Norwegian School of Economics
Bergen, Autumn 2022

Improving online clothing returns

For planet and profits

Alexander Frayne

Supervisor: Tor Wallin Andreassen

Master's Thesis in Strategy and Management

NORWEGIAN SCHOOL OF ECONOMICS

Preface

This master's thesis is written as part of my Master of Science in Economics and Business Administration program at the Norwegian School of Economics (NHH), where I am specializing in Strategy and Management. This study is a contribution to ongoing research projects at NHH, including the Digital Innovation for Sustainable Growth (DIG) program and the Business Model Innovation & Ecosystems for Seamless Transactions in Retail (BEST in Retail) program. DIG aims to promote digital transformation and innovation for sustainable growth, while BEST in Retail aims to generate knowledge, methods, and roadmaps to help the retail sector navigate the recent disruptive wave of digitalization. My participation in these programs has greatly benefited the research project, and I am deeply appreciative of the support I have received.

Acknowledgements

The author would like to thank Professor Tor Wallin Andreassen for his insight, suggestions, input, and feedback. His support was exceptionally valuable when things got tough. The author would also like to express gratitude to the participants who took the time to contribute through interviews, as well as to all the other individuals who have crossed my path and engaged in informal discussions, offering their views and input on the thesis and research topic.

Executive Summary

In this thesis, returns of clothing bought online are researched. The background of this is an increasing amount of clothing being ordered online and returned for various reasons. As many online retailers sell products with free, no-hassle returns, it is easy for consumers to order anything they want to try on or see in person and return as much as they like, with little to no personal consequences. However, this has a large cost for retailers who must deal with complex reverse logistics, as well as the environmental impact of shipping and sending products back and forth, which sometimes leads to items getting destroyed instead as it often is cheaper. Although there is a lot of research on the status of this and a lot of research on consumer behaviours, there is little connecting the two, examining what, in particular, can be done to mitigate the issue. This thesis attempts to answer what can be done to reduce the return rate of clothing bought online and the financial and environmental impact of returns for clothing bought online.

This thesis has a sequential exploratory design, starting with a qualitative stage using grounded theory development, followed by a quantitative stage testing the generated theory through an online questionnaire distributed to a large sample of U.S. consumers. In addition, it builds on existing research, such as The Theory of Planned Behaviour, in an attempt to produce a research model and find the independent variables that explain return volume.

Key findings are that returns handed back to a physical store can reduce the financial and environmental impact of returns compared to sending clothing back by mail. The research model, therefore, differs between these two types of returns. However, it does find that shopping volume is the strongest determinant for both sent and physical returns. Additionally, the knowledge among consumers on the topic seems to be low, and socially responsible consumers surprisingly have higher sent returns, but the same physical returns. Normative influences also explain return behaviour, however, for sent returns, it is personal norms, while for physical returns it is not, but rather one's subjective opinion of others' norms.

Table of contents

Table of Figures

Table of Tables

1. Introduction

1.1 Background

Worldwide e-commerce sales have experienced a significant surge in recent years, increasing from $1,336 billion U.S. dollars in 2014 to $5,211 billion U.S. dollars in 2021, with an estimated projection of $8,148 billion by 2026 (eMarketer, 2022). Clothing consistently ranks as the top category for online purchases in most countries, a trend that has persisted since at least 2018 (DC Velocity et al., 2021; Salesforce Research, 2021; Statista & We Are Social, 2018). In 2025, the estimated market value of the fashion e-commerce sector is projected to reach $1,207 billion U.S. dollars (Research and Markets & Statista, 2021).

The growth in online clothing sales has been driven by various factors, such as a broader online product range, expedited shipping, hassle-free returns, an improved user experience on smartphones, personalized content and recommendations, the influence of social influencers, and the pervasive presence of social media (Cullinane, 2019; Hull, 2012; Statista, 2021). The prevalence of smartphones and the increased availability of high-speed internet have made online shopping more accessible, convenient, and user-friendly. Consequently, consumers can now comfortably shop online during their daily commutes, with mobile devices occupying an ever-growing share of the platforms used for online purchases (PYMNTS, 2022). In the meantime, a recent study reveals that young British adults, on average, make five mobile purchases per week while commuting, with clothing ranking as the top category of interest (Kinetic, 2022). This trend is unsurprising, given that convenience is a significant factor influencing store choice, whether online or offline, for many consumers (Bellenger & Korgaonkar, 1980; Rohma & Swaminathan, 2004; Vignali & Reid, 2014). The convenience of browsing a wide variety of clothing items using search filters and AI recommender systems to refine the search, without the need to physically browse stores or wait in line, can be particularly appealing while commuting or in various other contexts. This appeal is further amplified when online retailers offer free shipping and a hassle-free returns service.

But what happens when you purchase clothing online that doesn't meet your expectations, and you wish to return it? What occurs when consumers order the same quantity of clothing to their homes as they would have taken to a changing room for trying on, knowing they can easily and freely return any unwanted items? As online shopping continues to gain popularity, and brands encourage consumers to order directly from their distribution centres, return policies become more lenient to attract convenience-seeking customers (McKinsey, 2022; Vignali & Reid, 2014). If one can order as much as they wish to try on and return items without consequences or charges, why would convenience-seeking consumers limit the size of their orders?

In line with the growth in online shopping, the number of returned products has also been steadily increasing. Research indicates that return rates vary, ranging from 10% for plain t-shirts (Cullinane & Cullinane, 2021) to as high as 90% for long dresses (Kristiansen, 2022). On average, return rates fall in the range of 20% to 50% across different product categories, however, these rates vary significantly among different demographic groups, such as age and gender (Cullinane & Cullinane, 2021; Cullinane, 2019; Kristiansen, 2022).

Studies suggest that the return process is often lengthy and intricate, resulting in an even larger carbon footprint associated with the returned item (Cullinane & Cullinane, 2021). Frequently, returned clothing embarks on extensive return journeys, spanning thousands of kilometres, as it is sent from one processing hub to another for resale (Kapner, 2023; Kristiansen, 2022). Furthermore, not all clothing items have profit margins high enough to justify the expense of returning, processing, and delivering them to new paying customers, as the additional income from resale may not offset these costs (Cullinane, 2019; Kristiansen, 2022; Manayiti & Edgecliffe-Johnson, 2022; Peiser, 2022).

1.2 Context

As both online clothing sales and their return rates continue to grow, it becomes increasingly important to understand how the climate footprint of this trend can be reduced. This significance is further underscored by The United Nations' Sustainable Development Goals, particularly goal #12, which aims to "Ensure sustainable consumption and production patterns" (2022). It has become evident that neither the current level of consumption nor its trajectory is sustainable for the planet, and it requires addressing (United Nations, 2022).

Clothing production has a significant negative climate impact due to carbon emissions, unsustainable resource usage, and the generation of substantial waste and pollution (Claudio, 2007; Ore, 2022). The European Union (EU) has introduced a new textile strategy to counter this adverse trend, aiming to transition the industry toward more sustainable practices and facilitating the initial steps toward circular business models (European Commission, 2022). These recent policies seek to limit waste and mitigate the impact of clothing, including efforts to discourage the destruction of unsold goods (European Commission, 2022; Kristiansen, 2022; Ore, 2022). However, the costs for businesses in managing excessive returns, as well as the associated climate impact, persist as challenges.

Given that this thesis is written in collaboration with DIG and BEST in Retail, the research also aims to explore the potential for technology to drive innovation and growth, in addition to providing concrete managerial insights for retailers and policymakers to address this issue effectively.

1.3 Research questions

What can be done to reduce both the financial impact on profits and the environmental impact on the planet, from consumers returning clothing bought online.

This thesis will try to answer the question: **What can be done to reduce both the financial impact on profits and the environmental impact on the planet, from consumers returning clothing bought online?** This can be formulated mathematically as the multiplication of two factors:

> *Number of returns × the environmental impact and financial impact of each return.*

Therefore, we get the following three ***possible*** research questions (RQs), where RQ1 addresses the number of returns, and RQ2 and RQ3 address the environmental and financial impact of each return.

RQ1*: What factors, if any, will have the largest impact on reducing consumers' return rate of clothing bought online?*

RQ2*: What factors, if any, will have the largest impact on reducing the carbon footprint of each returned online order?*

RQ3*: What factors, if any, will have the largest impact on reducing the financial cost of each returned online order?*

By answering RQ1 we attempt to find explanations on the current situation and what can be done to reduce the rate of returns. A reduction in returns will have a positive impact on the climate footprint of returns and a positive impact on retailers' profits. By answering RQ2 we attempt to further reduce the climate footprint of returns. And finally, by answering RQ3 we attempt to reduce the costs for the retailers covering the practice of returning goods.

As far as the actor involved, the thesis will attempt to address this for all actors involved, i.e., what can companies do to address this, what can policymakers do, and what can others do to address this.

1.4 Research design and methodology

The research design is a plan that determines how a research question will be answered and shapes the research process (Saunders et al., 2016). This thesis aims to address how returns of clothing purchased online can have a reduced impact on the climate and on retailers' profits, as well as how the return rate itself can be reduced. Due to the exploratory nature of this research and the limited amount of previous research on the topic, a sequential exploratory design will be employed. The first stage will be exploratory, with the goal of using grounded theory methodology to generate insights that can be applied in the second stage to answer the research questions. The second stage will be descriptive and employ a quantitative approach, specifically using a questionnaire to test the theories developed in the first stage and further answer the research questions. This is an abductive research approach, combining elements of induction and deduction (Saunders et al., 2016). This design and approach are well-suited because the first stage allows for a flexible, open, curious, and creative approach to gaining insights about the topic, exploring the phenomenon in depth, and developing a theory and model (Flick, 2013). The second stage is more deductive and suitable for testing the theory and model, evaluating its generalizability, and examining the relationship between the model's variables deductively (Saunders et al., 2016). The methods used in each stage and the findings from each stage will be presented fully in their respective chapters.

1.5 Thesis outline

Chapter 2 will introduce the general research approach and briefly review current research and literature related to the issue, linking it to the focus areas for the first research stage.

Chapter 3 will present the first research stage, utilizing grounded theory development to explore the topic areas identified in Chapter 2. The methodology and findings for this stage will be presented separately.

Chapter 4 will cover the second research stage, where the theory and findings from the previous stage are integrated to form a model and hypotheses. These will be quantitatively tested through a questionnaire. The methodology and findings for this stage will also be presented.

Finally, Chapter 5 will discuss the implications of the overall findings, address the research's limitations, and suggest potential areas for future research.

2. Research approach and background

The upcoming chapter will elucidate the research approach, summarize the current research findings, identify existing gaps, and highlight the areas that will be the primary focus of the first stage. To accomplish this, it will begin by presenting a framework for addressing the issue. Subsequently, it will review the existing literature and research pertaining to the components of the framework. Lastly, it will uncover the gaps in the current research and clarify the contributions made by this research.

2.1 Framework for tackling the research questions

This research aims to address the questions by loosely adapting a modified version of Cullinane and Cullinane's framework (2021). The adjustment is made to encompass not only environmental aspects but also improvements for retailers. The issue is disaggregated into sub-issues that are mutually exclusive and collectively exhaustive components, which, in turn, are further divided into sub-issues. A selection of these will be reviewed and investigated during the research in pursuit of answers to the research questions. This structure is presented in **Error! Reference source not found.** below.

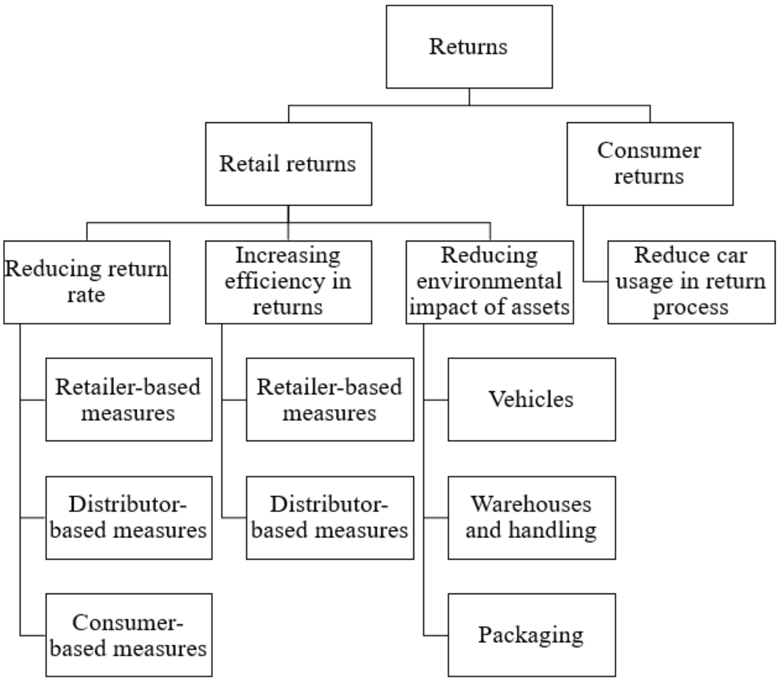


Figure 2-1: Cullinane and Cullinane's framework (2021)

A selection of these sub-issues will be explored throughout the research in both phases. The results of the answers and findings will contribute to the development of theories regarding consumer behaviour related to environmental actions. They will also offer managerial insights for retailers regarding potential strategies to mitigate and address the issue. Lastly, the research will address whether these efforts are sufficient or if policymakers need to intervene to effect substantial change.

2.2 Current status on returns in online clothes shopping

There are various areas and approaches to commence the investigation of this broad topic and problem. In the following section, the initial direction and area of focus will be discussed, defined, and justified, as this will lay the groundwork for the research's path.

Several reasons for the issue remaining uncorrected have been identified. Firstly, the research will examine consumers and the nature of returns. Secondly, businesses will be addressed, as they often react and adapt their behaviours in response to consumers' actions and desires.

2.2.1 Consumers – Nature of returns

Initially, the nature of returns will be examined, including what previous research says about the sources of returns, as this information is crucial for reducing the return rate. Extensive research and quantitative descriptive studies have been conducted on consumer returns from online shopping. A selection of this research will be presented below.

Saarijärvi et al. (2017) identified the drivers for returning clothing purchased online, categorizing them as planned or unplanned behaviour, as well as consumer-initiated or other-initiated. These reasons are varied but typically fall within the drivers categorized below in Table 2-1. A detailed explanation of these can be found in Table 8-1 in the appendix.

	Planned returning behaviour	Unplanned returning behaviour
Consumer-initiated	Benefit maximization driven Just trying out driven Money shortage driven	Money shortage driven Feeling driven
Other initiated		Competition driven Disconfirmation driven Order fulfilment driven Faded need driven Size chart driven Reclamation driven

Table 2-1: Unplanned and planned online returning behaviour

Furthermore, Kaushik et al. (2020) also compiled the reasons for returns and which research that supports these reasons. They provide a list of six main reasons: apparel attributes, disconformity, dissonance, service failure, opportunism, and perception. Each of these has sub-reasons. A comprehensive overview can be found in the appendix in Table 8-2.

Additionally, Cullinane and Cullinane (2021) reviewed several studies conducted from 2000 to 2020, indicating that the number of customers purchasing items they already planned to return has been increasing, ranging from 20% up to 45% for younger consumers (Accenture, 2018; IMRG, 2020; Metapack, 2020; Piron & Young, 2000). Other sources report that 58% of consumers intentionally buy more goods than they intend to keep (Salerno-Garthwaite, 2022).

Furthermore, these studies also reveal that many consumers prefer purchasing clothing from retailers with lenient policies (Narvar, 2019, 2022).. This preference often leads to 'bracketing,' wherein consumers order several items in multiple sizes and/or colours, increasing the likelihood of receiving something that matches their expectations of the product's attributes based on their initial perception when viewing it online (McKinsey, 2022; Narvar, 2022). Upon receiving the items, they have the option to return anything they do not wish to keep (Cullinane & Cullinane, 2021).

Narvar's reports from 2019 and 2022 confirm this trend of consumers buying more items than they intend to keep, with the incidence of bracketing increasing from 56% globally in 2019 to 63% in 2022 (Narvar, 2019, 2022). In contrast, McKinsey finds that bracketing accounts for approximately 15% of returns (McKinsey, 2022).. Earlier data from McKinsey indicates that 20% of returns result from receiving damaged products, 22% are due to the product looking different, 23% involve receiving the wrong product, and 35% are for other reasons (McKinsey, 2018). Additionally, return rates are on the rise and currently stand at a minimum of 20% for online retail (Kapner, 2023; Mull, 2023), potentially even higher for clothing (Jansson-Boyd, 2023), with some multi-brand retailers experiencing return rates as high as 54% (McKinsey, 2022).

These consumer behaviour patterns are costly for retailers, creating opportunities for specialized logistics companies to manage the complex and expensive process of reverse logistics (Cullinane & Cullinane, 2021; Cullinane, 2019; Kapner, 2023; Peiser, 2022).. The expense of handling a returned item is significant, with data indicating that it can amount to approximately 66% of the item's value (Kapner, 2023; Smith, 2022).. The precise cost varies, ranging from $21 to $46 (McKinsey, 2022). This underscores that many items may not be economically viable to process for return, as their selling price is unlikely to cover the costs of return processing.

To summarize, a significant portion of returns arises from "bracketing," wherein consumers order clothing online with the intention or awareness that they will be returning at least some of it.

2.2.2 Retailers – Challenges in Mitigating The Issue

How are businesses addressing this challenge? Cullinane and Cullinane (2021) have observed that an increasing number of goods purchased online are being bought cross-border, with clothing being the most frequently purchased product category (Frederick, 2015). Coupled with the fact that handling returned items requires up to 20% more space than their outbound journey (Salerno-Garthwaite, 2022), the entire process becomes highly intricate, time-consuming, and costly for retailers (Dickler, 2022; Kapner, 2023). Consequently, many retailers opt to engage specialized reverse logistics companies to manage this aspect of their operations (Cullinane & Cullinane, 2021; Cullinane, 2019; Kapner, 2023; Peiser, 2022). Furthermore, consumers often appear to have limited knowledge about the issue or a clear understanding of the reverse journey that items undertake after they are returned (Kapner, 2023; Kristiansen, 2022).

So why aren't businesses proactively sharing this information, educating consumers, and implementing paid return policies? In fact, an increasing number of retailers, up to 40% or more, are beginning to modify their free return policies by introducing return fees. This change aims to bring attention to the issue, reduce the problem, recover a portion of return costs, and discourage customers from returning goods (Dickler, 2022; Kapner, 2023; Mull, 2023; Ore, 2022; Peiser, 2022; Smith, 2022). However, the policy of easy, free, and hassle-free returns is a critical factor in attracting consumers to shop online. Retailers offering such policies can achieve higher customer loyalty and boost sales through more repeat purchases and larger order sizes (Gäthke et al., 2021; Kapner, 2023; McKinsey, 2018; Narvar, 2019, 2022). This creates a situation where those who do not follow suit lose a competitive edge (Ore, 2022). It appears that many are compelled to adopt the most lenient policy. While new data suggests a potential shift in industry standards, as of now, free and hassle-free returns remain the norm.

There are two other areas retailers could explore to mitigate the issue, as paid returns often only cover a small portion of the total return process costs (Dickler, 2022; McKinsey, 2022). Thus, there is mixed data on whether this will become the industry standard going forward or if it is just the case with some retailers. Furthermore, it is unclear how much paid returns will help. Therefore, other options must also be explored to further address the issue. This leads to the next two areas.

The first area is technology, which can help consumers better assess product attributes before making a purchase to avoid situations where there is a mismatch in expectations once an item arrives. This, along with standardized sizes, could assist customers in getting their purchase right the first time. Knowing the correct size is only one part of the equation, as two items with the same size or dimensions can fit differently and feel differently depending on the material and cut. Additionally, defining standard sizes is challenging, as there are too many different measurements on a body to fit all into standard sizes (Bogusławska-Bączek, 2010; LaBat, 2007). Therefore, technology and standardization would likely improve the chances of a customer getting it right but may never completely eliminate the issue, as aspects like touch and feel are currently not possible to simulate remotely using technology. This means that risk-averse customers would still likely benefit from purchasing more than they intend to keep, as long as this comes without a cost or at a low cost that doesn't deter the behaviour. However, technology can help increase the confidence consumers have that they are ordering the correct items, reducing their need to over-order. This aspect should be considered in conjunction with return policies, as overly lenient policies will still encourage consumers not to take the chance on items they think are correct, leading to excessive ordering.

At the same time, there is uncertainty regarding how much consumers care about the environmental impact of excessive returns.

2.2.3 Possibilities to remedy the situation: Technology and supply chain logistics

Researchers and businesses have identified several technological solutions that can assist consumers in ordering the correct items. For instance, technology such as avatars can enable users to visualize clothing on a digital twin before making a purchase. This solution varies in complexity, with some companies offering detailed full-body scans to create precise digital avatars (H&M Group, 2019, 2021; Marks and Spencer, 2019), These digital twins can try on clothing on behalf of the user. Other companies utilize simpler technology, leveraging users' mobile devices to scan and obtain approximate measurements for their avatars (MYSIZE, 2023).

Avatars and virtual fitting rooms of this nature can be accessible in the metaverse, on mobile phones, or through augmented or virtual reality (NeXR, 2022; Pauly, 2022). However, this approach necessitates digitalizing the clothing to be able to apply it to the avatar. Many fashion companies are embracing this digitalization trend, as similar processes are frequently employed in the design phase (Clo3D, 2022; H&M Group, 2021).

Additionally, artificial intelligence and machine learning can assist consumers in making informed choices by providing sizing recommendations based on user input, such as past purchases or other user data (Bahuleyan et al., 2022; Fearn, 2020; ZIZR, 2022).

Besides technology to assist consumers in making informed choices, retailers can also address the reverse logistics process to minimize the impact of each item that enters this process. Reverse logistics as a topic in itself is not widely researched, as found by Cullinane and Cullinane (2021), who show that Rubio et al (2008) found 186 articles in 26 journals between 1995 and 2005 addressing the environmental focus. Furthermore, they reference the work of Wang et al. (2017), who found sustainability to be a key theme in reverse logistics research between 1992 and 2015, primarily on social and economic sustainability, rather than environmental sustainability. In their seminal paper, Cullinane and Cullinane (2021) further propose measures for retailers and distributors to improve the efficiency of the returns process, which can reduce the environmental and cost impacts of returns, as part of the framework presented initially. Here, they suggest that technology can help increase the speed of product flow by using tracking, and returned packages could be shipped directly to new customers. In addition, alternative paths for stock that cannot be resold and better strategic choices for the location of regional and international return processing facilities are recommended.

2.3 Gaps in current research and way ahead

Following the above, I will present a conclusion of the current research and the gaps in the research, laying the path for this research study.

There is uncertainty about the different reverse logistics possibilities for retailers and logistics operators. For example, there may be a willingness from consumers to deliver returns to shops, which would help mitigate the cost and environmental impact (Narvar, 2019), but the extent of this possibility and the exact benefits are unclear.

It is not clear why companies are not using technology to a larger degree, compared to the number of tools available and the time these have existed. Is this due to the cost of the technology? Are they not aware? Or are they afraid it could lead to lower sales?

More information is needed about consumers. It is unclear how much awareness there is around the environmental costs of returns, whether consumers are unaware, indifferent, or if they care but behave otherwise because they cannot help themselves. What are consumers' attitudes toward the problem, and what shapes their attitudes and behaviour surrounding returns? Are larger deterrents needed, and would they help? It is also unclear if a policy change is required to tackle the issue if neither consumers nor businesses can address it effectively.

Another literal and figurative gap in the research is the empty quadrant in Saarijärvi's overview (2017), as seen in Table 2-1. Although not specified, one could hypothesize that social norms and the influence of others might be reasons that affect return behaviour.

Based on this, more data and information are needed from four different groups: logistics providers, retailers, consumers, and experts working in the intersection between technology, retail, and logistics.

3. Grounded theory development

3.1 Methodology

3.1.1 Design

Approach

The research approach in this phase was abductive, where deduction and induction are combined in an iterative process in which data is collected to explore a topic, identify themes, and generate a theory (Saunders et al., 2016). The grounded theory method used for this was qualitative, as it allowed greater possibilities to delve into depth, explore, and primarily collect non-numerical data.

Method

The Grounded Theory method was used in this stage, as it is well-suited for the above-explained approach and aims of this stage (Flick, 2013; Saunders et al., 2016).. Grounded Theory was developed as a method to analyse, interpret, and explain the meanings that individuals create to understand their daily experiences within specific contexts (Charmaz, 2014; Glaser & Strauss, 1967). It aims to develop a theory based on systematically analysing and interpreting data to generate new insights and understandings. The process of the grounded theory method is flexible, but there are guidelines for the process. Based on Glaser and Strauss (1967), Corbin and Strauss (2015), Charmaz (2014), Flick (2013), and Saunders et al. (2016), these steps will be briefly presented below as an introduction to the method.

The process starts with formulating a broad, exploratory, and open-ended research question. Next, data is collected from a broad range of sources. The data can be quantitative, qualitative, or both. The data is coded without preconceived theories or frameworks in mind, focusing on the data itself. Coding is done simultaneously with data collection, allowing for an iterative process of data collection and coding. As more data is collected and coded, constant comparisons are made with previous data to identify similarities, differences, and patterns. This comparative analysis helps in developing categories and concepts, which are the building blocks of grounded theory. Based on the emerging categories and concepts, one must purposively select additional data sources that will provide further insights and contribute to the theoretical development. This is called theoretical sampling, as it is driven by the need to explore and refine the emerging theory. Data is continuously collected and analysed until theoretical saturation is achieved. This occurs when the newly collected data no longer reveals new insights or contributes significantly to the development of the theory. This indicates that the theory has reached a point of comprehensive-ness. Throughout the research process, memos are written to document thoughts, ideas, and reflections on the emerging theory. Memos serve as a means to capture and preserve the analytical process, aiding in the development of the theory. Finally, the coded data, categories, and memos are analysed to develop a conceptual framework and theory that explains the phenomenon under investigation. This involves integrating and refining the categories, identifying relationships and connections between them, and generating a coherent and comprehensive theory. This theory is then assessed through member checking, peer debriefing, or seeking feedback from experts in the field.

It is important to note that the grounded theory method is a flexible and iterative process. Re-searchers continually move back and forth between data collection, analysis, and theory development, allowing for constant refinement and modification of the emerging theory. Now that the general guidelines are covered, the exact way it was conducted in the study will be presented in detail in the coming sections. Also, important to note is that published theory may be used before and during research (Saunders et al., 2016) to help inform the project in general terms. It should, however, not influence the analysis, how the data is coded, or which new cases to look at (Corbin & Strauss, 2015).

Strategy and objective

The primary objective of Grounded Theory Method is to generate a new theory that explains and illuminates a particular phenomenon or social process. The aim is to provide a deeper understanding of the researched area by systematically analysing the data, identifying patterns and categories, and constructing a theory that emerges from the data itself. This process picks up where Chapter 2 ended and is covered in the next section.

The primary objective of the Grounded Theory Method is to generate a new theory that explains and illuminates a particular phenomenon or social process. The aim is to provide a deeper understanding of the researched area by systematically analysing the data, identifying patterns and categories, and constructing a theory that emerges from the data itself. This process picks up where Chapter 2 ended and is covered in the next section.

3.1.2 Data collection

Data sources and sample selection

The initial plan was to try to gather data by speaking to people in many different roles within the topic area. The plan was to speak to people in four different areas: Firstly, businesses that sell clothing online to consumers, and thus deal with the issue regularly. Secondly, people working in logistics, also directly involved in the issue on a regular basis. Thirdly, experts on the issue who could combine knowledge about consumer behaviour, technology, industry knowledge on logistics, fashion, and e-commerce. Fourthly, consumers who regularly purchase clothing online. This data was to be supplemented with secondary data and previous research to generate an emerging theory.

The initial plan proved to be too ambitious and extensive, as well as challenging to gain access with the available resources, time, and budget. Two groups were not included in this stage: clothing retailers, and consumers. The reasons each of these two were dropped will be elaborated sequentially in the following two paragraphs.

Many attempts were made to contact businesses in textiles; unfortunately, all were unsuccessful. Over 20 companies were contacted, asking for a 60 or 30-minute interview. Many of these were not cold calls, but they were contacted via an introduction by a person familiar with the company, or via existing contacts of the author. For example, the DIG network, which included Virke Handel, a network of thousands of Norwegian retailers, was also used to reach out to companies to participate. A wide range of companies were contacted, including small and large companies, national and international companies, companies known for a strong eco-friendly profile, and companies criticized for having a poor eco-friendly profile. The contacted businesses all said it was due to a lack of resources and time. Despite these efforts, only one agreed to a 30-minute interview but later withdrew and said they did not have time after they received the consent form containing information about the research study and the discussion points for the interview. All those contacted showed little willingness to speak about the issue on record, even after being assured anonymity and that it would only last 30 minutes. Many did, however, acknowledge the importance of the topic and wished me luck "with the important work I was doing."

