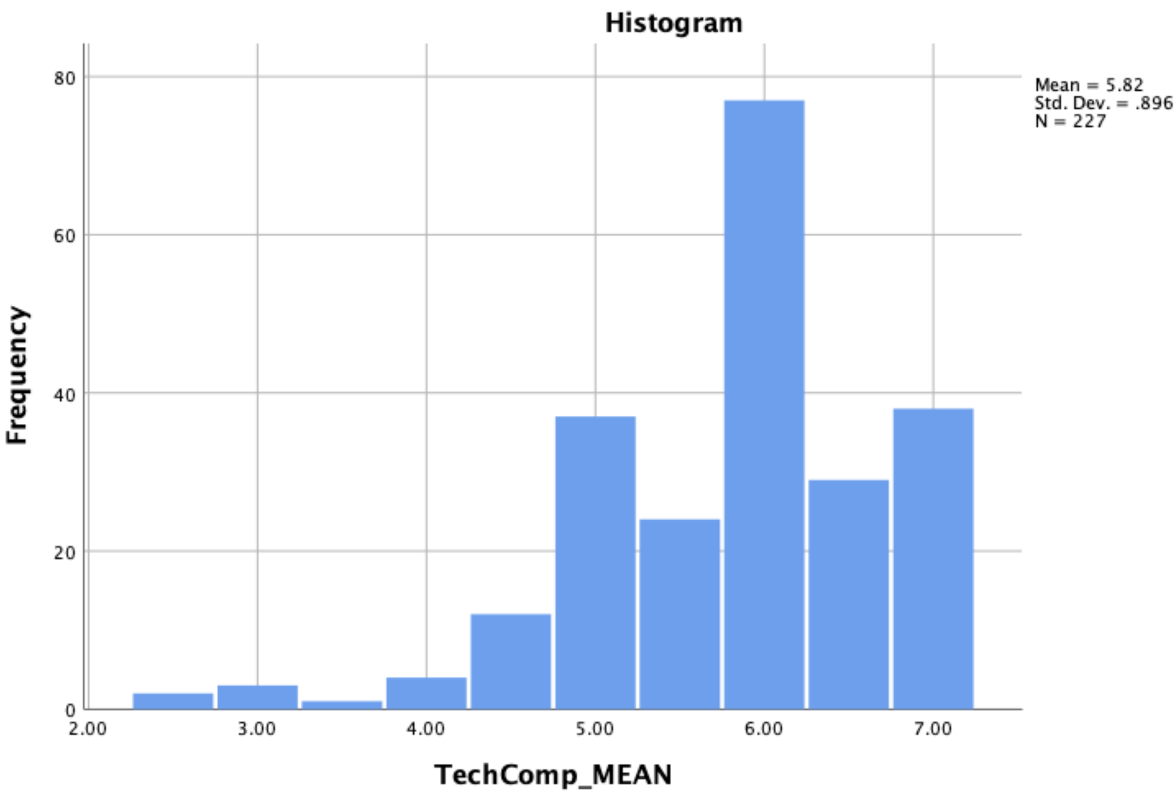
**Tests of Normality**

| | Kolmogorov-Smirnov[a] | | | Shapiro-Wilk | | |
|---|---|---|---|---|---|---|
| | Statistic | df | Sig. | Statistic | df | Sig. |
| Transparency item 1 and 2 - Shoes | .214 | 226 | .000 | .884 | 226 | .000 |
| Transparency item 1 and 2 - TP | .194 | 226 | .000 | .895 | 226 | .000 |
| TechComp_MEAN | .215 | 226 | .000 | .902 | 226 | .000 |
| PrivAware_MEAN | .078 | 226 | .002 | .980 | 226 | .003 |
| Autonomy_MEAN | .076 | 226 | .003 | .966 | 226 | .000 |
| Age | .489 | 226 | .000 | .481 | 226 | .000 |
| Gender | .353 | 226 | .000 | .666 | 226 | .000 |
| Education | .158 | 226 | .000 | .913 | 226 | .000 |
| Identrel_MEAN_Shoes | .135 | 226 | .000 | .914 | 226 | .000 |
| Identrel_MEAN_TP | .198 | 226 | .000 | .910 | 226 | .000 |

