Universiteit Leiden
The Netherlands

Informatica & Economie

The Impact of Sustainable Information on Customer Behaviour

Niels Cluistra

Supervisor:
Yingjie Fan

BACHELOR THESIS

Leiden Institute of Advanced Computer Science (LIACS)
www.liacs.leidenuniv.nl

09/08/2022

Abstract

Increasingly more products, which are ordered online, are being returned, which hurts the environment. In this thesis, the aim was to find a way to change customer behaviour concerning returning/ordering products. The research question therefore was: How can the online customer behaviour concerning returning or ordering products be influenced by adding sustainable information to the process? To research this, two surveys were constructed that show products and ask if the respondent would order/return them. The only difference was that one survey had an eco-label added to the product information. The data were analyzed using a chi-squared test, RFE and logistic regression. Results showed that when an eco-label is shown while ordering, more products are ordered. When an eco-label is shown on the package, fewer products are being returned. Logistic regression showed that age, favourite sport, hours spent on phone per day, conscientiousness and neuroticism influence behaviour when ordering. Furthermore, it showed that hours spent on phone per day, favourite holiday destination, environmental consciousness, openness and neuroticism influence behaviour when returning. Online retailers can benefit from these results getting a lower return rate.

Contents

1 Introduction

1.1 The situation

In the past years, the online market has grown in the Netherlands. In 2015 70% of the dutch citizens aged 12 years or older ordered products online, in 2019 this percentage had grown to 79% [CBS20]. Webshops offer free delivery and return to make sure customers order in their online store. These options have led to more returns of products. The dutch returned the highest percentage of their online orders in Europe in 2020 returning 13% of their orders [dpd20].

Offering these lenient return policies benefits retailers in terms of encouraging product purchase [JSF16]. However, while these return policies benefit retailers, they do not benefit the environment. In 2019 in the United States alone about 5.1 billion products were returned. This led to 5 billion pounds of waste from returns sent to landfills, which corresponds to more than three times the amount of waste the whole city of Seattle generates in a year. Not only does the waste hurt the environment, but also the CO^2 the transportation of these returns emits. In 2019 in the United States about 15 million metric tonnes of CO^2 were emitted transporting these returns [Opt19].

These numbers are expected to rise in the upcoming years making the problem even bigger. Optoro, a reverse logistics company, expects that in 2025 about 8 billion products will be returned resulting in 7.8 billion pounds of waste and 23 million tonnes of carbon emissions [Opt19]. To prevent these high numbers in the future, the behaviour of the consumers must be changed. In this thesis, the aim is to find a way to change this behaviour. This leads to the following research question:

How can the online customer behaviour concerning returning or ordering products be influenced by adding sustainable information to the process?

1.2 Thesis overview

This bachelor thesis made for Leiden Institue of Advanced Computer Science supervised by Yingjie Fan aims to find a solution to diminish carbon emissions by changing customer behaviour. This chapter contained the introduction. Section 2 will give some definitions of commonly used concepts in this thesis; Section 3 discusses related work; Section 4 explains how the experiment was performed; Section 5 describes the outcome of the experiments; Section 6 concludes the outcomes and gives options for further research.

2 Definitions

In this thesis, the aim is to nudge customers to a more sustainable choice by adding sustainable information to the process. In this chapter, the definitions of the complex concepts of nudging and sustainable information will be explained.

2.1 Nudging

To understand what nudging means. First, the concept of choice architecture must be explained. Choice architecture exhibits the fact that there are multiple ways to stage a choice for the decision-maker. The choice that is often made is dependent on how the choice is presented [JSD+12]. For example, the way the choice of becoming an organ donor is presented. In many states in the United States, you have to fill out a form to become an organ donor. Many people want to become organ donors, but they are too lazy to fill in these forms. By changing 'organ donor' into the default setting, more people become an organ donor [TS08].

So what is nudging? A nudge is any aspect of the choice architecture that amends people's behaviour in a way that is predictable and does not forbid other options. A nudge also cannot change economic incentives. For example, placing fruit at eye level is a nudge. However, forbidding people to eat unhealthy food is not [TS08].

2.2 Sustainable information

To understand what sustainable information is, first the meaning of sustainable must be explained. According to the Cambridge dictionary sustainable means "causing, or made in a way that causes, little or no damage to the environment and therefore able to continue for a long time" [dicnd]. Following this definition, sustainable information is information about how your choice could affect the environment. For example, when ordering your groceries at Albert Heijn (a Dutch chain of supermarkets), you are told that a certain choice is better for the environment.

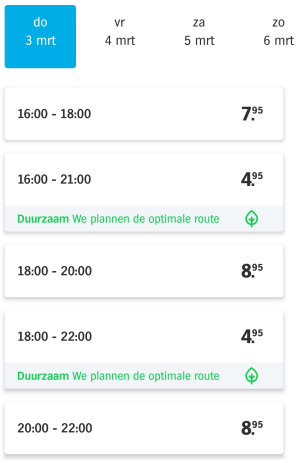


Figure 1: Albert Heijn shows which timeslot is better for the environment when you are ordering your groceries online

As can be seen in Figure 1, Albert Heijn also gives a monetary incentive to choose the more sustainable option. The prices of the sustainable options are €4,95 whereas the prices of the other options are €8,95 or €7,95. This means that it is not a real nudge. However, it is still a good example of what sustainable information is.

3 Related Work

As discussed in Section 1, the retailers are part of the problem. By installing lenient return policies, they encourage consumers to order more, which leads to more products being sent back. Customers also take advantage of these policies. For example, when ordering fashion items some customers order some products for fun to just try them on and then send them back because sending them back is free. They also order excess products to avoid service fees or dispatching [SSH17]. The environmental consequences of this customer behaviour are considerable, but customers do not show any tendency to change their behaviour. Furthermore, young people return the most products and as they will grow older and the new generations will likely copy this behaviour [CC21]. This unwillingness to change and copying of behaviour can explain why Optoro believes that in 2025 about 8 billion products will be returned in the United States [Opt19].

To tackle these problems customer behaviour needs to change. Study shows that the usage of labels affect decisions customers make [SW06] [SS09]. Sammer and Wüstenhagen [SW06] show that the EU eco-label does affect customers. They are willing to pay more for a product with an A-status than for a product with a C-status. Furthermore, Shen and Saijo [SS09] show that energy efficiency ranks presented on an energy label have a significant effect on the choice of customers. So the usage of labels affects the choices customers make. On the other hand, a study shows that although labels increase awareness and intentions, there has to be a clear connection between the certified product and an environmental outcome. When this is not the case, labels do not necessarily alter the purchasing behaviour of customers [MHKK+21]. So just putting a label on a package is not enough, if this label does not give a connection to the environmental outcome.

The way a package looks influences the purchase intentions of consumers. Each element of packaging influences the purchase intentions of customers significantly [WKA18]. Packaging shows a significant positive relation with brand loyalty [DMD14]. Brand loyalty is a measure of the unwillingness of a consumer to convert to using a competing service or product [Refnd]. So the question rises if changing your package by adding an eco-label does affect your brand loyalty. If this would be the case, online retailers probably will not be adding these labels to their packages for fear of losing clients. However, adding environmental propaganda to a package positively affects consumer behaviour. Furthermore, the decline in brand loyalty caused by the change of the package is alleviated by using a package with high environmental propaganda satisfaction [GCW+22].

The eco-labels can be seen as green nudges. The difference between 'normal' nudges and green nudges is that nudges aim to improve the welfare of an individual, whereas green nudges aim to reduce the impact of a negative environmental externality [CGJSK21]. In the case of my study, the added sustainable information is the green nudge. For this green nudge to be effective and ethical, it must be organized in a transparent way [Sch17]. Additionally, a study shows that for

communication to customers you could see the sustainability of products as a continuum to position products at different levels [TH21].