The reasons consumers were not included are twofold. Firstly, as the study was written in collaboration with DIG, it was important to recruit participants who could somewhat be generalized to the population. Therefore, consumers in the researchers' own network of friends, family, and fellow students were excluded from the sample from which participants were to be selected. As there was a budget for collecting data, the plan was to use a market-research firm to recruit desired participants for the data collection, which would happen in the form of one or more focus groups. Secondly, after several attempts, it became clear that the budget was not large enough to cover the expenses needed to conduct such a focus group. These two reasons led to the decision not to prioritize collecting qualitative data from this group but rather to collect quantitative data from them at a later stage, once a theory had been generated that could be tested.

This led to two remaining groups that were interviewed: the logistics operators and the experts. The first group was the logistics providers, and these were interviewed first as it allowed the possibility to gather and explore information without being influenced by any prior knowledge or inputs from the experts, as is desired in grounded theory practice (Flick, 2013). The first interviewee was found by using the Virke network, and after going through several links of people, we finally reached a person working in logistics who was said to work directly with online retailers, returns, sustainability, and future solutions. After interviewing this person, transcribing the interview, and analysing the data, a second person was recruited based on the emerging topics from the first interview that seemed most prominent. By using LinkedIn, combined with research and information on logistic companies' different departments, plans, and strategies, several possible candidates were identified. These were contacted, and one individual agreed to participate. Following this interview and subsequent data analysis, themes and topics were better solidified, and the data collection from this group was assessed to be sufficiently saturated. It was decided to continue to the next group of participants. Both participants had many years of experience working in the logistics sector and specific knowledge and experience with online retail, returns, and emerging technologies for facilitating improved future logistics to mitigate the issue.

The next group of participants were the experts. Here, the aim was to speak to individuals with expertise that spanned several areas and who could provide data that connected the themes from the logistics operators with the other areas not covered so far, which included consumers, the online retail industry, and the possibilities of technology to mitigate the problem. Through the DIG network, two participants were identified, working as advisors and consultants in leading positions at a top consulting firm. They had many years of experience and extensive knowledge in working with consumer industries, retail and fashion, logistics and operations, and how technology can play a role in improving businesses in these realms.

Participant	Area	Business	Sector
Interviewee 1	Logistics	Company A	Freight, shipping, forwarding
Interviewee 2	Logistics	Company B	Freight, shipping, forwarding
Interviewee 3	Expert on network	Company C	Consulting
Interviewee 4	Expert on network	Company C	Consulting

Table 3-1: Interviewees and their positions

Semi-structured interviews

As it was a qualitative abductive approach, semi-structured interviews were chosen to understand the topic in depth. This method is well-suited for an exploratory research phase where the goal was to delve into the topic and comprehend the interviewees' opinions and thoughts about the topics (Saunders et al., 2016). The interview questions were open-ended to gather the maximum amount of information. The approach was flexible, allowing for adaptation during the interview and gathering information about different sub-topics based on what the interviewee had to say. It provided room for asking follow-up questions, clarifying topics, and grasping the true meaning of the interviewees' opinions. The interviews lasted 60-90 minutes to allow full coverage of the topics.

However, the interviews were not entirely unstructured. Prior to each interview, a tailored interview guide was created, noting the different themes and topics that had emerged from previous interviews and data collection. This allowed for comparisons of data, finding commonalities and differences, and assessing the level of data saturation. Memos and notes from coding and working with the data were also used to shape the interview guides.

Interview guide

The interviewees received information beforehand about the themes I wanted to discuss with them. Each interview was adapted based on the previous interviews and the interviewee's role. See an example of such an interview guide in the Chapter 3 appendix.

The interview guide was adjusted before each interview based on the interviewee's profile. It was also modified depending on how my knowledge of the subject developed after each subsequent interview, following grounded theory principles.

Interview Process

The interview process was comprised of three stages. Pre-interview, interview, and post-interview. The process of each stage will be explained below.

First, the pre-interview phase. This involved contacting the interviewee, sharing preliminary information about the research study, and inquiring if they were interested in participating. In case of a positive response, an interview guide was tailored to the person's profile, building on previously emerged knowledge and themes. An information leaflet was then created and sent to them, providing details about the research study, the interview's purpose and topics, the researcher's identity, the significance of their participation, voluntariness, privacy measures, data handling, their rights, and their consent to anonymized data usage in the research. See an example of one in the Chapter 3 appendix. Subsequently, interviewees provided available timeslots, and the interview date and time were confirmed.

Next was the interview process. Since the interviewees were physically located elsewhere than the interviewer, all interviews were conducted as video calls over Microsoft Teams. While this reduced the possibility of establishing a physical connection, it offered the advantage of interviewing individuals from any location, with greater timing flexibility, reduced risk of biases based on appearances, and a lower threshold for the interviewee to participate (Saunders et al., 2016). It also simplified the recording and secure storage of interviews. Each interview commenced with an explanation of the key information from the leaflet. The interviewee was informed about the research's purpose, objectives, their invitation to participate, and expectations. Full anonymity and sensitive data handling were assured. The interviewee was then asked to consent to video recording, participate in the research, and their consent was recorded once more when the recording began. They were reminded that there were no right or wrong answers and that the goal was to gather their opinions and views. These measures were implemented to ensure interviewees felt comfortable sharing openly and understood the professional and respectful handling of the process. The interview proceeded with open and exploratory questions, supplemented by probing inquiries when necessary. Occasional specific and closed questions were utilized for clarification purposes. This mix of question types aimed to acquire a comprehensive understanding of the interviewee's perspective, opinions, and interpretations (Saunders et al., 2016). During the interview, notes were taken while the interviewee spoke, facilitating reference to earlier segments for clarification without interrupting longer monologues. After covering the topics outlined in the interview guide and any additional topics that arose, the interviewees were invited to add or clarify any information, discuss relevant points that were not addressed, or suggest topics they believed I should inquire about. These measures were taken to ensure the interviews were as comprehensive as possible. Finally, they were informed about the final stage of the interview process.

he final and third stage of the interview process involved rewatching the video recording and taking additional notes to ensure that nothing had been missed during the interview. Subsequently, the interview was transcribed. This process typically took about a week, and the transcriptions were then sent to the interviewees via email. They were asked to review the transcriptions and confirm if there were any errors, clarifications needed, or if there were any discrepancies between the transcription and their intended statements. All the interviewees approved the transcriptions without suggesting any amendments, comments, or changes. It took approximately 1-3 weeks to obtain this confirmation each time. Each interview was analysed and coded during this period, and no new interviews were conducted until the most recent one had been approved by the interviewee, fully analysed, and coded, in accordance with the grounded theory research method (Flick, 2013). All the participants seemed eager to participate and were positive contributors.

Other primary data sources

In addition to the semi-structured interviews, there were two other sources of primary data. The first was a panel discussion in Norwegian, roughly translated as "Tech and Textile – How Does the Future Sustainable Textile Industry Look?" This event took place on September 19, 2022, in Bergen as part of the innovation week "OPP." It featured five speakers and one moderator and lasted approximately two hours. The main topics covered environmental challenges in the fashion industry, the potential role of technology in addressing these challenges, and other factors influencing this interplay, such as consumer behaviour, trends, regulations, and more. Following the panel discussion, the author engaged in informal conversations with some of the panel members. The five speakers had diverse backgrounds, ranging from large established clothing companies to small start-ups with circular business models, technology experts, and industrial suppliers to the industry.

The second source of primary data comprises the numerous informal discussions held with individuals about the research topic throughout the study. Grounded theory does not solely rely on interviews for data collection, and informal conversations are considered a valid source of primary data, enhancing the variety of sources and strengthening data collection (Flick, 2013). From the beginning of the study in April 2022, until its conclusion in August 2023, I engaged with a wide range of people. This included colleagues of the authors from two different consulting jobs – one in technology and the other in supply chain – as well as informal meetings with clients in these roles. It is worth noting that none of the interviewees from the semi-structured interviews were employed or connected to any of the author's workplaces. These informal conversations also encompassed discussions with professors and fellow students at NHH, as well as with co-professors and fellow students while on exchange at ESADE. Additionally, family and friends were included. During these meetings, the topic of the master's thesis was frequently discussed, and most individuals were willing to share their perspectives on the issue, their opinions, and their thoughts on the findings I had gathered thus far. This data was documented in memos and notes, which were used to shape and test emerging theories and ideas. Throughout this research phase, a journal was maintained to continuously capture thoughts and ideas as they arose during work, helping to maintain direction and steady progress.

Secondary data

Throughout the study, secondary data was also used to complement the primary data sources. As argued by Glaser (1992, 1998, 2008), using other forms of data, such as quantitative data, can be compatible with Grounded Theory. However, it is essential to be mindful of how this impacts the research. For instance, using existing literature to find codes and forcing the collected data to align with these codes is not consistent with the grounded theory method. Thus, this approach was not employed in such a manner. Nonetheless, secondary data and existing literature served several other purposes in the study.

Primarily, secondary data was used to gather information in three distinct areas: fashion retailers with an online presence, consumers, and technology companies offering potential solutions to mitigate the issue. In addition to these three areas, the research incorporated secondary data, such as news articles about the industry and the issue to stay updated, as well as data about the interviewees and their companies. It also encompassed research papers providing data on consumer behaviours, whitepapers, and news reports related to e-commerce and the retail and fashion industry, some of which were presented in the introduction and chapter 2. In the later stages, as theories began to emerge, literature about consumer actions, such as the theory of planned behaviour, was employed to supplement and compare the findings with existing literature. This will be discussed in detail later.

The extent of the usage and analysis of secondary data, as well as the topics and types of sources, varied over time and were adapted based on several factors. First, secondary sources were used sparingly on topics that were intended to be explored through in-depth interviews. Conversely, secondary data was employed more extensively on topics that were not going to be covered to the same degree, such as consumer behaviours, including statistics on the extent of returns in retail, reasons for returns, ethical vs. non-ethical returns, and so on. Second, initially, secondary data was utilized sparingly for qualitative data and more for quantitative data. Third, due to time and resource constraints, secondary data was used to complement data collected from primary sources, enriching and saturating the data further. While it would have been preferable to achieve this primarily through primary sources, there was a trade-off, as secondary sources allowed for a broader range of coverage and facilitated comparisons, among other advantages. An example of this would be delving further into freight solutions and reverse logistics possibilities mentioned in two interviews, researching other companies to compare the different service offerings available without conducting additional interviews with personnel from those companies.

3.1.3 Data analysis

Data preparation

This subsection will cover the data preparation of the primary data collected through in-depth semi-structured interviews. The other sources of data were analysed with no prior data preparation.

The data preparation began with the post-interview stage, as described previously in the interview process subchapter. After rewatching the recording and taking additional notes, the recording was started from the beginning and transcribed using data software that allowed for easy manipulation of the speed and keyboard shortcuts to rewind, change speed, pause, and start the video while simultaneously writing the transcription. The transcription included notes about any relevant non-verbal communication, such as facial expressions, laughter, tone of voice, pauses, etc. The interview was transcribed verbatim, word-for-word, and included timestamps and notes of who spoke when. This process ensured that the full meaning of the interviewees could be analysed, and any relevant contextual information that might impact the meaning of the data was also taken into account (Saunders et al., 2016).

Coding and analysis

The coding and analysis process utilized the qualitative data analysis tool NVivo. The method used for coding was that of Strauss and Corbin, where the data first goes through open coding, followed by axial coding, and finally selective coding (Corbin & Strauss, 2015). These three stages will be elaborated upon below.

The open coding stage consists of reading the transcription line by line and trying to derive meaning from the data to provide labels (Saunders et al., 2016). The result is a large number of codes that are generated in vivo (as opposed to a priori), and they are constantly compared and adjusted if needed.

Following the initial coding stage, the next step in the Grounded Theory Method is axial coding, as outlined by Corbin and Strauss (2015). During this stage, the assigned codes were organized into a hierarchical structure, capturing the emerging relationships between them. To further validate these relationships, a comparative analysis was conducted, involving both existing and newly acquired data. In this regard, informal discussions proved valuable as a primary data source, especially since the number of semi-structured interviews was limited. The aim is to gather evidence that either confirms or disconfirms the hypothesized relationships and makes adjustments where applicable, thereby enhancing the empirical basis of the theory.

Between axial and open coding, adjustments were also made along the way to the codes and the existing coding scheme, as well as the hierarchical structure describing their relationships. Some codes were merged and rewritten, while others were split into two different codes. The process of coding the interviews was done after each interview before the next interview, but also when coding newer interviews, the codes from the previous interviews were also adjusted if needed. This process helped adapt the theory to the data as it emerged and as it was tested and compared iteratively.

The findings after each interview provided the groundwork for what new information was needed to collect in the next interviews to reach theoretical saturation (Flick, 2013). After some interviews, it was seen as necessary to collect more information on the same topic as was already discussed. In others, it was seen as more beneficial to expand the scope of the discussion and include broader or other topics again. This was assessed on a case-by-case basis and, to a certain degree, also steered by which interviewees were available to interview and what data was available. Sometimes secondary data sources were used to supplement the primary data from the interviewees, as mentioned earlier with the example of the logistics interviews.

Finally, in the last stage of selective coding, the major categories were integrated and connected to form a grounded theory. This process consisted of making it all come together. It was done in several stages. First, key central categories were selected, and these categories were connected to the other major categories from the axial coding, incorporating the codes from the upper levels of the hierarchy from the axial coding. The main central categories and their sub-branches were also chosen to incorporate key findings from the qualitative research phase, even if these were less obvious from the coding. Although difficult, this process was carried out by reading memos, reviewing the interviews and data sources, reviewing the codes, tutor brainstorming, drafting and drawing with pen and paper, and adjusting and editing. This result is presented in Chapter 3.3 Findings.

3.2 Research quality

This section explains the measures taken to ensure the overall quality of the research by evaluating the methods employed for data collection and analysis. In the context of Saunders et al. (2016), research quality is primarily determined by its reliability and validity. Reliability focuses on the replicability of study outcomes if conducted by a different researcher. Validity, on the other hand, encompasses the assessment of construct validity, which examines the measurement of the intended variable; internal validity, which evaluates causal relationships in the research; and external validity, which considers the generalizability of the results (Saunders et al., 2016). However, it is worth noting there is a split here between interpretivist and positivist researchers. Positivist researchers frequently employ reliability and validity as criteria to assess the quality of their own and others' research, while interpretivists may choose to either adapt these terms to evaluate their research or dismiss them as unsuitable (Guba & Lincoln, 1989, 2005; Lincoln et al., 2011). Consequently, qualitative research often aims to establish trustworthiness by employing the criteria of dependability, credibility, confirmability, and transferability (Lincoln & Guba, 1985;

Sinkovics et al., 2008). These trustworthiness concepts bear similarities to validity and reliability, but they are better suited for the grounded theory and qualitative approach employed in this part of the study. Therefore, in the context of qualitative research, Lincoln and Guba (1985) proposed alternative terms that will be used, using dependability for reliability, transferability for external validity, and credibility for internal validity. Given the qualitative nature of this research, Lincoln and Guba's definitions are considered more appropriate for assessing research quality and will be utilized accordingly. In addition to the aforementioned criteria, it is important to address how biases that could impact the quality of the research were handled. The way these four were tackled will be presented in the following parts.

3.2.1 Dependability

Dependability, parallel to the positivist concept of reliability, requires some adaptation in its interpretation to cover the traditional notion of repeatability in research. This adaptation is necessary because qualitative data collection, such as through semi-structured interviews and other sources, involves a flexible approach that necessitates researcher interaction and sensitivity to the context and the topic. Consequently, the value of qualitative research lies in the uniqueness of its findings, rather than perfect reproducibility by others (Lincoln & Guba, 1985; Saunders et al., 2016). Since data collection and analysis can be context-dependent, and grounded theory methodology allows for a fluid approach, it inherently reduces the potential for strict repeatability.

However, the study and research should still strive for dependability by ensuring that the processes are well-documented and transparent, providing clarity on how the research was conducted and how the findings were identified or generated. Achieving dependability in this research stage involved documenting every phase, sharing the process with my tutor, as well as other tutors and students involved in DIG and RaCE, and utilizing NVivo software for comprehensive documentation of data collection and analysis. The whole process is thoroughly explained in Chapter 3.1.

3.2.2 Biases

The concerns of dependability are also related to biases, which can affect the research process and findings, thus decreasing the dependability or reliability needed to generate findings that are reproducible. Especially interviewer bias, interviewee bias, and participation bias were attempted to be mitigated through several measures. Interviewer bias can occur if the interviewer behaves in a way that leads to biased responses from the interviewee. Interviewee bias can occur due to the intrusive nature of semi-structured interviews, where participants may avoid revealing information they find uncomfortable. These biases were mitigated in a number of ways, namely by improving rapport, enhancing the credibility of the interviewer, clarifying questions, building trust, and striving to reach an accurate understanding of the subject. Finally, participation bias can also happen as a selection bias, where those who agree to be interviewed have characteristics that make them different from those who do not, thus impacting the findings. This was attempted to be mitigated by using informal discussions to supplement the findings from individuals who may not typically agree to a full semi-structured interview. The full extent of the measures against these biases is elaborated in the different subsections of Chapter 3.2, both previously and below.

Other factors that could impact the findings by creating biases will be presented below. Firstly, to minimize the risk of interviewers' appearance having an influence on the interviewee's answers, neutral and professional clothing was worn during the interviews. Additionally, a blurred background effect was applied to the camera, and the interviews were conducted in a silent area without background noises or disturbances, with good natural light facing the interviewer. Next, each interview began with some informal chatting to build rapport before transitioning to a professional approach with a clear presentation of the research study. This was done to establish credibility, explain why the interviewee was chosen, and ensure that their data and participation were treated professionally and privately, all of which helped build trust and reduce uncertainty for the interviewee.

Furthermore, topics and themes for the interviews were prepared before each interview and sent to the participants. This promoted transferability, credibility, and reliability by providing the interviewee with the opportunity to prepare in advance (Saunders et al., 2016). During the interview, the questions were open and clearly phrased, with minimal jargon and theories. Leading questions were avoided, and attentive listening skills with notes on contextual data were used to ensure that the participant's full meaning could be understood, and that they felt assured that their message was getting through.

3.2.3 Transferability

By providing a comprehensive description of the research questions, design, context, findings, and interpretations, others are enabled to design similar projects for different research settings. Transferability, in this sense, focuses on the transfer of research design rather than statistical generalizations. It is important to acknowledge that qualitative studies using semi-structured interviews cannot be used to make statistical generalizations about an entire population when data are derived from a small non-probability sample (Saunders et al., 2016).

3.2.4 Credibility

The participants in this study were given the opportunity to review and amend the transcriptions of their interviews if they felt that any discrepancies existed between what they intended to convey and what was transcribed, or if any errors were identified. Additionally, during the interviews, an active listening approach was employed, along with the use of clarifying and probing questions to ensure a comprehensive understanding of the participants' true intentions and enhance the credibility of the study (Saunders et al., 2016). Furthermore, establishing trust and rapport with the interviewees was crucial in obtaining credible and frank responses. To foster this, participants were provided with ample information prior to the interviews and were treated professionally and respectfully. Their time and privacy were respected, and the research purpose and process were shared with them in a transparent manner.

The reflexive and inquiring nature of grounded theory methodology also helps to increase the credibility of the findings, as all data are used to form the emerging theory, and thus any negative cases or abnormalities are tested for and evaluated throughout the process of continuous comparison and analysis.

To strengthen the findings, data were collected from various sources throughout the research phase, and each of the topics explored in the semi-structured interviews was covered by two different people. Additionally, informal discussions were used as a supplementary primary data source. Thus, the research process and preliminary findings were discussed with a diverse range of individuals, including professors, students, professionals, and laymen, obtaining valuable input and enhancing the robustness of the study.

3.3 Findings:

3.3.1 Theory generation and discussion

After concluding the data collection and continuous analysis and comparison of the findings, several key findings regarding the theory emerged. These will be presented below.

One of the first findings is that consumers seem to be oblivious to the issue. Most of them are not aware of what happens when clothes are returned. They do not know how long the return journey is or the number of resources needed to get an item returned and back in stock somewhere available for a new potential buyer. Specifically, they are not aware of the financial cost involved, nor the physical and human resources needed. Nor are they aware of the environmental impact of returning clothing that must travel long distances on its return journey, or that sometimes clothing is destroyed as it is cheaper than sending it through the whole long process needed to make it available for a new buyer. This is quite interesting, since a parallel finding is that many consumers say they care about the environment, and eco-friendly behaviour from businesses is expected more and more by consumers today. This is also expected by businesses in shipping and logistics, where there is a large pressure from retailers on their freight partners to reduce emissions and find more environmentally friendly ways of transporting goods, both outbound and reverse.

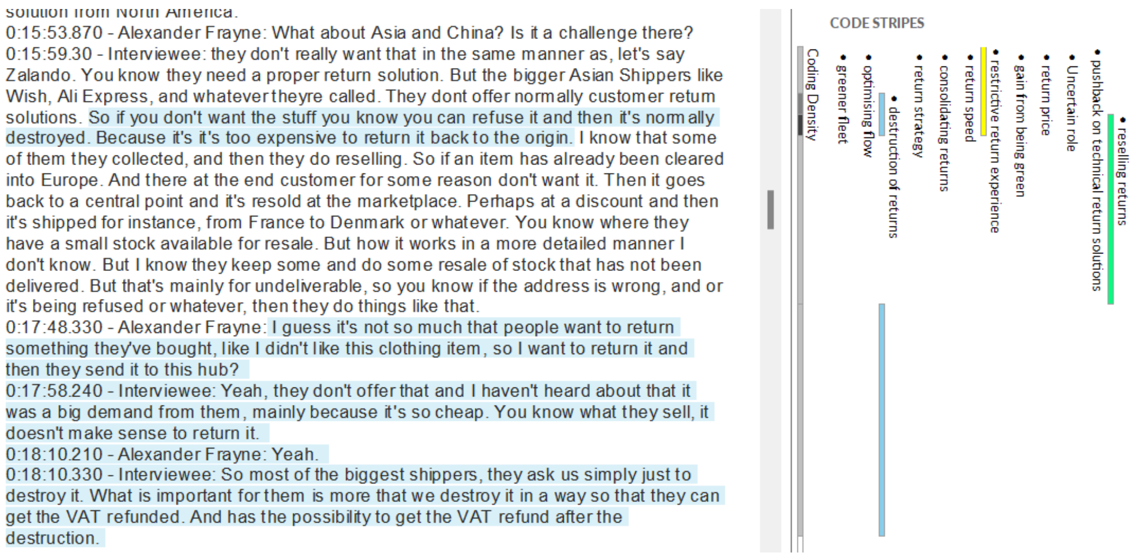


Figure 3-1: Example of open coding in NVivo[1]

A second finding is that retailers are not using technology to the extent it can be used to mitigate the problem. There is a wide array of various technologies that can help consumers find the correct item and, thus, reduce the need for excessive ordering. Yet, these technologies are only used by a small fraction of retailers, and when they are used, it's to a limited degree.

A potential reason for this is that, as long as retailers offer free shipping and free returns, they are actively encouraging their customers to order excessively. Retailers may hope that shoppers will buy or keep more than they need, or that this service offering will help them obtain more loyal and recurring customers. Retailers might expect a decreasing return rate over time as these recurring and loyal customers become more accustomed to their sizes in future purchases. Therefore, the extra cost of returning goods is considered a trade-off for gaining future loyal customers who may return less over time or become regular shoppers. As long as retailers promote these policies as selling points, it could be seen as somewhat counterintuitive for them to invest in technology that would encourage customers to add fewer items to their shopping cart, despite good intentions. Such technologies could help reduce bracketing, but might unintentionally reduce order sizes and make customers less inclined to order several different styles to see which fits best (and potentially keep more than they originally planned).

[1] Where the code "destruction of returns" is highlighted

So, are retailers avoiding using technology due to their lenient return policies? Evidence to the contrary is that there are more and more retailers who do not offer their customers completely free rein when it comes to orders and returns, and many of these do not use much technology either. They either do not use it at all or use basic and minimal-effort measures. From this, one could conclude that technology is not the limiting factor in helping consumers order more accurately and reduce return rates. However, there are other possibilities that technology is the issue. Proof of this can be found in the many cases where technology could help consumers get the item right, but very few retailers are using it. This could be due to the price for retailers to start using these technologies (such as body scanners, digital twins of people and clothing, and the use of virtual reality or the metaverse). Or it could be due to the uncertainty about how much an investment in such technology would lead to consumers with high return rates actually using it effectively to reduce their excessive ordering and returning.

However, there are players who are using these technologies, as mentioned in Chapter 2, for instance, H&M. This is somewhat evidence that the technology is affordable enough to be utilized. Nevertheless, one might think that others do not use them due to being smaller or having other characteristics than H&M (or other retailers who use this technology). However, H&M has many comparable competitors in size, budget, and revenue, who are not using even close to the same amount of technology. This suggests that the technology and its price does not seem to be the problem or the limiting factor. Rather, it is the fact that the majority of retailers do not have a sufficient incentive to deal with the problem at hand.

If retailers under the status quo do not have the incentive, this leads us to the question of whether consumers can be trusted to influence the retailers and create a large enough financial incentive for retailers to address the issue and offer better solutions. Alternatively, an outside force such as regulations may need to come into play to make retailers take responsibility for this issue, which negatively impacts the climate. If this were to happen, an interesting side note here is where consumer rights come into play. How do regulators balance consumer protection and rights in online shopping, such as the right to test and return goods, while also possibly limiting retailers from giving consumers so many rights that they abuse this freedom by ordering and returning clothing excessively?

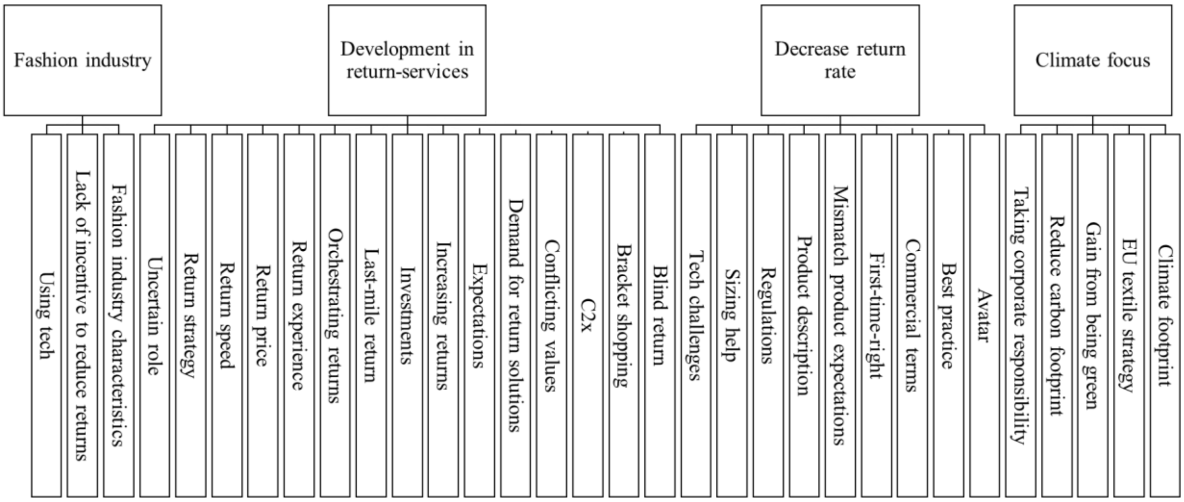


Figure 3-2: Result of axial coding: Final hierarchical structure and relationship of level 1 and level 2 codes[2]

A third finding is that reverse logistics could likely be more environmentally friendly if returns were handled at a physical local store, where they can be pre-processed to a certain degree, sorted, and eligible items could then be returned in a consolidated shipment backward up the supply chain. This would reduce the impact of the return in several ways. Goods would not travel long distances just to be destroyed. They could be pre-sorted at an early stage, saving facility space and shipping space at later stages of the reverse logistics journey. It also prevents shipping items that shouldn't be shipped. Furthermore, it allows products to potentially be sold again locally. It is also a possibility for stores to help consumers find the right item in-store, instead of the customer ordering more products online that may also need to be returned. There is evidence of several retailers offering incentives to return goods in stores, such as free returns (for those who otherwise charge processing fees) (McKinsey, 2022; Narvar, 2022).

A fourth key finding regarding reverse logistics is the complexity in the return process. This makes it difficult and costly for retailers to handle themselves, especially retailers who have customers in geographical areas far from those areas where they might have their own facilities for processing returns (for instance, Asian retailers selling to European consumers). But the case is the same for small retailers (regardless of geographical distance to consum-

[2] See appendix for level 3 and level 4 codes of this coding tree

ers) or anyone else who lacks local processing facilities. The result of this is a plethora of third-party logistics companies who specialize in returns. Such specialized logistics companies handle different parts of the reverse logistics chain, and as there are a vast number of retailers with different needs, the exact specializations of these logistics' companies vary almost as much as the various outbound logistics of each retailer. There seems to be a large potential to improve these reverse processes, as these emerging companies are evidence of, as well as the number of returns and complexity continues increasing. Many retailers are using such companies, which should, in theory, help reduce the impact of each return on profits and emissions. However, the return rate in itself would not be influenced by this, as such retailers typically only improve the efficiency of the returns, and not the rate of returns as such. There is a significant opportunity for reducing the impact from retailers who haven't fully optimized their return processes. This potential can be unlocked by partnering with a specialized company solely focused on this task.