a. Lilliefors Significance Correction

*Appendix 11: Test of Normality*

**TechComp_MEAN**




**One-way ANOVA**

## Descriptives

TechComp_MEAN

| | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean Lower Bound | 95% Confidence Interval for Mean Upper Bound | Minimum | Maximum |
|---|---|---|---|---|---|---|---|---|
| .00 | 76 | 5.8553 | .93743 | .10753 | 5.6411 | 6.0695 | 2.50 | 7.00 |
| 1.00 | 77 | 5.8831 | .92087 | .10494 | 5.6741 | 6.0921 | 3.00 | 7.00 |
| 2.00 | 74 | 5.7162 | .82794 | .09625 | 5.5244 | 5.9080 | 3.50 | 7.00 |
| Total | 227 | 5.8194 | .89639 | .05950 | 5.7021 | 5.9366 | 2.50 | 7.00 |

## Test of Homogeneity of Variances

| | | Levene Statistic | df1 | df2 | Sig. |
|---|---|---|---|---|---|
| TechComp_MEAN | Based on Mean | .052 | 2 | 224 | .949 |
| | Based on Median | .058 | 2 | 224 | .944 |
| | Based on Median and with adjusted df | .058 | 2 | 219.854 | .944 |
| | Based on trimmed mean | .118 | 2 | 224 | .889 |

## ANOVA

TechComp_MEAN

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 1.198 | 2 | .599 | .744 | .476 |
| Within Groups | 180.396 | 224 | .805 | | |
| Total | 181.595 | 226 | | | |

## Robust Tests of Equality of Means

TechComp_MEAN

| | Statistic[a] | df1 | df2 | Sig. |
|---|---|---|---|---|
| Welch | .802 | 2 | 149.145 | .450 |
| Brown-Forsythe | .746 | 2 | 222.380 | .476 |

a. Asymptotically F distributed.

# Post Hoc Tests

## Multiple Comparisons

Dependent Variable: TechComp_MEAN

Tukey HSD

| (I) ConditionTransp | (J) ConditionTransp | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval Lower Bound | 95% Confidence Interval Upper Bound |
|---|---|---|---|---|---|---|
| .00 | 1.00 | -.02785 | .14511 | .980 | -.3702 | .3145 |
| | 2.00 | .13905 | .14656 | .610 | -.2067 | .4848 |
| 1.00 | .00 | .02785 | .14511 | .980 | -.3145 | .3702 |
| | 2.00 | .16690 | .14609 | .489 | -.1778 | .5116 |
| 2.00 | .00 | -.13905 | .14656 | .610 | -.4848 | .2067 |
| | 1.00 | -.16690 | .14609 | .489 | -.5116 | .1778 |

*Appendix 12: ANOVA tech competence*

**Descriptives**

Transparency item 1 and 2 - Shoes

| | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean | | Minimum | Maximum |
|---|---|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound | | |
| .00 | 75 | 5.5867 | 1.15482 | .13335 | 5.3210 | 5.8524 | 1.50 | 7.00 |
| 1.00 | 77 | 5.7857 | .95431 | .10875 | 5.5691 | 6.0023 | 3.00 | 7.00 |
| 2.00 | 74 | 5.4797 | 1.09619 | .12743 | 5.2258 | 5.7337 | 2.00 | 7.00 |
| Total | 226 | 5.6195 | 1.07347 | .07141 | 5.4788 | 5.7602 | 1.50 | 7.00 |

**Test of Homogeneity of Variances**

| | | Levene Statistic | df1 | df2 | Sig. |
|---|---|---|---|---|---|
| Transparency item 1 and 2 - Shoes | Based on Mean | .636 | 2 | 223 | .530 |
| | Based on Median | .683 | 2 | 223 | .506 |
| | Based on Median and with adjusted df | .683 | 2 | 212.345 | .506 |
| | Based on trimmed mean | .674 | 2 | 223 | .511 |

**ANOVA**

Transparency item 1 and 2 - Shoes

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 3.654 | 2 | 1.827 | 1.594 | .205 |
| Within Groups | 255.621 | 223 | 1.146 | | |
| Total | 259.274 | 225 | | | |

**Robust Tests of Equality of Means**

Transparency item 1 and 2 - Shoes

| | Statistic[a] | df1 | df2 | Sig. |
|---|---|---|---|---|
| Welch | 1.760 | 2 | 147.034 | .176 |
| Brown-Forsythe | 1.589 | 2 | 216.559 | .207 |

a. Asymptotically F distributed.