Labels are effective [SW06][SS09], so eco-label can be effective as well. Packages design influences the choices of consumers [WKA18], so putting a label on a package could also influence the choices of consumers. Green nudges are effective when transparent and the communication is placed at a continuum [Sch17][TH21], so when the eco-label is transparent and placed at a continuum it could be effective. As a consequence of these facts, the following hypotheses arise: (they hypotheses are null hypotheses for a chi-squared test, so they are negative instead of positive)

H_0: *The usage of an eco-label on a package does not influence online customer behaviour concerning ordering products*

H_1: *The usage of an eco-label on a package does not influence online customer behaviour concerning returning products*

Also the following hypotheses were analyzed:

H_2: *The age of a respondent does not influence online customer behaviour concerning ordering/returning products*

H_3: *The income of a respondent does not influence online customer behaviour concerning ordering/returning products*

H_4: *The amount of products a respondent orders per month does not influence online customer behaviour concerning ordering/returning products*

Figure 2 describes the research framework. It shows the features a customer is built on in this research and the focus of this study, which is to check if the usage of an eco-label influences customer behaviour. Furthermore, it shows the other hypotheses that check whether age, income and number of products an individual orders per month influence customer behaviour concerning ordering/returning products.

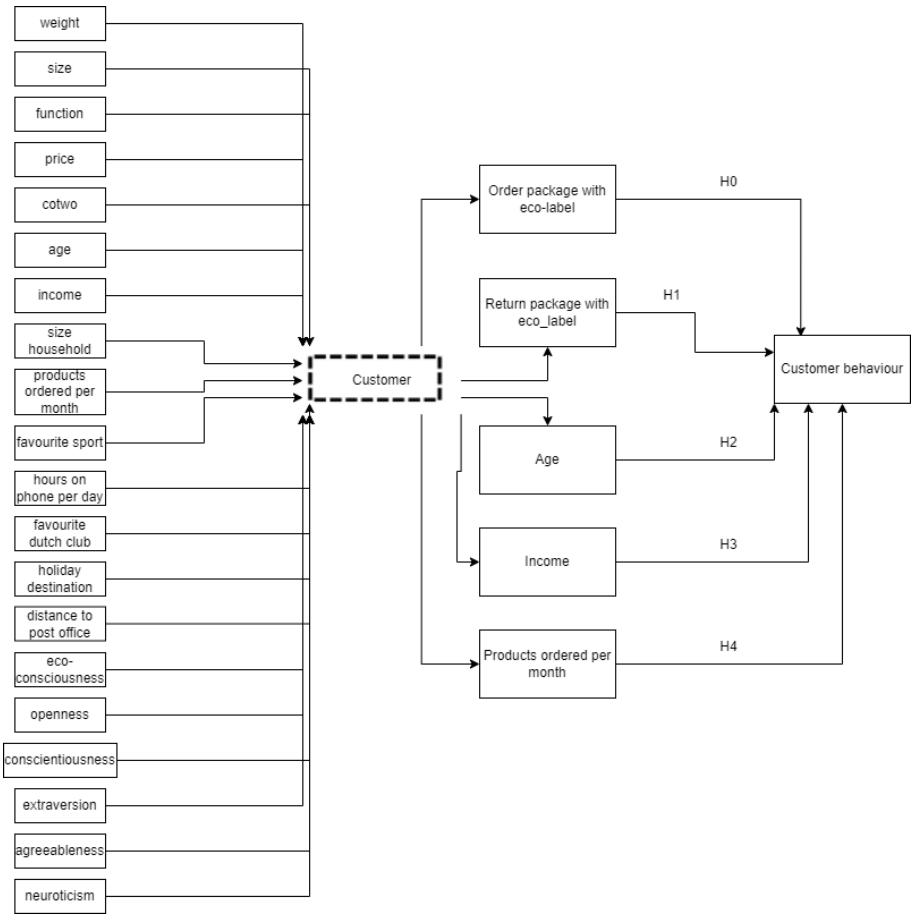


Figure 2: Research framework

4 Methodology

4.1 Experiment design

Two surveys were created to determine if people choose differently when sustainable information is added. The surveys show different products with their price, weight and length, width and height. Furthermore, a reason why you doubt about ordering or sending back the product was added. The surveys have one important difference. The first survey also includes how many grammes of CO^2 will be emitted by sending back the product. As this is the only different variable, the idea is that when more respondents would not send the product back when knowing how that would influence the environment, then adding sustainable information to the process helps.

The data was collected by sending the surveys as a link via WhatsApp. The surveys were sent to friends, family and contacts of my mother (to make sure the population was more diversified). Everyone received only one of the two surveys to make sure no one would submit answers to both surveys. Furthermore, people who knew what was the exact goal behind the two surveys were not invited to fill them in. This way, bias would be prevented. Otherwise, some people might answer 'No' more often than they would, when they do not know the goal, in the survey with label or answer

'Yes' more often than they would, when they do not know the goal, in the survey without label. This would have led to incorrect results. Moreover, when sending out the surveys the population was held in mind to make sure the population was diversified.

The products used in the surveys were the following for ordering products:

- Dumbbells
- A game
- A dreamcatcher
- A book
- An anti-bird mechanism
- A lego storage head
- A grilling iron
- A carpet
- A foam clay set
- An instax camera
- A planter
- A slide

The products used in the surveys were the following for returning products:

- A flashligt
- A book
- An inflatable doll
- A crate
- A wallet
- A sword
- Chains (for your bike)
- A phone set
- A powerbank
- A beamer
- Dumbbells
- A greenhouse

4.2 Building the surveys

The survey with label was built the following way:

1. An introduction of the research is given;
2. 12 different products are shown with the question if the participant would **order** the product
3. The price of the product is given. There are four different price ranges each containing three products. The price ranges were €5 - €7, €17 - €22, €47 - €52 and €97 - €102. This is done to see if people would react differently to different prices;
4. The size of the package is given by showing the length, width and height of the package. This is done to see if participants act differently when it comes to bigger/smaller packages;
5. The weight of the package is given (in kilograms). This is done to see if respondents would react in another way if the product weighs more/less;
6. The participants are told that delivering and returning is free, so there is no economic incentive to not order or return the product;
7. The respondents are told how many grams of CO^2 is emitted by returning the package. T These numbers were found by using the website *ecotransit.org*, where the weight of the package (retrievable from the website where I found the product) and the distance the package had to go needed to be entered. This distance differs for every package, but in this process, the distance between Bodegraven, a village in South Holland, and Utrecht was used, which is about 36 kilometres by car. Using this distance for calculating the amount of CO^2 emitted for every package made sure that the amounts were found using the same scale. This would lead to these amounts being on a continuum, which is a requirement for good communication to customers [TH21];
8. To make sure that there is a clear connection between the product and the environmental outcome, the amounts of CO^2 were connected to the time it takes in the shower to emit this amount of CO^2. Without this, the label does not necessarily alter the behaviour of customers [MHKK+21]. To do this, one gram of CO^2 emitted was equalled to showering for two minutes. This created a continuum to place the products on and made sure there was an understandable link between the CO^2 emitted and the environmental outcome of these emissions.
9. The respondents are given a reason why they would not order the product. For example, as seen in Figure 3 the survey tells the participant "It is nice to use dumbbells for doing sports at home, but you usually go to the gym". This indicates that they might not need the product;
10. 12 different products are shown with the question if the participant would **return** the product
11. Again the price (with the same price ranges), size, weight and amount of CO^2 emitted are given
12. The respondent is given a reason why he/she might want to return the product;