Furthermore, traditional logistics providers and freight forwarders are also working on being able to better handle reverse logistics to counter competition from smaller, specialized companies. This brings up the next key finding, related to the previous. The demand for improved ***reverse*** logistics is only a small fraction compared to the demand for improved ***outbound*** logistics (such as last-mile delivery and improved outbound shipping solutions). Again, this comes back to the demand from the retailers, who are in turn influenced by their customers, government, investors, and society. This influence to improve reverse logistics is evidently too small, and as previously stated, there seems to be a need for more involvement and pressure from consumers to influence a change. If not, government and policymakers may be needed to force the retailers to address the situation.

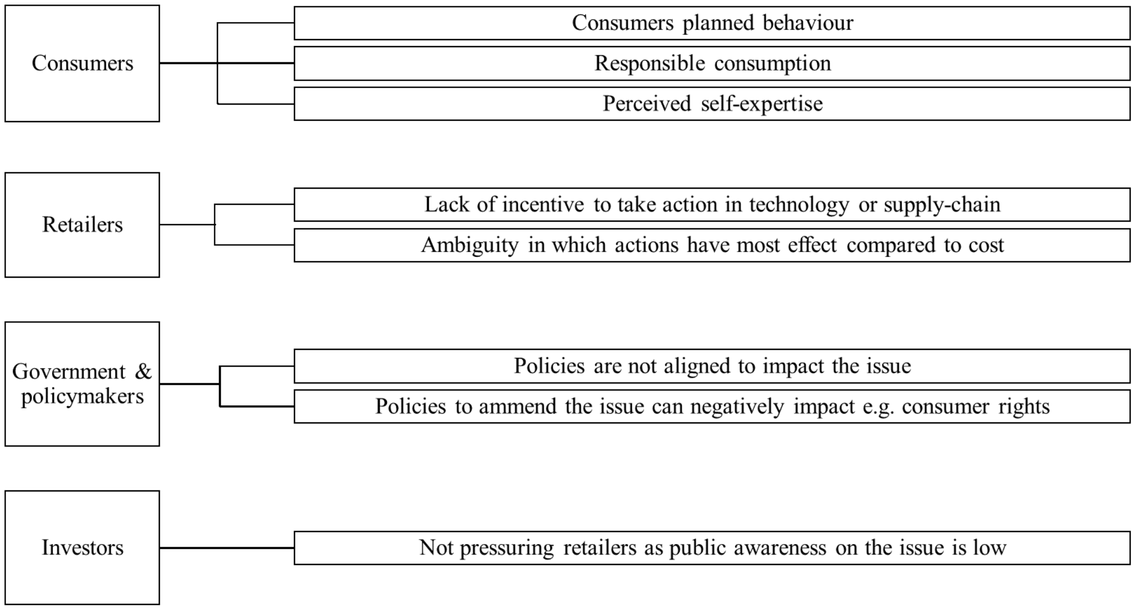


Figure 3-3: Result of selected coding.

3.3.2 A preliminary theory

The return rate can be reduced by applying technology and reverse logistics to a larger degree, as retailers are currently doing this to a small extent. Presumably, they do not have the incentives or pressure to do so, or because the cost outweighs the benefits. We can infer that this is the financial cost, and that the environmental cost is not fully accounted for in their cost calculation, as this cost is borne by the public, and not directly by the retailers. This is akin to the tragedy of the commons (Hardin, 1968). Thus, the retailers do not have sufficient incentives to use technology and smarter reverse logistics, which could help mitigate the impact of returning clothing.

This leads one to examine the consumers. Either they **do not care** and do not have the incentives to behave more responsibly, or they **do care** but are not aware of the problem or the extent of it, or they are not able to pressure the retailers to improve. To influence the retailers to act, these questions must be answered. This will be a part of the base theory: that the return rate is dependent on the consumers' planned behaviour, their interest in the environment, and having socially responsible consumption. Once we understand these variables and how they are connected, we can determine what can be done to make consumers care enough to influence the retailers, and if this is possible at all. Additionally, we need to assess if this change in consumer behaviour will be enough to compel retailers to implement the necessary changes. Alternatively, if neither the actions of retailers nor consumers can be influenced enough to make responsible choices for the planet, it may necessitate government or policymakers' interference to ensure responsible production and consumption of goods by consumers and retailers.

3.3.3 The next steps

Based on the findings presented above, the focus area for the next steps will be quantitative research on consumers, as they are key players in the system who can influence the retailers to take action. It is necessary to examine if consumers are responsible enough to demand a change and what, in turn, is needed to influence them to change their behaviour and demand more and better from the retailers. This is essential to understand to determine whether actions towards consumers will be sufficient, what factors influence consumers' returns the most, or if policymakers must take action towards consumers and retailers.

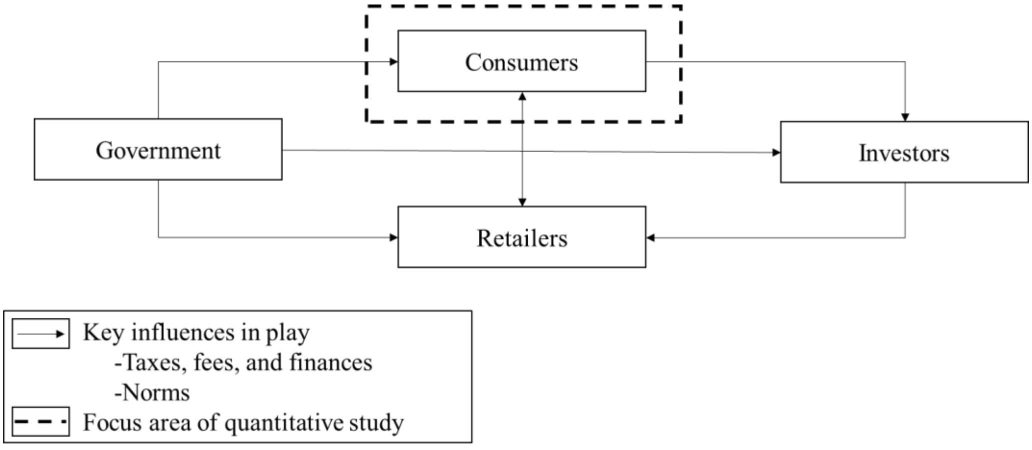


Figure 3-4: A model of the main parties influencing return rates

In the next phase of the research, consumers' planned behaviour, their interest in the environment and being responsible consumers, as well as their knowledge and interest in the subject, will be expanded upon and researched quantitatively. This will help understand the importance of each factor, the relationships between the variables, and their influence on the return rate.

3.3.4 Limitations

There are limitations in this phase of the research, particularly in the number of interviewees. Since only four people were interviewed, there is a considerable chance that additional interviews could have been useful to achieve a higher level of theoretical saturation (Corbin & Strauss, 2015). The decision not to continue with more interviews was made due to limitations in time, resources, and the availability of additional interviewees. With more time, additional interviews could have added further value to the grounded theory. As mentioned earlier, there were substantial difficulties in recruiting participants for interviews, despite significant efforts.

4. Questionnaire

This chapter will follow the outline below: It begins in Chapter 4.1 by building on the theory presented at the end of Chapter 3, focusing on consumers' planned behaviour, their interest in the environment, their responsibility as consumers, and their perceived knowledge of the subject. Furthermore, this section will elaborate on how these factors are broken down and their meanings, as well as thoughts on their relationships. Initially, theory related to consumer behaviour will be presented, followed by an explanation of how elements of this theory will be integrated into the proposed research model. Subsequently, the other variables from the results of the quantitative research will be covered. Finally, a research model will be presented using these variables, illustrating the suggested relationships between them. In the following section, hypotheses will be stated. Moving on to Chapter 4.2, the methodology for researching this quantitatively with a questionnaire will be explained in detail. In 4.3, the analyses and steps taken to ensure high-quality research will be elaborated upon. Finally, in 4.4, the findings will be presented.

4.1 Theory development, research model, and hypotheses

4.1.1 Generation from Grounded Theory and The Theory of Planned Behaviour

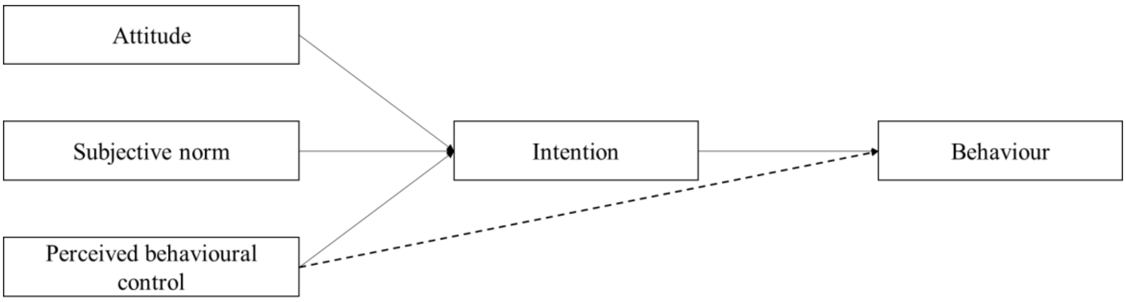


Figure 4-1: Theory of Planned Behaviour (Ajzen, 1985).

Behavioural intention is commonly predicted using multiattribute models, particularly with the Theory of Planned Behaviour (TPB: Ajzen, 1985, 1991). This theory is an expanded version of the Theory of Reasoned Action (TRA: Fishbein & Ajzen, 1975). The TPB model has been used to predict many different behaviours, including consumer actions (Armitage & Conner, 2001; Venkatesh et al., 2003). The model postulates that an individual’s planned behaviours can be explained through their subjective norms, intentions, attitude, and their perceived behavioural control. These three factors are thought to directly influence an individual's intentions, which in turn influence their behaviour (Ajzen, 1991). This theory will form the basis for theory development regarding the behaviours surrounding consumers ordering more items than they intend to keep. The three variables of this theory will be covered one by one, starting with subjective norms.

In short, a person's subjective norms can be described as social pressure to adhere and conform to certain behaviours. Subjective norm is included in the model, as several studies have shown it to have an influence on an individual's behavioural intentions and their own personal norms (Ajzen, 1991; Armitage & Conner, 2001; Moser, 2015).

The subjective norm can be split into two categories: the subjective descriptive and the subjective injunctive. Since subjective norm has been found to be a weak determinant of intention (Armitage & Conner, 2001), it is therefore seen as necessary to figure out if there is a difference between injunctive and descriptive subjective norm. Both injunctive and descriptive norms are related to an individual's beliefs about others' norms, specifically other people whom they look up to, admire, respect, or care about the opinion of. These people can be role models, family, friends, peers, or public figures they admire. Who exactly such a person is, will vary for each individual, but the key is that the individual in question cares about the opinion of this other person or persons. To simplify going forward, such a person will be referred to as a normative influencer from now on, for the sake of having a simple term that describes this person, as it could be a great many different people, persons, or groups, for each individual.

The difference between the subjective descriptive and subjective injunctive norms is as follows: The subjective descriptive norm describes the norms an individual believes their normative influencers have. For example, they may believe their normative influencers care about recycling rubbish and that throwing rubbish in nature is not acceptable. Since an individual can have many normative influencers, the norms of each normative influencer may vary, as will the beliefs an individual has about each normative influencer's norms.

On the other hand, the subjective injunctive norm describes what norms an individual thinks their normative influencer expects of them (the individual). For example, an individual may believe that their normative influencer(s) expect them to recycle plastic bottles. This belief is the subjective injunctive norm.

To fully illustrate the difference between the subjective descriptive and the subjective injunctive, an example will be presented. Person A has a normative influencer who will be called Person B. Person A believes that Person B cares a lot about the environment and that their normative influencer (Person B) therefore recycles their rubbish, drives an electric vehicle, and avoids disposable cups at their local coffee shop. These beliefs may or may not be true, but they are what Person A believes about Person B. These beliefs that Person A has about Person B are Person A's subjective descriptive norms. Person A also believes that Person B expects Person A to also recycle their waste. Whether these beliefs are true or not is not the question; what Person A believes that Person B expects Person A to adhere to are Person A's subjective injunctive norms. The subjective injunctive, together with the subjective descriptive, make up a person's subjective norms. In addition to subjective norm, an individual's behaviours are also influenced by their attitude. This is defined by Fishbein and Ajzen (1975:216) as "an individual's positive or negative feelings about performing the target behaviour". It can also be seen as the advantages an individual sees with certain types of behaviour.

Furthermore, an individual's behaviours are also influenced by their intention. This is defined by Fishbein and Ajzen (1975:288) as "the strength of one's intention to perform a specific behaviour." This is found to be a strong determinant of an individual's actual behaviours, and in a meta-study by Armitage and Conner (2001), it was found to explain 39% of the variation in intention.

Finally, their behaviour is affected by their perceived behavioural control. Ajzen (1991:183) defines this as "people's perception of the ease or difficulty of performing the behaviour of interest," and in this regard, it can be seen as comparable to self-efficacy.

Adapting The Theory of Planned Behaviour to suit the context

Other studies regarding environmentally friendly behaviour have shown that TPB does not always provide the best fit. For example, numerous studies have found a disconnect between an individual's attitudes toward environmentally friendly behaviour and their actual intentions (Gupta & Ogden, 2009; Kollmuss & Agyeman, 2002; Pickett-Baker & Ozaki, 2008). Many people have positive attitudes toward environmentally friendly behaviour, but intention does not always translate into a change in behaviours (Norberg et al., 2007). This can be explained, for instance, by cognitive dissonance. The disconnect in how people believe they should act and how they act is handled differently by different people (Kollmuss & Agyeman, 2002). So the inclusion of intention in TPB is in the hope that an individual’s intention to behave a certain way is strongly correlated with and can predict how they actually behave (Ajzen, 1985). Therefore, TPB is often used to predict behaviours that cannot be measured directly (Armitage & Conner, 2001; Venkatesh et al., 2003). Instead, one looks at the antecedents and an individual’s intentions. In this case, instead of looking at the intention of actual behaviour, it will be studied indirectly by rather looking at reported past behaviour.

As the behaviour in question is returning clothes bought online, and the previous stages of research found that consumers, in general, knew little about this, this has implications for which variables could be measured and used in the study. For consumers, there is a disconnect and ambiguity regarding whether returning clothes bought online is related to environmentally friendly behaviour and to what degree. It would make little sense to measure the attitude consumers have toward an activity they are hardly conscious of. For this reason, consumers' attitude toward returns was omitted from the research model, as asking respondents about their attitude toward returning clothes bought online would likely produce non-meaningful answers and a higher rate of random responses. Instead, a proposed variable of socially responsible consumption is suggested, which will be elaborated on later.

Perceived behavioural control was not included in the model for the same reasons, despite it being a part of TPB. The reason is similar to the one mentioned above: it would not make much sense to measure how much control someone perceives they have over a behaviour where they are not aware they are exerting control, as it is done automatically. Most people order what they want and return what they don't want, and asking how they perceive they have control over these actions would likely not produce meaningful answers. This is based on the previously mentioned research and the qualitative stage.

It's also worth mentioning about adapting the model, as per Ajzen's own words on TPB, that in some instances, for example, subjective norm alone may be sufficient to explain a behaviour, in others, it may only be attitude and perceived behavioural control, and in others, it may involve all three (Ajzen & Driver, 1992). He says “Note that the predictors in the theory of planned behaviour are assumed to be sufficient to account for intentions and actions, but that they are not always necessary in any given application. The relative importance […] is assumed to vary across behaviours and populations” (p.209-210).

4.1.2 Expanding the model

As the model has now been stripped down, with three variables removed, it will now be built upon in an attempt to enhance the explanation of return behaviour through the motivation and behaviours of consumers. This will be done by adding other variables based on other relevant theories and the findings from the qualitative part of this study, along with an explanation of the rationale behind the hypotheses.

Personal norms – Objective descriptive

Personal norms refer to an individual's sense of moral duty or obligation to engage in a particular behaviour (Ajzen, 1991). Although not explicitly included in TPB, Ajzen (1991) suggests that personal norms operate in a similar manner to subjective norms and directly influence one's intention to act, regardless of whether this is due to internal or external pressures. Generally, consumers' behaviours can be influenced by their own feelings of ethical obligation or moral duty (Gorsuch & Ortberg, 1983; Schwartz, 1973, 1977).

Personal norms are built upon the moral norm-activation theory of altruism by Schwartz (1973, 1977). According to this theory, individuals who are aware of negative consequences associated with certain actions and believe that their actions can help prevent those consequences will feel a sense of responsibility and moral obligation to engage in altruistic behaviour (Stern, 2000). In this particular context, personal norms can be defined as the individual consumer's internalized perceptions of duty or moral compulsion to engage in or abstain from a particular behaviour (Ajzen, 1991). In this case, it is excessive returning. In other words, it can help us confirm if people indeed lack awareness of how their returning behaviours may have negative consequences.

Ajzen additionally posited that personal moral obligations are anticipated to exert an influence on intention, in conjunction with the other variables within TPB. In the model, personal norms will be labelled as objective descriptive norms to differentiate them from and highlight the parallels to subjective norms.

Although excessive ordering and a high volume of returns, both in relative and absolute numbers, can be viewed as negative behaviour by some, it remains unclear how much consumers know about it. Therefore, measuring the respondents' personal norms is particularly relevant in this study. This measure is based on a person's personal norms concerning ordering and returning clothes bought online. It is therefore hypothesized that personal norms have a positive influence on sent returns and physical returns.

H_1: Objective descriptive norms have a negative influence on sent returns.

H_2: Objective descriptive norms have a negative influence on physical returns.

Subjective norms – Descriptive and injunctive

As described earlier, subjective norms have been showed in numerous studies by Ajzen and others (Ajzen & Driver, 1992; Ajzen et al., 2011; De Leeuw et al., 2015), to strongly influence behaviours and intentions. Subjective norms can be divided into subjective descriptive norms and subjective injunctive norms.

Subjective descriptive norms

The subjective descriptive norms are the norms a person believes another person has, where said other person is someone whom they look up to, respect, or admire. For instance, this can be that a person's beliefs about their parents are that their parents care about recycling product packaging. These beliefs are particularly related to pro-environmentally friendly behaviour, as social circles and peer pressure can often influence how one acts (De Leeuw et al., 2015; Kollmuss & Agyeman, 2002).

H_3: Subjective descriptive norms have a negative influence on sent returns.

H_4: Subjective descriptive norms have a negative influence on physical returns.

Subjective injunctive norms

Subjective injunctive norms pertain to the perceived norms individuals believe others hold towards them, especially individuals they look up to, respect, or admire. For example, a person may believe that their parents expect them to prioritize recycling product packaging. These expectations from others can potentially influence a person's own behaviours, which is why they are considered in this study.

H_5: Subjective injunctive norms have a negative influence on sent returns.

H_6: Subjective injunctive norms have a negative influence on physical returns.

Self-reported knowledge

An individual's knowledge about the topic has been found to impact behaviours (Amoako et al., 2020; Dumitrescu et al., 2011; Guerin & Toland, 2020) in some studies, while in others, it does not (Ajzen et al., 2011). This variable is particularly relevant to returns, as the qualitative phase revealed many indications of low knowledge and awareness among consumers regarding the topic. An individual's knowledge and interest in the topic are considered important for their behaviour, especially concerning the environmental aspect of this behaviour. Examining this variable will help determine the extent to which a person's knowledge and interest influence their return behaviour and measure how much knowledge they believe they possess on the subject. In addition, environmental knowledge is often an intervention as it is assumed to produce more environmentally responsible behaviour (Duerden & Witt, 2010). Therefore, it is relevant to examine its relationship with behaviours and compare it with, for example, normative influences.

This variable can also be viewed as capturing a portion of Ajzen's Attitude variable since one's interest and knowledge are likely to influence their attitude on a subject. Furthermore, based on the findings from the qualitative stage, it is likely that having more knowledge about the issue will impact returning behaviours. For these reasons, it is hypothesized that self-reported knowledge negatively influences sent returns and physical returns.

H_7: Self-reported knowledge has a negative influence on sent returns.

H_8: Self-reported knowledge has a negative influence on physical returns.

Socially responsible consumption

Webb et al. (2008) uses the definition provided by Mohr et al. (2001), which states that socially responsible consumption refers to how much a person bases "...his or her acquisition, usage, and disposition of products on a desire to minimize or eliminate any harmful effects and maximize the long-run beneficial impact on society" (p. 47). A similar measure is employed in current unpublished research by Andreassen, that assesses the extent to which a consumer engages in socially responsible consumption of goods with regard to environmental aspects.

Typically, socially responsible consumption encompasses three dimensions: CSR performance, consumer recycling behaviour, and avoidance/reduction of environmental impact from purchase and use (Prendergast & Tsang, 2019; Webb et al., 2008). These dimensions have been adapted and simplified to measure consumers' behaviours in relation to online shopping and returns for the purposes of this research. The inclusion of these dimensions provides an alternative way to measure consumers' general attitudes towards socially responsible consumption of clothing. In this context, it can be viewed as a means of incorporating the "attitude" element from TPB but under a different name, more accurately reflecting the attitudes being assessed.

Socially responsible consumption is divided into two categories: buying and returning, and product. Starting with buying and returning, it is likely that the degree to which a person practices socially responsible consumption in their purchasing and returning behaviours will also influence their actual behaviours concerning returning clothes bought online and in-store. Therefore, it is hypothesized that socially responsible consumption will have a direct impact on sent returns and physical returns.

H_9: Socially responsible consumption – buying and returning has a negative influence on sent returns.

H_{12}: Socially responsible consumption – buying and returning has a negative influence on physical returns.

The second part of socially responsible consumption is related to products that are consumed. Based on the qualitative stage, a person’s behaviours regarding the products they purchase, are likely to influence the return behaviour they have, due to the environmentally impact of returns.

H_{11}: Socially responsible consumption – products has a negative influence on sent returns.

H_{12}: Socially responsible consumption – products has a negative influence on physical returns.

Self-reported behaviour

As intention is removed from Ajzen’s model, self-reported behaviour is added as a replacement to assess how the other variables influence actual past behaviour (reportedly). Self-reported behaviour has also been used by Ajzen & Driver (1992) to measure actual behaviour in TPB-studies. In this study participants first estimated how much they expected to perform certain behaviour over the next 6 months, and later were asked 12 months later how much they estimated they had performed them over the past 12 months. For this reason, 6 months was used as a timescale for self-reported behaviour. Self-reported behaviour was collected on two areas, return volume and shopping volume. Shopping volume is a possible confounding variable for return volume, and therefore return volume must be adjusted for or take shopping volume into account. Return volume is further divided into two sub-variables, sent returns, and physical returns.

Shopping volume

This measured the extent to which a person had previously ordered clothing online. Respondents were asked to estimate their volume over the last six months, both in terms of the number of orders and the number of items ordered. One could argue that shopping volume indirectly influences returns; however, it is entirely possible for someone to have a high shopping volume and make no returns, or vice versa. Additionally, including it in the model allows us to determine its contribution to explaining the variation in return volume. Higher shopping volume means more opportunities for items to not meet the buyer's expectations. Therefore, the hypothesis is that shopping volume has a positive influence on sent returns and physical returns.

H_{13}: Shop volume has a positive influence on sent returns.

H_{14}: Shop volume has a positive influence on physical returns.

Return volume

This measures the portion of a person's shopping volume that they have returned. It's also an estimate covering the last six months and is divided into two measures: the amount returned to a physical store and the amount sent back by mail. This division was based on data from the pre-study phase, which indicated that returning clothes to a store where they can be processed, potentially bundled and sent back to a central warehouse, or sold in the store, may be more environmentally friendly. Respondents were asked about this to explore whether there was any correlation between the amount a person returned to a physical store and their norms, behaviours, and attitudes. It is believed that there could be significant differences in the factors influencing return volume among those with high and low return volumes. For example, those with higher scores on socially responsible consumption might have more in-store returns, which have a lower environmental impact.

Our independent variables are how we categorize the respondents. Therefore, we use normative influences, socially responsible consumption, and self-reported knowledge as our independent variables. We want to determine how these variables affect our dependent variable: self-reported behaviour (return volume).

4.1.3 Research model

Based on the theory development and hypotheses, the research model is presented below in figure 4-1.

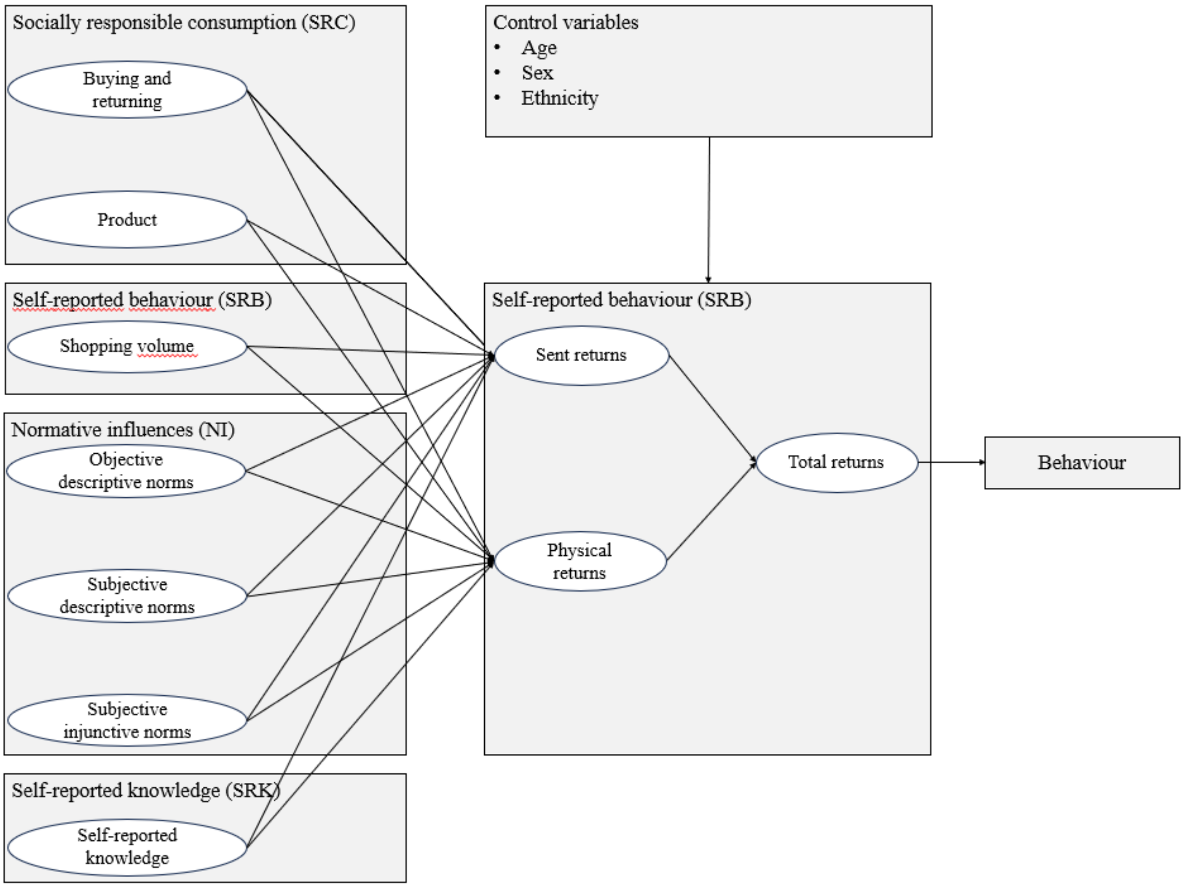


Figure 4-2: Research model

The parts of the model to the right of sent returns and physical returns, will not be assessed.

4.2 Method

4.2.1 Approach

The primary purpose of the questionnaire was to collect data to measure the variables, test the relationships in the theory, and confirm or disprove the hypotheses. Most of the variables were assessed using a Likert scale ranging from one to seven, where respondents could indicate their agreement level, from "totally disagree" to "totally agree." The questionnaire's structure organized items for each construct concurrently. Although this approach can potentially increase within-measure correlational systematic error, it typically reduces systematic error across measures (Podsakoff et al., 2003). Additionally, it minimizes the risk of confusing respondents (Viswanathan & Kayande, 2012).

