The Sustainability Liability Effect:

*How Consumers May Choose the Regular Product
Instead of the Sustainable One*

Colin Brustad and Andreas Sæther

Supervisor: Helge Thorbjørnsen

Master of Science in Economics and Business Administration,
Energy, Natural Resources and the Environment

NORWEGIAN SCHOOL OF ECONOMICS

Preface

This thesis is written as a part of our Master of Science in Business Administration degree at the Norwegian School of Economics (NHH). We are completing our degree in the profile Energy, Natural Resources and the Environment. This study is a continuation of the unpublished study by Skard et. al. (2021) for the Center for Sustainable Business at NHH, working to uncover disadvantageous biases towards sustainable products.

Abstract

This paper evaluates the potential problem of consumers believing that sustainable products with strength-related attributes are less effective compared to the competing non-sustainable products, also known as the sustainability liability effect. The concern is that the sustainability liability effect may result in consumers choosing a regular product instead of a sustainable one, which may have consequences such as lower sales of sustainable products, countering sustainable progress made within production and marketing.

A study was conducted at IKEA in Bergen, which consisted of a combined field experiment and survey where the goal was to identify if the sustainability liability effect impacted the choice of household cleaning products. Two identical oven cleaners, with the only difference being that one was indicated as sustainable and the other did not have any indication (therefore categorized as the "regular" product), were placed next to each other. Participants were asked to choose one of the products to clean an oven rack with a chance to win a 150 NOK gift card based on performance. After choosing a product, participants were instructed to complete a questionnaire instead of undertaking the cleaning challenge. Upon completing the questionnaire, which included questions based on the reasoning behind the choice of product, participants received the 150 NOK gift card.

101 participants were part of the study. Although the majority chose the sustainable product (61 out of 101), evidence of the sustainability liability effect was found when testing for the perceived effectiveness of the products and the subsequent impact on product choice. In other words, the average participant perceived the regular (versus sustainable) product to be more effective, and participants who perceived the regular (versus sustainable) product to be more effective had a higher likelihood of choosing the regular (versus sustainable) product.

Evidence of the sustainability liability effect impacting product choice should have implications for various stakeholders including sustainable producers and marketers. Suggestions include increased investment in R&D to investigate and implement behavioral interventions that could counter the sustainability liability effect such as product placement in stores. Future research on sustainability liability effect should investigate aspects such as a store setting, rural and urban areas, online shopping, and post-COVID-19.

Table of Contents

1. Introduction

1.1. Sustainability

As global temperatures continue to rise during the current climate crisis, sustainability has become a critical topic of discussion. The broad definition of sustainability is to meet the needs of the present without compromising the ability of future generations to meet their needs (United Nations, 2021a). It is common among businesses to use a more specific definition of sustainability to adhere to ESG metrics which comprise of environmental, social, and governmental criteria (The Investopedia Team, 2021). Such a definition of sustainability can be phrased as the complete plan of ethical action for an organization to become pro-environmental, prosocial, and traditional pro-economic (Lijo & Narayanamurthy, 2015).

The climate crisis was arguably catalyzed by the industrial revolution, which saw rapid technological progress that improved the efficiency of production. Today, an increasingly prosperous world along with cheaper goods and services has led to unprecedented levels of consumption. To access necessary materials through mining, deforestation, farming, and urbanization, humans have terraformed the world to support this lifestyle. The consequences have been grave, with CO_2, NO_X, and other byproducts disturbing the fragile ecosystems threatening animals, humans, and nature itself. A recent report published by the IPCC (2021) concludes that greenhouse gases from human activities have already been responsible for a global increase in temperature of about 1.1 degrees Celsius. Furthermore, if immediate, rapid, and large-scale actions are not taken, limiting warming to 1.5 or even 2.0 degrees Celsius will not be possible, and the 1.5-2.0 average temperature increase may be reached during the next 20 years (IPCC, 2021). Consequences will be devastating and include more extreme weather such as floods and droughts, sea-level rise, thawing of permafrost, and destruction of ecosystems (IPCC, 2021).

The UN's climate panel has listed 17 goals for sustainable development (United Nations, 2021b). Goals number 12 and 13 are the most relevant for this paper. Goal 12 emphasizes the necessity of sustainable production and consumption patterns, which is critical due to the problem of overconsumption. 1969 was the last year planet Earth's resources were sufficient to sustain the current generation's demand with future generations in mind. Between 2000

and 2017, humanity increased its material footprint by 70% (United Nations, 2021b). Another measure of overconsumption is Earth Overshoot Day, which marks the date when humanity's demand for ecological resources in a given year exceeds what the Earth can regenerate in that year (Earth Overshoot Day, 2021). In 2021, Earth Overshoot Day was July 29th, which means that every day between July 30th and December 31st, humanity borrowed resources from future generations. To sustain today's global consumption, 1.7 planet Earths would be necessary. Goal number 13, climate action, emphasizes the importance of taking care of the world through sustainable behavior (United Nations, 2021b). The Stockholm Resilience Centre (2015) has made a list of nine different measurements that impact the environment. Minimizing consumption is just a small part of reducing the global footprint. Companies work continuously to further reduce their climate footprint through innovation. New technology and more efficient, long-lasting products all contribute to reducing emissions and waste.

1.2. Sustainable Production

Sustainability is now viewed as a competitive advantage and is the basis of many new innovations (Nidumolu et al., 2009). An approach that allows producers to develop a sustainable competitive advantage is the adoption of traditional marketing and production tactics that include environmental issues regarding four P's - product strategy, pricing strategy, place distribution, and promotion (Papadas & Avlonitis, 2014).

For product strategy, common tactics include environmentally friendly packaging and ingredients, recyclable and reusable content, and greening of the production process such as integrating the use of renewable energy (Papadas & Avlonitis, 2014). An example of a company that has taken steps to gain a sustainable competitive advantage within product strategy is Patagonia, as it prioritizes the use of environmentally friendly materials in its products (Haugland & Nysveen, 2021a; Patagonia, 2021a).

Pricing strategy concerns the justification of potentially higher prices due to environmental costs. Tactics to encourage consumption despite the price premium includes the use of promotional prices for sustainable products or the increase of regular product prices to make sustainable products seem relatively cheaper (Papadas & Avlonitis, 2014). Patagonia has used promotional prices to stimulate the sale of sustainable products (Haugland & Nysveen, 2021a; Patagonia, 2021b).

When it comes to place distribution, green tactics include optimizing the location of production and distribution channels, selecting green distribution channels, digitalizing distribution programs, and forming a joint commitment between suppliers and distributors (Papadas & Avlonitis, 2014). Looking once again at Patagonia, the company gives a description of the environmental profile of factories, farms, and mills used to produce clothing (Haugland & Nysveen, 2021a; Patagonia, 2021c).

Finally, when tackling promotion, the effectiveness of a green marketing strategy may be improved through advertising environmental appeals and claims, and publicizing environmental efforts on the product packaging (Papadas & Avlonitis, 2014). An example is to mark commodities with external product labels that certify certain sustainability or ethical criteria. Product labels by both NGOs and governmental organs have increased significantly in recent years. In 1990 there were 12 product labels (Delmas et al., 2013), whereas currently there are 455 (Ecolabel Index, 2021). These labels span 25 industry sectors, including food, clothing, cosmetics, furniture, tourism, energy (Ecolabel Index, 2021).

2. Consumer Behavior for Sustainable Products

If all else is equal, consumers should favor a sustainable product versus a non-sustainable (regular) product due to the removal of negative externalities. The assumption is that a non-sustainable product may have negative externalities such as pollution, whereas a sustainable product has less externalities as the environmental footprint is reduced. Therefore, if attributes such as price and quality of the products are equal, the consumer should choose the sustainable product to minimize social marginal cost (Dacanay et al., 2011).