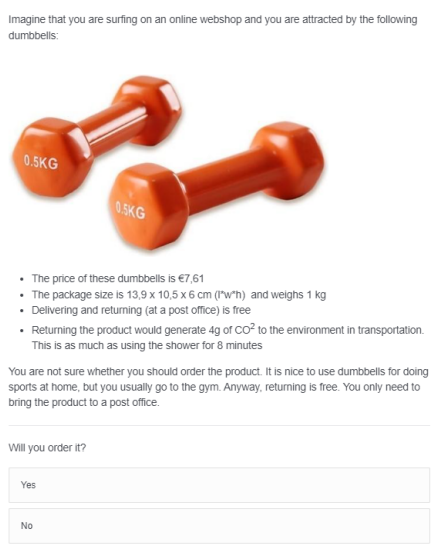
13. Personal questions were added to the survey. These questions were asked to find out if subgroups act differently. For example, if people who are older order more or fewer products;

14. Questions about the environmental consciousness of the respondents were asked. These questions were copied from a Turkish study [EY15] Only the questions about environmental awareness were used;

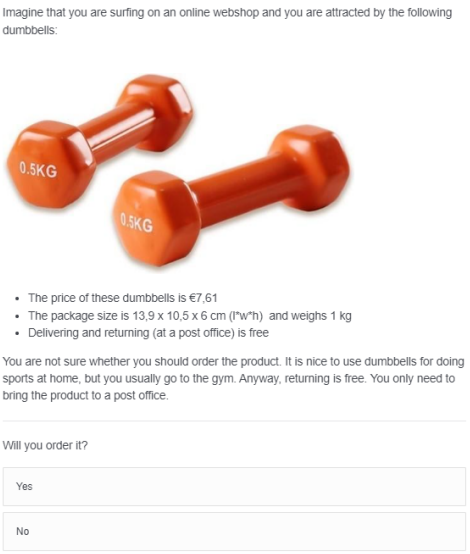
15. Big five personality test statements were added. These statements were copied from a real test on this subject [ptnd]. These were the following statements:

    - I have a rich vocabulary;
    - I have a vivid imagination;
    - I have excellent ideas;
    - I am always prepared;
    - I pay attention to details;
    - I get chores done right away;
    - I am the life of the party;
    - I feel comfortable around people;
    - I start conversations;
    - I am interested in people;
    - I sympathize with others' feelings;
    - I have a soft heart;
    - I am relaxed most of the time;
    - I seldom feel blue;
    - I get upset easily.

The survey without a label is built the same as the survey with label. The only difference is that in this survey the amount of CO^2 emitted when returning the package is not told, so as the name indicates, there is no eco-label in the survey without label. Figures 3 and 4 respectively show a question about ordering products and returning products from the two different surveys.

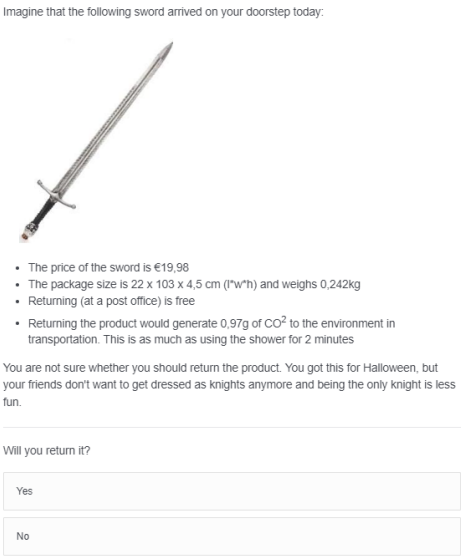


(a) Question with eco-label

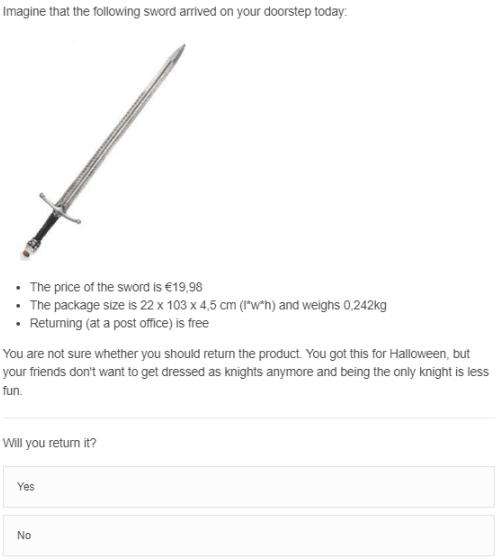


(b) Same question without eco-label

Figure 3: Example of a question about ordering in the different surveys



(a) Question with eco-label



(b) Same question without eco-label

Figure 4: Example of a question about returning in the different surveys

4.3 Analyzing the results

The surveys had a different number of participants. The survey with an eco-label got 34 answers and the survey without an eco-label got 32 answers. The results of the surveys were importable from Qualtrics, the software for making the surveys, as a .CSV file. This file was then loaded into Python as a data frame using the Pandas package. After cleaning the data, a chi-squared test was conducted as well as Recursive Feature Elimination to analyze which features affect what choices the respondents make. Furthermore, logistic regression is applied to analyze the data. Moreover, a

small analysis is performed on what respondents would do with the product when they would not return it.

4.3.1 Chi-squared test

As explained, the surveys only differ in the fact that one survey has an eco-label and the other does not. A chi-squared test tests whether the findings of a certain series are statistically correct assuming your hypothesis is true. In the case of this research the chi-squared test tests if the difference between the answers of the respondents can be traced back to the eco-label. This test is performed for both ordering products and returning products. The formula for the chi-squared test is the following:

$$\chi_c^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

C stands for degrees of freedom, which is calculated by $(r-1)(c-1)$, where r stands for rows and c stands for columns. In this case, there was one row with two numbers. The first number was the sum of the number of times the participants indicated they would order/return the product. The second number was the sum of the number of times the respondents indicated they would not order/return the product. So the degrees of freedom were equal to zero because there was one row and two columns. O_i and E_i mean the observed values and the expected values. The expectation was that there would be no difference, so the expected values were the values from the survey without a label. The observed values were the values from the survey with a label. The alpha level was set at 5%.

Furthermore, the hypotheses mentioned at the beginning of this section about age, income and amount of products ordered per month are also analyzed using the chi-squared test. To perform this analysis, the following process was conducted:

1. As the hypothesis states that age/income/products ordered per month (from now on called the features) do not influence the behaviour, the E_i (expected values) are the distribution of the sum of the number of times the respondents indicate they would (not) order/return the products because the expectation (of the hypothesis) is that for example, the behaviour of younger people does not differ from the behaviour of the whole group;

2. The subgroups of the features are cut in half. For example with the feature age, the data of the subgroups 15-30 and 30-45 and the subgroups 45-60 and 60+ are combined. These data are both used as O_i (observed values);

3. Firstly, the first half of the observed data is used in the chi-squared test and then the second half is used. After this process, two p-values are found. If both of these p-values are lower than the alpha-level (which is set at 5%), then the feature has a significant effect on the behaviour. So there could be concluded that the different behaviour of a subgroup is due to being in that subgroup.

4.3.2 Recursive Feature Elimination

Recursive Feature Elimination, RFE, repeatedly constructs a model and then chooses the best or worst performing feature. RFE sets these features aside and continues with the operation until all

features in the dataset are exhausted [Li17]. In this research, the features used by the RFE operator were the personal questions, environmental consciousness questions and personality trait questions. Thus the features found have the most influence on the choices made by the participants of the survey.

After these features were found, for each of these features a graph was made indicating the average percentage of the respondents within a subgroup of the feature indicated they would not order/return the product. This way the differences between these subgroups can be shown.