## Post Hoc Tests

**Multiple Comparisons**

Dependent Variable: Transparency item 1 and 2 - Shoes

Tukey HSD

| (I) ConditionTransp | (J) ConditionTransp | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| .00 | 1.00 | -.19905 | .17370 | .487 | -.6089 | .2108 |
| | 2.00 | .10694 | .17543 | .815 | -.3070 | .5208 |
| 1.00 | .00 | .19905 | .17370 | .487 | -.2108 | .6089 |
| | 2.00 | .30598 | .17429 | .187 | -.1052 | .7172 |
| 2.00 | .00 | -.10694 | .17543 | .815 | -.5208 | .3070 |
| | 1.00 | -.30598 | .17429 | .187 | -.7172 | .1052 |

*Appendix 13: ANOVA transparency shoes*

### Descriptives

Transparency item 1 and 2 - TP

| | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean | | Minimum | Maximum |
|---|---|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound | | |
| .00 | 76 | 5.9079 | .86298 | .09899 | 5.7107 | 6.1051 | 3.00 | 7.00 |
| 1.00 | 77 | 5.8377 | .91206 | .10394 | 5.6307 | 6.0447 | 3.00 | 7.00 |
| 2.00 | 74 | 5.5405 | 1.02955 | .11968 | 5.3020 | 5.7791 | 1.00 | 7.00 |
| Total | 227 | 5.7643 | .94574 | .06277 | 5.6406 | 5.8880 | 1.00 | 7.00 |

### Test of Homogeneity of Variances

| | | Levene Statistic | df1 | df2 | Sig. |
|---|---|---|---|---|---|
| Transparency item 1 and 2 - TP | Based on Mean | .832 | 2 | 224 | .437 |
| | Based on Median | 1.251 | 2 | 224 | .288 |
| | Based on Median and with adjusted df | 1.251 | 2 | 221.681 | .288 |
| | Based on trimmed mean | 1.107 | 2 | 224 | .332 |

### ANOVA

Transparency item 1 and 2 - TP

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 5.687 | 2 | 2.843 | 3.242 | .041 |
| Within Groups | 196.454 | 224 | .877 | | |
| Total | 202.141 | 226 | | | |

### Robust Tests of Equality of Means

Transparency item 1 and 2 - TP

| | Statistic[a] | df1 | df2 | Sig. |
|---|---|---|---|---|
| Welch | 2.971 | 2 | 148.097 | .054 |
| Brown-Forsythe | 3.232 | 2 | 217.409 | .041 |

a. Asymptotically F distributed.

## Post Hoc Tests

### Multiple Comparisons

Dependent Variable: Transparency item 1 and 2 - TP

Tukey HSD

| (I) ConditionTransp | (J) ConditionTransp | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| .00 | 1.00 | .07023 | .15143 | .888 | -.2870 | .4275 |
| | 2.00 | .36735* | .15294 | .045 | .0065 | .7282 |
| 1.00 | .00 | -.07023 | .15143 | .888 | -.4275 | .2870 |
| | 2.00 | .29712 | .15245 | .128 | -.0626 | .6568 |
| 2.00 | .00 | -.36735* | .15294 | .045 | -.7282 | -.0065 |
| | 1.00 | -.29712 | .15245 | .128 | -.6568 | .0626 |

*. The mean difference is significant at the 0.05 level.

*Appendix 14: ANOVA transparency TP*

### Descriptives

PrivAware_MEAN

| | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean | | Minimum | Maximum |
|---|---|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound | | |
| .00 | 76 | 4.7193 | 1.31887 | .15128 | 4.4179 | 5.0207 | 1.00 | 7.00 |
| 1.00 | 77 | 4.8874 | 1.15800 | .13197 | 4.6246 | 5.1503 | 2.33 | 7.00 |
| 2.00 | 74 | 4.6577 | 1.17686 | .13681 | 4.3850 | 4.9303 | 1.33 | 7.00 |
| Total | 227 | 4.7562 | 1.21863 | .08088 | 4.5969 | 4.9156 | 1.00 | 7.00 |