2.1. Consumer Attitude Has Taken a Positive Turn

There is evidence that consumers perceive sustainable products as superior to non-sustainable products. Cotte & Trudel (2009) explored consumer preferences for sustainable and ethical consumption and found that consumers believe it is important that products exhibit some degree of ethicality. As such, consumers are willing to pay a premium for ethical goods and demand a discount for non-ethical goods (Cotte & Trudel, 2009). Olsen et al. (2014) also argue that consumers are often willing to pay more for sustainable products compared to non-sustainable products, and therefore attitude towards sustainability is trending in a positive direction. Furthermore, Deloitte conducted a survey on the importance of

sustainability in the United Kingdom, and found that 78% of participants claimed they had made at least one change in their lifestyle to become more sustainable during the previous year (Deloitte, 2021). In a Norwegian survey by Sparebank1 Østlandet (2018), customers were asked if during the past year, there has been a change in importance when it comes to products having sustainable attributes. The data unveiled the rising importance of sustainability among consumers, as 33% claimed that sustainable products have become "of large importance", and another 33% of participants responded that it has become "more important to some degree" (SpareBank1 Østlandet, 2018).

2.2. Individual Differences

There is evidence of individual differences in consumer behavior in relation to sustainable products, which can be categorized by age, gender, education, and personality traits.

When it comes to age, there seems to be a particularly positive attitude towards sustainable products and consumption among younger generations compared to older generations (Brown & Robertson, 2011; Kumar et al., 2012; Agarwal & Kasliwal, 2017). A suggested reason for differences in sustainable behaviors between age cohorts is that consumption habits vary based upon the time period an individual grew up in, and the important external events experienced during the formative or coming-of-age years (Brown & Robertson, 2011). In the case of sustainability, younger generations have been significantly more exposed to environmental knowledge and events related to climate change, and therefore there is a higher likelihood that this is reflected in their values, attitude, and behavior (Brown & Robertson, 2011). However, there is also counterevidence suggesting little or no differences in environmental behavior between generations. This includes willingness to pay (Agarwal & Kasliwal, 2017), purchasing decisions (Sarti et al., 2018), and overall environmental concern (Gray et al., 2019).

For gender, studies have found that men are less likely to be environmentally friendly in their attitudes, choices, and behaviors compared to women (Davidson & Freudenburg, 1996; Lee & Holden, 1999). Women also show greater concern and willingness to take action to benefit the environment, and this is consistent and robust across age groups and countries (Cottrell, 2003; Dietz et al., 2002; Levin 1990; Zelezny et al., 2000). In addition, Brough et al. (2016) explore reasons behind the gap between sustainable behavior and gender. Through a series of studies, evidence points towards that the concepts of "greenness" and "femininity" are

cognitively linked and therefore consumers who engage in more sustainable behavior are categorized by others as more feminine and even perceive themselves as more feminine (Brough et al., 2016). Furthermore, the willingness of men to engage in sustainable behavior can be influenced by threatening or affirming their masculinity (Brough et al., 2016).

Studies show that a higher level of education has a positive effect on consumer attitude and behavior when it comes to sustainability. For instance, a higher level of education can lead to lower CO_2 emissions (Balin, 2021), higher rate of recycling (Pelau & Catalina, 2018), and increased green behavior in the workplace (Fawehinmi et al., 2020). In a set of surveys, Meyer (2015) finds that higher levels of education causes individuals to be more concerned with social welfare and subsequently behave in a more environmentally friendly manner. This translates to the observation that individuals with higher education levels tend to be more environmentally friendly.

Sustainable consumer behavior may also differ due to personality traits, such as the degree a person identifies as "green". Previous research suggests that behavior may be clustered in a way that reflects similar types of behavior with respect to environmental commitment (Whitmarsh & O'Neill, 2010). For instance, three clusters that may affect one another are purchasing decisions, habits, and recycling (Barr et al., 2005; Thøgersen & Ölander, 2006). There is evidence that these spillover effects in behavior are attributed to self-identity. In other words, if a person identifies as being environmentally conscious, it will contribute to more environmentally friendly behavior, including when it comes to the clusters of purchasing decisions, habits, and recycling (Whitmarsh & O'Neill, 2010). Sarti et al. (2018) identified three customer segments in relation to green identity - collectivists, individualists, and indifferents. Collectivists make sustainability a personal priority, individualists are engaged in some way, while indifferents are not concerned about personal and environmental stewardship (Sarti et al., 2018).

2.3. The Sustainability Liability Effect

2.3.1. Definition

As discussed in section 2, all else being equal, the rational consumer should choose the sustainable product (versus a non-sustainable product) to minimize social marginal cost (Dacanay et al., 2011). However, when presented with a choice between a sustainable and

non-sustainable product, it is observed that consumers tend to prefer the non-sustainable product (Luchs et al., 2010). Luchs et al. (2010) argue that the positive effect of product sustainability on consumer preferences is reduced when attributes such as strength and effectiveness are valued, which may even result in preferences for less sustainable product alternatives. This is known as the "sustainability liability". In other words, a product marketed as sustainable may be viewed as less effective in achieving its task compared to a product that is not marketed as sustainable. For instance, a consumer may perceive that a sustainable laundry detergent is less effective at cleaning clothes compared to a regular laundry detergent. Therefore, indicating the product as sustainable may serve as a liability to the product itself as it impacts consumer choice negatively, and in turn, may reduce sales.

To further understand the dynamics of products where the sustainability liability effect may occur, it is important to classify products into two main categories - products with gentleness-related attributes, and products with strength-related attributes.

2.3.2. Products with Gentleness-Related Attributes

Products associated with attributes that provide consumers with benefits such as safety and health are classified as part of the gentleness category (Luchs et al., 2010). For example, products such as baby shampoo and body lotion may deliver gentleness-related benefits (Haugland & Nysveen, 2021b). Other attributes that may be associated with gentleness products are "friendly" and "protective" (Gildea, 2001). In other words, these products are often for cleaning but likely in relation to a person's body, meaning that a person would want them to be gentle on the body compared to products that are meant for cleaning other surfaces which instead are desired to be strong, tough, and effective.

All else being equal, sustainable products with gentleness-related attributes should be perceived as superior compared to non-sustainable products in the same category, as the sustainable product has less of a negative externality associated with it. A rational consumer would therefore select the superior good, being the sustainable product with gentleness-related attributes. Furthermore, consumers may infer that a sustainable product with gentleness-related attributes may perform better compared to a non-sustainable product of the same category. Luchs et al. (2010) conducted an initial survey to test this perception and found that almost half of respondents viewed products with positive ethical attributes,

such as sustainability, to perform better in other attributes such as safety, health, and gentleness.

Additional studies by Luchs et al. (2010) find support for the hypotheses that consumers associate higher ethicality with gentleness-related attributes, and that sustainability enhances product preferences to a greater extent when gentleness-related attributes are valued. These results therefore find that sustainable products are associated with gentleness, whereas non-sustainable products are associated with strength (Haugland & Nysveen, 2021b). The underlying reasoning is that when a product includes a positive attribute such as sustainability, this positivity often extends to other attributes of the product such as gentleness. This effect might originate from the context of the social judgment that there is a conflicting association between strength and ethicality, and subsequently a benefitting association between gentleness and ethicality. This is relevant in the case of sustainability, as consumers do not have perfect information of the products they are choosing between, and therefore may infer the quality of its attributes, such as inferring gentleness-related benefits when observing a sustainable product through association (Luchs et al., 2010).

Skard et al. (2020) further explore the investigation by Luchs et al. (2010) of the interaction between sustainability, gentleness, and consumer preferences. Here, the focus is divided into product-related green attributes and non-product-related green attributes. Product-related green attributes are directly linked to the product itself, for instance ingredients, and are defined as green core attributes. Non-product-related green attributes are indirectly linked to the product, for instance the bottle or packaging of the product, and are defined as green peripheral attributes. Three out of four studies conducted by Skard et al. (2020) support the key hypothesis that in gentleness product categories, consumers infer higher functional product quality when the product has a green core attribute compared to no attribute. Interestingly, the four studies did not support the similar hypothesis that consumers infer higher functional product quality when the product with gentleness-related attributes has a green peripheral attribute. However, the report concludes that there is a sustainable asset effect (which is that sustainable attributes may lead to the consumer perceiving that the product has superior functionality, therefore the opposite of the sustainability liability effect) in gentleness-dependent categories, although for core attributes only.