4.3.3 Logistic Regression

Logistic regression is a classification algorithm in machine learning. It is used for predicting the probability of categorical dependent variable [Li17]. In the case of this study, it uses the answers to the personal questions, personality test and questions about ecological consciousness to predict the choice a person makes. The choice is to (not) order/return the product. The logistic regression is only performed on the data from the survey with an eco-label. The reason for this is that the goal of this research is to find out how people act when an eco-label is added. So why individuals act how they act when there is an eco-label matters the most to this research.

To perform this regression, first, the data had to be pre-processed. The price (in euros), size (in cm^3), weight (in kilograms), functionality (sports, entertainment, useful, kitchenware and decoration) and amount of CO^2 emitted (in grams) were added to the data. Moreover, the answers to the question about environmental consciousness were combined into one score. These questions were a statement and the participant had to answer how often they do it. For example, 'I throw used batteries in waste collection boxes'. Then the participant had the following choices:

- Never
- Rarely
- Sometimes
- Often
- Every time

These choices all got a score from one to five. One being 'Never' and five being 'Every time'. These scores were summed together for every respondent to get an environmental consciousness score per participant. These scores were added as a column to the data.

Furthermore, the data from the big five personality test were turned into scores. The five traits of this test are 'Openness', 'Consciousness', 'Extraversion', 'Agreeableness' and 'Neuroticism'. For each of these traits, five questions were added. These questions consisted of a statement and then the respondent had to answer the extent to which they agreed with this statement. For example, 'I have a rich vocabulary. Then the participant had the following options:

- Totally disagree
- Disagree

- Neutral
- Agree
- Totally agree

These choices all got a score from one to six. One being 'Totally disagree' and six being 'Totally agree'. These scores were summed together per participant per trait and were then added to the data as columns.

After this process, the regression was conducted by Python using the package statsmodels. The regression is conducted for both the ordering and the returning of the products. The results consist of six scores given to all the features that are put into the regression analysis. These are the following scores: Coefficient, Standard Error, z, $P > |Z|$, [0,025 and 0,0975].

The coefficient is the value of the intercept. It shows how a change in each variable affects the independent variable. The standard error is the standard deviation of the coefficient. The z stands for the z-score and is a measurement of the precision of the coefficient. The most important score is the $P > |Z|$ score. This score measures how likely your coefficient is measured through the regression model by chance. For example, if the $P > |Z|$ score of age is equal to 0,456, it means that there is a 45,6% change the age variable does not affect the dependent variable, which is returning/ordering the product in this case. [0,025 and 0,975] both measure values of the coefficients within 95% of the data. It could also be within two standard deviations. The values outside these values are generally considered outliers [McA20].

When the $P > |Z|$ score of a feature is lower than the alpha-level, which is 5% (0,05), there can be said that the feature influences the behaviour. However, if you remove a feature, the coefficients for other features may change. For this reason, after every regression, the values with a $P > |Z|$ score over 0,3 are removed because these features have a bad score, which may influence other features to perform worse than they could. If a feature has a $P > |Z|$ score below 0,05, but after removing other features has a score above 0,05, this feature is still counted as a feature that influences the behaviour.

5 Results

The populations of the surveys were the following:

	Number of respondents
Total	34
Age	
15-30	23
31-45	2
46-60	5
60+	4
Income	
20.000 <	21
20.000-35.000	3
35.000-50.000	4
> 50.000	5
Won't tell their income	1
Products ordered per month	
0	4
1-3	23
4-6	6
7-9	0
> 10	1
Size of household (incl. respondent)	
1	9
2	10
3	4
4	7
> 4	4

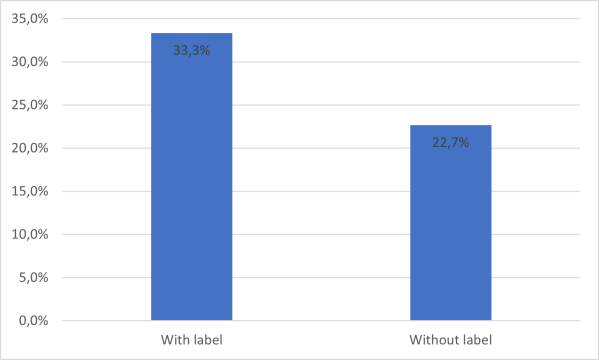
Table 1: The population of the survey with label

	Number of respondents
Total	32
Age	
15-30	20
31-45	0
46-60	9
60+	3
Income	
20.000 <	17
20.000-35.000	5
35.000-50.000	3
> 50.000	5
Won't tell their income	2
Products ordered per month	
0	0
1-3	31
4-6	0
7-9	0
> 10	1
Size of household (incl. respondent)	
1	4
2	9
3	7
4	1
> 4	11

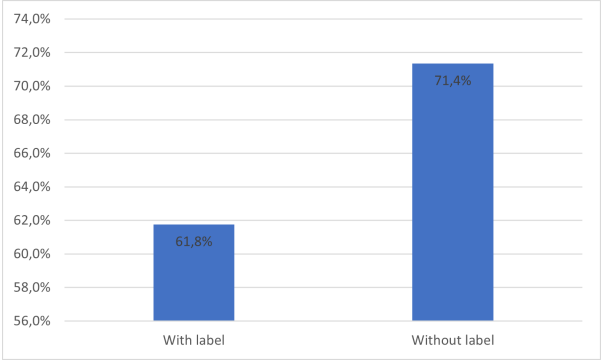
Table 2: The population of the survey without label

5.1 Chi-squared test and RFE

In this section, the results from the experiments will be shown starting with the chi-squared test. The p-value and test-statistic will be shown. A p-value smaller than the alpha level means that the hypothesis will be rejected and there is a connection. The test statistic assesses the consistency of the sample data with the null hypothesis. It is a single value that indicates the variation between your sample and the null hypothesis. The more extreme this value becomes, the larger the differences between your data and the null hypothesis will be [Frond]. The test statistic and p-value of the test conducted on the data about orders are respectively 29,70 and $5{,}05e^{08}$. The test statistic and p-value of the test conducted on the data about returning are respectively 21,17 and $4{,}21e^{06}$. Figures 5a and 5b show what percentage of the respondents would not order/return the product. This is expressed in percentages instead of absolute numbers because the number of respondents differs per survey. The results are shown separately because people could behave differently when it comes to ordering or returning products.



(a) Order the product; p-value of $5.05e^{08}$



(b) Return the product; p-value of $4.21e^{06}$

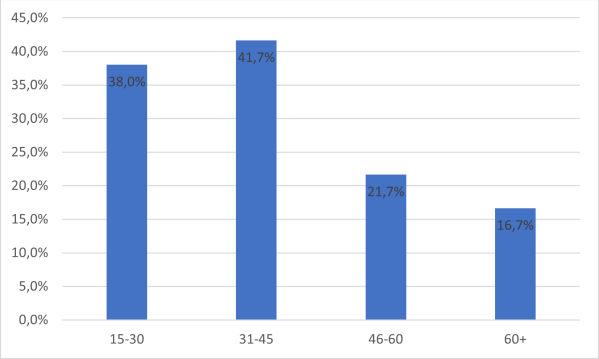
Figure 5: Average percentage of respondents that would order/return the product

The features RFE had found to perform the best were age, income and number of products ordered online per month. Of these features, multiple bar charts have been made showing the average percentages of how many respondents within a subgroup of the feature would not order/return the product. The Figures 6a and 6b, which are shown underneath, indicate how many respondents aged a certain age would respectively order or return the product. The Figures only used the data from the survey with label, because this data shows the behaviour when sustainable information is added and this is the behaviour that is researched. This also applies to Figures 7 and 8.