### Test of Homogeneity of Variances

| | | Levene Statistic | df1 | df2 | Sig. |
|---|---|---|---|---|---|
| PrivAware_MEAN | Based on Mean | .397 | 2 | 224 | .673 |
| | Based on Median | .346 | 2 | 224 | .708 |
| | Based on Median and with adjusted df | .346 | 2 | 214.229 | .708 |
| | Based on trimmed mean | .424 | 2 | 224 | .655 |

### ANOVA

PrivAware_MEAN

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 2.148 | 2 | 1.074 | .722 | .487 |
| Within Groups | 333.475 | 224 | 1.489 | | |
| Total | 335.623 | 226 | | | |

### Robust Tests of Equality of Means

PrivAware_MEAN

| | Statistic[a] | df1 | df2 | Sig. |
|---|---|---|---|---|
| Welch | .779 | 2 | 148.805 | .461 |
| Brown-Forsythe | .722 | 2 | 220.856 | .487 |

a. Asymptotically F distributed.

## Post Hoc Tests

### Multiple Comparisons

Dependent Variable: PrivAware_MEAN
Tukey HSD

| (I) ConditionTransp | (J) ConditionTransp | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| .00 | 1.00 | -.16815 | .19729 | .671 | -.6336 | .2973 |
| | 2.00 | .06164 | .19926 | .949 | -.4085 | .5318 |
| 1.00 | .00 | .16815 | .19729 | .671 | -.2973 | .6336 |
| | 2.00 | .22979 | .19863 | .480 | -.2388 | .6984 |
| 2.00 | .00 | -.06164 | .19926 | .949 | -.5318 | .4085 |
| | 1.00 | -.22979 | .19863 | .480 | -.6984 | .2388 |

*Appendix 15: ANOVA Privacy awareness*

### Descriptives

Identrel_MEAN_Shoes

| | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean | | Minimum | Maximum |
|---|---|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound | | |
| .00 | 76 | 5.34 | 1.369 | .157 | 5.03 | 5.65 | 1 | 7 |
| 1.00 | 77 | 5.31 | 1.280 | .146 | 5.01 | 5.60 | 1 | 7 |
| 2.00 | 74 | 4.93 | 1.238 | .144 | 4.64 | 5.21 | 1 | 7 |
| Total | 227 | 5.19 | 1.305 | .087 | 5.02 | 5.36 | 1 | 7 |

### Test of Homogeneity of Variances

| | | Levene Statistic | df1 | df2 | Sig. |
|---|---|---|---|---|---|
| Identrel_MEAN_Shoes | Based on Mean | .309 | 2 | 224 | .734 |
| | Based on Median | .176 | 2 | 224 | .838 |
| | Based on Median and with adjusted df | .176 | 2 | 216.956 | .838 |
| | Based on trimmed mean | .209 | 2 | 224 | .811 |

### ANOVA

Identrel_MEAN_Shoes

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 7.947 | 2 | 3.973 | 2.361 | .097 |
| Within Groups | 377.024 | 224 | 1.683 | | |
| Total | 384.971 | 226 | | | |

### Robust Tests of Equality of Means

Identrel_MEAN_Shoes

| | Statistic[a] | df1 | df2 | Sig. |
|---|---|---|---|---|
| Welch | 2.452 | 2 | 149.148 | .090 |
| Brown-Forsythe | 2.363 | 2 | 222.701 | .096 |

a. Asymptotically F distributed.

## Post Hoc Tests

### Multiple Comparisons

Dependent Variable: Identrel_MEAN_Shoes
Tukey HSD

| (I) ConditionTransp | (J) ConditionTransp | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| .00 | 1.00 | .037 | .210 | .983 | -.46 | .53 |
| | 2.00 | .416 | .212 | .123 | -.08 | .92 |
| 1.00 | .00 | -.037 | .210 | .983 | -.53 | .46 |
| | 2.00 | .380 | .211 | .173 | -.12 | .88 |
| 2.00 | .00 | -.416 | .212 | .123 | -.92 | .08 |
| | 1.00 | -.380 | .211 | .173 | -.88 | .12 |

*Appendix 16: ANOVA High Identity-relevance*

### Descriptives

Identrel_MEAN_TP

| | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean | | Minimum | Maximum |
|---|---|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound | | |
| .00 | 76 | 2.48 | 1.418 | .163 | 2.16 | 2.80 | 1 | 7 |
| 1.00 | 77 | 2.62 | 1.283 | .146 | 2.33 | 2.91 | 1 | 6 |
| 2.00 | 74 | 2.39 | 1.180 | .137 | 2.12 | 2.67 | 1 | 6 |
| Total | 227 | 2.50 | 1.296 | .086 | 2.33 | 2.67 | 1 | 7 |