As indicated by both Luchs et al. (2010) and Skard et al. (2020), there is support for a sustainability asset effect when it comes to products with gentleness-related attributes. As

such, since there is not a sustainability liability problem with the gentleness product category, but instead a benefit to these types of products being marketed as sustainable, there is not a continued focus on products with gentleness-related attributes in this paper. It is however important to include this section describing the category of products with gentleness-related attributes, as they are discussed in multiple studies in comparison to products with strength-related attributes, and provide a clearer picture of the specificity of the problem.

2.3.3. Products with Strength-Related Attributes

Products with strength-related attributes can be seen as the opposite of products with gentleness-related attributes. Products in the strength category are often associated with attributes such as power and durability (Luchs et al., 2010). Examples of products with strength-related attributes include laundry detergents and car tires (Haugland & Nysveen, 2021b). Products with strength-related attributes may be associated with terms such as effectiveness, toughness, and getting the job done.

Multiple studies find that there is a sustainability liability effect for products with strength-related attributes. Luchs et al. (2010) conducted five studies to test if the sustainability liability effect exists for products in the strength category, and found it to be present. One of the studies concludes that consumers associate a lower ethicality with strength-related product attributes, inferring that non-sustainable products are more associated with strength compared to sustainable products. Lin & Chang (2012) present the hypothesis that green products are perceived as less effective than regular products, which is supported by the results from their studies. This is further strengthened by Luchs et al. (2012), where they investigate a trade-off between a product with superior functional performance versus a product with superior sustainable performance. Specifically, in study 1 of Luchs et al. (2012), shoes were the product used to investigate this trade-off, and support was found for the hypothesis that consumers will select the product with superior functional performance instead of the more sustainable product. This is relevant since shoes can be categorized as a product with strength-related attributes, as shoes are supposed to endure surfaces, the outdoors, and weather. In a similar study, Newman et al. (2014) found that consumers are less likely to purchase a sustainable product when they perceive that the company intentionally made the product better for the environment compared to when the same environmental benefit occurred as an unintended side effect. Four experiments were conducted, and experiments 1-3 used home goods such as cleaning detergents and dish soap

for measuring consumer preferences (Newman et al., 2014). The findings are of importance since these products are in the strength category, therefore underlining the fact that the sustainability liability effect exists for products with strength-related attributes. The Newman et al. (2014) report is also relevant as sustainable cleaning detergents are often intentionally marketed as sustainable, for instance through eco-labels and using the color green (Pancer et al., 2017). Pancer et al. (2017) find further evidence of the sustainability liability effect, this time when looking at laundry detergents, dish soaps, and toothpastes. Laundry detergent and dish soap products are determined to be in the strength category, whereas toothpaste can be argued as a product in the gentleness or strength category depending on its perception, as associated attributes may include health, safety, but also effectiveness and power. Across the three studies conducted by Pancer et al. (2017), results concluded that environmental cues in isolation, for instance the color green without an eco-label and vice versa, cause consumers to view the product as less effective, and are therefore prompted to choose the alternative they view as superior. Skard et al. (2020) focus on both green core attributes, such as ingredients, and green peripheral attributes of products, such as packaging. Products representing the strength-dependent category were drain openers and hand sanitizers. Findings from Skard et al. (2020) were consistent with previous reports, as a sustainability liability effect was present in the strength-dependent category for both core and peripheral attributes. Finally, Skard et al. (2021) further investigate the sustainability liability effect by focusing on the attitude-behavior gap, meaning that consumers that view sustainable products positively may not necessarily purchase them when in competition with regular products. Their results indicate that the attitude-behavior gap along with the sustainability liability effect is present for products with strength-related attributes. These findings are further discussed in section 2.3.7.

2.3.4. Consequences of the Sustainability Liability Effect Being Present in Products with Strength-Related Attributes

The sustainability liability effect being present in products with strength-related attributes is a problem as it inhibits the goal of sustainable measures reducing negative environmental externalities. Consequences of the sustainability liability effect include lower sales of sustainable strength-dependent products, as well as a higher per dosage use of a sustainable product leading to overuse.

Today, strength-dependent products such as cleaning detergents are more relevant than ever. The COVID-19 pandemic caused a global shock increase in demand for cleaning products, as consumers became more conscious of hygienic behavior, disinfecting and sanitizing surfaces more frequently in private homes and public spaces. The global household cleaning products market size was an estimated 220 billion USD in 2020, and is projected to rise to a 320 billion USD market in 2028 (Fortune Business Insights, 2021). Laundry detergents, a product with strength-related attributes, make up 53.7% of the American market share (Fortune Business Insights, 2021). There is cause for optimism when it comes to drivers for growth, with sustainable products being a key factor. The rise of environmentally conscious consumers has increased demand for sustainable and natural cleaning products which should further boost market growth (Fortune Business Insights, 2021). The total value of environmentally friendly cleaning products is projected to amount to 73 billion USD in 2021 globally, with a forecasted CAGR of 8.5% from 2021 to 2026, resulting in a potential market size of 110 billion USD in 2026 (Smithers, 2021). This is far superior to the overall market growth for all cleaning products globally, which is projected at a CAGR of 4.1% between 2021 and 2026 (Smithers, 2021). Significant growth has also occurred in the United States market. Looking at the overall consumer packaged goods market, meaning goods that are used roughly on a daily basis, products marketed as sustainable delivered 54.7% of growth between 2015 and 2019, despite only having a 16.1% overall market share (Kronthal-Sacco & Whelan, 2021).

However, there is evidence that strength-related products marketed as sustainable continue to struggle with sales today. Of all household cleaning items, only skincare products, which are a part of the gentleness category, have a sustainable product market share higher than 20% in the United States as of 2020 (Kronthal-Sacco & Whelan, 2021). All strength-dependent household cleaning products have a sustainable product market share of less than 20%, including floor cleaners, dish detergents, household cleaners, laundry care products, and laundry detergents. In fact, laundry care products and laundry detergents have a less than 5% sustainable market share in the United States (Kronthal-Sacco & Whelan, 2021).

It is concerning that non-sustainable products still take up the majority of the market share within strength-dependent household cleaning products, as it shows a hesitancy from consumers to act on their growing environmentally conscious mindset. This is a significant challenge to sustainable initiatives in motion, and may disincentivize continued

environmental innovation within this marketplace as sustainable products are outmatched by regular products.

Sustainable products with strength-related attributes are also subject to overuse. Lin & Chang (2012) find that since green products are viewed as less effective than regular products, consumers use a greater amount per dose of a green product than a regular product. Furthermore, environmentally conscious consumers surprisingly use more of a sustainable product per dose compared to less conscious consumers (Lin & Chang, 2012). Since the sustainability liability effect is found to increase the dosage consumption of sustainable products, it leads to overconsumption, and therefore counters a key point of sustainable products as they are meant to minimize environmental impact. This aspect of the sustainability liability effect has therefore been directly addressed by producers, retailers, and advertisers by calling for consumers to follow the recommended dosage of the product (Kiwi, 2021; Gulbrandsen, 2017).

2.3.5. Behavioral Reasons for the Sustainability Liability Effect, Other Reasons for Product Choice

To solve the problem of the sustainability liability effect, its underlying reasons must be addressed. As discussed in section 2.3.3, the consensus theory of the sustainability liability effect is that the consumer perceives the sustainable product with strength-related attributes to be less effective compared to the competing regular product.