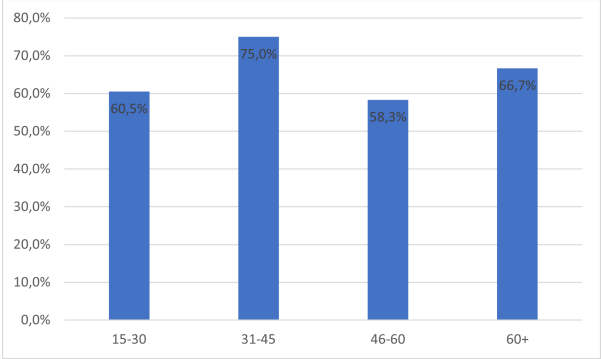
The data was split in half at the age of 45 to calculate the p-value. So one part consisted of 15-45 and the other of 45-60+. These age groups were compared with the normal distribution of the data to check if age does affect the choice of the respondents. The results are shown in Table 3.

	Ordering		**Returning**	
	Aged below 45	Aged above 45	Aged below 45	Aged above 45
Test statistic	30.40	247.40	27.02	243.02
p-value	$3{,}52e^{08}$	$9{,}59e^{09}$	$2{,}01e^{07}$	$8{,}62e^{055}$

Table 3: The test statistics and p-values for ordering and returning products by age



(a) Order the product



(b) Return the product

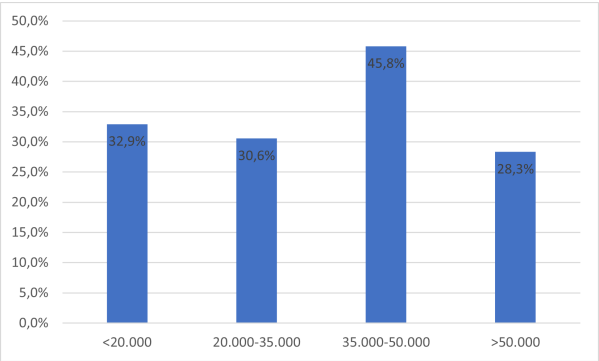
Figure 6: Average percentage of respondents that would order/return the product by age

In Figures 7a and 7b the same is shown. However, in these graphs, the results are ordered by income. The income is in euros. The option 'Rather not tell/I don't understand the question is not shown in these charts, because only one respondent chose this option as shown in Table 1. This would then lead to only 0% or 100% as scores, which probably would not be representable if there would have been more people choosing this option. So this subgroup was left out.

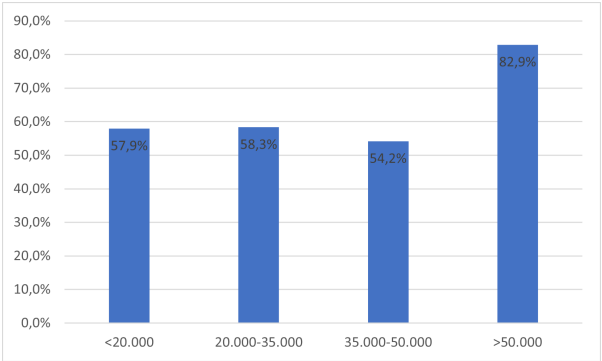
The data was split in half at the income of €35.000 to calculate the p-value. So one part consisted of the data to the income of €35.000 and the other part of the data from the income of €35.000. These income groups were compared with the normal distribution of the data to check if income does affect the choice of the respondents. The results are shown in Table 4.

	Ordering		**Returning**	
	Income below 35.000	Income above 35.000	Income below 35.000	Income above 35.000
Test statistic	33.29	244.29	33.12	245.12
p-value	$7{,}95e^{09}$	$4{,}57e^{055}$	$8{,}66e^{09}$	$3{,}01e^{055}$

Table 4: The test-statistics and p-values for ordering products by income



(a) Order the product



(b) Return the product

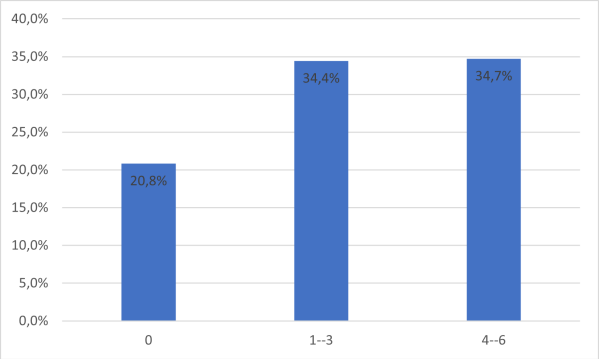
Figure 7: Average percentage of respondents that would order/return the product by income

In Figures 8a and 8b the same is shown. However, in these graphs, the results are ordered by how many products the respondent orders per month. Here the options '7-9' and '10+' are not shown. Not a single one of the participants chose the option '7-9' packages, which means there was no data on this option. Furthermore, only one participant was indicating that he/she orders more than ten products in a month. This would then lead to only 0% or 100% as scores, which probably would not be representable if there would have been more people choosing this option. So also this subgroup was left out. These numbers are also shown in Table 1.

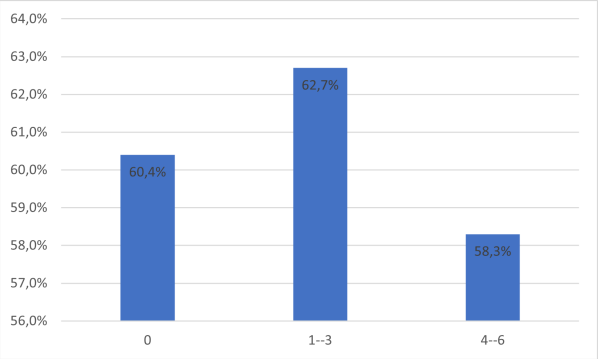
The data was split in half at ordering four products per month to calculate the p-value. So one part consisted of people ordering less than five products per month and the other of people ordering more than four products per month. These groups were compared with the normal distribution of the data to check if the number of products you order per month does affect the choice of the respondents. The results are shown in Table 5.

	Ordering		**Returning**	
	Less than 4 products	More than 4 products	Less than 4 products	More than 4 products
Test statistic	17.89	315.89	17.93	314.93
p-value	$2{,}33e^{05}$	$1{,}14e^{070}$	$2{,}29e^{05}$	$1{,}84e^{070}$

Table 5: The p-values for ordering products by products ordered per month



(a) Order the product



(b) Return the product

Figure 8: Average percentage of respondents that would order/return the product by products ordered per month

5.2 Logistic regression

When a $P > |Z|$ score is smaller than 0,05, 0,01 or 0,001, this is shown in Table 6 by * (smaller than 0,05), ** (smaller than 0,01) and *** (smaller than 0,001). When a previous score is above 0,3, this score is removed and this is shown in the table by not showing a new coefficient. When the $P > |Z|$ score of a feature is lower than the alpha level at a certain point, but after removing other features the $P > |Z|$ score is higher than the alpha level, this feature is still seen as a feature that influences the behaviour. Furthermore, Table 6 shows the coefficients of the features that influence the behaviour. A logistic regression model looks like the following formula: $log(p/1-p) = a + bX_1 + cX_2 + (restoftheequation)$. The a stands for the intercept/constant. Furthermore, the X stands for the predictor variable and b and c are the coefficients that correspond with these variables [Jan21]. If the coefficient is positive, the higher the score of the feature, the more likely the respondent does want to order/return the product. If the coefficient is negative, it means that the higher the score of the feature, the less likely the respondent does want to order/return the product. For example, age, if the coefficient is positive, it would mean that the older you are the more likely you want to order/return the product. The scale of the features is shown in Table 7.