For these reasons, the structure commenced with introductory questions about shopping and return volume. Subsequently, it presented items for each construct logically grouped and ordered to minimize respondent confusion. To further mitigate within-measure correlational systematic error, two response instability checks were implemented as an alternative mechanism (Van der Veld & Saris, 2004). Furthermore, specific variables were grouped into low and high categories to examine differences among groups of respondents in their respective variables, where applicable. Matrix structures were employed for items using the Likert scale, allowing users to view the scale without the need to scroll up and down. Although Dillman et al. (2011) have raised concerns that this approach may be too advanced for participants, it was considered less relevant, as the respondents are typically experienced in completing questionnaires. Consequently, prioritizing user-friendliness and layout were deemed higher priorities, as it was less likely that respondents would find this feature too advanced.

4.2.2 Data collection

To test the hypotheses, an online questionnaire was conducted among US consumers in the period of 22. March 2023 and 29. March 2023. The questionnaire was distributed by Prolific, a market research firm, who in turn collected answers from a pool of 120 000 users, where a sample was recruited from a pool of 40 000+ active US users of Prolific’s questionnaires in the past 12months (Prolific, 2022). The respondents were offered a reward of 0.4 £ for completing the 4-minute questionnaire. Based on the budget available to pay the participants and the Prolific fee, the questionnaire lies on Prolific's site until the maximum number of participants is reached, given the budget. This produced 1391 respondents.

Data sources and sample

A wide range of filters can be added to pre-screen which participants can answer the questionnaire once it is published on Prolific. Three prescreeners were applied to narrow down the sample, based on questions Prolific members have answered about themselves. The following prescreeners were used:

1. "What is your nationality?"
2. "What is your date of birth?"
3. "How often (on average) do you shop online?"

For the first question, the study only included participants who had selected "United States" as their response. For the second question, the study only considered participants aged between 18 and 100 years old. Only the participants' age was visible; their date of birth was not disclosed. For the third question, participants were presented with several response options: "Don't know/Rather not say," "Never," "Once in a few months or longer," "About once a month," "Several times a month," "About once per week," and "More than once per week." Those who had chosen "Never" or "Don't know/rather not say" were excluded from participating in the questionnaire. While this exclusion may potentially reduce the generalizability of the study, it was done to increase the number of relevant participants who actually shop online. Additionally, the Prolific software indicated that excluding participants who chose these two options resulted in only a minor reduction of around 120 possible respondents from a total pool of 120,000 users.

As a result of applying these three filters, the study had a final pool of 17,913 eligible participants out of the 121,615 members on Prolific (Prolific, 2022).

USA was chosen as a sample because it shares some comparable similarities with Norwegian consumers. It was not possible to collect data from a large enough and diverse sample of Norwegian consumers due to budget constraints. While the resemblance is not perfect, it was considered the best trade-off between collecting a convenience sample of Norwegian consumers and gathering a larger and more representative sample of US consumers.

4.2.3 Research design

Measures and creation of questionnaire items

The research model contains nine constructs. The complete questionnaire is presented in its entirety in Appendix 8.3.2, while the individual items and a simplified overview are provided in Table 4-1 below. The questionnaire items were designed based on Ajzen (2006), and other TBP questionnaires, such as Gkargkavouzi (2019), de Leeuw et al. (2015), and more. Please refer to Appendix 4 for a comprehensive list of sources for each item and measurement. Additionally, since this thesis is a collaborative effort with DIG and BEST in Retail, parts of the questionnaire also incorporated items and constructs from unpublished research conducted in those organizations. These items were further grounded in the qualitative stage and relevant literature mentioned earlier.

Measure	**Questionnaire items**
SRB – Shop volume	"How many times have you ordered clothes online? (i.e., the number of orders in the last 6 months, to the best of your recollection)" "How many items of clothing have you ordered online? (in the last 6 months, to the best of your recollection)"
SRB – Sent returns	"How many items of clothing ordered online have you returned by sending back? (in the last 6 months, to the best of your recollection)"
SRB – Physical returns	"How many items of clothing ordered online have you returned by handing back to a physical store? (in the last 6 months, to the best of your recollection)"
SRC – Shopping and returns	"When online returns are easy, I shop a little extra" "I often shop a little extra online to get free shipping" "Overall, I find returning online goods to be easy and hassle free"

	"If online webshops offer free returns I will order more than I need" "If online webshops require payment from me to return goods, for example 5 USD, it will make me think more carefully about what I buy"
SRC – Product	"I prefer to shop for products that can last a while" "I often check where the product comes from before I buy it" "I often check how the product is manufactured before I shop" "I am careful about the correct handling of product packaging when I recycle"
Normative influences – Objective descriptive norms	"I feel a moral obligation to only order what I intend to keep, when ordering clothes online" "I do not feel a moral obligation to only order what I intend to keep, when ordering clothes online. " "I feel guilty ordering clothing online if I know I will be returning some or all of it"
Normative influences – Subjective descriptive norms	"Most likely, people who are important to me will only order clothing online they intend to keep." "Most likely, people who are important to me will not only order clothing online they intend to keep" "Most likely, people who I respect and look up to, will order clothing online, knowing that they will return some/all of it"
Normative influences – Subjective injunctive norms	"Most likely, people who are important to me will recommend me (for environmental purposes) to only order clothing I intend to keep" "I think people who are important to me expect me to not order

	clothing online if I know that I will be returning some/all of it."
Self-reported knowledge	"Ordering more items of clothing online than one intends to keep, with the intention of returning some/all, has a negative effect on the environment" "I spend a lot of time reading about sustainability in general" "Based on my previous online shopping, I have gained considerable insight into sustainability" "Carbon emissions cause serious environmental problems, such as climate change" "In general, I consider myself more interested in sustainability than the average online shopper" "Overall, I act in an environmentally responsible way"

Table 4-1: Measures and items for questionnaire

Below each measure and item will be elaborated upon.

Shopping volume

This is one of the independent variables, and consisted of two items where the respondent estimates their shopping volume in the past six months. 6 months was used as a period as it was found to be the best balance between having a long enough time period that irregularities between months would be somewhat smoothed out, and a short enough time period that respondents are able to accurately enough estimate the correct amount. In addition this time frame has been used in other TPB-studies before for reporting past behaviour (Ajzen & Driver, 1992).

Return volume

This measure can be further split into two sub-measures: sent returns and physical returns, where the latter are returns handed back to a physical store and the former are returns sent and shipped back. Each sub-measure consists of one item. For both items the respondent was asked to estimate the number of items returned in the last six months. Physical returns were added as a result of the preceding qualitative phase, to add nuance to total return volume, and find differences in total return volume as well as test the theory that consumers who score higher on socially responsible consumption also have more physical returns. These items were made based on the theory generation and qualitative stage.

Socially responsible consumption – Shopping and returns

As the topic is new and no questionnaires about consumer consumption in relation to returns were found, items for this measure were sourced from alternative places. The items were found and used from current research projects by BEST in Retail by Andreassen (personal communication, email 16.01.2023). This measure was to understand the current behaviour of the respondents in regard to shopping and returns items bought online, in a more general sense than the questions about the actual volume.

Socially responsible consumption – Product

Four items were used for this measure. These items also came from current research projects by BEST in Retail by Andreassen et al., the measures are similar as those used by Collado et al. (2019), Gkargkavouzi et al. (2019), and Webb et al. (2008). These items were used to measure the attitudes surrounding purchasing of products when it comes to socially responsible consumption. Although like the previous measure, this measure focuses on the product, while the previous measure focuses on the act of shopping or returning.

Personal norms – Objective descriptive

This measure consists of two items based on research by Petschnig et al. (2014), Jansson (2011), Gkargkavouzi (2019), Ateş (2020), Harland et al. (1999), Abrahamse and Steg (2009), van der Werff & Steg (2015, 2016). A lot has already been said about objective descriptive norms (i.e., personal norms), to summarize shortly, the items are to measure what moral obligations a person feels when it comes to ordering clothing online, knowing some of it will be returned.

Norms – Subjective norms: descriptive and injunctive

This measure consists of two sub-measures, the first is subjective descriptive, while the second is subjective injunctive. First is based on de Leeuw (2015), Ajzen (2006), Gkargkavouzi et al. (2019), Gao et al. (2017), Albayrak et al. (2013), and Hong & Tam (2006) while second is based on items from research by Gkargkavouzi (2019), Gao et al. (2017), Hong & Tam (2006), Ajzen (2006), and Mathieson (1991). Both are as mentioned in the theory generation section thought to impact behaviours, in the sense that people one looks up to or admire, or one's social circle, will influence one's own actions and behaviours. These two measures measure firstly the personal norms the respondent believes others have, and secondly what norms the respondent thinks others except the respondent to have or follow.

Self-reported knowledge

This consisted of six items, adapted based on research by van der Werff & Steg (2015, 2016), Coyle (2005), Arcury and Jonhson (1987), Gkargkavouzi (2019) and Collado et al (2019). As mentioned previously, an individual's self-reported knowledge on the topic and on sustainability in general is thought to influence their behaviours in this area, and is measured by these items.

4.2.4 Research quality

Reliability and reducing the risk of errors and biases

When designing the questionnaire, various factors were considered to make sure that the results are trustworthy. In particular, the possibility of systematic errors and biases, such as participant bias, observer bias, and survey errors, were carefully considered, as they can compromise the reliability of the results (Saunders et al., 2016). The following sections will cover how the questionnaire design was done to mitigate these risks and ensure high data quality. This is particularly important as respondents were paid to answer the survey, and there was considerable risks of random responses and other non-sensical answers in an attempt to earn the reward with minimum effort and time spent answering or thinking.

Participant error

Participant error can occur if the respondents misunderstand the questions, or if they do not understand the issue and cannot relate to the topic and see no personal relevance. This was addressed by adding an information text before the participants could join the questionnaire, and another information text in the start after joining. The second text was followed by a comprehension check question to ensure participants had read and comprehended the information before continuing. Any participants who failed this check were discarded from the sample[3]. In addition to these two tests, the questions were reformulated for clarity after feedback from two rounds of pilot tests.

[3] See instructional manipulation check 2 in appendix 8.3 – Data cleaning details

Participant errors can occur if respondents misunderstand the questions or fail to comprehend the issue, finding it unrelated and lacking personal relevance. To address this, an information text was added before participants could access the questionnaire, along with another information text at the beginning after joining. The second text was followed by a comprehension check question to ensure participants had read and understood the information before proceeding. Any participants who failed this check were excluded from the sample (see instructional manipulation check 2 under 2.1 Data collection and sample). Additionally, the questions were reformulated for clarity based on feedback from two rounds of pilot tests.

Participant bias

Participant bias can take various forms, including social desirability bias, recall bias, acquiescence bias, central tendency bias, random response bias, item order effect bias, mood bias, and demand characteristic bias. Each of these will be discussed below.

Social desirability bias is a tendency for respondents to adjust their answers in favour of what they perceive as socially acceptable (Maccoby & Maccoby, 1954). This can occur subconsciously. To mitigate this potential bias, two steps were taken. First, participants were assured of their anonymity in both information texts. Second, they were explicitly told that there were no right or wrong answers and that the study was solely interested in their personal opinions, encouraging them to respond honestly to the best of their ability.

Acquiescence bias involves respondents generally and subconsciously agreeing with the statements they are asked about, potentially skewing results to be more positive than they truly are. While reducing this bias is challenging, awareness of its presence is important (Lydeard, 1991). The use of reversed scale questions was employed to assess acquiescence bias, and the results showed a moderately low balance between agree/disagree for such questions. However, it is important to note that measuring acquiescence bias can be difficult, and there may be some level of uncertainty associated with it.

Central tendency bias is characterized by respondents consistently rating items around the midpoints of scale questions. This tendency can manifest when respondents are uncertain about how to rate or become fatigued or disinterested during the questionnaire, leading them to avoid extreme ends of the scale. To mitigate this bias, two strategies were employed. First, the questionnaire was kept as concise as possible, with elements like demographic questions removed. This reduction aimed to prevent respondents from becoming disengaged due to questionnaire length, as well as to avoid purpose creep.

Secondly, the questionnaire underwent three rounds of pilot testing to assess statement clarity and relevance. While some statements were initially found to be challenging, adjustments were made to enhance clarity. Additionally, the testing evaluated whether respondents generally avoided the scale's lower or higher ends.

Random response bias can also stem from factors related to central tendency bias, with respondents sometimes selecting responses randomly. This risk is especially pertinent in paid surveys where respondents may be incentivized to complete them quickly. To address this concern, two measures were implemented.

First, attention checks, as mentioned in Appendix 8.3 Data cleaning details, were included to motivate respondents to carefully read and respond to the statements. It was made clear that respondents would not receive a reward if they failed these checks, encouraging them to engage sincerely rather than select responses randomly.

However, there still exists a risk that some respondents may only read the questions to identify attention checks and subsequently choose responses randomly. To reduce this risk, the second measure was taken, which involved optimizing statement clarity, comprehensibility, and relatability. This aimed to make it effortless for respondents to respond sincerely rather than randomly.

Item order effect bias can arise when respondents' answers are influenced by the sequence of items in the questionnaire. This bias can be reduced by randomizing the order of statements for each respondent. However, this was not implemented as it is generally considered to require more effort and mental capacity from respondents, potentially increasing the likelihood of random responses (Saunders et al., 2016). Consequently, the order of items was not randomized, introducing a potential risk of item order bias.

The questionnaire's item ordering was thoughtfully chosen to minimize biases. Initially, questions solicited quantitative behaviours followed by personal norms. This sequencing aimed to encourage respondents to provide honest answers without feeling judged or overly influenced by social expectations. Subsequently, questions about respondents' perceptions of others' norms and expectations were presented. These were placed later to prevent respondents from first contemplating external expectations, which could have heightened social desirability bias. Finally, questions about self-perceived knowledge were positioned last, as they were perceived as easy to answer and less likely to be influenced by preceding questions. While the risk of previous questions colouring the responses here was considered low (though not zero), this order was deemed the most reasonable based on the available trade-offs.

Mood bias can occur when a participant's emotional state influences the responses they provide. The ordering of questions, particularly those related to how respondents believe others perceive them, can impact their mood. To address this, the questionnaire was thoughtfully structured as previously explained. However, fully assessing and mitigating the extent of mood bias is challenging, as external factors beyond the questionnaire's control may also influence respondents' moods. These external factors can include confusion, boredom, irritation, and enthusiasm, all of which were minimized to the greatest extent possible through the previously mentioned measures.

Demand characteristics bias can manifest in various ways. One way is if the participants are aware of the research's intended use and consequently tailor their responses to serve their own interests. For example, they might alter their answers if they believe that policy changes may result from the recommendations derived from the survey responses. To mitigate this bias, participants were not informed about the specific purposes for which the research results would be utilized, apart from being told it was for a master's thesis.

In summary, among the participant biases considered in this study, social desirability bias, acquiescence bias, and random response bias were deemed the most critical.

Researcher error and bias

Observer error was minimized by directly importing data into SPSS to eliminate the potential for errors resulting from manual data entry. Observer bias, which can arise when answers are subjectively interpreted based on open-ended questions, was reduced by employing an online questionnaire with closed-ended questions.

Survey error, including various forms such as survey scope error, purpose creep error, sampling error, response rate refusal, and item nonresponse error, was also taken into account. The first two were addressed through careful questionnaire design based on the theoretical model and predefined constructs. Purpose creep, in the form of unnecessary demographic questions, was avoided since all demographic data about the respondents was provided by Prolific, eliminating the need for additional demographic inquiries in the questionnaire. This approach helped keep the questionnaire concise and reduce respondent fatigue.

Regarding the response rate, it wasn't a concern as the questionnaire was distributed through Prolific, allowing anyone in the sample who wished to participate to do so. A small number of respondents opened the questionnaire but did not complete it within the time limit[4], and these responses were discarded (n = 8). Calculating the response rate was deemed meaningless because the questionnaire remained open until it reached the maximum number of respondents within the budget. However, the possibility of nonresponse bias exists, especially if the 8 non-completers and the 121 excluded from the total pool differed significantly from the 1391 who completed the questionnaire (Armstrong & Overton, 1977). Additionally, those who are not on Prolific may differ from those who are, potentially introducing bias. Mitigating or calculating nonresponse bias in online questionnaires is challenging, but it is considered an acceptable trade-off for the convenience and other advantages of online data collection methods.

In summary, the wording and design of the questionnaire were carefully crafted to minimize the impact of biases and errors on the collected data.

Validity

For a questionnaire, content validity, construct validity, and external validity are the most relevant (Saunders et al., 2016), and will be elaborated below.

Content validity refers to the extent to which the questionnaire adequately covers the entire range of content or topics it is meant to assess. It ensures that the items or questions within the tool are relevant and representative of the construct or concept being measured (Saunders et al., 2016). The model has been designed to include all the variables needed to explain the dependent variable, but since the model is created based on various other theoretical sources and the qualitative phase, there is a chance that the independent variables do not sufficiently explain the dependent variable or that the model is incorrectly formulated. Great efforts have been made to prevent this, such as the extensive qualitative stage, as well

[4] 36 minutes, as set på Prolific

as the other underlying theories and literature used in generating the theory and research model to be tested. These steps are hopefully sufficient to ensure an adequate model and variables.

Construct validity is particularly important when developing and using questionnaires. It refers to the extent to which the questionnaire accurately assesses the underlying theoretical construct or concept it is intended to measure (Saunders et al., 2016). To maintain construct validity, several steps were taken. Firstly, the measurements and items measuring the underlying latent variables of the model are based on previous research and their measurements. Secondly, the questionnaire underwent a pilot test before modification and then another pilot test. Finally, the author's supervisor also contributed in sparring sessions on the design of the questionnaire and provided feedback over several iterations.

External validity concerns the generalizability of research findings to a broader population or different settings beyond the specific conditions of the study (Saunders et al., 2016). In other words, it assesses whether the results can be applied to real-world situations and other groups or contexts. Although the sample collected is close to a representative sample of the US population, there is a risk that the participants are not representative since they are all individuals who complete online surveys in exchange for money, and thus share a commonality not shared with those in the population who do not do this. However, the size of the sample and the range of different demographics included should help increase the generalizability of the results.

4.3 Data analysis

4.3.1 Data preparations and data cleaning

The survey data was exported from Qualtrics, cleaned in Excel and analysed in SPSS and STATA. When offering rewards for questionnaire responses, there is a risk of receiving careless responses. To reduce the risk of obtaining careless responses, several efforts were taken. These efforts were used to remove respondents who likely responded carelessly, not to their best effort, unreasonably fast, or misunderstood the questions. Among these steps were three instructional manipulation checks[5] and two questions checking for response instability[6]. In total n = 419 respondents were removed, leaving n = 904 in the final sample. Below is a summary of the step taken, for a full explanation and justification for each, see section "Data cleaning details" in Appendix Chapter 8.3.

What is checked	**Number failed**	**% of full answers (n=1323)**
Sum failed at least one check	**N=419**	**31.67%**
Instructional manipulation check 1	N=73	5.52%
Instructional manipulation check 2	N=2	0.15%
Instructional manipulation check 3	N=13	0.98%
Illogical shopping or return volume 1	N=41	3.10%
Illogical shopping or return volume 2	N=15	1.13%
Response instability check 1	N=118	8.92%
Response instability check 2	N=275	20.79%
Respondent speed	N=64	4.84%

Table 4-2: Summary of all data quality measures

After cleaning the data in excel, the dataset was analysed using the software SPSS and STATA for structural equation modelling (SEM).

The demographics of the remaining respondents was close to representative for the U.S population in general (U.S Census, 2020), as seen below. The left column shows the quanti-

[5] Where respondents are asked to press a certain answer to prove they are reading the instructions and questions

[6] Where respondents answer significantly different on two questions asking the same or exact opposite thing as an earlier question

ty for a perfectly representative sample with the given sample size. The right column shows the quantities for the sample achieved. The are two considerable differences. The first are the differences in the sample compared to a representative sample is the higher proportion of people aged 25-44, and lower proportion aged 60 and above. This difference decreases the generalizability of the data, however online shopping is less common for those aged 60 years and above (McKinsey, 2022), which helps mitigate this discrepancy. The second is the difference in ethnicity, where a representative sample would have a higher proportion of black respondents, and lower proportion of white respondents. In absolute numbers this discrepancy is not very large, only 5% or approximately n = 54, however this signifies the black respondents are almost half the number of what they should have been to be representative.

		Representative sample N (%)	**Sample used N (%)**
Sex	Female	459 (51%)	450 (50%)
	Male	445 (49%)	454 (50%)
Age	18-24	127 (14%)	94 (10%)
	25-34	158 (17%)	235 (26%)
	35-44	146 (16%)	223 (25%)
	45-54	155 (17%)	147 (16%)
	55-59	77 (9%)	85 (9%)
	60-150	241 (27%)	120 (13%)
Ethnicity	Asian	49 (5%)	59 (7%)
	Black	115 (13%)	61 (7%)
	Mixed	28 (3%)	45 (5%)
	Other	52 (6%)	25 (3%)
	White	661 (73%)	713 (79%)

Table 4-3: Sample demographics vs a representative sample

The next step was to test the assumptions required to perform statistical analyses. This was followed by first transforming data if required to satisfy the assumptions needed. Finally, analyses were performed to assess the previously presented research model and developed conceptual framework. These steps will be presented below, where step two and three (testing assumptions and collecting results) will be presented consecutively for each of the dependent variables.

4.3.2 Assumptions and tenatative results for multivariate analyses

To conduct analyses of multivariate models, certain assumption must be satisfied. These assumptions encompass conditions such as normality, homoscedasticity, linearity, and the absence of autocorrelation and multicollinearity, as outlined by (Black et al., 2010; Field, 2013) Also, there mustn't be significant unusual points, such as outliers, leverage points, or highly influential points (Fox, 2015).

As there are two dependent variables in the model, sent returns and physical returns, each must satisfy the assumptions along with their relationship with the independent variables. The following section will first handle the variable sent returns and its relationship to the independent variables presented in the model. Later, physical returns will be assessed. All results of the regressions will report the standardized beta-coefficient, of each of the independent variables to assess, compare and rank the amount of variance explained by each of the independent variables.

Part 1 – Sent returns

Run 1 – Testing of assumptions

First, shipped returns was plotted against each independent variable in a scatterplot to visually assess their relationship. The scatterplots against each individual independent variable all showed relationships with no signs of non-linearity, approximately equal variance across the range of values (homoscedasticity), and no significant unusual points. There was one exception, where it was plotted against items purchased, here there were slight tendencies of heteroscedasticity, where the variance is possibly increasing along the x-axis. To quantitatively check for homoscedasticity, a Breusch pagan test was conducted in STATA. This produced a p-value of >0.001, rejecting the null-hypothesis of homoscedasticity, confirming the presence of heteroscedasticity (Astivia & Zumbo, 2019).

Next, the relationship between the dependent variable and the independent variables collectively was assessed by plotting the studentized residuals against the unstandardized predicted values, see Appendix 8.3. This plot showed no signs of non-linearity, it did however have clear heteroscedasticity with increasing variance, and several significant unusual points. A histogram of the standardized residuals was also produced. Although it had a mean of approximately 0 (-9.71E-17) and a standard deviation of close to 1 ($\sigma = 0.994$), indicating normality, it showed a non-normal distribution with positive kurtosis (a high peak). A Normal P-P plot also showed a non-normal distribution of the standardized residuals, as did a Normal Q-Q plot.

Next, the data was tested for unusual points, this included 3 tests. The data was tested for significant outliers, leverage points, and highly influential points. First tested was outliers, where the number of cases where the studentized residuals were ± 3 standard deviations. This is SPSS's default test for diagnosing outliers, and it produced n = 19 outliers. Fox (2015) argues that assessing studentized deleted residuals is a better way of finding outliers, by looking at cases where the standard deviation is ± 3, so this was also done, and also produced n = 19 outliers. Second, were diagnosing for any leverage points. According to Huber (1981), leverage values of 0.5 and above are dangerous, and values between 0.2 and 0.5 are risky. The highest value found was 0.100, thus no cases were identified as leverage points. Thirdly, highly influential points were searched for my assessing Cook's distance for each case. Here, all cases except n = 1 was above the recommended threshold of 1 (Fox, 2015), this case overlapped with previously identified outliers. Hence, 19 cases in total were identified as unusual points.

Following this, data was tested for independence of residuals, where the Durbin-Watson test for first order autocorrelation was utilized. This produced a score of 1.890, where scores between 1 and 3 are satisfactory (Field, 2013).

Finally, multicollinearity was checked with two assessments. First the variance inflation factors (VIF) were all well within the acceptable range of < 10 (Hair, 2009; Hammervold, 2020) (VIF = 1.047 – 1.987). Next, the inter-construct correlations were assessed, where all values were below the recommended level of 0.7 (Field, 2013). These results mean that no two variables are so closely correlated that the individual effect contribution from each of them would be too difficult to distinguish apart (Fox, 2015).

To summarize, the assumptions satisfied were that of no non-linear relationships, independence of residuals (no issues with autocorrelation), and no issues with multicollinearity. The assumptions not satisfied, ere unusual points, normality of residuals, and homoscedasticity. What remains of options then is to proceed with the analyses in spite of these shortcomings, or to transform the data in hopes to satisfy the assumptions (Fox, 2015). There is empirical evidence that the assumption of normality is less important for large samples (n > 200) as discussed by e.g. (Black et al., 2010). As the sample size here is n = 904, it could possibly be ignored, as it mainly suffers from positive kurtosis, and minimal skewness. That would then leave the assumptions of no unusual points and homoscedasticity the only issues. These can possible be managed by transforming of the data (Fox, 2015). Therefore, it was decided to both do the analyses without transforming the data, and to transform the data then do the analyses, before finally comparing the results of both to assess what findings are more likely trustworthy and robust.

Run 1 – Analysis 1

Starting with the analyses without transforming the data, the model was plotted into STATA to use its structural equation modelling software package. After creating the model in STATA, the analyses failed to provide any results due to not satisfying the assumptions or the model not fitting the data. Structural equation modelling is less flexible to data not conforming to the underlying assumptions (Osborne, 2012; Ullman & Bentler, 2012, p.165), so this does indicate that the issues described above may have a significant severity. Data could be transformed to satisfy the assumptions, then run through SEM in STATA, but this would mean there is no “baseline” to compare the results after transforming the data with the findings from the analyses of the untransformed data. Due to this, the analyses following will be to proceed with multiple linear regression, instead of SEM.

Linear regression was done in SPSS, measuring the relationship between the dependent sent returns (untransformed) and the independent variables (socially responsible consumption, normative influences, and self-reported knowledge. In addition, control variables included shopping volume, age, gender, sex, and ethnicity. The overall model fit gave an adjusted R^2 score of 0.462. Weak but significant relationships were found for the control variables age ($\beta = -0.08$, $p < 0.001$) and sex ($\beta = 0.079$, $p < 0.001$). Worth noting is the lower and upper bound of the 95% confidence interval for age was -0.037 and -0.003, implying a very weak relationship. A strong relationship was found with shopping volume ($\beta = 0.653$, $p < 0.001$). And medium strength negative relationships for Norms – Objective descriptive ($\beta = -0.113$, $p < 0.001$), and medium strength positive relationship with Socially responsible consumption - Shopping and returns ($\beta = 0.133$, $p < 0.001$).

Run 2 – Testing of assumptions

For the second run, the data were transformed following guidelines from Fox (2015). This comprised of filtering the unusual points, to see if this would sufficiently satisfy the assumption of homoscedasticity and normality of the residuals. The n = 19 unusual points identified in run 1 were filtered out, leading to a sample size of n = 885. The assumptions were re-assessed following the exact same procedure as run 1. The results of the assumption testing were almost identical to run 1, with the exception of the following: There were identified a total of n = 20 outliers. The same plots that previously showed heteroscedasticity and non-normality still exemplified these issues, but to a lesser extent than run 1. The rest of the tests and assumptions had the same findings. Furthermore, the linear regression found the same independent variables to have significant relationships with the dependent variable proved almost the same results as run 1. The overall model had an R^2 score of 0.444. Socially responsible consumption – Shopping and returns had a positive relationship (β = 0.207, $p < 0.001$) Norms – Objective descriptive had a negative and significant relationship (β = -0.117, $p < 0.001$). Shopping volume had a positive and significant (β = 0.554, $p < 0.001$). In addition, Norms – Subjective descriptive have a small negative relationship (β =-0.075, $p = 0.022$, 95% CI [-0.207, -0.016]), however as the confidence level almost crosses zero, this relationship is almost negligible.

As the results were very similar to run 1, the data were transformed further, by filtering out the n = 20 additional unusual cases found in run 2. This equals to 4.31% of the sample size. The results after this transformation will be considered the findings of run 2.

The testing of assumptions was done again and led to similar findings as previously. The assumptions met previously were still satisfied. Furthermore, the scatterplots that previously showed heteroscedasticity now exhibited it to a far lesser degree, almost not at all. The data did however still provide a significant test result in the Breusch Pagan test, implying some heteroscedasticity. The plots assessing normality now showed the residuals having an approximate normal distribution. In short, based on the plots, all the assumptions are met. However, the Breuch-Pagan test still had a $p < 0.001$ implying some heteroscedasticity.