However, explanations of the underlying reasoning as to why there is a perception of lower effectiveness vary. As briefly outlined in section 2.3.2, Luchs et al. (2010) speculate that sustainability is an attribute associated with the product, and therefore influences the perception of other attributes of the product as well. In society, there is a prevalence of sociocultural messages that frame ethicality and strength as conflicting, meaning that there is a trade-off between the two attributes. In other words, the more ethical a product, the less perceived strength or effectiveness it has. This social context is then transferred to product judgments by consumers, especially in the case of sustainability, since consumers do not have perfect information regarding how strong or effective a product is, and therefore uses sustainability as a factor to infer this (Luchs et al, 2010). Pancer et al. (2017) add to this, arguing that environmental cues in isolation activate competing functionality and ethicality attributes when the consumer evaluates the product. An alternative explanation proposed by

Luchs et al. (2010) is that consumers know that brands operate under budget, product development, and manufacturing constraints, and therefore may infer that a green product is superior in the sustainability attribute but inferior in other attributes. A similar explanation is provided by Newman et al. (2014), who states consumers' lay theories of resource allocation as a reason, meaning that consumers infer that the company diverted resources away from effectiveness, strength, and functionality to enhance the product's sustainability attribute. Lin & Chang (2012) also discuss consumers' lay theories as an explanation, stating that consumers often rely on lay theories in forming inferences about missing or unavailable information. Another reason put forth by Luchs et al. (2012) is that emotions play a part in the decision-making process, and concludes that consumers may prefer perceived functionality attributes over sustainability attributes due to feelings of distress.

There are also other variables that may affect the consumers' perception of sustainable products with strength-related attributes, steering them towards non-sustainable products instead. Standard economic theory is suggested by Luchs et al. (2010) as an explanation, as sustainable products are generally more expensive and less accessible. Sustainable products are often sold at a 20-25% price premium compared to regular products (Lin & Chang, 2012). If sustainable products are priced higher while suffering from limited distribution, the consumer is more likely to pick the regular product. In addition, self-proclaimed sustainability statements without backing from trusted product labels increases consumers' perceived risk towards purchasing sustainability-related products, whilst increasing the difficulty for consumers to identify the product's benefits (Chen & Chang, 2012; Mishra et al., 1998). This leads to suboptimal purchasing decisions, especially for individuals who would prefer to purchase sustainable products (Darnall & Aragon-Correa, 2014). Furthermore, customers believe that companies disclose information that favors their products (Cai et al., 2017; Oates et al., 2008), while also exaggerating environmental claims (Shahrin et al., 2017). Olsen et al. (2014) suggest that the higher the number of green claims on a product, the more confused or skeptical a consumer may become, and in turn the less likely the sustainable product is chosen in competition with a non-sustainable product. The feeling of skepticism refers to the fact that a consumer may become aware of the possibility that the brand is portraying a green image when in fact the brand does not engage in environmental efforts. This is known as "greenwashing" and has become a prominent phenomenon in recent years (Olsen et al., 2014). Luchs et al. (2012) argue that firms should

be aware that consumers may be more skeptical of green marketing tactics due to greenwashing, leading to lower sales.

2.3.6. Possible Countermeasures to the Sustainability Liability Effect

Several measures, including behavioral interventions and nudges, have been suggested by various studies to attenuate the sustainability liability effect. Behavioral interventions are measures designed to alter the behavior of a person during a decision-making process, and include everything besides price incentives, regulations, and information disclosure (Ekström, 2021a). A nudge falls under the umbrella of behavioral interventions and is defined as any aspect of choice architecture that alters people's behavior in a predictable way, but without restricting elements of choice or changing economic incentives (Thaler & Sunstein, 2008, p. 6).

2.3.6.1. Social Signaling, Social Influence, Social Desirability, and the Self-Prophecy Effect

Taking advantage of behavioral aspects such as social signaling, social influence, social desirability, or the self-prophecy effect through nudges can be used to counter the sustainability liability effect.

Social signaling is related to asymmetric information about people, and subsequently that there is a need for people to send signals to indicate their "type". For instance, a potential employee knows his level of productivity, but a future employer cannot know it for certain. The candidate can therefore share private information, such as a university degree, as a signal that verifies a certain level of productivity (Ekström, 2021b). In the case of sustainability, a consumer may send a signal of being environmentally conscious by choosing to purchase a sustainable product rather than a regular one. Observability of behavior, especially prosocial behavior, may have an effect on choice due to social signaling, as it enables the consumer to signal his or her "type" to others (DellaVigna, 2009). People want to be viewed as environmentally conscious although they may not necessarily be that, and therefore feel the need to choose the sustainable product instead of the non-sustainable product when they are observed, as they send a signal to others that they care about the environment. As such, ethical concerns about sustainability may come from an internal desire to satisfy one's ego and behaving in line with how one wishes to be perceived. There are two versions of this - reward-seeking, where a person seeks to gain social approval of their peers, and

punishment-avoiding, where a person avoids guilt (Cotte & Trudel, 2009). Furthermore, Luchs et al. (2010) argue that due to social signaling, when respondents answer questions about themselves, also known as the "self condition", respondents may present themselves in a better light than what is true. This is evident when it comes to sustainability, as consumers tend to overreport sustainable purchases and underreport less ethical purchases (Sarti et al., 2018). However, when respondents answer questions predicting what the average participant would respond, also known as the "other condition", there is evidence that social signaling is reduced and therefore answers are more truthful (Luchs et al., 2010). Luchs et al. (2010) demonstrated this, where results indicated that respondents reported a significantly higher preference for a sustainable laundry detergent in the self condition compared to the other condition.

Stakeholders such as producers or retailers can take advantage of social signaling and observability by presenting a situation of social influence, meaning influencing the consumers to make the choice that the producer or retailer wishes. For instance, retailers may introduce social influence by placing competing sustainable and non-sustainable products in a more observed part of a store. Luchs et al. (2010) conducted a study to investigate how consumers' choices between a sustainable and non-sustainable product with strength-related attributes are affected when observed by others. Hand sanitizers were used in a cafeteria, where one was labeled as sustainable and the other was not. Findings showed that when respondents were not observed by others, they tended to choose the non-sustainable hand sanitizer, but when the respondents were observed by others, they tended to choose the sustainable hand sanitizer. Therefore, the introduction of social influence through observability increased prosocial behavior and subsequently resulted in more use of the sustainable hand sanitizer.

Two other psychological aspects that can be exploited to nudge consumers towards sustainable behavior are social desirability and the self-prophecy effect. Social desirability is the tendency for people to present themselves in a favorable fashion (Holden & Passey, 2009). The self-prophecy effect is that people predict their future behavior being biased in a favorable direction, but is then followed by behavior that is consistent with this biased prediction (Bodur et al., 2015). Brands can take advantage of these two psychological effects by utilizing prediction requests in advertisements as a nudge to positively affect consumers' perception of sustainable products (Bodur et al., 2015). An advertisement with a prediction

request asks consumers to predict their future behavior. An example of a prediction request in an environmental setting could be the question: "Ask yourself... will you help the world become more sustainable?". Social desirability is triggered as consumers want to present themselves in a positive ethical light, therefore answering "Yes" to the question, and the self-prophecy effect is triggered as this causes consumers to actually have a higher likelihood of engaging in sustainable behavior, such as buying a sustainable product instead of a non-sustainable one, since they predicted that they would do so. Prediction requests in advertisements can therefore be a powerful tool in attenuating the sustainability liability effect as they can quickly reach a large portion of the consumer market with relatively low cost and effort (Bodur et al., 2015).

2.3.6.2. Other Behavioral Interventions

Other types of behavioral interventions such as a strength guarantee or superior aesthetic design may also be effective in curbing the sustainability liability effect.

As discussed in section 2.3.3, the key takeaway from Newman et al. (2014) is that when a company makes an environmental enhancement to a product, consumers are less likely to purchase the product if they learn that the enhancement is intended compared to a product where it is an unintended side effect. This implies that producers should not promote their products with strength-related attributes as sustainable, but instead subtly include environmental enhancements to their products. However, this seems illogical as there is no competitive advantage to invest in sustainable research & development if consumers do not value sustainable claims, meaning there is a high cost and no additional revenue. Furthermore, there is evidence that green new product introductions increase consumers' attitude of a brand, meaning that the more green new products a brand introduces, the better consumers view the brand, and the more likely consumers are to purchase products from the brand (Olsen et al., 2014). Therefore, it does not make sense to disregard sustainable claims.