	Model 1	Std. Error	Model 2	Std. Error	Model 3	Std. Error	Model 4	Std. Error
constant	-0.824		-0.403		-0.301		-0.442	
weight	-89.434	123.295						
size	0	0						
function	0.119	0.081	0.122	0.073	0.122	0.073	0.123	0.073
price	-0.007	0.005	-0.003	0.003	-0.003	0.003		
cotwo	22.368	30.822						
age	0,387*	0.179	0,347*	0.162	0.273	0.145	0.274	0.145
income	-0.167	0.0156	-0.143	0.139				
size household	-0.092	0.092						
products ordered per month	-0.199	0.187	-0.234	0.168	-0.221	0.167	-0.22	0.167
favourite sport	-0,087*	0.038	-0,073*	0.033	-0,074*	0.033	-0,074*	0.033
hours on phone per day	0,494**	0.179	0,412**	0.149	0,465***	0.141	0,461**	0.141
favourite dutch club	0.047	0.081						
holiday destination	-0.061	0.099						
distance to post office	-0.378	0.236	-0.258	0.212	-0.241	0.210	-0.24	0.210
eco-consciousness	0.027	0.033						
openness	0.105	0.068	0.093	0.058	0.084	0.057	0.079	0.057
conscientiousness	-0.161	0.088	-0.124	0.071	-0,138*	0.070	-0,139*	0.070
extraversion	0.023	0.077						
agreeableness	-0.05	0.076						
neuroticism	-0.12	0.099	-0.133	0.072	-0,147*	0.070	-0,154*	0.070
Pseudo R-Squared	0.068		0.058		0.056		0.054	

Table 6: The coefficients for ordering products; * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$

Feature	Scale
weight (in kg)	0.044 - 18
size (in cm3)	184.8 - 143.276
function	sports; entertainment; decoration; useful; kitchenware
price (in euros)	5,87 - 99,99
cotwo (in grams)	0.178 - 72
age	15-30; 31-45; 46-60 ; 60+
income	< 20.000; 20.000-35.000; 35.000-50.000; > 50.000
size household	1; 2; 3; 4; > 4
products ordered per month	0; 1-3; 4-6; 7-9; 10+
favourite sport	football; tennis; ice skating; basketball; athletics; hockey; korfball; volleyball; I don't like sports
hours on phone per day	< 1; 1-2; 3-4; 4-5; > 5
favourite dutch club	Feyenoord; Ajax; PSV; AZ; I don't have a favourite dutch club
holiday destination	camping; hotel; B&B; cottage; all-inclusive resort
distance to post office	$< 1000m$; 1000m-5000m; 5000m-10.000m; $> 10.000m$
eco-consciousness	8-40
openness	3-15
conscientiousness	3-15
extraversion	3-15
agreeableness	3-15
neuroticism	3-15

Table 7: The scale of the features

When a $P > |Z|$ score is smaller than 0,05, 0,01 or 0,001, this is shown in Table 8 by * (smaller than 0,05), ** (smaller than 0,01) and *** (smaller than 0,001). When a previous score is above 0,3, this score is removed and this is shown in the table by not showing a new coefficient. All scores below the alpha level are shown in bold. When the $P > |Z|$ score of a feature is lower than the alpha level at a certain point, but after removing other features the $P > |Z|$ score is higher than the alpha level, this feature is still seen as a feature that influences the behaviour. Furthermore, Table 8 shows the coefficients of the features that influence the behaviour. Table 7 shows the scale of the features.

	Model 1	Std. Error	Model 2	Std. Error	Model 3	Std. Error
constant	1.507		1.442		1.555	
weight	-117.259	124.448				
size	0	0				
function	0.156	0.081	0.110	0.073	0.111	0.073
price	0.004	0.005				
cotwo	29.32	31.110				
age	0.239	0.189	0.232	0.183	0.195	0.176
income	-0.172	0.156	-0.154	0.148	-0.164	0.147
size household	-0.118	0.092	-0.132	0.084	-0.132	0.084
products ordered per month	-0.096	0.178				
favourite sport	-0.04	0.037	-0.031	0.034		
hours on phone per day	-0,344*	0.169	-0,334*	0.159	-0,365*	0.156
favourite dutch club	0.002	0.080				
holiday destination	0,234*	0.095	0,229*	0.090	0,239**	0.090
distance to post office	0.149	0.230				
eco-consciousness	-0,067*	0.033	-0,064**	0.024	-0,073**	0.023
openness	0,163*	0.067	0,174**	0.062	0,165**	0.061
conscientiousness	-0.129	0.083	-0.143	0.073	-0.132	0.072
extraversion	-0.089	0.078	-0.101	0.058	-0.088	0.056
agreeableness	-0.014	0.078				
neuroticism	0,315**	0.101	0,332**	0.091	0,337**	0.091
Pseudo R-Squared	0.089		0.083		0.081	

Table 8: The coefficients for returning products; * $p < 0,05$; ** $p < 0,01$

6 Discussion and conclusion

6.1 Discussion

What do these results imply? The test statistic and p-value of the test conducted on the data about orders are respectively 29,70 and $5{,}05e^{08}$. The test statistic and p-value of the test conducted on

the data about returning are respectively 21,17 and $4{,}21e^{06}$. These values are found by performing a chi-squared test in Python with degrees of freedom equal to zero and an alpha level set at 5% (0,05). In this research, this p-value indicates whether the difference between people's choice in the survey with an eco-label and the survey without an eco-label is significantly different. The p-values are smaller than the alpha level, $5{,}05e^{08}$ to 0,05 and $4{,}21e^{06}$ to 0,05. This means that the null hypotheses can be rejected. So the usage of an eco-label on a package does influence customer behaviour concerning ordering or returning packages.

Looking at Figures 5a and 5b, it can be seen that the eco-label influences customers differently when it comes to returning and ordering products. Figure 5a shows that generally more respondents would order the product in the survey with an eco-label, whereas Figure 5b shows that generally more participants would return the product in the survey without an eco-label. A cause of this difference could be the low amount of respondents or that people would identify a notification about the environmental impact of returning the product before even ordering the product as annoying. In conclusion, the usage of an eco-label has a positive influence when it comes to returning packages.

The hypothesis was that age does not influence online customer behaviour concerning returning or ordering products. However, the results do not show this. Figure 6a shows that people aged over forty-five generally would not order the product more often than people from other age groups. An explanation for this could be that these respondents are less used to ordering products. Younger people, respondents in age groups 15-30 and 31-45, generally would not order the product less often than participants from the other age groups. An explanation for this could be that these respondents are more used to ordering products. Figure 6b shows a smaller difference between the age groups. It does indicate that persons aged 31-45 are the most likely to return the product even though an eco-label is shown. This finding does not suit the expectation. These conclusions can be made because the p-values found using the chi-squared tests are all lower than the alpha level as shown in Table 3. So the hypothesis can be rejected and as explained age does influence online customer behaviour.

When it comes to income, the hypothesis was that income does not influence online customer behaviour considering ordering or returning products. The results show something else. The p-values found using the chi-squared tests are all lower than the alpha level as shown in Table 4. So the hypothesis can be rejected and income does influence online customer behaviour. Figure 7a shows that there is not a big difference between the subgroups when it comes to ordering products. It does indicate that However, it does indicate that respondents earning between €35.000 and €50.000 will order the product the most often. Figure 7b shows that people with an income over €50.000 per year would return the products the most often. Whereas the other subgroups score almost equally. An explanation could be that they can afford a more luxurious item that does fulfil their needs. Another explanation could be that people with a lower income are more used to buying products that do not completely fulfil their needs because they are not able to afford products that can.

The hypothesis was that the amount of products you order per month does not influence online customer behaviour when it comes to ordering/returning products. This hypothesis can be rejected. This conclusion can be made because the p-values found using the chi-squared tests are

all lower than the alpha level as shown in Table 5. Figure 8a shows that people who never order products are generally the most likely to not order the product. It does not show that participants who order 4-6 products a month would order more products than respondents that order 1-3 products per month. Figure 8b also does show that people who order the most products are generally more likely to return the products shown in the survey, however, the differences between the subgroups are at most about 4%. So the conclusion that people who order more products per month are more likely to return a product when an eco-label is shown, cannot be made.