Run 2 – Results

The final results for run 2, where n = 39 unusual cases were filtered out, were again similar to previous findings. The overall model had an R^2 score of 0.446. Socially responsible consumption – Shopping and returns had a positive relationship (β = 0.212, $p < 0.001$) Norms –

Objective descriptive had a negative and significant relationship (β = -0.115, $p < 0.001$). Shopping volume had a positive and significant relationship (β = 0.607, $p < 0.001$). In addition to these, results also found relationships with age and Norms – Subjective descriptive, these were however very close to the confidence interval and almost negligible. Age had a negative relationship (β = -0.121, p < 0.001, 95% CI [-0.020, -0.006]). In addition, Norms – Subjective descriptive had a small negative relationship (β =-0.065, $p = 0.022$, 95% CI [-0.144, -0.001]).

Run 3 – Testing of assumptions

For the third run, data was this time transformed following a different procedure from Fox (2015). After some trial and error, the dependent variable sent returns was transformed using a square root function, to deal with the positive kurtosis and increasing homoscedasticity. The square root was taken for all the values, and then used as the dependent variable for the rest of the analyses in this third run.

First, all previously mentioned scatterplots were created, they all showed no signs of non-linearity, approximate equal variance (homoscedasticity), and very few unusual points. The plot for the independent variables collectively plotted against the dependent variable, showed small signs of heteroscedasticity, this did appear to possibly be a consequence of using a Likert scale on many of the independent variables. A histogram of the residuals and normal P-P and Q-Q plots showed a normal distribution.

The various tests for unusual points found n = 12 significant outliers, no leverage points, and no highly influential point. The Durbin-Watson test produced a score of 0.633, showing independence of residuals and no issues with autocorrelation. The assumption of multicollinearity was also satisfied with acceptable VIF values (VIF = 1.041 – 1.987), and all inter-construct correlations under < 0.6, well within the acceptable range.

In short, all the assumptions were satisfied, with the possible exception of homoscedasticity, but this seemed to be borderline. A Breusch-Pagan test was therefore conducted, which confirmed the hypothesis of heteroscedasticity ($p < 0.001$).

Run 3 – Results

The final results for run 3, where no cases were filtered out, but the dependent variable was transformed with a square root function, were again similar to previous findings. The overall model had an R^2 score of 0.461. Socially responsible consumption – Shopping and returns had a positive relationship ($\beta = 0.267$, $p < 0.001$) Norms – Objective descriptive had a negative and significant relationship ($\beta = -0.119$, $p < 0.001$). Shopping volume had a positive and significant relationship ($\beta = 0.594$, $p < 0.001$). Sex has a positive relationship ($\beta = 0.085$, $p = 0.011$). In addition to these, results also found relationships with Age and Norms – Subjective descriptive, these were however very close to the confidence interval and almost negligible. Age had a negative relationship ($\beta = -0.092$, $p = 0.007$, 95% CI [-0.011, -0.002]). In addition, Norms – Subjective descriptive had a small negative relationship ($\beta = -0.066$, $p = 0.039$, 95% CI [-0.089, -0.002]).

Part 2 – Physical returns

This part will assess the dependent variable Physical returns against the independent variables in the previously mentioned model, which are the same independent variables as were tested against sent returns.

Run 4 – Testing of assumptions

First, all previously mentioned scatterplots and graphs were created, they all showed no signs of non-linearity, approximate equal variance (homoscedasticity), and very few unusual points. The plot for the independent variables collectively plotted against the dependent variable, showed small signs of heteroscedasticity, this did appear to possibly be a consequence of using a Likert scale on many of the independent variables. A histogram of the residuals and normal P-P and Q-Q plots showed a normal distribution.

The various tests for unusual points found n = 16 outliers, no leverage points and no highly influential point. The Durbin-Watson test produced a score of 2.003, showing independence of residuals and no issues with autocorrelation. The assumption of multicollinearity was also satisfied with acceptable VIF values (VIF = 1.041 – 1.987), and all inter-construct correlations under < 0.6, well within the acceptable range.

In short, all the assumptions were satisfied, with the possible exception of homoscedasticity, but this seemed to be borderline. A Breusch-Pagan test was therefore conducted, which confirmed the hypothesis of heteroscedasticity ($p < 0.001$).

Run 4 – Results

The final results for run 4, are based on the dependent variable physical returns. Some trial and error was done in attempting to transform the variable or filter cases, but this did not show any improvement or change, and was therefore not done or kept as the final data for run 4. Run 4 is therefore the original data for physical returns. The overall model had an R^2 score of 0.175. Shopping volume had a positive and significant relationship (β = 0.381, $p < 0.001$). The remaining variables with significant relationships all had confidence intervals that were close to crossing zero. Norms – Subjective descriptive had a negative relationship (β = -0.111, $p = 0.005$, 95% CI [-0.202, -0.036]). Socially responsible consumption – Shopping and returns had a positive relationship (β = 0.088, $p = 0.007$, 95% CI [0.031, 0.191]). Age had a negative relationship (β = -0.08, $p = 0.02$, 95% CI [-0.015, -0.001].

4.4 Findings and summary of results

This section will first present the findings from the various runs of transforming the data prior to testing and retesting the assumptions and findings. Following this it will present the final findings and support of the hypotheses.

4.4.1 Construct validity

As mentioned, several steps were taken to ensure construct validity. To test this quantitatively, a Chronbach's alpha score was calculated for each of the measures. The threshold recommended for acceptable scores is 0.7 (Saunders et al., 2016). All constructs were above the threshold, except for socially responsible consumption – product, which produced a score of 0.69. As this partially consists of self-constructed items, and is relatively close to the threshold value of 0.7, this is often deemed as an acceptable reasons for items with scores between 0.6 and 0.7 (Thrane, 2018), and it was therefore kept in the study.

Construct	**Alpha**	**Items**
Objective descriptive norms	0.79	2
Subjective descriptive norms	0.76	2
Subjective injunctive norms	0.73	2
Self-reported knowledge	0.83	6
Socially responsible consumption – Product	0.69	4
Socially responsible consumption – Shopping and returns	0.71	5

Table 4-4: Chronbach's alpha for constructs

4.4.2 Descriptive statistics

Table 4-5 below shows the descriptive statistics of the variables from the model. Socially responsible consumption, normative influences and self-reported knowledge all used one to seven Likert scale items, and have means ranging from 4,17 to 4,58. Self-reported behaviour (sent returns, physical returns, items ordered) all have quite diverse ranges, means and positive skewness and kurtosis. This will be commented upon later in the analysis chapter, as will the transformed version of sent returns.

Construct	N	Minimum	Maximum	Mean	Std. Deviation	Skewness	Kurtosis
Sent returns	904	0	50	1,42	3,49	7,195	75,863
Sent returns (transformed)	904	0	7,1	0,71	0,96	1,762	5,084
Physical returns	904	0	20	0,53	1,48	5,940	52,726
Items ordered	904	0	100	9,00	9,95	3,468	18,649
SRC[7], shopping and returns	904	1	7	4,58	1,17	-0,244	-0,277
SRC, product	904	1	7	4,45	1,16	0,049	-0,422
Objective descriptive norms	904	1	7	4,37	1,79	-0,307	-0,946
Subjective descriptive norms	904	1	7	4,51	1,38	-0,048	-0,240
Subjective injunctive norms	904	1	7	4,17	1,44	-0,017	-0,125
Self-reported knowledge	904	1	7	4,38	1,17	-0,175	-0,363

Table 4-5: Descriptive statistics

4.4.3 Testing of hypotheses

This section will first summarize the findings from the 4 runs of transforming the data, testing assumptions, and calculating results. Then, a summary of the hypothesis support will be presented in Table 4-6.

Regression 1

This regression tested the independent variables on the dependent variable sent returns. There were done three runs of varying data transformation, testing of assumptions and linear regression. Each run showed similar results and findings, however run 3 satisfied the assumptions of multiple linear regression best, and for this reason was found to be the most accurate results. Furthermore, the transformation was not seen as too much or skewing the results too much, as the other runs with similar results (and no transformation) had similar findings. In addition to this, the transformation closely follows the procedure of Fox (2016). The full procedure of this testing is previously elaborated in Chapter 4.3.2.

[7] Socially responsible consumption

For the model as a whole, an explanation of 46.1% of the variance was found jointly through the variables, The individual determinants from strongest effect to weakest effect, were the following: shopping volume (β[8] = 0.594, $p < 0.001$), socially responsible consumption – shopping and returns (β = 0.267, $p < 0.001$), objective descriptive norms (β = -0.119, $p < 0.001$). In addition, weak but significant relationships were found for sex (β = 0.085, $p = 0.011$), age (β = -0.092, $p = 0.007$), and subjective descriptive norms (β =-0.066, $p = 0.039$). However, the 95% confidence interval was very close to zero on these three and thus their effects are almost negligible. Thus, support was found for H1 and H13. Although H9 found a strong and significant relationship, the direction was the opposite of the hypothesis, hence there was not found support for it.

Regression 2

For the model as a whole, an explanation of 17.5% of the variance was found jointly through the variables. The individual determinants from strongest effect to weakest effect, were the following: Shopping volume (β = 0.381, $p < 0.001$), subjective descriptive norms (β =-0.111, $p = 0.005$) and socially responsible consumption – shopping and returns (β = 0.088, $p = 0.007$). Objective descriptive norms was close to significant, but the 95% CI crossed zero (β = -0.076, $p < 0.054$, 95% CI [-0.127, 0.001]). In addition, a weak but significant relationship was found for age (β = -0.092, $p = 0.007$). However, the 95% confidence interval was very close to zero on the latter and thus its effects are almost negligible (95% CI [-0.015, -0.001]). Thus, support was found for H4 and H14. Although H10 found a significant relationship, the direction was the opposite of the hypothesis, hence there was not found support for the hypothesis.

[8] All reported beta-coefficients are standardized beta-coefficients unless otherwise stated.

No.	IV	Direction	DV	Standardized β	*p*
H1	Objective descriptive norms	-	Sent returns	-0.119	<0.001
H2	Objective descriptive norms	-	Physical returns	-0.076	0.054
H3	Subjective descriptive norms	-	Sent returns	-0.066	0.039
H4	Subjective descriptive norms	-	Physical returns	-0.111	0.005
H5	Subjective injunctive norms	-	Sent returns	0.011	0.743
H6	Subjective injunctive norms	-	Physical returns	0.023	0.592
H7	Self-reported knowledge	-	Sent returns	-0.012	0.733
H8	Self-reported knowledge	-	Physical returns	0.031	0.464
H9	Socially responsible consumption - Shopping and returns	-	Sent returns	0.267	<0.001
H10	Socially responsible consumption - Shopping and returns	-	Physical returns	0.088	0.007
H11	Socially responsible consumption - Product	-	Sent returns	0.03	0.321
H12	Socially responsible consumption - Product	-	Physical returns	-0.048	0.2
H13	Shopping volume	+	Sent returns	0.594	<0.001
H14	Shopping volume	+	Physical returns	0.381	<0.001

Table 4-6: Summary of hypotheses results

Control variables

Some of the control variables showed significant effects, albeit weak, on the dependent variable, however these effects were close to zero, especially considering their 95% confidence intervals, and are therefore almost negligible. See overview in Table 4-7 below of the variables in question.

				95% CI, unstandardized	
Control variable	**Dependent variable**	**Standardized β**	**p**	**Upper**	**Lower**
Age	Sent returns	-0.092	0.007	-0.011	-0.002
Sex	Sent returns	0.085	0.011	0.038	0.287
Age	Physical returns	-0.08	0.020	-0.015	-0.001

Table 4-7: Control variables

5. Discussion

5.1 Implications of the findings

5.1.1 Part 1 – Grounded theory development

In the first part, four key findings were identified. In short, it was found that consumers have low knowledge about returns, the environmental cost, and the retailer cost. Retailers are not using existing technological tools to the extent they could, presumably due to a mismatch between the cost of implementing such tools and not implementing them. Furthermore, they have conflicting interests, and implementing return rate reduction practices can have negative side effects elsewhere, which they do not want to risk, again presumably due to the reward. Additionally, the complexity in reverse logistics runs deep, and the majority of improvements could be achieved by outsourcing the task to specialized companies. However, this is more related to the impact of each return than the return rate in itself. Finally, it was found that physical returns are likely more environmentally friendly than sending returns back by mail, given a retailer has a physical store with the capacity to process the return, sort and bundle it before sending it further back up the supply chain, and that the consumer doesn’t take a polluting means of transport, which otherwise wouldn’t have been done if they didn’t return (for example, a return while doing a routine drive that would have happened anyway, or taking public transport, or walking or cycling to return it).

All in all, this signifies that although there are measures retailers can take to reduce the financial and environmental impact, there are likely not large enough financial incentives to take the risk of investment due to its potential to reduce sales or have other adverse effects. For this reason, retailers can likely not be trusted to take the steps required to enact the change needed. In addition, due to the game-theory aspect of potentially losing competitive power by going against the rest of the market with measures to reduce and discourage excessive ordering, retailers risk missing out or having other actors undercut the efforts. In other words, central governance would be a possible mechanism, for instance, with cross-border policies, that ensure all retailers align on the issue.

5.1.2 Part 2 - Questionnaire

The model shows varying degree of fit with the data. For sent returns, the model has a medium-high degree of explanation with 46.1% of the variance explained. Compared to other TPB research this is somewhat on par (Armitage & Conner, 2001). For physical returns, the model has a low degree of explanation with only 17.5% of the variance explained. This can indicate that perhaps there is another variable omitted from the model, which influences return behaviour for both sent returns and physical returns, but which is more prominent and has a stronger relationship with physical returns. For instance, it could be related to the perception of effort or cost for returning goods, and encompass aspects like how easy it is for someone to return items compared to just keeping them, how far away they live from a post office or the physical store where they could return the item, what type of item or clothing they bought, the size of the item, what means of transportation they have available, or how many “hoops” a retailer makes them jump through to get a shipping label and approval to refund the item, et cetera.

The main variables significantly influencing sent returns to the largest degree were found to be the number of items a person has ordered, their degree of socially responsible consumption in regard to shopping and return behaviour, and their objective descriptive norms (i.e., personal norms). The main variables significantly influencing physical returns most were found to be the number of items returned and subjective descriptive norms. There were minor influences from some of the other variables too, for both types of returns. Below, the influence of each independent variable will be discussed.

Normative influences

Normative influences encompassed objective descriptive norms, subjective descriptive norms, and subjective injunctive norms. These showed varying degrees of influence on physical and sent returns. Objective descriptive norms were found to have a significant relationship with sent returns, with a standardized beta coefficient of -0.119. Subjective descriptive norms were found to have a weak but statistically significant relationship with sent returns, with a standardized beta of -0.066.

On the contrary, for physical returns, subjective descriptive norms had a significant relationship. The standardized beta coefficient was -0.111. The other norms had no significant relationship. The difference in relationships to sent returns and physical returns is quite interesting, especially the fact that subjective descriptive norms had such a weak relationship with sent returns, while, on the contrary, they were the only significant norms to influence physical returns. Furthermore, the standardized beta coefficient for subjective descriptive on physical returns was almost the exact same as objective descriptive norms were for sent returns (-0.111 vs. -0.119).

These findings could imply that overall, our norms influence the way we act when it comes to returning clothing but do not explain the majority of our behaviour. Also, if we return items physically in a store, our thoughts about others' norms are the main normative influence (i.e., subjective descriptive norms), while, on the contrary, if we return by sending clothing back by mail, our own personal norms have that same impact (i.e., objective descriptive), while the norms we think others expect of us (subjective descriptive) take the back seat and only have a very minor influence, if any. In any case, the subjective injunctive norms (i.e., what we think others expect of us) were not found to be influential in either type of returns.

All this could be an indication of a larger root cause, the fact that shopping online can be done without peers observing one's behaviour, hence leading to their norms being less influential than when returning in a store and having to possibly explain it to or defend it to another human being. When one sends items back in the mail, no one knows if you ordered 20 items just to try them on or post on social media and are sending 19 of them back. While in a store, this behaviour would likely be less socially acceptable, and few people would likely do this regularly.

Socially responsible consumption

The findings on this theme were particularly interesting. Socially responsible consumption was divided into two variables, one for products and one for shopping and returns. Both were hypothesized to have a negative relationship with returns, i.e., consumers who scored higher on socially responsible consumption would have fewer returns. First of all, there was found no significant relationship between the product variable of socially responsible consumption and returns. Second of all, the dimension for shopping and returns had a positive and significant relationship, with a higher coefficient than norms. This is the opposite direction of what was hypothesized.

Interestingly enough, persons with higher degrees of socially responsible consumption (shopping and returns) also seemed to have higher returns as the β-coefficient was positive, not negative as hypothesized. Perhaps there is a belief that it is more socially responsible to order excessive amounts all in one go and then return what is not wanted, rather than ordering one item at a time and possibly generating more returns and also more outbound orders. Again, the findings indicate there may be a lack of knowledge among consumers, as well as causal ambiguity surrounding what behaviour has the lowest impact.

Self-reported knowledge

Self-reported knowledge showed no significant relationship with either sent returns or physical returns. This is interesting, as there are cases where this has been found to be an influential predictor of behaviour previously (Duerden & Witt, 2010). There are several possible explanations for this. There may be a difference between the knowledge people believe they have and report they have, compared to what they actually know about the topic. The questions only tested self-reported knowledge and not actual knowledge. In hindsight, the items could have had a more quiz-like approach to test the actual knowledge of the respondents; however, this approach could also have its issues, as some respondents might possibly search for the correct answer online, making it difficult to assess the quality and authenticity of the answers. Nevertheless, having respondents self-report their knowledge also has its downsides; it is not perfect either.

Another possibility is that the items and scale, although based on the literature, did not have specific enough questions about returns and asked too general questions. Although the questions are valid and backed by the literature, the majority of them were more general about sustainability rather than specifically about returning clothing online. This could lead to the results where there is no clear relationship between self-reported knowledge and returns if the knowledge is too general. This does, however, imply that general knowledge about sustainability and the environment is not a strong enough predictor for return behaviour or sustainable return behaviour. In other words, it indicates that the topic is too narrow, and knowledge about it is not common, even among those who report that they themselves have high knowledge about sustainability.

To conclude regarding self-reported knowledge, the findings show no clear relationship, implying that there possibly is a mismatch between what people think they know and what they actually know, or that the knowledge people have is too general, and this topic is not well-known enough for those with generally good knowledge to be aware of the issue. Lastly, it could also mean that some people do know, while others do not, but this knowledge just is not enough to make people change their behaviour, possibly because they are not held accountable, or they do not care either way.

5.2 Contributions

5.2.1 Theoretical implications

To assess the theoretical contributions, the findings, and implications of these will be compared to the initial research question and the gaps in research identified in Chapter 2.3.

The aim of this thesis was to better understand what can be done to reduce the impact of returning clothing bought online, by examining two aspects: the return rate and the impact of each return, both environmental and on retailers' profits. This resulted in the following three research questions:

RQ1*: What factors, if any, will have the largest impact on reducing consumers' return rate of clothing bought online?*

RQ2*: What factors, if any, will have the largest impact on reducing the carbon footprint of each returned online order?*

RQ3*: What factors, if any, will have the largest impact on reducing the financial cost of each returned online order?*

RQ1 was thoroughly assessed with a questionnaire answered by a large and almost representative[9] sample of U.S consumers, based on a model building on existing theory, and on new findings from qualitative grounded theory. Here, it was found that normative influences from the theory of planned behaviour explained some of the behaviour. Furthermore, different norms explained sent returns compared to physical returns, but to approximately the same degree. For sent returns, personal norms were the best normative influence to explain

[9] With the exception of the sample having a slightly higher proportion respondents aged 25-44 and fewer aged 60+, as well as a slightly higher proportion with a white ethnicity, and fewer with a black ethnicity, compared to a representative sample, see more details in section 4.3.1.

the behaviour among the norms. While for physical returns, it was a person's subjective understanding of others' norms that had this effect. Besides this, the results found that knowledge on the subject may be low, incorrect, or otherwise unrelated to behaviour, as no relationship was found. However, socially responsible consumers regarding shopping and returns surprisingly had a higher number of sent returns. This also possibly implies ambiguity in knowledge about the actual impact of returns. However, the strongest predictor of return volume was shopping volume, so regardless of a person's norms, values, influences from others, and knowledge on the subject, those who shop more seemingly have higher returns as well. All in all, this led to the independent variables explaining 46.1% of the variance in sent returns and 17.5% of the variance in physical returns. In comparison, Armitage and Conner (2001) found in their meta-study of 185 studies that TPB explains 39% of the variance in behavioural intention. Based on this, the model for sent returns could be considered decently successful. It also indicates that there are likely other variables not sufficiently covered in this thesis, outside of TPB and what was found in the grounded theory phase, that influence behaviour, particularly physical returns, but likely also sent returns.

RQ2 and RQ3 were mostly addressed in Chapter 3, and in part Chapter 2.

Both the financial impact of returns and the environmental impact of returns are closely connected, as both are related to how far an item is sent, the number of links in the return chain, how it is packaged in terms of material used, excess space taken up during transport, the number of items discarded due to not arriving in good enough condition, and how quickly it can be routed to a new customer. All these factors impact both the environmental and financial aspects. They can both be mitigated by more localized processing centres, avoiding the costly and complex return process back to central distribution hubs far from the consumers. Specifically, by returning items to local stores that have the possibility to resell them directly and possibly help the consumer find a new item on the spot, without purchasing many more items online that may also need to be returned. However, the environmental impact of this depends on how consumers travel to a physical store, for instance, by not using high emission means of transport for the sole purpose of returning. Besides this, returns will have a certain unavoidable impact that cannot be mitigated without reducing the number of returns altogether.

5.2.2 Practical implications

The findings in this thesis have practical implications for retailers, policymakers, and environmental conservation and sustainability groups.

For retailers, a key implication is that different measures can be taken to reduce both the return rate and the impact of returns. This can help them reduce their scope 3 emissions and reduce the cost of accepting customer returns. Firstly, by increasing and investing more in the technological tools available, they can assist customers in getting their purchases right the first time. Secondly, by encouraging customers to return items physically in-store, either by communicating the environmental benefits or using incentives, they can cut return costs and increase the chance of a customer finding the correct item, rather than having to lose the entire sale and the cost of the return. Thirdly, the findings do not provide evidence that increased buying volume leads to consumers better understanding their size or preferences for future buying; in other words, return volume does not decrease as a consumer buys more; it keeps increasing. Therefore, by understanding that increased purchases will likely lead to increased returns, reverse logistics operations can be scaled and optimized to avoid both excess capacity, which can be costly, and undercapacity, which can lead to delays in getting returns of seasonal stock back on shelves for resale before it becomes obsolete. Finally, by informing customers more about the issue and the financial and environmental costs of returns, there is a possibility that this will help raise awareness of the cause and make selected consumers more conscious of their shopping and returning behaviour. Although the model here did not specifically find support that knowledge would impact behaviour, there is still a possibility that it could indeed do so, given that the measurement of knowledge used in this research may not have aligned well enough with actual knowledge on the subject[10].

For policymakers, one key implication is that retailers that only aim to maximize profits with no concern for the environment may very well do the bare minimum when it comes to reducing the impact of returns. This is due to the ambiguity regarding how well measures actually reduce return rates without negatively affecting total sales volumes and whether a measure to reduce return rates and thus the costs of returns will outweigh the risk of a pos-

[10] This is discussed further in Chapter 5.1.2

sible reduction in sales and customer loyalty. A second implication is that consumers have few constraints stopping them from excessive ordering and returning, even those with knowledge of sustainability and socially responsible consumers. Also, knowledge is likely lacking regarding the consequences and costs of returning clothing bought online. Informing consumers could possibly help; however, as mentioned above, there is uncertainty surrounding the effectiveness of such an action.

To address both of these implications, a possible action could be to work on cross-border policies, for instance, ratified by the European Union, to influence market mechanisms and force retailers to implement measures that will, in turn, impact consumer behaviours. Specifically, this could involve financial incentives, such as reducing consumer rights to freely return items purchased online or implementing incrementally higher fees for returning higher quantities of clothing to address excessive ordering.

For environmental and sustainability groups, one key implication is the discrepancy between the behaviours of environmentally responsible consumers and returns, as well as subjective norms only being significant for physical returns and not sent returns. This could imply that the guilt of less sustainable behaviour is easier to deal with when no one else is looking, and increased consumerism and online shopping are likely to exacerbate and increase the problem. A second implication concerns knowledge on the topic, similar to the points mentioned above.

5.3 Limitations

The limitations of each stage have been laid out in their respective chapters; below, they will be briefly repeated and somewhat elaborated on.

In Chapter 3, the main limitation was the number of subjects interviewed. Although the data collected indicated that for each of the two themes explored, saturation was possibly reached, it is not unlikely that additional interviews could have provided more insights into the addressed themes.

In Chapter 4, there were several limitations. Firstly, the data was borderline in terms of satisfying all the assumptions required for conducting multiple linear regression, specifically the assumption of no unusual points or outliers and homoscedasticity. Regarding the data not quite meeting the assumptions required to conduct multiple linear regression, there are possibly other models that would have been better suited to model and represent the relationship between the variables. For instance, different combinations of grouping the return volumes could have better captured the data. For instance, dividing them into two groups using a median split and comparing the high and low groups. This could have provided data that better satisfied the assumptions, as outliers and the increasing variance for return volume (i.e., heteroscedasticity) would have been eliminated by grouping all the data into high or low groups. However, this approach comes with drawbacks, one being that the large number of data points close to the median but only marginally different would be classified as different from each other as the outliers farthest from the median. This would reduce the richness and natural variation in the data dramatically. This is only one drawback, as other drawbacks are also elaborated by DeCoster et al. (2011), Grace-Martin (2018), and Osbourne (2012). An implication of this is that the theory possibly only holds true for the "average" consumer who has shopping and return volumes close to the average consumer, while the model is not true for and fails to explain the behaviour of the outlier consumers who have exceptionally large deviations from the rest of the population concerning shopping and return volumes. Finally, even in a normal distribution, any sample will by nature on average have approximately 1 in every 300 data points located outside of 3 standard deviations from the average (Frost, 2019) and thus be categorized as an outlier (Osborne, 2012). Thus, with a sample of ~900, a certain number of outliers should be expected and does not automatically invalidate the results of multiple regression, despite two assumptions (a normal distribution and not having unusual points in the data) seemingly not being met.

To build on this, respondents were asked to estimate their return and shop volumes. It is likely that a person who has shopped for 50 items in the past 6 months will estimate this volume with a larger margin of error than a person who has purchased 3 items in the past 6 months. The latter person would have an easier time recalling the correct number, as the number of instances of shopping online would be fewer. If this is indeed the case, it could explain why some of the data exhibited heteroscedasticity, especially among outliers with high shop volume or return volume. In other words, the wording of the question naturally results in different variances across the range of answers. The drawback is that this complicates the statistical analyses, even though it provides richer and more accurate data than reducing the data to ordinal variables, instead of scale variables, by dividing them into groups of two or more categories, such as low and high or very low, low, medium, high, very high, or similar. Again, this was not done due to the mentioned drawbacks.

Secondly, the sample was not entirely representative. The sample only contains persons who complete surveys for money, and many of these have done a large number of surveys. As such, the survey is naturally biased and does not contain data from the portion of the population that does not complete surveys for money. Additionally, after the survey was finished, and demographics and data were collected for the respondents, the number of surveys the respondents had previously completed and had approved was also displayed in the dataset. The average number was 1704 approved surveys in total, ranging from 1 to 6647, with a standard deviation of 1487. In other words, many of the respondents could almost be seen as professional survey responders. This was also part of the reason why a large number of answers were filtered out in the data cleaning process, as described in detail in Appendix 8.3, in the section "Data Cleaning Details." Despite the high number of surveys done by many of the respondents, the author still received messages from respondents saying they thought the topic was very interesting and thanked for participating. This indicates that many likely took it seriously and answered honestly to the best of their ability. In addition to this, survey publishers on Prolific (like the author) must approve the answers from the respondents, using instructional manipulation checks. If respondents fail such tests, they are penalized and will lose their membership on the site if their submissions are rejected several times (Prolific, 2023). In other words, although the respondents may be very experienced, there are a series of checks to ensure that any negligent respondents are filtered out of the final data set and sample. This is done both by the checks in place performed by the author and by the many checks the respondents likely have been through when answering all their previous surveys that have been approved. However, there will always be a risk that a certain number of them get past all the checks and manage to game the system. On the other hand, too many checks could also filter out valid answers, so there is a balance to be struck. As mentioned, due to the large sample size, the author has erred on the side of caution and attempted to be relatively strict in filtering out possible random responses from negligent participants, as elaborated in the mentioned sections of this thesis.

5.4 Future research

Several interesting results have emerged from this study that can serve as a foundation for future research. Additionally, any research will inherently have weaknesses due to design trade-offs and decisions regarding budget and resource allocation. The study has strived to ensure high reliability and validity of the results; however, future research could replicate the study in different populations and with more representative samples to test the generalizability and replicability of the findings.