### Test of Homogeneity of Variances

| | | Levene Statistic | df1 | df2 | Sig. |
|---|---|---|---|---|---|
| Identrel_MEAN_TP | Based on Mean | 1.174 | 2 | 224 | .311 |
| | Based on Median | .449 | 2 | 224 | .639 |
| | Based on Median and with adjusted df | .449 | 2 | 213.371 | .639 |
| | Based on trimmed mean | .954 | 2 | 224 | .387 |

### ANOVA

Identrel_MEAN_TP

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 2.067 | 2 | 1.033 | .613 | .543 |
| Within Groups | 377.433 | 224 | 1.685 | | |
| Total | 379.500 | 226 | | | |

### Robust Tests of Equality of Means

Identrel_MEAN_TP

| | Statistic[a] | df1 | df2 | Sig. |
|---|---|---|---|---|
| Welch | .669 | 2 | 148.789 | .514 |
| Brown-Forsythe | .615 | 2 | 219.776 | .542 |

a. Asymptotically F distributed.

## Post Hoc Tests

### Multiple Comparisons

Dependent Variable: Identrel_MEAN_TP
Tukey HSD

| (I) ConditionTransp | (J) ConditionTransp | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| .00 | 1.00 | -.143 | .210 | .774 | -.64 | .35 |
| | 2.00 | .088 | .212 | .909 | -.41 | .59 |
| 1.00 | .00 | .143 | .210 | .774 | -.35 | .64 |
| | 2.00 | .231 | .211 | .518 | -.27 | .73 |
| 2.00 | .00 | -.088 | .212 | .909 | -.59 | .41 |
| | 1.00 | -.231 | .211 | .518 | -.73 | .27 |

*Appendix 17: ANOVA Low Identity-relevance*

**Descriptives**

Autonomy_MEAN

| | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean Lower Bound | 95% Confidence Interval for Mean Upper Bound | Minimum | Maximum |
|---|---|---|---|---|---|---|---|---|
| .00 | 76 | 5.4918 | .82420 | .09454 | 5.3034 | 5.6801 | 3.13 | 7.00 |
| 1.00 | 77 | 5.3377 | .86009 | .09802 | 5.1424 | 5.5329 | 3.00 | 6.88 |
| 2.00 | 74 | 5.1419 | .94595 | .10997 | 4.9227 | 5.3611 | 2.00 | 6.88 |
| Total | 227 | 5.3254 | .88522 | .05875 | 5.2097 | 5.4412 | 2.00 | 7.00 |

**Test of Homogeneity of Variances**

| | | Levene Statistic | df1 | df2 | Sig. |
|---|---|---|---|---|---|
| Autonomy_MEAN | Based on Mean | .074 | 2 | 224 | .929 |
| | Based on Median | .060 | 2 | 224 | .942 |
| | Based on Median and with adjusted df | .060 | 2 | 209.567 | .942 |
| | Based on trimmed mean | .068 | 2 | 224 | .934 |

**ANOVA**

Autonomy_MEAN

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 4.607 | 2 | 2.304 | 2.992 | .052 |
| Within Groups | 172.491 | 224 | .770 | | |
| Total | 177.099 | 226 | | | |

**Robust Tests of Equality of Means**

Autonomy_MEAN

| | Statistic[a] | df1 | df2 | Sig. |
|---|---|---|---|---|
| Welch | 2.898 | 2 | 148.457 | .058 |
| Brown-Forsythe | 2.984 | 2 | 219.675 | .053 |

a. Asymptotically F distributed.

## Post Hoc Tests

**Multiple Comparisons**

Dependent Variable: Autonomy_MEAN

Tukey HSD

| (I) ConditionTransp | (J) ConditionTransp | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval Lower Bound | 95% Confidence Interval Upper Bound |
|---|---|---|---|---|---|---|
| .00 | 1.00 | .15411 | .14189 | .524 | -.1807 | .4889 |
| | 2.00 | .34988* | .14331 | .041 | .0118 | .6880 |
| 1.00 | .00 | -.15411 | .14189 | .524 | -.4889 | .1807 |
| | 2.00 | .19577 | .14285 | .358 | -.1413 | .5328 |
| 2.00 | .00 | -.34988* | .14331 | .041 | -.6880 | -.0118 |
| | 1.00 | -.19577 | .14285 | .358 | -.5328 | .1413 |

*. The mean difference is significant at the 0.05 level.