Multiple studies suggest similar behavioral interventions related to marketing the product as both sustainable and effective. In one of their studies, Luchs et al. (2010) compared eco-tires and non-sustainable tires and found the main point to be that participants who observed a sustainable tire with a strength guarantee responded that the tire was significantly more likely to be a bestseller than participants who observed a sustainable tire with an availability guarantee. This shows that explicitly advertising the sustainable product as effective or strong

attenuates the sustainability liability effect. Lin & Chang (2012) presents a similar conclusion, which is that when the perceived effectiveness of a product is increased by explicit information about product effectiveness, the discrepancy between green and regular product per dose usage disappears. In other words, consumers who used larger doses of the sustainable product to begin with, end up using the same dose as a regular product following guarantee statements of the sustainable product's effectiveness since the perceived effectiveness of the sustainable product is increased. Luchs et al. (2012) also conclude that there must be some level of perceivable functionality along with the product being sustainable, even for environmentally conscious consumers. Specifically, it is critical to reassure that the product meets a minimum acceptable level of functional performance along with being sustainable. Furthermore, for brands that are interested in promoting sustainable products for the mass market where consumers are not as dedicated to sustainability, it is particularly important to market the product as effective to improve consumer confidence. Skard et al. (2020) further endorse this strategy, suggesting that when introducing a strength-dependent product with improved sustainability characteristics, such as implementing green packaging, it is important to highlight that the core attributes such as the ingredients are the same, reassuring the consumer that the product's effectiveness remains intact.

A non-promotion-based behavioral intervention to the sustainability liability effect is superior aesthetic design of sustainable products. Luchs et al. (2012) explore the theory that superior product aesthetic design has a disproportionately positive effect on consumer confidence and choice likelihood for sustainable products. An example of a superior aesthetic design is when a product's design is more visually appealing, which was identified by the consumers during an experiment by Luchs et al. (2012). Therefore, brands seeking to attenuate the sustainability liability effect could conduct market research to find the optimal design of a sustainable product, which in turn should increase consumer confidence and choice likelihood.

2.3.7. Call for Future Research Regarding the Sustainability Liability Effect

There is a lack of data on the sustainability liability effect in a field experiment setting. Skard et al. (2020) specifically state that future research should validate their findings using experimental designs in the field, measuring actual purchase decisions. This is due to one of their limitations being that the research was conducted in a laboratory environment. Most

other reports discussed throughout section 2 also lack data from field experiments, as questionnaires or laboratory experiments have been the main method of data collection.

In a current study and unpublished report, Skard et al. (2021) begin to investigate this aspect of the sustainability liability effect as they ask whether consumers are willing to choose green products in a field experiment setting. In particular, the gap between the consumers' positive view of sustainable products and actual purchase decisions is explored through two studies. Study 1 investigated the attitude of participants towards sustainable products with a survey. Participants were asked to select between a sustainable and regular cleaning product with strength-dependent attributes when imagining that they had to clean a plate. Results indicated a positive attitude towards sustainable products, with 67% out of 88 participants rating the sustainable product as having a higher quality than the regular product. Study 2 measured the behavior aspect by recording the same choice in a field experiment setting, where participants actually had to clean a dinner plate using either the sustainable or regular cleaning product. Here, 40% of 58 participants chose the green product, meaning a 27% drop from the hypothetical choice, demonstrating the sustainability liability effect and the attitude-behavior gap (Skard et al., 2021).

3. Household Cleaning Products

This section tackles the sustainability challenges and possible solutions of household cleaning products, which can be characterized as products with strength-related attributes.

3.1. Challenges

Several cleaning products throughout history have had unforeseen consequences that in turn have caused severe damage. Non-biodegradable and toxic cleaning products were used in a time when there was little knowledge about the consequences of releasing unprocessed toxic chemicals into nature (Hahn et al., 2019). An example is carbon tetrachloride, which was used in dry cleaners, but has since been phased out due to environmental and safety concerns. Exposure over time can affect the central nervous system, damaging the kidneys and liver, and even result in death (Hahn et al., 2019). Another example is the use of CFCs in solvent-cleaning applications (Kim et al., 2016). CFCs damage the ozone layer, increasing the amount of the dangerous UV radiation that reaches Earth's surface. Increased UV exposure can cause skin cancers, cataracts, and immune system damage in humans (WHO,

2021). Animals, marine organisms, and plant life are also vulnerable to exposure (WHO, 2021). It became critical to the point where the Montreal Protocol agreement was established as an international effort to phase out CFCs (EPA, 2021).

Other elements such as packaging can also pose immense challenges. For household cleaning products, a common form of packaging has been plastic containers. Most plastic is not biodegradable and often ends up in the ocean, posing a threat to aquatic animals and plants. Sea turtles often mistake transparent plastic for jellyfish and starve to death because they cannot process the plastic through their intestines (WWF, 2021). In fact, plastic is so prevalent in the ocean that in the Pacific Ocean, there is an area of accumulated plastic known as the Great Pacific Garbage Patch. The area is estimated to contain approximately 80,000 tons of plastic (Lebreton et al., 2018). Furthermore, plastic decomposes into microplastic in the ocean. Microplastic is defined as any plastic particle smaller than 5 mm in length (Kooi et al., 2016). Microplastic can become part of a food chain as small sea organisms consume it (The Ocean Clean Up, 2021), resulting in fish contaminated with microplastic (Kooi et al., 2016).

3.2. Sustainable Measures

As highlighted in section 1.2, in addition to protecting the environment and combating climate change, it is in the interest of businesses to become more sustainable to remain relevant and competitive (Lüdeke-Freund et al., 2019). By addressing social and environmental needs, a company creates superior value to its consumers compared to competitors (Lüdeke-Freund et. al, 2019).

Various sustainable measures have been implemented by sustainable household cleaning brands. An example is Klar, a Norwegian company focused on sustainable cleaning products ranging from laundry detergents to liquid soap (Klar, 2021). Klar (2021) lists the sustainable attributes of its household cleaning products, including ingredients without unnecessary chemicals, zero microplastics, vegan products that are not tested on animals, easily degradable cleaning detergents, packaging consisting of 100 % recycled plastic, and higher concentration of products for lower per dose use. Avoiding unnecessary chemicals, microplastics altogether, and testing products and animals help reduce the company's own negative externalities of cleaning products. Jørgensen & Pedersen (2018) argue that companies must address activities that cause negative externalities, which is what Klar has

done through these measures. Making detergents easily degradable also reduces the negative externality of harming the environment if the product is disposed of before being fully used. This approach to consumerism is known as "cradle-to-grave", where producers consider the impact of a product during each stage of its life-cycle, from extracting natural resources to the consumers' disposal of the product (Bocken et al., 2016). The use of 100% recycled plastic attends to other entities' negative externalities, as it creates a purpose for disposed plastic and thereby reduces plastic that reaches and harms the environment, as discussed in section 3.1. This also addresses the sustainable strategy of closing resource loops, which is that the loop between post-use and production is closed through recycling, resulting in a circular flow of resources when it comes to packaging (Bocken et al., 2016). A higher concentration of products addresses the sustainable strategy of slowing resource loops, which is redesigning products to extend their life, subsequently extending the utilization period of a product, thereby resulting in a slowdown of the flow of resources (Bocken et al., 2016).

However, due to the sustainability liability effect discussed in section 2.3, the sustainable measures and marketing of household cleaning products may not reduce the negative impact on the environment, as consumers may still choose to purchase non-sustainable household cleaning products instead. Furthermore, the higher concentration of detergents may not lead to lower per dose use, as discussed in section 2.3.4. Investigating whether the sustainability liability effect is present in the choice between sustainable and regular household cleaning products is therefore the motivation for an experiment, which is discussed in section 4.

4. Motivation for Experiment

4.1. Background

As mentioned in sections 2.3.3 and 2.3.7, Skard et al. (2021) conducted two studies to investigate the attitude-behavior gap in relation to the sustainability liability effect. However, there were a few elements of the experiment that could be improved, including sampling and the observation of participants during the experiment.