Furthermore, as explained earlier, there are potential extensions to the research model, particularly to explain the variation in physical returns. The study collected limited qualitative data directly from consumers, and more insights into their behaviour could possibly be uncovered through qualitative exploratory studies, such as focus groups, to gain a deeper understanding of potential independent variables affecting return volume. Additionally, this approach could shed light on the behaviours of average consumers as well as those with behaviours that could be described as outliers, characterized by exceptionally high or low return or purchase volumes. These findings could then be tested in extended versions of the research model proposed in this thesis, with an aim to increase explanatory power, as hypothesized.

There are several potential variables that could be explored in greater depth. Firstly, the cost or effort of returns could be assessed to understand how easy it is to return an item compared to keeping it. This variable could include dimensions related to how difficult a retailer makes it to return an item, such as requiring forms to be filled out or pre-approval from customer service. It could also assess the perceived effort required from the customer, such as the travel distance to return an item, the need for extra packaging, and the type of item being returned.

Another variable worth deeper exploration is knowledge. As mentioned, the measurement of knowledge in this study was self-reported, making it challenging to determine whether the lack of a relationship is due to low actual knowledge, consumers overestimating what they know, or the measurement being too general in assessing sustainability knowledge rather than knowledge specific to returns of clothing bought online. Therefore, a more precise measurement of knowledge, along with a stronger measure for knowledge, could be developed. This could include one facet testing general sustainability knowledge and another assessing reported knowledge specifically related to returns of clothing bought online.

Lastly, the self-reported shop and return volume could be improved through a longitudinal study that collects actual behavioural data, potentially reducing or mitigating the issue of heteroscedasticity.

6. Conclusion

Overall, the research aimed to reduce the impact of clothing bought online by addressing both the return rate and the impact of returns. These aspects were initially explored through qualitative research, followed by quantitative research. The study also incorporated existing literature and research, combining newer findings with well-established theories.

In terms of the return rate, the research culminated in a research model with sent returns and physical returns as dependent variables. These models were quantitatively tested using a questionnaire, and the independent variables explained 46.1% and 17.5% of the variance in sent returns and physical returns, respectively. For sent returns, the strongest predictor was the number of items purchased, followed by socially responsible consumption related to shopping and returns. Surprisingly, socially responsible consumption was associated with higher returns, and personal norms had a negative relationship with sent returns. For physical returns, the number of items purchased was the most influential predictor, followed by subjective norms, which refer to what individuals believe others' norms to be. In summary, it became evident that norms alone might not be sufficient to address the issue, and policies or other mechanisms should be implemented to encourage both retailers and consumers to reduce the return rate.

Regarding the financial and environmental impact of returns, the study found that returning items to local stores presented an opportunity for improvement. This approach reduces the cost and environmental impact compared to extensive journeys through multiple processing centres over long distances, which often result in items being destroyed due to the high cost of such return processes. To encourage this practice, greater financial incentives could be introduced. Additionally, technology could play a significant role in helping consumers order correct sizes, provided there are stronger financial incentives for retailers to invest in and apply such technology.

In conclusion, raising consumer awareness is essential, and additional mechanisms beyond norms must be employed to address the issue, targeting both consumers and retailers.

7. References

Abrahamse, W., & Steg, L. (2009). How do socio-demographic and psychological factors relate to households' direct and indirect energy use and savings? *Journal of economic psychology*, *30*(5), 711-720.

Accenture. (2018). Returns - The Value Conundrum. https://www.accenture.com/_acnmedia/pdf-95/accenture-postal-vision-2020-returns-slideshare.pdf

Ajzen, I. (1985). *From intentions to actions: A theory of planned behavior*. Springer.

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2), 179-211.

Ajzen, I. (2006). Constructing a theory of planned behavior questionnaire. In: Amherst, MA.

Ajzen, I., & Driver, B. L. (1992). Application of the theory of planned behavior to leisure choice. *Journal of leisure research*, *24*(3), 207-224.

Ajzen, I., Joyce, N., Sheikh, S., & Cote, N. G. (2011). Knowledge and the prediction of behavior: The role of information accuracy in the theory of planned behavior. *Basic and applied social psychology*, *33*(2), 101-117.

Albayrak, T., Aksoy, Ş., & Caber, M. (2013). The effect of environmental concern and scepticism on green purchase behaviour. *Marketing Intelligence & Planning*, *31*(1), 27-39.

Amoako, G. K., Dzogbenuku, R. K., & Abubakari, A. (2020). Do green knowledge and attitude influence the youth's green purchasing? Theory of planned behavior. *International Journal of Productivity and Performance Management*, *69*(8), 1609-1626.

Arcury, T. A., & Johnson, T. P. (1987). Public environmental knowledge: A statewide survey. *The Journal of Environmental Education*, *18*(4), 31-37.

Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *British journal of social psychology*, *40*(4), 471-499.

Armstrong, J. S., & Overton, T. S. (1977). Estimating nonresponse bias in mail surveys. *Journal of marketing research*, *14*(3), 396-402.

Astivia, O. L. O., & Zumbo, B. D. (2019). Heteroskedasticity in Multiple Regression Analysis: What it is, How to Detect it and How to Solve it with Applications in R and SPSS. *Practical Assessment, Research, and Evaluation*, *24*(1), 1.

Ateş, H. (2020). Merging theory of planned behavior and value identity personal norm model to explain pro-environmental behaviors. *Sustainable Production and Consumption*, *24*, 169-180.

Bahuleyan, H., Lasserre, J., Lefakis, L., & Shirvany, R. (2022). Knowing When You Don't Know in Online Fashion: An Uncertainty-Aware Size Recommendation Framework. Recommender Systems in Fashion and Retail: Proceedings of the Third Workshop at the Recommender Systems Conference (2021),

Bellenger, D. N., & Korgaonkar, P. K. (1980). Profiling the recreational shopper. *Journal of Retailing*, *56*(73), 77-92.

Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis: A global perspective*. Pearson.

Bogusławska-Bączek, M. (2010). Analysis of the contemporary problem of garment sizes. 7th International Conference on Textile Science (TEXSCI 2010), Liberec, Czech Republic,

Charmaz, K. (2014). *Constructing Grounded Theory*. Sage.

Claudio, L. (2007). Waste Couture: Environmental Impact of the Clothing Industry. *Environmental Health Perspectives*, *115*(9), 448-454. https://doi.org/https://doi.org/10.1289/ehp.115-a449
Clo3D. (2022). *Clients*. Clo3D. Retrieved 09.06.2022 from https://www.clo3d.com/en/company/clients
Collado, S., Staats, H., & Sancho, P. (2019). Normative influences on adolescents' self-reported pro-environmental behaviors: The role of parents and friends. *Environment and Behavior*, *51*(3), 288-314.
Corbin, J., & Strauss, A. (2015). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (4 ed.). Sage.
Coyle, K. (2005). Environmental literacy in America: What ten years of NEETF/Roper research and related studies say about environmental literacy in the US. *National Environmental Education & Training Foundation*.
Cullinane, S., & Cullinane, K. (2021). The Logistics of Online Clothing Returns in Sweden and How to Reduce its Environmental Impact. *Journal of Service Science and Management*, *14*(1). https://doi.org/10.4236/jssm.2021.141006
Cullinane, S. L., Browne, M., Karlsson, E., & Wang, Y. (2019). Retail Clothing Returns: A Review of Key Issues. *Contemporary Operations and Logistics*, *1*. https://doi.org/https://doi.org/10.1007/978-3-030-14493-7_16
DC Velocity, Censuswide, & Logistyx Technologies. (2021). *Most popular cross-border online shopping product categories worldwide in 2021*. Retrieved 23.10.2022 from https://www.statista.com/statistics/509438/leading-cross-border-online-shopping-categories-worldwide/
De Leeuw, A., Valois, P., Ajzen, I., & Schmidt, P. (2015). Using the theory of planned behavior to identify key beliefs underlying pro-environmental behavior in high-school students: Implications for educational interventions. *Journal of environmental psychology*, *42*, 128-138.
DeCoster, J., Gallucci, M., & Iselin, A.-M. R. (2011). Best practices for using median splits, artificial categorization, and their continuous alternatives. *Journal of experimental psychopathology*, *2*(2), 197-209.
Dickler, J. (2022). *Free returns may soon be a thing of the past as retailers roll out stricter policies*. CNBC. Retrieved 20.05.23 from Free returns may soon be a thing of the past as retailers roll out stricter policies
Dillman, D. A. (2011). *Mail and Internet surveys: The tailored design method--2007 Update with new Internet, visual, and mixed-mode guide*. John Wiley & Sons.
Duerden, M. D., & Witt, P. A. (2010). The impact of direct and indirect experiences on the development of environmental knowledge, attitudes, and behavior. *Journal of environmental psychology*, *30*(4), 379-392.
Dumitrescu, A. L., Wagle, M., Dogaru, B. C., & Manolescu, B. (2011). Modeling the theory of planned behavior for intention to improve oral health behaviors: the impact of attitudes, knowledge, and current behavior. *Journal of oral science*, *53*(3), 369-377.
eMarketer. (2022). *Retail e-commerce sales worldwide from 2014 to 2026*. eMarketer. Retrieved 23.10.2022 from https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/
European Commission. (2022). *Questions and Answers on EU Strategy for Sustainable and Circular Textiles*. Retrieved 12.08.2022 from https://ec.europa.eu/commission/presscorner/detail/en/QANDA_22_2015
Fearn, J. (2020). *5 Ways Technology Can Minimize Product Return Rates*. Total Retail. Retrieved 07.06.2022 from https://www.mytotalretail.com/article/5-ways-technology-can-minimize-product-return-rates/

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. sage.
Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behaviour: An introduction to theory and research* (Vol. 27).
Flick, U. (2013). *The SAGE handbook of qualitative data analysis*. Sage.
Fox, J. (2015). *Applied regression analysis and generalized linear models*. Sage Publications.
Frederick, J. (2015). Online Retail Cross-Border Sales: The Global Trend That's Here to Stay. *PFSweb*. https://www.pfscommerce.com/PDF/whitepapers/Online-Retail-Cross-Border-Sales-FINAL.pdf?x18118
Frost, J. (2019). *Guidelines for Removing and Handling Outliers in Data*. Statistics by Jim. Retrieved 17.07.2023 from https://statisticsbyjim.com/basics/remove-outliers/
Gao, L., Wang, S., Li, J., & Li, H. (2017). Application of the extended theory of planned behavior to understand individual's energy saving behavior in workplaces. *Resources, Conservation and Recycling*, *127*, 107-113.
Gäthke, J., Gelbric, K., & Chen, S. (2021). A Cross-National Service Strategy to Manage Product Returns: E-Tailers' Return Policies and the Legitimating Role of the Institutional Environment. *Journal of Service Research*, *25*(3), 402–421. https://doi.org/https://doi.org/10.1177/10946705211989440
Gkargkavouzi, A., Halkos, G., & Matsiori, S. (2019). Environmental behavior in a private-sphere context: Integrating theories of planned behavior and value belief norm, self-identity and habit. *Resources, Conservation and Recycling*, *148*, 145-156.
Glaser, B. (1992). *Basics of Grounded Theory Analysis*. Sociology Press.
Glaser, B. (1998). *Doing Grounded Theory: Issues and Discussions*. Sociology Press.
Glaser, B. (2008). *Doing Quantitative: Grounded Theory*. Sociology Press.
Glaser, B., & Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. AldineTransaction.
Gorsuch, R. L., & Ortberg, J. (1983). Moral obligation and attitudes: Their relation to behavioral intentions. *Journal of personality and Social Psychology*, *44*(5), 1025.
Grace-Martin, K. (2018). *3 situations when it makes sense to categorize a continuous predictor in a regression model*. The Analysis Factor. Retrieved 14.08.2023 from https://www.theanalysisfactor.com/3-situations-when-it-makes-sense-to-categorize-a-continuous-predictor-in-a-regression-model/
Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Sage.
Guba, E. G., & Lincoln, Y. S. (2005). Paradigmatic controversies, contradictions, and emerging confluences.
Guerin, R. J., & Toland, M. D. (2020). An application of a modified theory of planned behavior model to investigate adolescents' job safety knowledge, norms, attitude and intention to enact workplace safety and health skills. *Journal of safety research*, *72*, 189-198.
Gupta, S., & Ogden, D. T. (2009). To buy or not to buy? A social dilemma perspective on green buying. *Journal of consumer marketing*, *26*(6), 376-391.
H&M Group. (2019). *H&M and the Perfect Fit*. H&M Group. Retrieved 09.06.2022 from https://hmgroup.com/our-stories/hm-and-the-perfect-fit/
H&M Group. (2021). *Taking sustainable fashion to a new level with tech*. H&M Group. Retrieved 09.06.2022 from https://hmgroup.com/our-stories/taking-sustainable-fashion-to-a-new-level-with-tech/
Hair, J. F. (2009). Multivariate data analysis.
Hammervold, R. (2020). Multivariate analyser med stata.

Hardin, G. (1968). The tragedy of the commons: the population problem has no technical solution; it requires a fundamental extension in morality. *science*, *162*(3859), 1243-1248.

Harland, P., Staats, H., & Wilke, H. A. (1999). Explaining proenvironmental intention and behavior by personal norms and the Theory of Planned Behavior 1. *Journal of applied social psychology*, *29*(12), 2505-2528.

Hong, S.-J., & Tam, K. Y. (2006). Understanding the adoption of multipurpose information appliances: The case of mobile data services. *Information systems research*, *17*(2), 162-179.

Huber, P., & Ronchetti, E. (1981). Robust Statistics, Wiley: New York. In: USA.

Hull, S. (2012). *2012 in Online Fashion*. Drapers. Retrieved 23.10.2022 from https://www.drapersonline.com/topics/digital-and-tech/2012-in-online-fashion

IMRG. (2020). *IMRG Returns Review—2020*. https://www.imrg.org/wp-content/uploads/aee8f77eb4d47ccae59293c2518ea67641e5a89d.pdf

Jansson-Boyd, C. (2023). *The Surprising Reason Online Shopping Can Be Less Satisfying*. Psychology Today. Retrieved 03.05.2023 from https://www-psychologytoday-com.cdn.ampproject.org/c/s/www.psychologytoday.com/us/blog/consumer-psychology/202302/the-surprising-reason-online-shopping-can-be-less-satisfying?amp

Jansson, J. (2011). Consumer eco-innovation adoption: assessing attitudinal factors and perceived product characteristics. *Business Strategy and the Environment*, *20*(3), 192-210.

Kapner, S. (2023). *Retailers Clamp Down on Returns*. Wall Street Journal. Retrieved 29.05.2023 from https://www.wsj.com/articles/online-retailers-tighten-return-policies-to-boost-profits-9bf6ccc2

Kaushik, V., Kumar, A., Gupta, H., & Dixit, G. (2020). Modelling and prioritizing the factors for online apparel return using BWM approach. *Electronic Commerce Research*, *22*(3), 843-873. https://doi.org/10.1007/s10660-020-09406-3

Kinetic. (2022). *Capturing The Mobile Pound*. Retrieved 23.10.2022 from https://kineticww.com/capturing-the-mobile-pound/

Kollmuss, A., & Agyeman, J. (2002). Mind the gap: why do people act environmentally and what are the barriers to pro-environmental behavior? *Environmental education research*, *8*(3), 239-260.

Kristiansen, R. K. (2022). *Dette skjer med klærne du sender tilbake: – Helt vanvittig*. TV2.no. Retrieved 27.04.2022 from https://www.tv2.no/a/14554847/

LaBat, K. L. (2007). Sizing standardization. *Sizing in Clothing: Developing Effective Sizing Systems for Ready-to-Wear Clothing, Cambridge: Woodhead Publishing Limited*, 88-107.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. sage.

Lincoln, Y. S., Lynham, S. A., & Guba, E. G. (2011). Paradigmatic controversies, contradictions, and emerging confluences, revisited. *The Sage handbook of qualitative research*, *4*(2), 97-128.

Lydeard, S. (1991). The questionnaire as a research tool. *Family practice*, *8*(1), 84-91.

Maccoby, E. E., & Maccoby, N. (1954). The interview: A tool of social science. *Handbook of social psychology*, *1*(1), 449-487.

Manayiti, O., & Edgecliffe-Johnson, A. (2022). Goods returned by US consumers surged 78% in 2021. https://www.ft.com/content/c926c427-83e2-4fa7-8165-8466b6b037a7

Marks and Spencer. (2019). *Marks & Spencer And Founders Factory Invest In 3D Digital Fit Company*. Marks and Spencer. Retrieved 08.06.2022 from https://corporate.marksandspencer.com/media/pressreleases/5c2f8d617880b21084450f5e/marks-and-spencer-and-founders-factory-invests-in-3d-digital-fit-company

Mathieson, K. (1991). Predicting user intentions: comparing the technology acceptance model with the theory of planned behavior. *Information systems research*, *2*(3), 173-191.

McKinsey. (2018). *The State of Fashion 2018*. McKinsey. https://www.mckinsey.com/~/media/McKinsey/Industries/Retail/Our%20Insights/Renewed%20optimism%20for%20the%20fashion%20industry/The-state-of-fashion-2018-final.ashx

McKinsey. (2022). *The State of Fashion 2022*. McKinsey. https://www.mckinsey.com/industries/retail/our-insights/state-of-fashion#/download/%2F~%2Fmedia%2Fmckinsey%2Findustries%2Fretail%2Four%20insights%2Fstate%20of%20fashion%2F2023%2Fthe-state-of-fashion-2023-holding-onto-growth-as-global-clouds-gathers-vf.pdf

Metapack. (2020). *The Post Purchase Experience*. https://info.metapack.com/rs/700-ZMT-762/images/Metapack-The-Post-Purchase-Experience.pdf

Mohr, L. A., Webb, D. J., & Harris, K. E. (2001). Do consumers expect companies to be socially responsible? The impact of corporate social responsibility on buying behavior. *Journal of consumer affairs*, *35*(1), 45-72.

Moser, A. K. (2015). Thinking green, buying green? Drivers of pro-environmental purchasing behavior. *Journal of consumer marketing*.

Mull, A. (2023). *The Free-Returns Party Is Over*. Retrieved 14.05.2023 from https://www.theatlantic.com/technology/archive/2023/05/free-online-shopping-returns-retailer-policy-changes/673975/

MYSIZE. (2023). *MysizeID app*. MysizeID. Retrieved 30.05.2023 from https://mysizeid.com/mysizeid-app/

Narvar. (2019). *State of Returns Report 2019*. Narvar. https://see.narvar.com/rs/249-TEC-877/images/narvar-state-of-online-returns-global-consumer-study-2019.pdf

Narvar. (2022). *State of Returns Report 2022*. Narvar. https://see.narvar.com/rs/249-TEC-877/images/2022%20State%20of%20Returns%20Report%20-%20102822.pdf

NeXR. (2022). *H&M And NeXR Launch The Future With The First Virtual Fitting Room*. NeXR-Technologies. Retrieved 29.05.2023 from https://www.nexr-technologies.com/virtual-fitting-hm-success-case/

Norberg, P. A., Horne, D. R., & Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of consumer affairs*, *41*(1), 100-126.

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.

Ore, J. (2022). *Free online returns cost retailers millions. Now they want you to pay for it*. CBC. Retrieved 20.05.23 from https://www.cbc.ca/radio/costofliving/online-free-returns-1.6654967

Osborne, J. W. (2012). *Best practices in data cleaning: A complete guide to everything you need to do before and after collecting your data*. Sage publications.

Pauly, S. (2022). *NeXR Becomes Technology Partner for Virtual Fitting at H&M Thailand*. Business Wire. Retrieved 29.05.2023 from https://www.businesswire.com/news/home/20220912006012/en/NeXR-Becomes-Technology-Partner-for-Virtual-Fitting-at-HM-Thailand

Peiser, J. (2022). *The age of free online returns is ending*. Washington Post. https://www.washingtonpost.com/business/2022/12/09/free-returns-holiday-shopping/

Petschnig, M., Heidenreich, S., & Spieth, P. (2014). Innovative alternatives take action–Investigating determinants of alternative fuel vehicle adoption. *Transportation Research Part A: Policy and Practice*, *61*, 68-83.

Pickett-Baker, J., & Ozaki, R. (2008). Pro-environmental products: marketing influence on consumer purchase decision. *Journal of consumer marketing*, *25*(5), 281-293.

Piron, F., & Young, M. (2000). Retail borrowing: insights and implications on returning used merchandise. *International Journal of Retail & Distribution Management*, *28*(1), 27-36. https://doi.org/https://doi.org/10.1108/09590550010306755

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of applied psychology*, *88*(5), 879.

Prendergast, G. P., & Tsang, A. S. (2019). Explaining socially responsible consumption. *Journal of consumer marketing*, *36*(1), 146-154.

Prolific. (2022). *Prolific Audience Checker*. Retrieved 18.12.2022 from https://app.prolific.co/audience-checker

Prolific. (2023). *Approvals, rejections & returns*. Retrieved 15.08.2023 from https://researcher-help.prolific.com/hc/en-gb/articles/360009092394-Approvals-rejections-returns#h_01HA4QHJX23295WNNZN56C18SD

PYMNTS. (2022). *Mobile Devices Are Tool of Choice for One-Third of Consumers' Online Shopping*. Retrieved 23.10.2022 from https://www.pymnts.com/news/mobile-payments/2022/mobile-devices-are-tool-of-choice-for-one-third-of-consumers-online-shopping/

Research and Markets, & Statista. (2021). *Fashion e-commerce market value worldwide from 2021 to 2025*. Retrieved 23.10.2022 from https://www.statista.com/statistics/1298198/market-value-fashion-ecommerce-global/

Rohma, A. J., & Swaminathan, V. (2004). A typology of online shoppers based on shopping motivations. *Journal of Business Research*, *57*(7), 748-757. https://doi.org/https://doi.org/10.1016/S0148-2963(02)00351-X

Rubio, S., Chamorro, A., & Miranda, F. J. (2008). Characteristics of the research on reverse logistics (1995–2005). *International journal of production research*, *46*(4), 1099-1120.

Saarijärvi, H., Sutinen, U.-M., & Harris, L. C. (2017). Uncovering consumers' returning behaviour: a study of fashion e-commerce. *The International Review of Retail, Distribution and Consumer Research*, *27*(3). https://doi.org/https://doi.org/10.1080/09593969.2017.1314863

Salerno-Garthwaite, A. (2022). *Bracketing: Fashion's hidden returns problem*. Vogue Business. Retrieved 29.05.2022 from https://www.voguebusiness.com/consumers/bracketing-fashions-hidden-returns-problem

Salesforce Research. (2021). *Most popular ways to purchase online worldwide 2021, by selected categories*. Retrieved 23.10.2022 from https://www.statista.com/statistics/1274399/favorite-ways-to-shop-online-category-worldwide/

Saunders, M., Lewis, P., & Thornhill, A. (2016). *Research Methods For Business Students* (7 ed.). Pearson.

Schwartz, S. H. (1973). Normative explanations of helping behavior: A critique, proposal, and empirical test. *Journal of Experimental Social Psychology*, *9*(4), 349-364.
Schwartz, S. H. (1977). Normative influences on altruism. In *Advances in experimental social psychology* (Vol. 10, pp. 221-279). Elsevier.
Smith, K. A. (2022). *Why Free Returns Could Soon Be A Thing Of The Past*. Forbes. Retrieved 20.05.23 from https://www.forbes.com/advisor/personal-finance/retailers-free-online-returns/
Statista. (2021). *Fashion eCommerce report 2021*. https://www.statista.com/study/38340/ecommerce-report-fashion/
Statista, & We Are Social. (2018). *Share of internet users who have purchased selected products online in the past 12 months as of 2018*. Retrieved 23.10.2022 from https://www.statista.com/statistics/276846/reach-of-top-online-retail-categories-worldwide/
Stern, P. C. (2000). New environmental theories: toward a coherent theory of environmentally significant behavior. *Journal of social issues*, *56*(3), 407-424.
Thrane, C. (2018). *Kvantitativ metode: en praktisk tilnærming*. Cappelen damm akademisk.
U.S Census. (2020). *United States Census Bureau*. Retrieved 02.05.2023 from https://data.census.gov/table/ACSDP1Y2015.DP05?q=United+States&y=2020
Ullman, J. B., & Bentler, P. M. (2012). Structural equation modeling. *Handbook of Psychology, Second Edition*, *2*.
United Nations. (2022). *Ensure sustainable consumption and production patterns*. Retrieved 23.05.2022 from https://sdgs.un.org/goals/goal12
Van der Veld, W., & Saris, W. E. (2004). Separation of error, method effects, instability, and attitude strength. *Studies in public opinion: Attitudes, nonattitudes, measurement error, and change*, 37-59.
Van der Werff, E., & Steg, L. (2015). One model to predict them all: Predicting energy behaviours with the norm activation model. *Energy Research & Social Science*, *6*, 8-14.
Van der Werff, E., & Steg, L. (2016). The psychology of participation and interest in smart energy systems: Comparing the value-belief-norm theory and the value-identity-personal norm model. *Energy Research & Social Science*, *22*, 107-114.
Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478.
Vignali, G., & Reid, L. (2014). Analysing consumer motivation towards purchasing fashion online. *International Journal of Business and Globalisation*, *13*(2), 133–152. https://doi.org/10.1504/IJBG.2014.064131
Viswanathan, M., & Kayande, U. (2012). Commentary on "common method bias in marketing: Causes, mechanisms, and procedural remedies". *Journal of Retailing*, *88*(4), 556-562.
Wang, J.-J., Chen, H., Rogers, D. S., Ellram, L. M., & Grawe, S. J. (2017). A bibliometric analysis of reverse logistics research (1992-2015) and opportunities for future research. *International Journal of Physical Distribution & Logistics Management*.
Webb, D. J., Mohr, L. A., & Harris, K. E. (2008). A re-examination of socially responsible consumption and its measurement. *Journal of Business Research*, *61*(2), 91-98.
Zaller, J., & Feldman, S. (1992). A simple theory of the survey response: Answering questions versus revealing preferences. *American journal of political science*, 579-616.
ZIZR. (2022). *FAQ*. ZIZR. Retrieved 19.12.2022 from https://zizr.id/info/faq

8. Appendix

8.1 Chapter 2 appendix

Reason for returning	Category of online returning behaviour	Description
Product has defects	Reclamation driven	Consumers return items due to the product having defects, e.g. bad sewing or inappropriate materials
The wrong product has been delivered	Order fulfilment driven	The wrong product has been delivered, e.g. wrong colour, size or item
The product is found faster from another outlet The product is found for a cheaper price from another outlet	Competition driven	The customer finds the same product faster or cheaper from another outlet while waiting for the order to be delivered
An unanticipated negative feature The material differs from what was expected A different hue than expected Misleading product description Misleading product pictures	Disconfirmation driven	The product does not meet the expectations of the customer that have been created when ordering the product, e.g. the material is not what the customer anticipated or expected or the product has some negative feature that was not visible in the product pictures, e.g. a rip or tear etc.
The size of the product is too big or too small	Size chart driven	The size of the product is not right even though the customer ordered exactly his or her size
The customer's perception of the fit is not right The product does not match the customer's style The feeling of the product is not right	Feeling driven	For some reason, the customer does not feel 'right' when wearing the product
The customer cannot afford to keep the products The customer cannot exceed their pre-defined budget The customer regrets spending too much money	Money shortage driven	The customer does not eventually have the money to keep all the ordered products or he/she is now willing to spend that much money on the ordered products
The customer's need for the product has faded away	Faded need driven	At the time of the product delivery, the customer realizes that he/she does not actually need the ordered product after all
Ordering many sizes of the same product with the intention to keep only one Ordering the same product in many different colours with the intention to keep only one Ordering alternative products for the same need with the intention not to keep all of them	Benefit maximization driven	The customer orders multiple products with the intention to keep only one or a few of them
Ordering the product to just try it out for fun Ordering the product to try it before purchasing it from another outlet	Just trying out driven	The customer orders products with no intention to keep any of them

Table 8-1: Categories of online returning behaviour

(Saarijärvi et al., 2017)

Table 2:

Main reasons	Sub-reasons	References
Apparel attributes (AA)	Fit and size variation (AA1)	Kim and Damhorst, [46], Shin [87]
	Colour variation (AA2)	Yu et al. [103], Kim and Damhorst [46]
	Style variation (AA3)	Meng [61], Kasambala et al. [42], Sampaio et al. [86]
	Thickness variation (AA4)	Our contribution
	Texture variation (AA5)	Cho and Workman [16], Yu et al. [103]
	Stretchability variation (AA6)	Our contribution
	Fabric variation (AA7)	Yu et al. [103], Nitse et al. [67]
Disconformity (DC)	Misleading information (DC1)	Forbes et al. [23], Román [83]
	Defects (DC2)	Forbes et al. [23], Donaldson [19]
	Unexpected features (DC3)	Our contribution
	Quality issue (DC4)	Mukhopadhyay and Setaputra [64], Li et al. [52], Walsh et al. [100]
Dissonance (DN)	Found cheaper price (DN1)	McConnell et al. [60], Powers and Jack [76]
	Comparison with offline pricing (DN2)	Our contribution
	Peer influence (DN3)	Our contribution
	Impulsive purchase leads to cognitive dissonance (DN4)	Spears [91], Kang and Johnson [41], Lee [51]
	Found a better product (wisdom of purchase) (DN5)	Powers and Jack [76], Walsh et al. [100]
	Unplanned buying (DN6)	Saleh [84], Zhang [105]
Service failure (SF)	Receiving partial order (SF1)	Forbes et al. (2006), Román [83]
	Wrong product delivery (SF2)	Holloway and Beatty [35], Hong and Cha [36], Lee [51]
	Late delivery (SF3)	Feng et al. [22], Jiang et al. [39]
	Packaging error (SF4)	Singh [89]
	Could not process transaction using card or online (SF5)	Our contribution
	Customer unavailable (SF6)	Our contribution
	Customer has no cash (SF7)	Our contribution
Opportunism (OP)	Knowledge of return policy (OP1)	Harris [33], Lantz and Hjort (2012), Zhang [105]
	Lenient return policy (OP2)	Hjort [34], Lantz and Hjort [50]
	Past experience (OP3)	Zhang [105]
	Use and return (OP4)	Our contribution
	Purchase for trial buying (OP5)	Bullens et al. [8], Pei and Paswan [73], Zhang [105]
Perception (PP)	Counterfactual thinking leads to regret (PP1)	Roese [81], Liao et al. [53]
	Mental imagery discrepancy (PP2)	Zhang [105]
	I can get better deal after sometime on the same product may be on sale or other (PP3)	Powers and Jack [76]
	Bad reputation of retailer lead to return (PP4)	Kim et al. [47]; Walsh et al. [100]
	Value for money (PP5)	Our contribution

Table 8-2: main reasons and sub-reasons for inline apparel return (Kaushik et al., 2020)

8.2 Chapter 3 appendix

Participant 2, Logistics, Company B

○ What do you do?
○ How does your company handle return shipments?
○ Are returns difficult to handle? Why/why not?
○ Do you have data and statistics on your returns?
  • Volume, where it comes from, where it's going to, type of product/industry/consumer
  • How much data is typically registered on these things? Is it the same for la-

belled end labelless shipments?