*Appendix 18: ANOVA Consumer autonomy*

### Univariate analysis

**Levene's Test of Equality of Error Variances[a]**

Dependent Variable: Transparency item 1 and 2 - Shoes

| F | df1 | df2 | Sig. |
|---|---|---|---|
| .814 | 2 | 223 | .445 |

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + ConditionTransp + TechComp_MEAN + ConditionTransp * TechComp_MEAN

**Tests of Between-Subjects Effects**

Dependent Variable: Transparency item 1 and 2 - Shoes

| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|---|
| Intercept | Hypothesis | 111.700 | 1 | 111.700 | 748.744 | .001 | .997 |
| | Error | .321 | 2.152 | .149[a] | | | |
| ConditionTransp | Hypothesis | .289 | 2 | .144 | .127 | .881 | .001 |
| | Error | 249.932 | 220 | 1.136[b] | | | |
| TechComp_MEAN | Hypothesis | 4.637 | 1 | 4.637 | 4.081 | .045 | .018 |
| | Error | 249.932 | 220 | 1.136[b] | | | |
| ConditionTransp * TechComp_MEAN | Hypothesis | .600 | 2 | .300 | .264 | .768 | .002 |
| | Error | 249.932 | 220 | 1.136[b] | | | |

a. .995 MS(ConditionTransp) + .005 MS(Error)

b. MS(Error)

*Appendix 19: Univariate Tech and Transparency Shoes*

**Levene's Test of Equality of Error Variances[a]**

Dependent Variable: Transparency item 1 and 2 - TP

| F | df1 | df2 | Sig. |
|---|---|---|---|
| .708 | 2 | 224 | .494 |

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + ConditionTransp + TechComp_MEAN + ConditionTransp * TechComp_MEAN

**Tests of Between-Subjects Effects**

Dependent Variable: Transparency item 1 and 2 - TP

| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|---|
| Intercept | Hypothesis | 88.784 | 1 | 88.784 | 1633.601 | .000 | .999 |
| | Error | .128 | 2.355 | .054[a] | | | |
| ConditionTransp | Hypothesis | .101 | 2 | .050 | .061 | .941 | .001 |
| | Error | 182.651 | 221 | .826[b] | | | |
| TechComp_MEAN | Hypothesis | 13.549 | 1 | 13.549 | 16.393 | .000 | .069 |
| | Error | 182.651 | 221 | .826[b] | | | |
| ConditionTransp * TechComp_MEAN | Hypothesis | .013 | 2 | .006 | .008 | .992 | .000 |
| | Error | 182.651 | 221 | .826[b] | | | |

a. .995 MS(ConditionTransp) + .005 MS(Error)

b. MS(Error)

*Appendix 20: Univariate Tech competence and transparency TP*

**Levene's Test of Equality of Error Variances[a]**

Dependent Variable: PrivAware_MEAN

| F | df1 | df2 | Sig. |
|---|---|---|---|
| .510 | 2 | 224 | .601 |

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + ConditionTransp + TechComp_MEAN + ConditionTransp * TechComp_MEAN

**Tests of Between-Subjects Effects**

Dependent Variable: PrivAware_MEAN

| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|---|
| Intercept | Hypothesis | 54.782 | 1 | 54.782 | 88.231 | .010 | .977 |
| | Error | 1.272 | 2.049 | .621[a] | | | |
| ConditionTransp | Hypothesis | 1.233 | 2 | .617 | .424 | .655 | .004 |
| | Error | 321.265 | 221 | 1.454[b] | | | |
| TechComp_MEAN | Hypothesis | 11.740 | 1 | 11.740 | 8.076 | .005 | .035 |
| | Error | 321.265 | 221 | 1.454[b] | | | |
| ConditionTransp * TechComp_MEAN | Hypothesis | .879 | 2 | .440 | .302 | .739 | .003 |
| | Error | 321.265 | 221 | 1.454[b] | | | |

a. .995 MS(ConditionTransp) + .005 MS(Error)

b. MS(Error)