The results of the logistic regression are shown in Tables 6 and 8 for both ordering and returning. When it comes to ordering products, Table 6 shows that the following features influence the behaviour:

- Age
- Your favourite sport
- The number of hours you spend on your phone per day
- Conscientiousness, which is a big five personality trait
- Neuroticism, which is also a big five personality trait

As shown in Table 6, the coefficient of age is positive, which means that the older you are the more likely you are to order the product. These findings are not in line with the findings in Figure 6a. An explanation for this could be the way the data is inserted into the regression and the small number of 'old' people that filled in the survey as shown in Table 1, which could mean that some of the findings from Figure 6a are not representable for the whole population.

To perform the logistic regression, favourite sports were turned into the following numbers:

1. Football
2. Tennis
3. Ice skating
4. Basketbal
5. Athletics
6. Hockey
7. Korfball
8. Volleyball
9. I don't like sports

This is used as a sort of scale in the regression. As shown in Table 6, the coefficient for favourite sport is negative, which means the lower the score of your favourite sport, the more products you will order. So for example, football fans are more likely to order products than volleyball fans, when an eco-label is shown.

Table 6 shows that the hours spend on your phone per day has a positive coefficient. This means that the more hours you spend on your phone per day, the more products you will order when an eco-label is shown. The amount of hours you spend on your phone per day has a positive coefficient as shown in Table 6. This means that the more hours you spend on your phone per day, the more likely you are to order the product even though an eco-label is shown.

In the big five personality test, conscientiousness is a trait that is made up of goal-directed behaviour, high levels of thoughtfulness and good impulse control. Highly conscientious persons analyse their behaviour to find out how their behaviour affects others. Furthermore, they regularly plan [Tho21]. The coefficient of conscientiousness is negative (as shown in Table 6). This means that the more conscientious a person is, the fewer products he/she will order.

Neuroticism is a trait that is distinguished by emotional instability, moodiness and sadness. Highly neurotic persons experience mood swings, irritability and anxiety [Tho21]. Table 6 shows that neuroticism has a negative coefficient. This indicates that the more neurotic an individual is, the fewer products he/she will order.

When it comes to returning products, Table 8 shows that the following features influence the behaviour:

- The number of hours you spend on your phone per day
- Your favourite type of holiday destination
- Your eco-consciousness
- Openness, which is a big five personality trait
- Neuroticism, which is also a big five personality trait

Table 8 indicates that the number of hours you spend on your phone per day has a negative coefficient. This means that the more hours you spend on your phone per day, the less likely you are to return the product when an eco-label is shown.

Just like the feature 'favourite sport', the options of the feature 'holiday destination' were turned into numbers. These were the following:

1. Camping
2. Hotel
3. BB

4. Cottage

5. All-inclusive resort

The coefficient has a positive score (see Table 8). This indicates that the higher the ranking of your favourite holiday destination, the more products you will return when an eco-label is used. So for example, individuals who like to go camping the most will return fewer products than those who like to go to an all-inclusive resort the most. With this feature (and the favourite sport feature in the ordering process) the furthest options are taken as an example, because of those in the middle the differences in behaviour probably will be less notable.

As expected, your environmental consciousness has a negative coefficient. This indicates that the more environmentally conscious someone is, the less likely this person will return the product when an eco-label is shown.

Openness is characterized by individuals who have imagination and insight. Furthermore, a big part of openness is creativity [Tho21]. It has a positive coefficient, which means that the more open a person is the more likely this person is to return a product even though an eco-label is shown. Table 8 shows that neuroticism also has a positive coefficient. This indicates that the more neurotic a person is, the more products this person will return.

In line with the earlier conclusions that an eco-label influences people differently when it comes to ordering and returning products, the coefficients of features that influence behaviour both in returning and ordering are positive/negative for ordering and then the other way around for returning. Hours spent on your phone per day has a positive coefficient for ordering and a negative coefficient for returning and neuroticism has a negative coefficient for ordering and a positive coefficient for returning.

In conclusion, to answer the research question, online customer behaviour concerning returning products can be influenced by adding sustainable information to the process. The addition of this sustainable information has a positive effect on the number of packages that customers return. Online retailers could use these eco-labels to have fewer products returned, which is better for the environment and their profit margin. When fewer products are being returned, they would also have to return less money to their customers. So who benefits from this conclusion? Firstly, everyone benefits when less CO^2 will be emitted. However, online retailers can benefit the most from these results. If they use this eco-label on their packages, they would have 13,4% fewer returns ($\frac{new-old}{old} * 100 = \frac{61,8-71,4}{71,4} * 100 = \%13,4$) (See Figure 5b for these percentages). In 2021 in the US, \$1,05 trillion of the total sales were online and about \$218 billion of this total was returned [SAP22]. Most of these returns cannot be sold again. When this eco-label would be used the online retailers in the U.S. would gain roughly \$29,2 billion ($13,4\% * 218$). These online retailers would however have to make sure that their calculations of the amount of CO^2 are correct. Otherwise, it could turn into a real drama when people find out the numbers are not correct and they are being lied to.

Moreover, online retailers could also use the eco-label when people are ordering a product, because then their sales will rise by %46,7 ($\frac{new-old}{old} * 100 = \frac{33,3-22,7}{22,7} * 100 = \%46,7$)(See Figure 5a for these percentages) Using this label at the process of ordering is much harder, because at the moment of

ordering you do not know the home address of the person ordering, so you also do not know the distance the package will travel. This way you do not know what amount of CO^2 will be emitted. For this reason, the eco-label probably will not be used by online retailers concerning ordering products.

6.2 Conclusion

Increasingly more products, which are ordered online, are being returned. This process of returning hurts the environment. In this thesis, the aim was to research if there is a way to change customer behaviour concerning returning/ordering products. The research question therefore was:

How can the online customer behaviour concerning returning or ordering products be influenced by adding sustainable information to the process?

Studies by other researchers indicated that usage of labels is effective [SW06][SS09] and package design influences the choices of customers [WKA18]. Furthermore, green nudges are effective when they are transparent and the communication is placed at a continuum [Sch17] [TH21].

To research this research question an experiment was set up. Two surveys were constructed showing products and the question if the respondent would or would not order/return the product. These products were from different price levels, sizes and weights. The only difference between the surveys was that one survey also showed an eco-label telling how many grams of CO^2 would be emitted if the product was sent back and how many minutes of showering this would be. The results of these surveys were analysed by a chi-squared test and by the usage of Recursive Feature Elimination.

The chi-squared test tests whether the findings are statistically correct assuming your hypothesis is true. This research, it tested if the difference between the answers of the participants of the surveys can be traced back to the added sustainable information, the eco-label. The test statistic and p-value of the test conducted on the data about orders are respectively 29,70 and $5{,}05e^{08}$. The test statistic and p-value of the test conducted on the data about returning are respectively 21,17 and $4{,}21e^{06}$. RFE found features that influence customer behaviour. The results of these features were shown as bar charts indicating what percentage of the subgroup would not order/return the product. Furthermore, Figures 5a and 5b show what percentage of participants would not order/return the product. Furthermore, logistic regression was used to find the underlying reasons why individuals behave like they do when it comes to ordering and returning products.

The features that RFE found to affect the choices of customers the most were age, income and the number of products a person orders online per month. Figure 6a shows that persons aged over forty-five are generally less likely to order the product than persons from another age group. Moreover, Figure 7b shows that participants with an income over €50.000 per year choose to return the product shown in the survey more often. Furthermore, Figure 8a shows that when a participant never orders products online, he/she is less likely to order the product shown in the survey.