- Cross border challenges in relation to returns, how does your company work with this?
- What can be done to reduce carbon footprint of returns? What makes it easier for you to deal with returns and optimize flow?
- EUs new strategy for textiles, are you aware of this? Do you notice this having any effect on you?
- Do you notice shops have different strategies around making it easy or difficult for consumers to return clothes?
- Do many retailers cover the return cost for the final customer? Is this trend changing?
- How much are businesses working with return logistics?
- Is your company doing anything towards a possible shift to a more circular economy?
- Do you notice that speed of returns is important for companies? Is it a metric/KPI you use?
- Anything else you want to add?

Table 8-3: Example of interview guide (interview 2)

Do you wish to participate in the research project «Greener logistics in online shopping»?

This is a question to you to participate in a research project where the aim is to find out how return logistics in e-retail can be improved. In this document you will get information about the goal for the project, and what participating will mean for you.

Aim

Part of the aim with this thesis is to investigate how online shopping can become greener, especially regarding returns, particularly in clothing sold B2C. To investigate this, I will speak to several players in this ecosystem, and wish to view this challenge from the point of view of three different groups:

1. People or firms who work with logistics, supply chain, reverse logistics, or similar
2. People or firms who sell clothing B2C through online channels
3. Consumers who buy clothing through online channels

I want to understand how product returns happen today and the impact it has on the environment and on the bottom line of the those involved. And I want to understand which factors affect the return rate and which measures (especially digital tools/instruments aimed at consumers) can help reduce the climate footprint from consumer returns. Finally, I want to be able to present recommendations to decision-makers with influence in the three groups, which can be related to the UN's sustainability goal no. 17 "Responsible production and consumption."

Who is responsible for the research project?

The master's thesis is written by Alexander Frayne, a student at the Norwegian School of Economics, in collaboration with the research projects Digital Innovation for Growth (DIG) and Best In Retail. The supervisor is Tor Wallin Andreassen, Professor in Innovation at the Norwegian School of Economics.

Why have you been asked to participate?

As I am early in the project, I want to understand more about return logistics, and am therefore interested in an unstructured and informal interview, where we can talk a bit back and forth. One of the things I'm interested in is learning a bit about how Postnord works with this, how Postnord handles returns and why return logistics seems to be so difficult and complicated for many players. I want to better understand which things would make returns easier to handle, and what affects the journey a return goes through and what could reduce the climate footprint of a returned item (except fewer returns) from your point of view. I know you work in international logistics, and to hear about some of the cross-border challenges in relation to returns, and how you work with this, would also be very interesting.

I don't expect you to have answers and solutions to all of the above challenges, but I am interested in hearing your thoughts and opinions on these topics.

What does participating mean for you?

If you choose to participate, it involves an interview that will last approx. 1 hour. I will take notes along the way and make an audio recording of the interview which will later be transcribed. The transcribed interview will be sent to you afterwards so that you can read over and approve it, or make adjustments if needed. The transcription will be anonymised so that your name will not be stored together with or in the document, and no other personal information about you will be collected or stored. Any identifying information that may come forth in the interview will be censored or redacted.

Participation is voluntary
Participation in the project is voluntary. If you choose to participate, you can withdraw your consent at any time without giving any reason. All your personal data will then be deleted. There will be no negative consequences for you if you do not want to participate, or later choose to withdraw.

Your privacy - how we store and use your information
We will only use the information about you for the purposes we have described in this document. We will treat the information confidentially and in accordance with the privacy regulations. Only Alexander Frayne will have access to the audio recording, and it will be deleted once transcription is complete. The transcription will be available to Tor Andreassen, but only after you have approved it, and it will be anonymised. If any direct quotes from the transcription end up being used in the thesis, they will be anonymised and will not be possible to trace back to you. All saved files are stored in a secure area with two-factor authentication so that no one else has access.

What happens to your personal data when the research project ends?
All data about you that is not anonymised will be deleted at the end of the project, if it has not already been deleted at an earlier time.

Your rights
As long as you can be identified in the data material, you have the right to:
- access to the information we process about you, and to be given a copy of the information
- to have information about you corrected that is incorrect or misleading
- to have personal data about you deleted
- to send a complaint to the Norwegian Data Protection Authority about the processing of your personal data

If you have questions about the study, or want to know more about or exercise your rights, please contact:

Alexander Frayne: alexander.frayne@student.nhh.no
Tor W. Andreassen: tor.andreassen@nhh.no

Kind regards,
Alexander Frayne
Norwegian School of Economics

Figure 8-1 Example of information leaflet for participants (Participant 2)

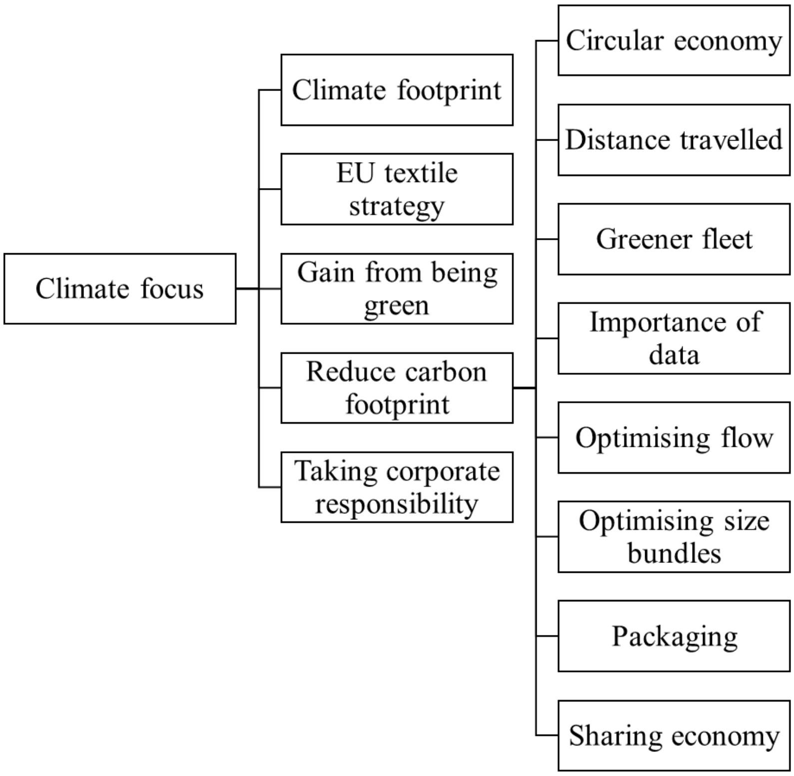


Figure 8-2 Level 2 to 3 codes for the level 1 code "Climate focus"

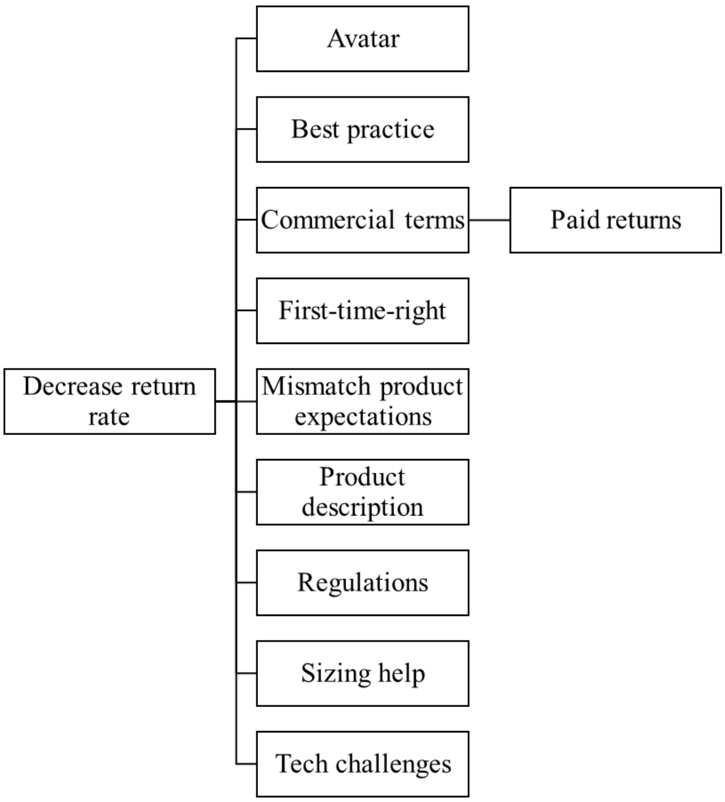


Figure 8-3: Level 2 to 3 codes for the level 1 code "Decrease return rate"

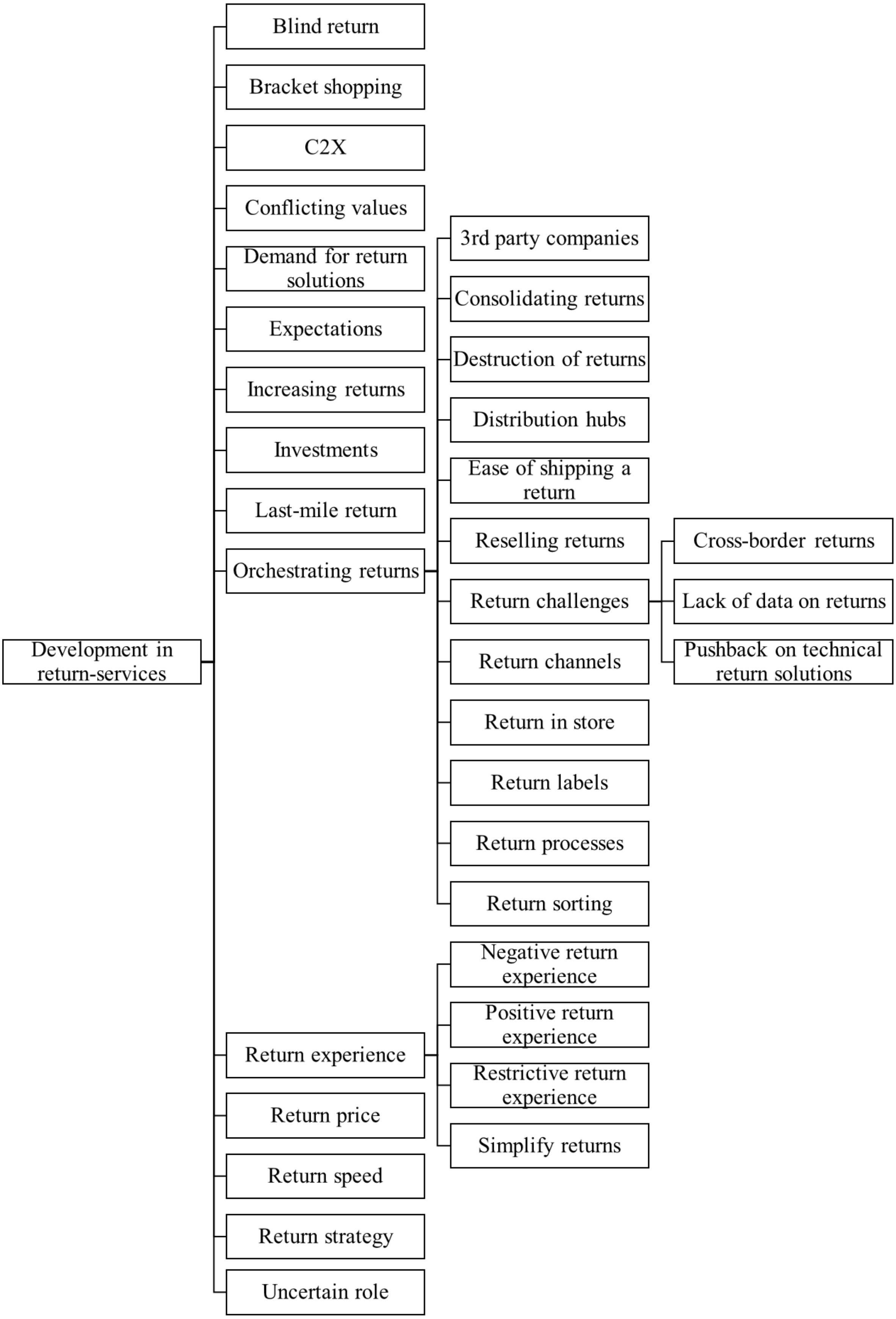


Figure 8-4: Level 2 to 4 codes for the level 1 code "Development in return services"

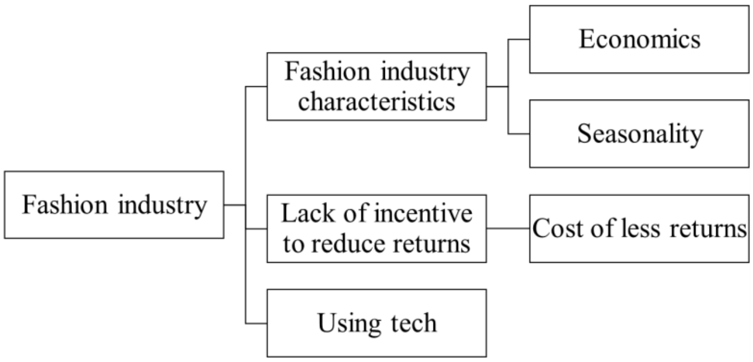


Figure 8-5: Level 2 to 3 codes for the level 1 code "Fashion industry"

8.3 Chapter 4 appendix

8.3.1 Questionnaire – Landing page on Prolific

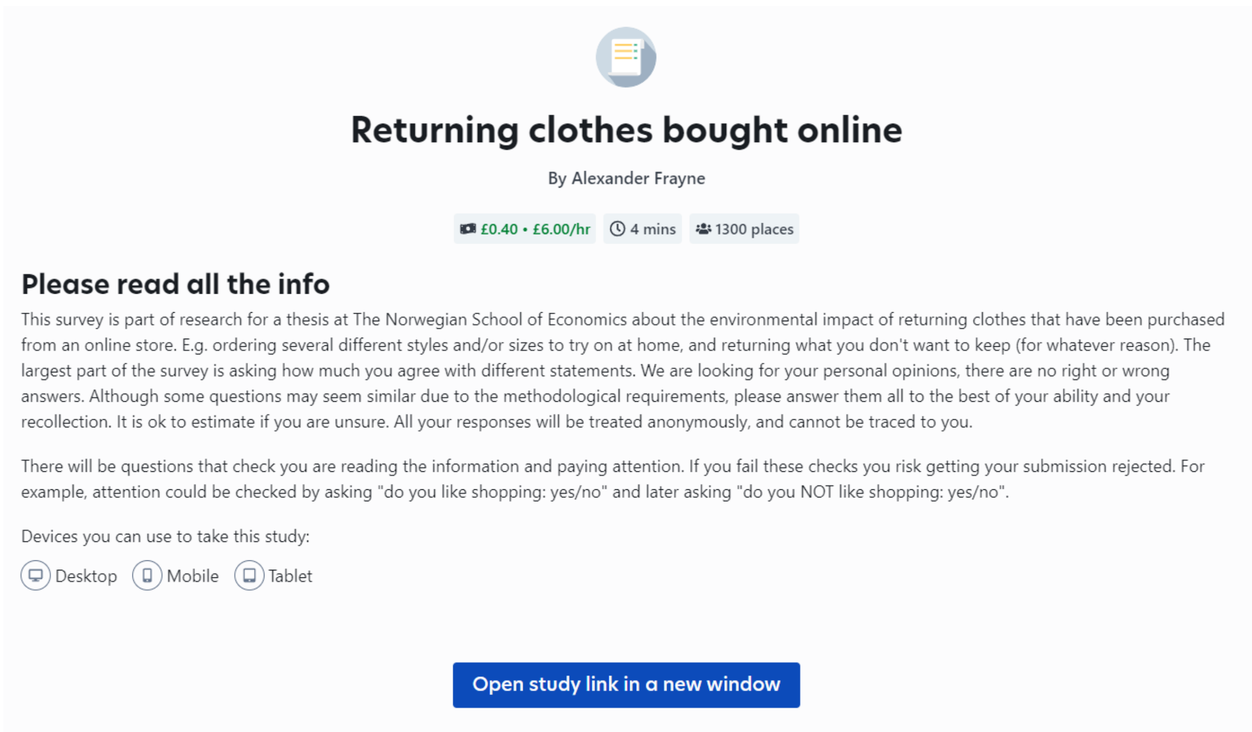


Figure 8-6: Questionnaire – Landing page on Prolific

8.3.2 Questionnaire – Qualtrics

NHH

What is your prolific ID?

→

NHH

This survey is about the environmental impact of returning clothes that have been purchased from an online store.
For example: ordering several different styles and/or sizes to try on at home, and returning what you don't want to keep (for whatever reason).
Please think of your behaviour in **the last 6 months** when you answer this survey.

Please continue if you consent to participating in the survey and allowing us to use your data in our research.

Based on the information above, how many of the last months should you think of when answering the questions in this survey?

1 2 3 4 5 6 7 8 9 10 11 12

← →

NHH

How many times have you ordered clothes online? (i.e. the **number of orders** in the last 6 months, to the best of your recollection)

How many **items of clothing** have you ordered online? (in the last 6 months, to the best of your recollection)

How many **items of clothing** ordered online have you returned by sending back? (in the last 6 months, to the best of your recollection)

How many **items of clothing** ordered online have you returned by **handing back to a physical store**? (in the last 6 months, to the best of your recollection)

→

NHH

To what degree do you agree or disagree with the following statements?

	1 Totally disagree	2	3	4 Neither agree nor disagree	5	6	7 Totally agree
When online returns are easy, I shop a little extra	○	○	○	○	○	○	○
I often shop a little extra online to get free shipping	○	○	○	○	○	○	○
Please click "7 Totally agree" on this statement to show you are paying attention	○	○	○	○	○	○	○
I prefer to shop for products that can last a while	○	○	○	○	○	○	○

	1 Totally disagree	2	3	4 Neither agree nor disagree	5	6	7 Totally agree
I often check where the product comes from before I buy it	O	O	O	O	O	O	O
I often check how the product is manufactured before I shop	O	O	O	O	O	O	O
I am careful about the correct handling of product packaging when I recycle	O	O	O	O	O	O	O
Overall, I act in an environmentally responsible way	O	O	O	O	O	O	O

To what degree do you agree or disagree with the following statements?

	1 Totally disagree	2	3	4 Neither agree nor disagree	5	6	7 Totally agree
I feel a moral obligation to only order what I intend to keep, when ordering clothes online	O	O	O	O	O	O	O
I do **not** feel a moral obligation to only order what I intend to keep, when ordering clothes online	O	O	O	O	O	O	O
I feel guilty ordering clothing online if I know I will be returning some or all of it	O	O	O	O	O	O	O
Most likely, people who are important to me will only order clothing online they intend to keep	O	O	O	O	O	O	O

	1 Totally disagree	2	3	4 Neither agree nor disagree	5	6	7 Totally agree
Most likely, people who are important to me will **not** only order clothing online they intend to keep	O	O	O	O	O	O	O
Most likely, people who I respect and look up to, will order clothing online, knowing that they will return some/all of it	O	O	O	O	O	O	O
Most likely, people who are important to me will recommend me (for environmental purposes) to only order clothing I intend to keep	O	O	O	O	O	O	O
I think people who are important to me expect me to **not** order clothing online if I know that I will be returning some/all of it.	O	O	O	O	O	O	O

To what degree do you agree or disagree with the following statements?

	1 Totally disagree	2	3	4 Neither agree nor disagree	5	6	7 Totally agree
Ordering more items of clothing online than one intends to keep, with the intention of returning some/all, has a negative effect on the environment	O	O	O	O	O	O	O
I spend a lot of time reading about sustainability in general	O	O	O	O	O	O	O
Please click "1 Totally disagree" on this statement to show you are paying attention	O	O	O	O	O	O	O

	1 Totally disagree	2	3	4 Neither agree nor disagree	5	6	7 Totally agree
Based on my previous online shopping, I have gained considerable insight into sustainability	○	○	○	○	○	○	○
Carbon emissions cause serious environmental problems, such as climate change	○	○	○	○	○	○	○
In general, I consider myself more interested in sustainability than the average online shopper	○	○	○	○	○	○	○

← →

8.3.3 Data cleaning details

First is a summary of them, in table 4-1. In total n = 419 were discarded, and n = 904 were kept. Below is first a summary of the tests and checks in Table 4-1, repeated below from Chapter 4, with an overview of how many did not pass. This is followed by the reasoning for each test and check and an elaboration.

N=2 withdrew consent. n = 11 didn't complete the first question. n = 73 failed IMC1 and didn't "move on" and thus didn't see or answer any more questions. n = 13 passed IMC1 but didn't answer any more questions. n = 14 answered the next 4 questions and none after that. n = 111 didn't complete the last question. Thus, we are left with 1323 respondents who completed the questionnaire, out of 1436 who started the questionnaire. Below shows which percentage of these who failed the different attention checks. All those who didn't do the attention checks obviously didn't pass them, as they didn't get that far in the questionnaire. The calculated percentage is therefore based on the total number of respondents who actually completed the full questionnaire (n = 1323). The 419 respondents who failed at least one attention check were discarded from the sample, leaving n = 904 in the final sample.

What is checked	**Number failed**	**% of full answers (n=1323)**
Sum failed at least one check	**N=419**	**31.67%**
Instructional manipulation check 1	N=73	5.52%
Instructional manipulation check 2	N=2	0.15%
Instructional manipulation check 3	N=13	0.98%
Illogical shopping or return volume 1	N=41	3.10%
Illogical shopping or return volume 2	N=15	1.13%
Response instability check 1	N=118	8.92%
Response instability check 2	N=275	20.79%
Respondent speed	N=64	4.84%

Table 4-1: Summary of data quality measures

Below, each of these filters and checks will be explained.

Instructional manipulation checks

Three instructional manipulation (IMC) checks were placed in the questionnaire. These aid in identifying respondents whom are not reading the text in the questionnaire, such as the instructions, general information, statements, and questions (Oppenheimer et al., 2009).

1. The participant was instructed to think of their behaviour for the last 6 months when answering the questionnaire. Then they were asked how many months they should think of when answering the questionnaire. Participants had to answer “6” to pass the test. 5.5% failed this test (N=76).

2. This question was a scale item on the “To what extent do you agree or disagree with the following statement” and the statement read “Please click "7 Totally agree" on this statement to show you are paying attention.” It was the third statement listed. Those who did not answer “7” failed the test. 7.6% failed this test (N=106).

3. This was similar to the one above, but the statement read “Please click "1 Totally disagree" on this statement to show you are paying attention.” It was the 22nd statement. Those who did not answer “1” failed the test. 7.3% failed the test (N=101).

Illogical shopping or return volumes

All answers were tested to check if their responses made sense on certain of the questions. This was also done to ensure respondents were reading the text and understood the questions. In particular, the questions asking for the shopping and return volume from the respondent. Two efforts here filtered out respondents who have been inattentive and given non-sensical answers or misunderstood the questions.

1. Q4 asked how many orders were made in the last 6 months. Q5 asked how many items of clothing were ordered in the last 6 months. If the number of clothing items was not equal to or greater than the number of orders, the respondents’ answers were discarded, as it is not possible to order less items of clothing than number of orders. 2.5% of respondents failed this check (N=35).

2. Q6 and Q7 asked for the return volume in number of items, in the last 6 months. If this number was greater than the number of items ordered in the last 6 months (answered in Q5), the respondents answers were discarded. 1.2% failed this check (N=17).

Response instability checks

Two statements were repeated in the questionnaire to check for response instability. Response instability occurs when a respondent answers the same questions differently when asked twice. In the case of the statements where respondents were asked to indicate their degree of agreement or disagreement, two of these statements were repeated. However, the second time these statements were presented, they were reversed. The wording of the statement remained identical, but an additional "not" was added to reduce any potential misunderstanding. For example, if a respondent answered "7 - Totally agree" to a statement, they were expected to answer "1 - Totally disagree" when the same statement was reversed.

This repetition was introduced to ensure that respondents were carefully reading and comprehending the statements. Zaller and Feldman (1992) argue that discrepancies in responses can occur even when respondents are attentive and understand the statements. The pilot tests supported this rationale. Consequently, if a respondent's second response differed from what was logically expected by a maximum of one point, they passed the check.

To clarify, if the initial answer was 7, the expected answer on the reversed version of the same statement was 1 or 2. Similarly, an initial answer of "6" required a response of "1," "2," or "3" on the reversed version. An initial response of 5 necessitated a response of "2," "3," or "4" on the reversed version, and so on. In other words, the sum of the reversed and original versions of the statement should ideally equal 8 when added together. However, a sum of 6 or 9 was also accepted to account for variability that may arise when reconsidering the same question twice, as these statements could be challenging to answer. Below are the results of this check.

1. The first asked “I feel a moral obligation to only order what I intend to keep, when ordering clothes online” and later asked “I do **not** feel a moral obligation to only order what I intend to keep, when ordering clothes online”. Here, 15.5% of the respondents failed the check (N=215)

2. Later came the statement “Most likely, people who are important to me will only order clothing online they intend to keep” and the reversed version “Most likely, people who are important to me will **not** only order clothing online they intend to keep”. Here, 26% of the respondents failed the check (N=361)

Respondent speed test

The final check to discard responses from respondents not paying full attention was based on the time taken to complete the questionnaire. The time required to complete the questionnaire should exceed the time needed to read and understand it, along with the time required to consider each response. A time threshold of 105 seconds was established for this purpose. The rationale behind this threshold will be explained below. But first, it's important to justify why this check is necessary. The fastest respondent completed the questionnaire in just 48 seconds while still passing all the previously mentioned checks. This highlights the need to include this additional check.

Now, let's delve into the explanation of the 105-second threshold. The time limit should account for: a) the time needed to read and comprehend the questionnaire, b) the time required to think about each response, and c) the time needed for navigation between pages, clicks, and typing. Regarding a), we can assume that most people read at a rate of around

350 words per minute. However, some skilled speed-readers can reach speeds of 600-700 words per minute. While it's safe to assume that very few respondents fall into this category, some may read at around 500 words per minute. Therefore, the absolute minimum time required to read the entire questionnaire is estimated at 64 seconds. For b), if we assume that each answer takes a minimum of 0.5 seconds, the 32 questions and statements would require 16 seconds to answer. As for c), we can reasonably estimate an additional 10 seconds, considering there are 5 text-box answers, and respondents must navigate between pages and click multiple times before completing the survey.