*Appendix 21: Univariate Tech competence and Privacy Awareness*

**Levene's Test of Equality of Error Variances[a]**

Dependent Variable: Identrel_MEAN_Shoes

| F | df1 | df2 | Sig. |
|---|---|---|---|
| .344 | 2 | 223 | .709 |

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + ConditionTransp + Transp12_Shoes + ConditionTransp * Transp12_Shoes

**Tests of Between-Subjects Effects**

Dependent Variable: Identrel_MEAN_Shoes

| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|---|
| Intercept | Hypothesis | 164.195 | 1 | 164.195 | 128.651 | .006 | .984 |
| | Error | 2.674 | 2.095 | 1.276[a] | | | |
| ConditionTransp | Hypothesis | 2.538 | 2 | 1.269 | .751 | .473 | .007 |
| | Error | 371.831 | 220 | 1.690[b] | | | |
| Transp12_Shoes | Hypothesis | 2.225 | 1 | 2.225 | 1.317 | .252 | .006 |
| | Error | 371.831 | 220 | 1.690[b] | | | |
| ConditionTransp * Transp12_Shoes | Hypothesis | 3.150 | 2 | 1.575 | .932 | .395 | .008 |
| | Error | 371.831 | 220 | 1.690[b] | | | |

a. .983 MS(ConditionTransp) + .017 MS(Error)

b. MS(Error)

*Appendix 22: Transparency and identity-relevance Shoes*

**Levene's Test of Equality of Error Variances[a]**

Dependent Variable: Identrel_MEAN_TP

| F | df1 | df2 | Sig. |
|---|---|---|---|
| 1.207 | 2 | 224 | .301 |

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + ConditionTransp + Transp12_TP + ConditionTransp * Transp12_TP

**Tests of Between-Subjects Effects**

Dependent Variable: Identrel_MEAN_TP

| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|---|
| Intercept | Hypothesis | 37.453 | 1 | 37.453 | 1047.425 | .000 | .974 |
| | Error | .986 | 27.569 | .036[a] | | | |
| ConditionTransp | Hypothesis | .019 | 2 | .009 | .006 | .994 | .000 |
| | Error | 377.232 | 221 | 1.707[b] | | | |
| Transp12_TP | Hypothesis | .042 | 1 | .042 | .025 | .876 | .000 |
| | Error | 377.232 | 221 | 1.707[b] | | | |
| ConditionTransp * Transp12_TP | Hypothesis | .147 | 2 | .073 | .043 | .958 | .000 |
| | Error | 377.232 | 221 | 1.707[b] | | | |

a. .984 MS(ConditionTransp) + .016 MS(Error)

b. MS(Error)

*Appendix 23: Transparency and identity-relevance TP*

**Reliability**

**Reliability Statistics**

| Cronbach's Alpha | Cronbach's Alpha Based on Standardized Items | N of Items |
|---|---|---|
| .845 | .861 | 8 |

**Inter-Item Correlation Matrix**

| | I feel a sense of choice and freedom in the choice I made | I feel that my decision reflected what I really want | I feel my choice expresses who I really am | I felt in control of my choice | I felt that my choices belonged to me | My choice reflected my preferences | Perceived Autonomy 4 Reversed | The choice I made were free from external influence |
|---|---|---|---|---|---|---|---|---|
| I feel a sense of choice and freedom in the choice I made | 1.000 | .540 | .463 | .644 | .675 | .496 | .342 | .489 |
| I feel that my decision reflected what I really want | .540 | 1.000 | .489 | .533 | .578 | .539 | .209 | .369 |
| I feel my choice expresses who I really am | .463 | .489 | 1.000 | .397 | .505 | .455 | .109 | .392 |
| I felt in control of my choice | .644 | .533 | .397 | 1.000 | .545 | .434 | .321 | .455 |
| I felt that my choices belonged to me | .675 | .578 | .505 | .545 | 1.000 | .540 | .363 | .489 |
| My choice reflected my preferences | .496 | .539 | .455 | .434 | .540 | 1.000 | .208 | .373 |
| Perceived Autonomy 4 Reversed | .342 | .209 | .109 | .321 | .363 | .208 | 1.000 | .261 |
| The choice I made were free from external influence | .489 | .369 | .392 | .455 | .489 | .373 | .261 | 1.000 |