The sampling of the studies is a potential weakness for Skard et al. (2021) due to the number of participants and the representation of the general population. The survey had 88 participants and the field experiment had 58 participants. Skard et al. (2021) expressed concern that the sample sizes for both studies were too small. With a too low sample size,

external validity may be threatened as the findings of the studies may not be able to be applied to the general population. The external validity may also be threatened as there is not a proper representation of the general population (Saunders et al., 2016, p. 204). This is due to the studies being conducted with students as participants, meaning that the ranges of age and occupation are narrower compared to the general population.

The observation of participants during the field experiment may have also had an unintended effect on the results, particularly due to social signaling. Participants were observed and instructed throughout a significant portion of the experiment, although they were not directly observed during the choice of product. However, the feeling of being observed may have lingered when choosing a product. As such, social signaling may have increased the choice likelihood of the sustainable product, as participants want to portray themselves as more generous and fair-minded than they necessarily are (DellaVigna, 2009). As discussed in section 2.3.6.1, Luchs et al. (2010) investigated the social signaling bias and found that participants who were observed (versus non-observed) acted more sustainably.

Skard et al. (2021) wanted to conduct a similar study where these weaknesses were reduced or eliminated. As Skard and her research assistant Landsvik are part of the Centre for Sustainable Business at the Norwegian School of Economics, it was natural to start a dialogue with them about a new and improved experiment investigating the sustainability liability effect. The Centre for Sustainable Business at the Norwegian School of Economics conducts research, teaching, and outreach on the behaviors, business models, and technologies that promote sustainable business, with a focus within three main areas - sustainable and circular business models, sustainable consumption and behavior, and sustainability management (Centre for Sustainable Business, 2021). The main points of emphasis and improvement for the new experiment were to include a larger and more representative sample, and further reduce or remove the social signaling bias. Other tweaks are highlighted in section 7, including the incentive structure for participants and an updated survey. As such, the previous unpublished studies conducted by Skard et al. (2021) have been used as the foundation for a new and improved study on the sustainability liability effect, specifically when looking at household cleaning products with strength-dependent attributes.

4.2. What: To Test if the Sustainability Liability Effect Impacts Product Choice

Before redesigning and improving on the studies Skard et al. (2021) conducted, it is important to question what the reason for a new study is to begin with. The main reason for the study is to investigate whether there is a sustainability liability effect present that impacts the product choice for consumers. More specifically, the study will test whether the sustainability liability effect impacts the choice of a household cleaning product with strength-dependent attributes in a field experiment setting.

The findings of the study will be of interest for several reasons, including that a study in this specific setting has not been conducted before, the potential to compare results with Skard et al. (2021), and the possibility of exploring whether the continued trend of consumers viewing sustainable products more positively may outweigh the sustainability liability effect.

As Skard et al. (2020) outlined, there is little research on the sustainability liability effect using experimental designs in the field, with studies usually being conducted as surveys or in a controlled laboratory setting. The main field experiment study on the sustainability liability effect was study number 5 by Luchs et al. (2010), which aimed to validate studies 1-4 that were controlled laboratory experiments. As discussed in section 2.3.6.1, Study 5 investigated consumers' behavior in using sustainable versus non-sustainable hand sanitizer when being observed (versus non-observed), and found that social signaling was present. Skard et al. (2021) also used a field experiment design in one of their studies, as outlined in section 2.3.7, although improvements are necessary as discussed in section 4.1.

The potential to compare results with Skard et al. (2021) is also relevant, as the findings may replicate the results of Skard et al. (2021) due to the similar methodological structure of the experiment discussed in section 7.2. However, results are not fully comparable, as key differences include location, sample size and structure, and products. This is further discussed in sections 8.2 and 10.1.

As introduced in section 2.1, there has been a trend of consumers both viewing sustainable products more positively and acting more sustainably in recent years (Olsen et al., 2014; Deloitte, 2021; Sparebank1 Østlandet, 2018). In fact, Olsen et al. (2014) note that although

green products have been argued to be of inferior quality historically, consumers' attitudes towards green new products may overcome this bias as attitudes seem to have taken a significant positive turn. Therefore, although Skard et al. (2021) demonstrated that there is an attitude-behavior gap for consumers regarding sustainable products with strength-related attributes, almost 40% of participants still chose the sustainable product in an incentivized field experiment setting, suggesting a continued trend of a positive mindset when it comes to sustainable products which may outweigh the sustainability liability effect. Findings from the general public in a field experiment setting could strengthen the support for this trend.

4.3. Why: Multilayered Reasoning

There is a multilayered structure to why this experiment is conducted. The core reason is to test if the choice of household cleaning product (sustainable versus non-sustainable) in an incentivized field experiment setting results in participants selecting the non-sustainable product. The second layer is to find out why consumers may choose the non-sustainable product, and if this is mainly due to the sustainability liability effect. The third layer is to verify that the sustainability liability effect may counter sustainable marketing tactics, and serve as a call for producers to begin initiating new measures that in turn negate the sustainability liability effect. The outer layer is for the new sustainable measures to have a positive impact on the environment as the consumption of sustainable products increases and non-sustainable products decreases. Figure 1 shows the layered structure to the justification of the study. In short, the core and first layers concern the study itself, whereas the second and outer layers serve as reflections made depending on the results of the study, all of which are discussed in sections 8-10.

Figure 1: Visualization of the multilayered justification for the study. The core layer is product choice, followed by the sustainability liability effect, sustainable marketing, and finally the impact on the environment.

5. Research Question

Following the analysis of existing literature on sustainability, consumer behavior, products with strength-related attributes, and studies indicating the existence of the sustainability liability effect, the research question of the study is presented as follows:

Does the sustainability liability effect impact the choice of a product with strength-related attributes?

More specifically, the study will be conducted in a field experiment setting and focus on household cleaning items.

6. Research Model and Hypotheses

This section will discuss the research model and the hypotheses of the studies. The research model is illustrated in Figure 2. Although the sustainability liability effect is not a part of the figure, it is investigated as it is the phenomenon that leads to a lower perceived effectiveness of a sustainable (versus regular) product.
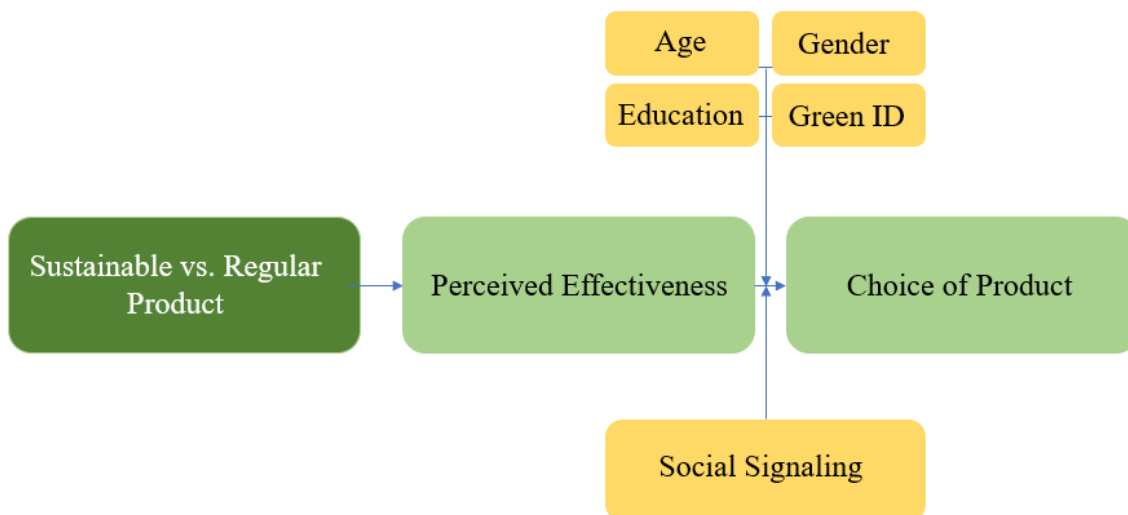
Figure 2: The conceptual research model for the study on the sustainability liability effect.