Logistic regression showed that age, the participants' favourite sport, the number of hours the participant spends on his/her phone per day, the participants' level of conscientiousness and the

participants' level of neuroticism influence behaviour when it comes to ordering products. Furthermore, it shows that the number of hours the participant spends on his/her phone per day, the participants' favourite holiday destination, the participants' environmental consciousness, the participants' level of openness and the participants' level of neuroticism influence behaviour when it comes to returning products.

In conclusion, adding sustainable information on the packages makes sure fewer products are being returned and has a positive effect on the profit margin of online retailers as well as on the environment.

6.3 Future research

Future research on this subject can be done by making sure the products in the survey are products everyone can imagine they would buy. Some of the respondents of these surveys indicated that imagining buying these products was a bit hard because they would never buy these products themselves. Furthermore, the surveys did not have a high number of participants. There were 34 participants in the research with an eco-label and 32 in the research without an eco-label. This low number of participants probably influenced the results because most participants were students, so not the whole population was used for this research. Future research could improve this research by having more participants so outliers have less influence on the results. Furthermore, future research could ask for the gender and highest level of education of the participants to check if these features may influence their behaviour.

References

[CBS20] CBS. Welk percentage van nederlanders koopt online? en wat... https://longreads.cbs.nl/nederland-in-cijfers-2020/welk-percentage-van-nederlanders-koopt-online-en-wat/#:~:text=Nederlanders%20kopen%20steeds%20vaker%20en,en%20accommodaties%20werden%20online%20gekocht., 2020. Accessed on 15-7-2022.

[CC21] Sharon Cullinane and Kevin Cullinane. The logistics of online clothing returns in sweden and how to reduce its environmental impact. *Journal of Service Science and Management*, 14:72–95, 2021.

[CGJSK21] Fredrik Carlsson, Christina Gravert, Olof Johansson-Stenman, and Verena Kurz. The use of green nudges as an environmental policy instrument. *Review of environmental economics and policy*, 15 (2):216–237, 2021.

[dicnd] Cambridge dictionary. Sustainable meaning in cambridge english dictionary. https://dictionary.cambridge.org/dictionary/english/sustainable, n.d. Accessed on 16-7-2022.

[DMD14] Manilall Dhurup, Chengedzai Mafini, and Tshepiso Dumasi. The impact of packaging, price and brand awareness on brand loyalty : evidence from the paint retailing industry : original research. *Acta Commercii*, 14:1–9, 2014.

[dpd20] dpd. Nederland blijft europees koploper retourneren en het aantal retouren groei. https://www.dpd.com/nl/nl/2020/01/08/nederland-blijft-europees-koploper-retourneren-en-het-aantal-retouren-groeit-2/ 08-01-2020. Accessed on 15-7-2022.

[EY15] Beytullah Eren and Muhammad Yaqub. Environmental consciousness survey of university students. 11 2015.

[Frond] Jim Frost. Test statistic: Definition, types formulas. https://statisticsbyjim.com/hypothesis-testing/test-statistic/, n.d. Accessed on 8-8-2022.

[GCW+22] Chao Gu, Jiangjie Chen, Wei Wei, Jie Sun, Chun Yang, Liao Jiang, Jingyue Hu, Baiwan Lv, Shuyuan Lin, and Qianling Jiang. The impact of reusable tableware packaging combined with environmental propaganda on consumer behaviour in online retail. *PLoS ONE*, 17 (3):e0264562–e0264564, 2022.

[Jan21] Dina Jankovic. A simple interpretation of logistic regression coefficients. https://towardsdatascience.com/a-simple-interpretation-of-logistic-regression-coefficients-e3a40a62e8cf, 2021. Accessed on 8-8-2022.

[JSD+12] EJ Johnson, SB Shu, Benedict Dellaert, C Fox, DG Goldstein, G Haeubl, RP Larrick, JW Payne, E Peters, D Schkade, B Wansink, and EU Weber. Beyond nudges: Tools of a choice architecture. *Marketing letters*, 23 (2):487–504, 2012.

[JSF16] Narayan Janakiraman, Holly A. Syrdal, and Ryan Freling. The effect of return policy leniency on consumer purchase and return decisions: A meta-analytic review. *Journal of Retailing*, 92 (2):226–235, 06-2016.

[Li17] Susan Li. Building a logistic regression in python, step by step. https://towardsdatascience.com/building-a-logistic-regression-in-python-step-by-step-becd4d56c9c8, 2017. Accessed on 21-7-2022.

[McA20] Tim McAleer. Interpreting linear regression through statsmodels .summary(). https://medium.com/swlh/interpreting-linear-regression-through-statsmodels-summary-4796d359035a, 2020. Accessed on 5-8-2022.

[MHKK+21] Julia Meis-Harris, Celine Klemm, Stefan Kaufman, Jim Curtis, Kim Borg, and Peter Bragge. What is the role of eco-labels for a circular economy? a rapid review of the literature. *Journal of cleaner production*, 306:127–134, 2021.

[Opt19] Optoro. Impact report. https://info.optoro.com/hubfs/Optoro%202019%20Impact%20Report.pdf, 2019. Accessed on 16-7-2022.

[ptnd] pesonality testing.info. The big five personality test (bfpt). https://sites.temple.edu/rtassessment/files/2018/10/Table_BFPT.pdf, n.d. Accessed on 18-7-2022.

[Refnd] Oxford Reference. Brand loyalty - oxford reference. https://www.oxfordreference.com/view/10.1093/acref/9780198736424.001.0001/acref-9780198736424-e-0207#:~:text=A%20measure%20of%20the%20consumers,a%20competing%20product%20or%20service., n.d. Accessed on 18-7-2022.

[SAP22] SAP. Smart tech is tackling the true cost of returns to the retail industry. https://www.forbes.com/sites/sap/2022/03/23/smart-tech-is-tackling-the-true-cost-of-returns-to-the-retail-industry/?sh=21a215826c67, 2022. Accessed on 7-8-2022.

[Sch17] Christian Schubert. Green nudges: Do they work? are they ethical? *Ecological economics*, 132:329–342, 2017.

[SS09] Junyi Shen and Tatsuyoshi Saijo. Does an energy efficiency label alter consumers' purchasing decisions? a latent class approach based on a stated choice experiment in shanghai. *Journal of environmental management*, 90 (11):3561–3573, 2009.

[SSH17] Hannu Saarijärvi, Ulla-Maija Sutinen, and Lloyd C. Harris. Uncovering consumers' returning behaviour: a study of fashion e-commerce. *The International review of retail, distribution and consumer research*, 27 (3):284–299, 2017.

[SW06] Katharina Sammer and Rolf Wüstenhagen. The influence of eco-labelling on consumer behaviour – results of a discrete choice analysis for washing machines. *Business strategy and the environment*, 15(3):185–199, 2006.

[TH21] Linda Lisa Maria Turunen and Minna Halme. Communicating actionable sustainability information to consumers: The shades of green instrument for fashion. *Journal of cleaner production*, 297:126605, 2021.

[Tho21] Thomas. What are the big 5 personality traits? https://www.thomas.co/resources/type/hr-guides/what-are-big-5-personality-traits, 2021. Accessed on 6-8-2022.

[TS08] Richard H. Thaler and Cass R. Sunstein. *Nudge: Improving Decisions About Health, Wealth, and Happiness*, chapter 1, 11. Yale University Press, New Haven London, 1 edition, 2008.

[WKA18] Sidrah Waheed, Marium Mateen Khan, and Nawaz Ahmad. Product packaging and consumer purchase intentions. *Market Forces Research Journal*, 13 (2):97–114, 2018.