*Appendix 24: Reliability Consumer autonomy*

**Reliability Statistics**

| Cronbach's Alpha | Cronbach's Alpha Based on Standardized Items | N of Items |
|---|---|---|
| .813 | .814 | 4 |

**Item-Total Statistics**

| | Scale Mean if Item Deleted | Scale Variance if Item Deleted | Corrected Item-Total Correlation | Squared Multiple Correlation | Cronbach's Alpha if Item Deleted |
|---|---|---|---|---|---|
| Transparency_Shoes1 | 17.12 | 7.259 | .613 | .583 | .773 |
| Transparency_Shoes2 | 17.18 | 6.535 | .696 | .654 | .732 |
| TransparencyTP1 | 17.00 | 7.717 | .570 | .530 | .792 |
| TransparencyTP2 | 17.03 | 7.636 | .655 | .611 | .756 |

*Appendix 25: Reliability transparency combined*

**Reliability Statistics**

| Cronbach's Alpha | Cronbach's Alpha Based on Standardized Items | N of Items |
|---|---|---|
| .817 | .831 | 3 |

**Item-Total Statistics**

| | Scale Mean if Item Deleted | Scale Variance if Item Deleted | Corrected Item-Total Correlation | Squared Multiple Correlation | Cronbach's Alpha if Item Deleted |
|---|---|---|---|---|---|
| Transparency_Shoes1 | 10.57 | 5.901 | .656 | .543 | .771 |
| Transparency_Shoes2 | 10.62 | 5.051 | .791 | .651 | .633 |
| Transparency_Shoes3 | 11.25 | 4.623 | .606 | .414 | .847 |

*Appendix 26: Reliability Transparency shoes*

**Reliability Statistics**

| Cronbach's Alpha | Cronbach's Alpha Based on Standardized Items | N of Items |
|---|---|---|
| .783 | .803 | 3 |

**Item-Total Statistics**

| | Scale Mean if Item Deleted | Scale Variance if Item Deleted | Corrected Item-Total Correlation | Squared Multiple Correlation | Cronbach's Alpha if Item Deleted |
|---|---|---|---|---|---|
| TransparencyTP1 | 10.92 | 4.440 | .600 | .488 | .732 |
| TransparencyTP2 | 10.94 | 4.178 | .764 | .605 | .583 |
| TransparencyTP3 | 11.52 | 3.560 | .555 | .367 | .821 |

*Appendix 27: Reliability Transparency TP*

**Reliability Statistics**

| Cronbach's Alpha | Cronbach's Alpha Based on Standardized Items | N of Items |
|---|---|---|
| .791 | .792 | 2 |

*Appendix 28: Reliability Tech competence*

### Reliability Statistics

| Cronbach's Alpha | Cronbach's Alpha Based on Standardized Items | N of Items |
|---|---|---|
| .797 | .795 | 3 |

### Item-Total Statistics

| | Scale Mean if Item Deleted | Scale Variance if Item Deleted | Corrected Item-Total Correlation | Squared Multiple Correlation | Cronbach's Alpha if Item Deleted |
|---|---|---|---|---|---|
| I am aware of the privacy risks related to using this website | 9.04 | 8.118 | .538 | .301 | .825 |
| I follow the news about information privacy risks | 9.80 | 5.994 | .675 | .501 | .686 |
| I keep myself updated about information privacy risks and possible solutions to ensure my information privacy | 9.70 | 5.520 | .736 | .553 | .614 |

*Appendix 29: Privacy awareness*

### Reliability Statistics

| Cronbach's Alpha | Cronbach's Alpha Based on Standardized Items | N of Items |
|---|---|---|
| .704 | .702 | 4 |

### Item-Total Statistics

| | Scale Mean if Item Deleted | Scale Variance if Item Deleted | Corrected Item-Total Correlation | Squared Multiple Correlation | Cronbach's Alpha if Item Deleted |
|---|---|---|---|---|---|
| Identity relevance | 12.77 | 10.452 | .531 | .452 | .614 |
| IdentRelTP_ExpressYourself2 | 13.00 | 11.208 | .457 | .435 | .660 |
| IdentRelShoes_Sayalot1 | 10.51 | 9.897 | .558 | .450 | .595 |
| IdentRelShoes_ExpressYourself2 | 9.88 | 11.693 | .413 | .401 | .685 |

*Appendix 30: Reliability Identity-relevance*