In Figure 2, the dark green box represents the independent variable of the sustainable versus regular product. Another way of interpreting the independent variable is the degree that a product is marketed as sustainable. The light green boxes represent the mediating and dependent variables, with the mediating variable being the perceived effectiveness of the products, and the dependent variable being the choice of product. Moderating variables are presented in the yellow boxes, and include social signaling, age, gender, education, and green identity.

6.1. Main Hypotheses Regarding the Sustainability Liability Effect

These hypotheses explore the sustainability liability effect in relation to the choice of sustainable versus regular products with strength-dependent attributes.

6.1.1. Effectiveness Self Condition

As discussed extensively in section 2.3, consumers may perceive sustainable products with strength-related attributes as less effective compared to competing regular products. This perception has been found by various studies, including Luchs et al. (2010), Lin & Chang (2012), Luchs et al. (2012), Newman et al. (2014), Pancer et al. (2017), Skard et al. (2020), and Skard et al. (2021). Behavioral reasons include that consumers may perceive a trade-off between strength and ethicality for the product (Luchs et al., 2012) due to sociocultural messages (Luchs et al., 2010; Pancer et al., 2017), infer that companies on a budget restriction sacrifice effectiveness for superior sustainability attributes (Luchs et al., 2010; Newman et al., 2014), or form lay theories due to missing information (Newman et al., 2014; Lin & Chang, 2012; Chen & Chang, 2012; Mishra et al., 1998; Darnall & Aragon-Correa, 2014).

Therefore, when participants indicate their own preferences, which is the "self condition" (Luchs et al., 2010) as discussed in section 2.3.6.1, they will perceive the regular product with strength-related attributes as more effective than the sustainable product with strength-related attributes.

H1: Consumers will perceive the regular (versus sustainable) product as more effective.

6.1.2. Effectiveness Other Condition

The hypothesis in this section is related to section 6.1.1 and hypothesis 1 in the sense that the underlying notion is that consumers will perceive the regular (versus sustainable) product as more effective. Here, participants will indicate their answers in the "other condition", which is what they believe the average consumer would respond (Luchs et al., 2010). For this study, the "other condition" is presented as the majority of the other participants of the experiment.

As previously discussed in section 2.3.6.1, Luchs et al. (2010) find that in the self condition, respondents may present themselves in a better light due to social signaling (DellaVigna, 2009), whereas in the other condition, respondents may reveal more honest answers since it is not a reflection of how they present themselves but instead how they present others. Therefore, the social signaling effect is reduced or removed in the other condition. Luchs et al. (2010) continue by highlighting the importance that respondents answer in the other condition when conducting market research to reveal as true responses as possible.

Also outlined in section 2.3.6.1, when it comes to social signaling and sustainability, consumers wish to be viewed as environmentally conscious even though they may not necessarily share this value, as being environmentally conscious is generally seen as positive by society. Therefore, the self condition answer in section 6.1.1 may be affected by social signaling being a factor in influencing respondents to indicate that the sustainable product is more effective (or that the sustainable and regular products are equally effective). However, in the other condition, social signaling is reduced or removed, and therefore true responses will exacerbate the perception that the regular product with strength-related attributes is more effective than the sustainable product with strength-related attributes.

H2: The effect postulated in H1 will be stronger when participants respond in the other condition (versus self condition).

6.1.3. Choice of Regular versus Sustainable Product

As discussed in sections 2.3 and 6.1.1, existing literature argues that consumers perceive the regular product to be more effective than the sustainable product (in the setting where the products have strength-related attributes and in reality do not differ in effectiveness). As consumers believe this to be true, in a rational setting where all else is equal, the consumer

would choose the product that is more effective. Therefore, consumers would choose the regular (versus sustainable) product if they believe that the regular product is more effective.

As discussed in sections 2.3.3 and 2.3.7, Skard et al. (2021), investigate the choice of sustainable versus regular products in an incentivized setting and seem to find this pattern. In their study, the majority of consumers chose the regular (versus the sustainable) product. Luchs et al. (2010) also find a majority preference for the regular (versus the sustainable product) with hand sanitizers.

Since the new experiment is based on the studies of Skard et al. (2021) and is developed in collaboration with Skard and Landsvik, the argument presented is that the majority of participants will choose the regular (versus sustainable) product.

H3: The choice share of the regular product will be higher than that of the sustainable product.

6.1.4. Choice Other Condition

As discussed in sections 2.3.6.1 and 6.1.2, when responding in the other condition (versus making the actual choice of product which can be categorized as the self condition), social signaling will be reduced or removed, which in turn should reveal more honest answers. Participants will therefore be affected by the perception that the regular product is more effective while being less subject to social signaling, and therefore be more likely to choose the regular (versus sustainable) product.

H4: The effect postulated in H3 will be stronger when participants respond in the other condition (versus actual choice).

6.1.5. Hypothetical Choice in Store

Participants will be asked to state which product (sustainable versus regular) they would buy in a store. If this is compared to the actual choice of product, an argument is that since participants imagine that they would be observed in a store, due to social signaling, participants would realize that it would be more likely that they would pick the sustainable product compared to if the choice was made in an isolated setting (which is the actual choice presented by H3). Therefore, the argument presented is that a higher percentage of

participants would indicate that they would buy the sustainable (versus regular) product in the hypothetical store setting compared to the choice made in the isolated setting of the study.

H5: The effect postulated in H3 will be weaker when participants respond in the hypothetical choice setting (versus actual choice).

6.1.6. Link Between Sustainability Liability Effect and Choice of Product

This section ties the previous hypotheses together (H1-H5) by stating that there is a link between the perceived effectiveness of a product and the choice made. The reasoning is that if a product with strength-related attributes is marketed as sustainable (versus not marketed as sustainable), the more likely a consumer may perceive the sustainable (versus regular) product as less effective, and the less likely a consumer chooses the sustainable (versus regular) product. Variations of this line of reasoning have been explored by Luchs et al. (2010), Newman et al. (2014), Pancer et al. (2017), Skard et al. (2020), and Skard et al. (2021).

H6: The effect of sustainability (regular versus sustainable product) on choice share (H3) will be mediated by beliefs about product effectiveness: The less a participant associates the sustainable product with efficiency (H1), the lower the choice share of the sustainable product.

6.2. Other Hypotheses

These hypotheses explore the moderating variables, including demographics such as age, gender, and education, as well as consumers' perception of their own green identity.

6.2.1. Age

As discussed in section 2.2, there is evidence that age is a factor in sustainable behavior, with younger generations seemingly having a more positive attitude towards sustainable products than older generations (Brown & Robertson, 2011; Kumar et al., 2012; Agarwal & Kasliwal, 2017). However, there is also evidence of little to no differences in environmental behavior depending on age (Agarwal & Kasliwal, 2017; Sarti et al., 2018; Gray et al., 2019).

To further elaborate, the demonstrated attitude-behavior gap by Skard et al. (2021) discussed in section 2.3.7 is also evidence of younger generations not acting as sustainable as expected,

as participating students stated environmental intentions, but tended to choose regular (versus sustainable) products when faced with an incentivized choice. The key here is the monetary incentive. Since younger generations do not enjoy the financial stability as older generations, and due to a general price premium for sustainable products, younger generations may not be able to fulfill their intended behavior of purchasing sustainable products. There is however evidence that younger generations are willing to pay a price premium upon reaching financial capabilities in the future (Kumar et al., 2012).

As the price of the sustainable and regular products in the study is implied to be equal, the ability to pay a price premium should not affect the choice of product. Therefore, the seemingly positive attitude towards sustainable products by younger generations (compared to older generations) may translate into a higher choice likelihood for the sustainable product. Therefore, as age increases, the likelihood of choosing a sustainable product decreases.

H7: The effect of sustainability on choice (H3) will be moderated by age: The choice share of the sustainable (versus regular) product will decrease as age increases.