In summary, the absolute minimum time required to complete the questionnaire while still comprehending the questions and providing truthful and well-considered responses is approximately 90 seconds. However, if we adjust a) to a more typical reading speed of 350 words per minute, it will take 115 seconds in total (89 seconds to read, 16 seconds to answer, and 10 seconds to navigate). Taking an average of 90 seconds and 115 seconds, we arrive at 102.5 seconds. To establish a clear threshold, this figure was rounded up to 105 seconds, equivalent to one minute and 45 seconds, which happens to be half of the median time used. Consequently, 105 seconds was set as the minimum threshold for completing the questionnaire while ensuring full understanding and providing accurate responses. A total of 144 respondents, or 10.4% of the participants, completed the questionnaire in less than 105 seconds, and their responses were subsequently discarded.

8.3.4 Adapted measures and sources for questionnaire items

Measure	My Item	Source[11]
Demographics	What is your prolific ID	N/A
N/A	[Infotext]	N/A
Instructional manipulation check	Based on the information above, how many of the last months should you think of when answering the questions in this survey?	N/A
SRB - Shop volume	How many times have you ordered clothes online? (i.e., the number of orders in the last 6 months, to the best of your recollection)	Qualitative phase

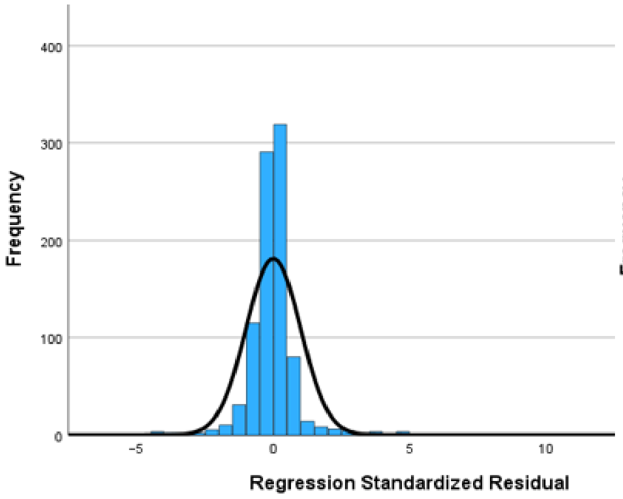
[11] Sources noted «Andreassen (2023) are from current unpublished research from Andreassen (personal communication, e-mail 16.01.2023)

SRB - Shop volume	How many items of clothing have you ordered online? (in the last 6 months, to the best of your recollection)	Qualitative phase
SRB - Sent returns	How many items of clothing ordered online have you returned by sending back? (in the last 6 months, to the best of your recollection)	Qualitative phase
SRB - Physical returns	How many items of clothing ordered online have you returned by handing back to a physical store? (in the last 6 months, to the best of your recollection)	Qualitative phase
N/A	**[Headline for remaining questions:] To what degree do you agree or disagree with the following statements?**	N/A
SRC - Shopping and returns	When online returns are easy, I shop a little extra	Andreassen (2023)
SRC - Shopping and returns	I often shop a little extra online to get free shipping	Andreassen (2023)
SRC - Shopping and returns	Overall, I find returning online goods to be easy and hassle free	Venkatesh et al (2012)
SRC - Shopping and returns	If online webshops offer free returns I will order more than I need	Andreassen (2023)
SRC - Shopping and returns	If online webshops require payment from me to return goods, for example 5 USD, it will make me think more carefully about what I buy	Andreassen (2023)
Instructional manipulation check	Please click "7 - Totally agree" on this statement to show you are paying attention	N/A
SRC - Product	I prefer to shop for products that can last a while	Andreassen (2023), Webb et al (2008)
SRC - Product	I often check where the product comes from before I buy it	Andreassen (2023), Gkargkavouzi (2009), Webb et al (2008)
SRC - Product	I often check how the product is manufactured before I shop	Andreassen (2023), Gkargkavouzi (2009), Webb et al (2008)
SRC - Product	I am careful about the correct handling of product packaging when I recycle	Andreassen (2023), Collado et al (2019), Webb et al (2008)
NI - Objective descriptive norms	I feel a moral obligation to only order what I intend to keep, when ordering clothes online	Gkargkavouzi (2009), Petschnig et al. (2014), Jansson (2011), Ateş (2020), Harland et al. (1999)
Response instability check	I do not feel a moral obligation to only order what I intend to keep, when ordering clothes online	Question above
NI - Objective descriptive norms	I feel guilty ordering clothing online if I know I will be returning some or all of it	van der Werff & Steg (2015, 2016), Petschnig et al. (2014), Ateş (2020), Harland et al. (1999)
NI - Subjective descriptive norms	Most likely, people who are important to me will only order clothing online they intend to keep	Gkargkavouzi (2009), Albayrak et

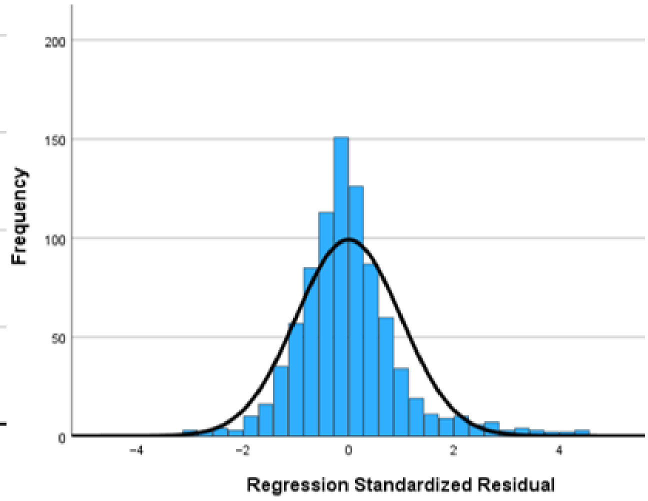
		al. (2013), Ajzen (2006), de Leeuw (2015)
Response instability check	Most likely, people who are important to me will not only order clothing online they intend to keep	Question above
NI - Subjective descriptive norms	Most likely, people who I respect and look up to, will order clothing online, knowing that they will return some/all of it	Gkargkavouzi 2009, Hong & Tam (2006), Ajzen (2006), Mathieson (1991), Gao et al. (2017)
NI - Subjective injunctive norms	Most likely, people who are important to me will recommend me (for environmental purposes) to only order clothing I intend to keep	Hong & Tam (2006), Mathieson (1991)
NI - Subjective injunctive norms	I think people who are important to me expect me to not order clothing online if I know that I will be returning some/all of it.	Gkargkavouzi 2009 Hong & Tam (2006)
Self-reported knowledge	Ordering more items of clothing online than one intends to keep, with the intention of returning some/all, has a negative effect on the environment	van der Werff & Steg (2016), Qualitative stage
Self-reported knowledge	I spend a lot of time reading about sustainability in general	Gkargkavouzi 2009
Instructional manipulation check	Please click "7 - Totally agree" on this statement to show you are paying attention	N/A
Self-reported knowledge	Overall, I act in an environmentally responsible way	Andreassen (2023), Gkargkavouzi (2009), van der Werff & Steg (2015, 2016)
Self-reported knowledge	Based on my previous online shopping, I have gained considerable insight into sustainability	Andreassen (2023)
Self-reported knowledge	Carbon emissions cause serious environmental problems, such as climate change	Coyle (2005), Gkargkavouzi (2009), van der Werff & Steg (2015, 2016), Arcury & Johnson (1987)
Self-reported knowledge	In general, I consider myself more interested in sustainability than the average online shopper	Andreassen (2023), Gkargkavouzi (2009)

Table 8-4: Adapted measures and sources for questionnaire items

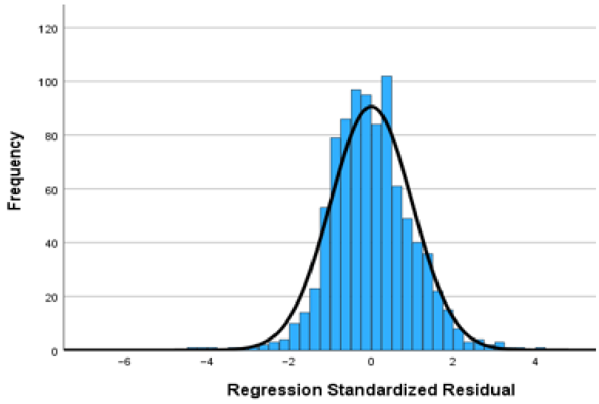
8.3.5 Histogram, Normal Q-Q plot and P-P plot, scatterplots



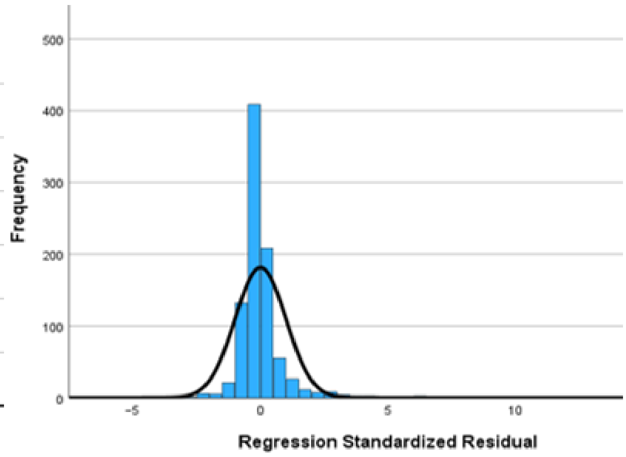
Run 1 – Sent returns, untransformed



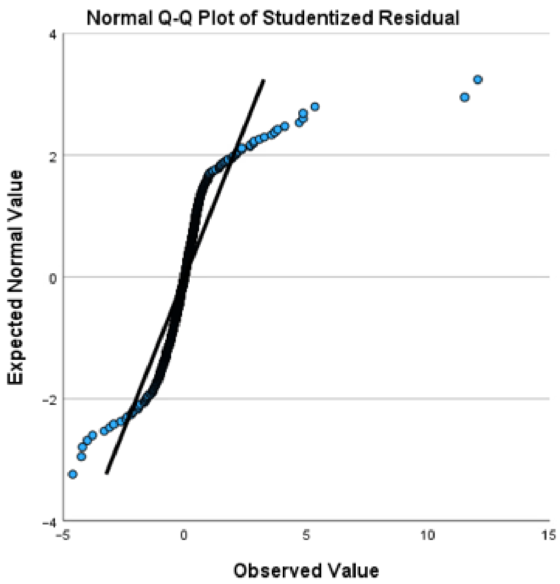
Run 2 – Sent returns, filtered



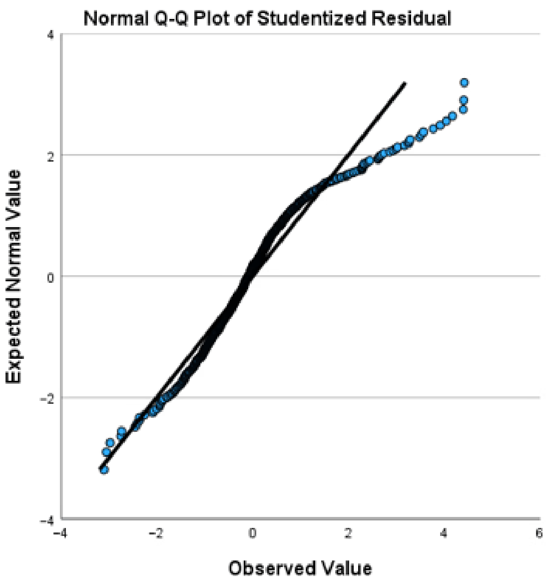
Run 3 - Sent returns, transformed



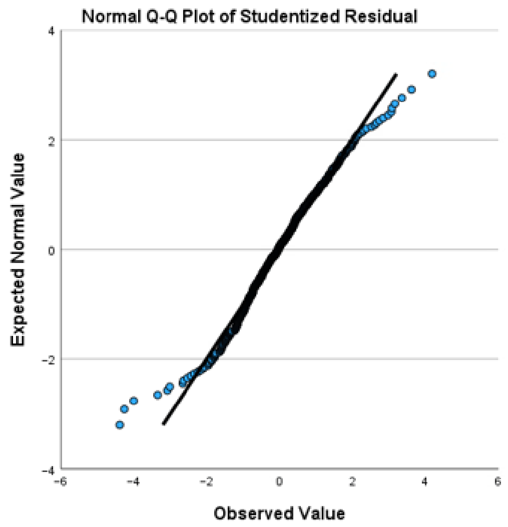
Run 4 – Physical returns, untransformed



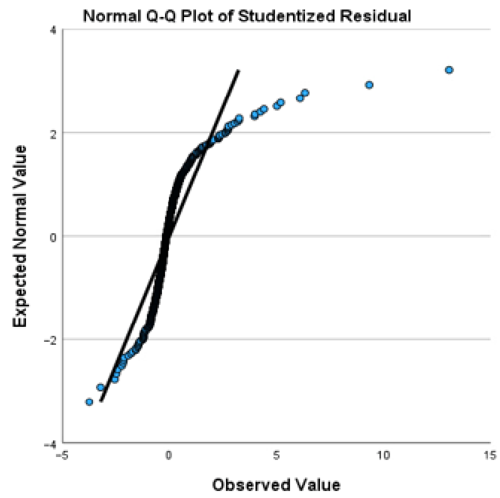
Run 1 – Sent returns, untransformed



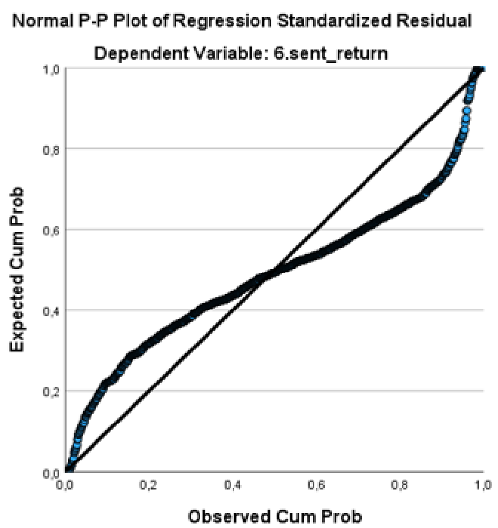
Run 2 – Sent returns, filtered



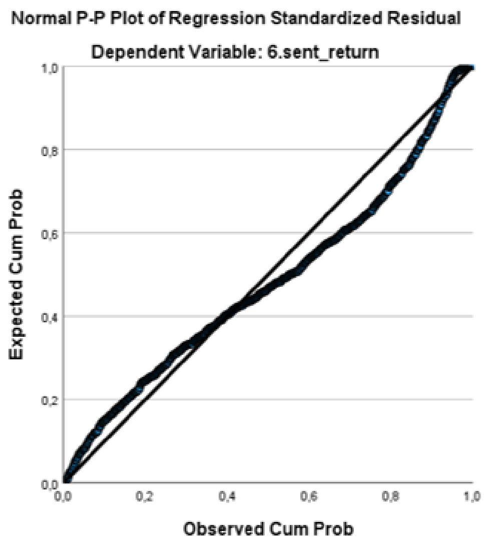
Run 3 - Sent returns, transformed



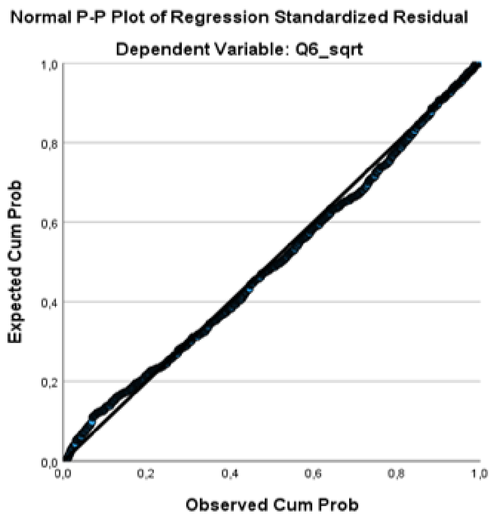
Run 4 – Physical returns, untransformed



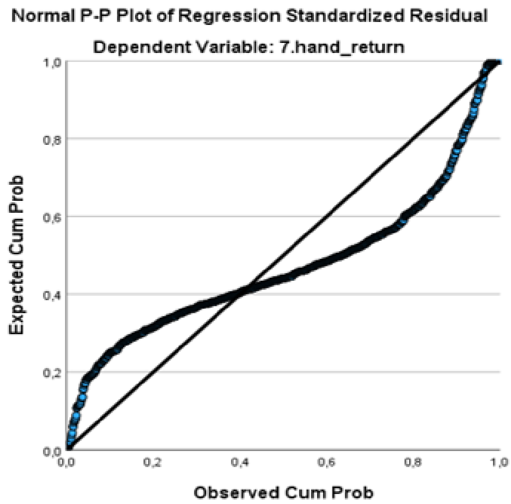
Run 1 – Sent returns, untransformed



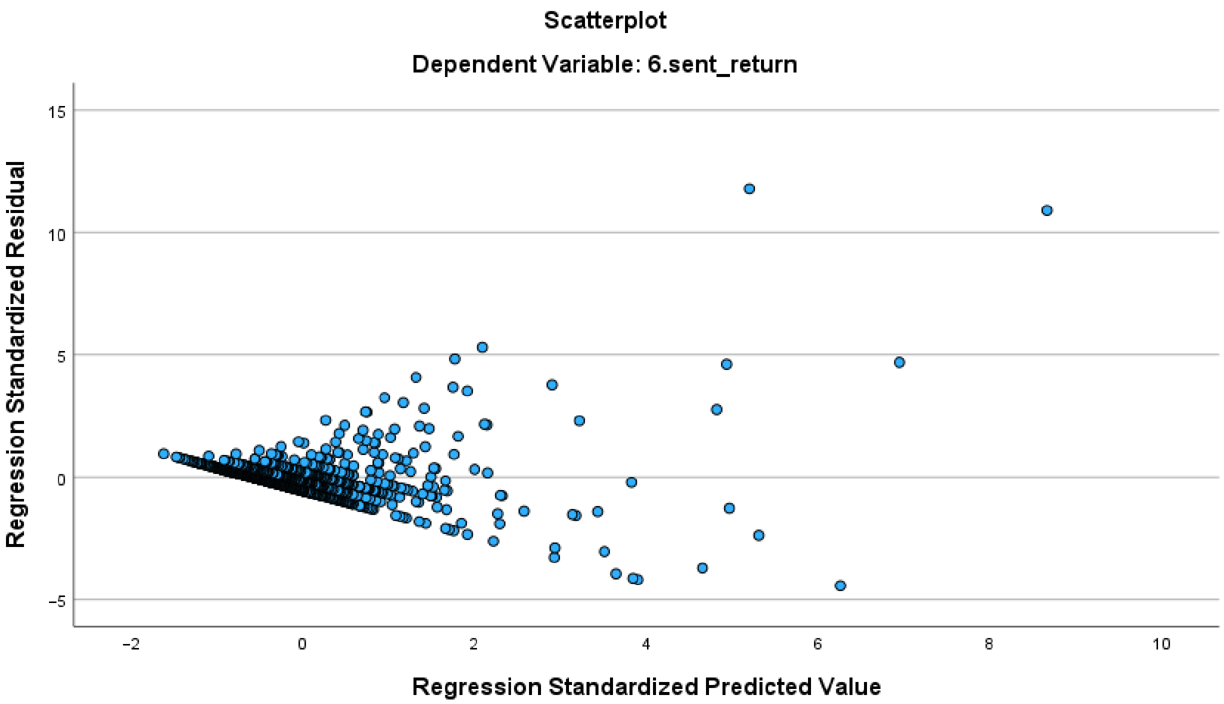
Run 2 – Sent returns, filtered



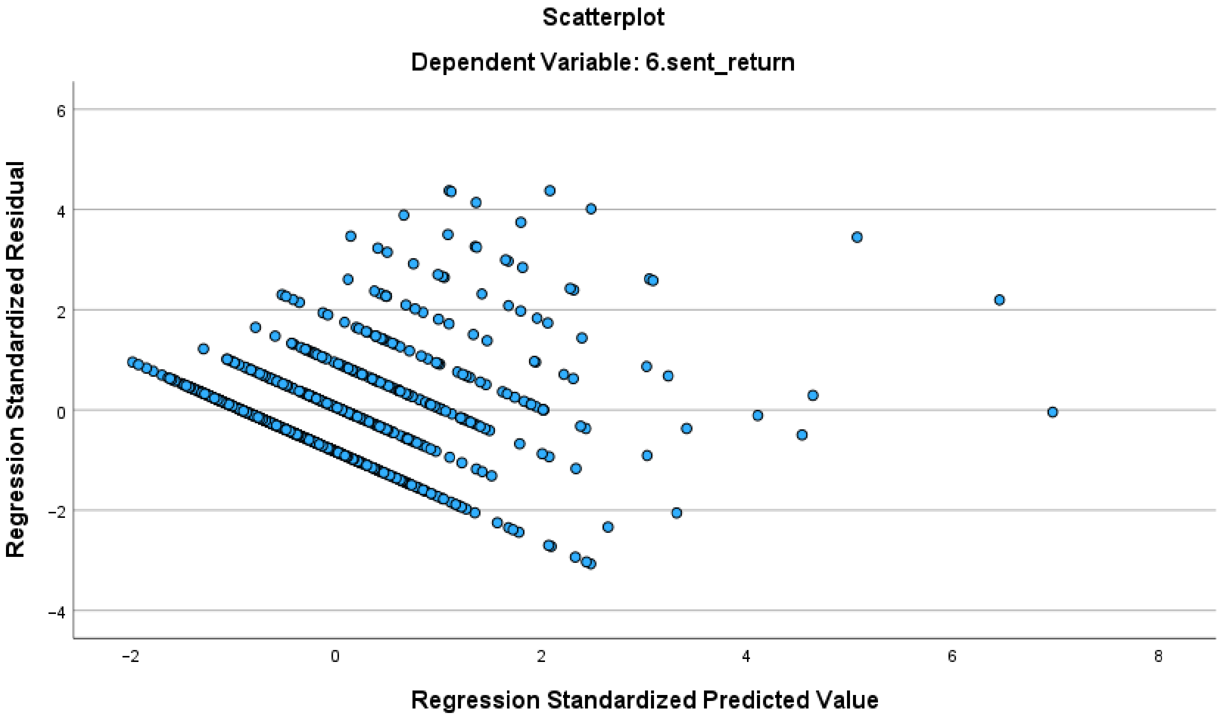
Run 3 - Sent returns, transformed



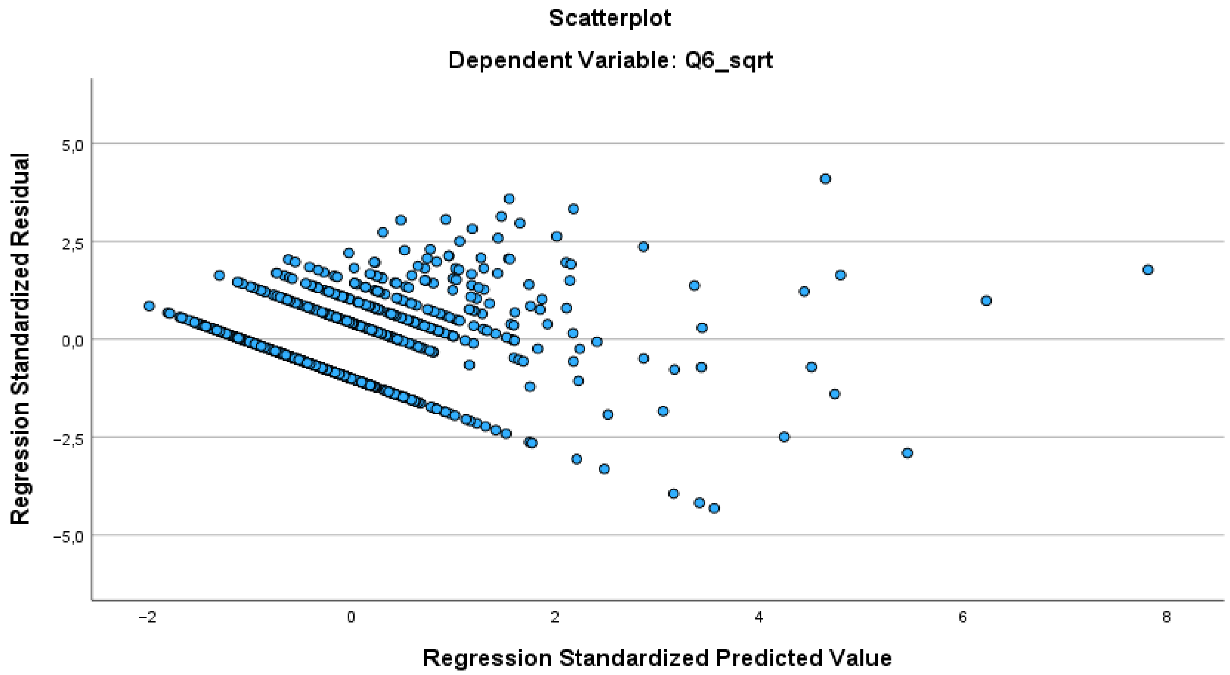
Run 4 – Physical returns, untransformed



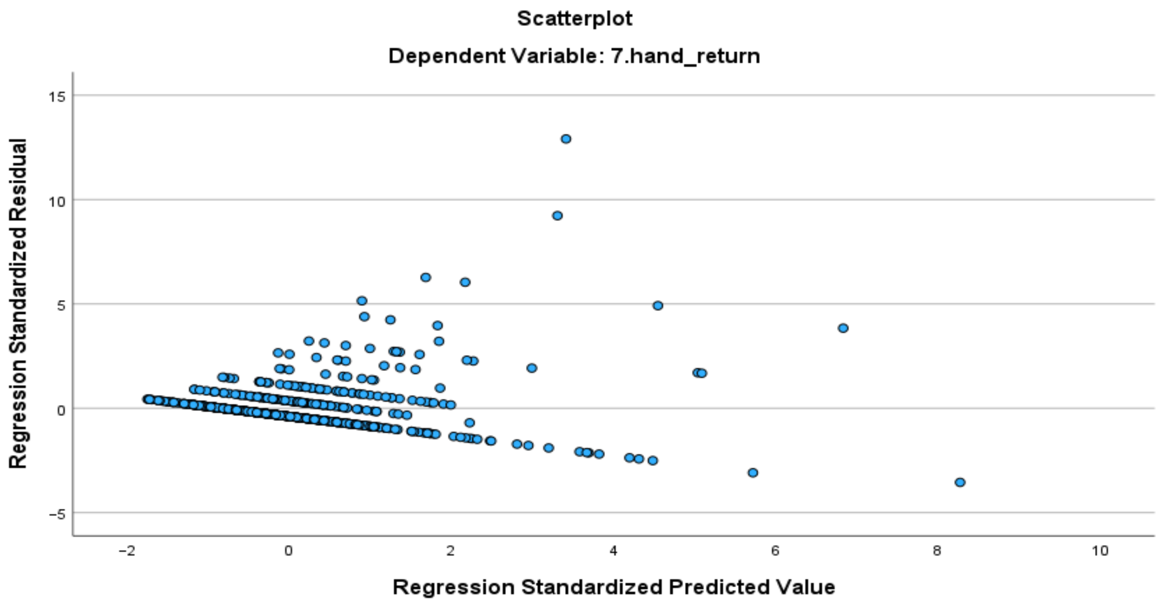
Run 1 Scatterplot



Run 2 Scatterplot



Run 3 scatterplot



Run 4 scatterplot

8.3.6 Regression results

Variable	Standardized Coefficients		95,0% Confidence Interval for B	
	Beta	p	Lower Bound	Upper Bound
Age	-0,092	0,007	0,0	-0,002
Sex_M/F	0,085	0,011	0,038	0,287
Ethnicity	-0,051	0,136	-0,092	0,013
Shopping volume	0,594	<,001	0,052	0,062
SRC, shopping and returns	0,267	<,001	0,176	0,261
SRC, product	0,03	0,321	-0,024	0,074
Objective descriptive norms	-0,119	<,001	-0,097	-0,03
Subjective descriptive norms	-0,066	0,039	-0,089	-0,002
Subjective injunctive norms	0,011	0,743	-0,038	0,053
Self-reported knowledge	-0,012	0,733	-0,064	0,045

Table 8-5 Regression 1, sent returns

Variable	Standardized Coefficients		95,0% Confidence Interval for B	
	Beta	p	Lower Bound	Upper Bound
Age	-0,08	0,02	0,0	-0,001
Sex_M/F	0,055	0,098	-0,03	0,356
Ethnicity	-0,005	0,882	-0,087	0,075
Shopping volume	0,381	<,001	0,047	0,066
SRC, shopping and returns	0,088	0,007	0,031	0,191
SRC, product	-0,048	0,2	-0,155	0,033
Objective descriptive norms	-0,076	0,054	-0,127	0,001
Subjective descriptive norms	-0,111	0,005	-0,202	-0,036
Subjective injunctive norms	0,023	0,592	-0,062	0,11
Self-reported knowledge	0,031	0,464	-0,065	0,142

Table 8-6 Regression 2, physical returns