6.2.2. Gender

As outlined in section 2.2, there is evidence of a gap in sustainable behavior between men and women, with women acting more sustainably (Davidson & Freudenburg, 1996; Lee & Holden, 1999; Cottrell, 2003; Dietz et al., 2002; Levin 1990; Zelezny et al., 2000; Brough et al., 2016). Furthermore, in the unpublished study conducted by Skard et al. (2021), they found a difference in the choice of sustainable versus regular products between genders. In short, there was an indication of the sustainability liability effect for male participants, whereas the sustainability asset effect was mitigated for female participants (Skard et al., 2021).

Since there is evidence of a gender gap in sustainable behavior, the argument presented is that women are more likely to choose the sustainable product compared to men.

H8: The effect of sustainability on choice (H3) will be moderated by gender: The choice share of the sustainable (versus regular) product will be higher for females (versus males).

6.2.3. Education

There is evidence that a higher level of education has a positive effect on sustainable attitude and behavior (Balin, 2021; Pelau & Catalina, 2018; Fawehinmi et al., 2020; Meyer, 2015), as highlighted in section 2.2.

With this in mind, the argument presented is that given a higher level of education, an individual will exhibit more environmentally friendly behavior, which in turn will increase the choice likelihood for the sustainable (versus regular) product.

H9: The effect of sustainability on choice (H3) will be moderated by education: The choice share of the sustainable (versus regular) product will increase as the level of education increases.

6.2.4. Green Identity

The degree a person identifies as "green" or environmentally conscious affects sustainable attitude, behavior, and consumption (Whitmarsh & O'Neill, 2010; Barr et al., 2005; Thøgersen & Ölander, 2006; Sarti et al., 2018), as noted in section 2.2. In an unpublished study, Landsvik (2021) investigated the link between a consumer's own green consumption values and product choice, and overall results indicated that the choice likelihood of a sustainable product increased the more consumers identified with having green consumption values.

On the other hand, there is evidence that environmentally conscious consumers may not necessarily act more environmentally friendly than consumers that are not environmentally conscious. As discussed in section 2.3.4, Lin & Chang (2012) found that environmentally conscious consumers use higher doses of various green products (including ones with strength-related attributes such as hand sanitizers and glass cleaners) relative to less environmentally conscious consumers. Lin & Chang (2012) speculates that environmentally conscious consumers may be susceptible to overuse since they are satisfied with the eco-friendly effort they have already made by choosing to purchase a green product.

Since the use of the sustainable and regular product is not measured in the study, but rather the choice between the two, the argument presented is that participants who identify

themselves as more green have a higher choice likelihood of the sustainable (versus regular) product.

H10: The effect of sustainability on choice (H3) will be moderated by green identity: The choice share of the sustainable (versus regular) product will increase the more participants identify themselves as green.

7. Methodology

This chapter will present the methods used to answer the research question and hypotheses, including research design, research strategy, sampling, experimental design, data collection, data analysis, and research ethics.

7.1. Research Design

A research design is an overall plan on how to answer the research question, including setting clear objectives derived from the research question, specifying on how data will be collected and analyzed, and discussing ethical concerns, limitations, and weaknesses (Saunders et al., 2016, p. 163). In this paper, limitations and weaknesses such as validity and reliability concerns are discussed in section 9.

The main objective of the study is to investigate whether the sustainability liability effect has an impact on the choice of a sustainable versus regular product with strength-dependent attributes in a field experiment setting. Another objective of the study is to find possible reasons behind the choice of sustainable versus regular products, including consumer perception of effectiveness, social signaling, demographics, degree of green identity, and aesthetic design.

The first methodological choice is regarding the choice of a quantitative, qualitative, or mixed methods research design, and this decision should be based on the research question presented in section 5 (Saunders et al., 2016, p. 164). Considering the research question, the study is explanatory in nature, with it attempting to determine a causal relationship between variables (being the indication of sustainability of a product, mediating beliefs of product effectiveness, and the choice of product) and the emphasis is on a situation or problem (being the possible presence of the sustainability liability effect) (Saunders et al., 2016, p. 176). A

quantitative approach complements the explanatory nature of the study (Saunders et al., 2016, p. 176). This is also in line with previous studies on the sustainability liability effect, which have been mainly quantitative, including the most recent ones by Skard et al. (2020) and Skard et al. (2021). Furthermore, the research study has a deductive approach, as it begins with a theory based on academic literature before a research strategy is designed (which is detailed in section 7.2) to test the theory (Saunders et al., 2016, p. 145). The research design is also mono method quantitative (since it is a use of a single quantitative data collection technique with the web-based survey tool Qualtrics, as further discussed in section 7.4.5), conducted through a field experiment which includes a survey, and cross-sectional (as it represents a snapshot in time) (Saunders et al., 2016, p. 164).

7.2. Research Strategy

The research strategy used in this study is an experiment with an incorporated survey. An experiment is an ideal form of research strategy since the purpose of an experiment is to investigate the relationship between an independent and dependent variable (Saunders et al., 2016, p. 178).

The type of experiment conducted is a field experiment, as it takes place outside the laboratory in a real-world setting (more on this in section 7.4.1) (APA Dictionary of Psychology, 2021a). The advantage of a field experiment is a stronger external validity, as its findings can likely be generalized to other relevant settings or groups, particularly compared to a controlled laboratory experiment (Saunders et al., 2016, p. 204). On the other hand, internal validity, which is the extent to which findings can be attributed to the interventions rather than any flaws in the research design (Saunders et al., 2016, p. 718), is likely weaker compared to a controlled laboratory setting, as in a field experiment one may not be able to account for and control all variables that affect the dependent variable. A thorough discussion of internal and external validity can be found in section 9.1.

The experiment itself is a type of within-subjects design, as there is only a single group rather than separate control and experimental groups, and every participant is exposed to the same intervention being an incentivized choice (which is detailed in section 7.4) (Saunders et al., 2016, p. 181). A within-subjects design was chosen to maintain similarity to the study by Skard et al. (2021) and for practical reasons. The same type of within-subjects design is used

by Skard et al. (2021), and a similar design therefore contributes to the ability of comparing results. A within-subjects design is also practical as it requires fewer participants, which is critical as the field experiment is relatively time-consuming per participant (more on this in section 7.3). A weakness to the study's variation of the within-subjects design is that there is no pre-intervention observation or measurement, as participants are only exposed to the planned intervention which is the product being marketed as sustainable (or not marketed as sustainable). This is discussed in section 9.1.

A questionnaire is incorporated at the end of the experiment. The goal of the survey is to collect data that can help determine if the sustainability liability effect is present, what variables may affect the choice of sustainable versus regular product, and categorize participants based on individual differences. This survey strategy is useful for multiple reasons, including that surveys are often used to determine and describe relationships between variables in a deductive research approach (Saunders et al., 2016, p. 181), which is the approach of this study as indicated in section 7.1. Furthermore, the survey strategy is helpful as it allows one to collect quantitative data that one can analyze quantitatively through descriptive and inferential statistics (Saunders et al., 2016, p. 181). Utilizing a survey is also advantageous due to it allowing easy comparison among participants. It is also comparatively easy to explain to and understand for participants (Saunders et al., 2016, p. 181). Furthermore, the concern of social signaling may be eliminated when using a questionnaire where participants are not observed, as participants are less likely to answer in a certain manner to satisfy social desirability and social signaling or appease the researchers (Dillman et al. 2014).

However, there are drawbacks to using a survey as a strategy that should be addressed. To begin, when designing the questionnaire, the following stages must take place for a question to be valid and reliable: The researcher is clear about data requirements and designs the question → the respondent decodes the question as the researcher intended → the respondent answers the question → the researcher decodes the answer in the way the respondent intended (Saunders et al., 2016, p. 450). In particular, the second step of this process is of concern, as the responses to the questionnaire are intended to not be observed by the researchers during the experiment, and therefore there is not an opportunity to provide clarifications without possibly contaminating the answer through removing the feeling of isolation for the participant and thereby introducing social signaling. To counter this potential

issue, simple and clear language is used in the questionnaire, questions are accompanied with clarifying pictures of the sustainable and regular product, previous questions from the unpublished research studies by Skard et al. (2021) and Landsvik (2021) are used and altered with their approval, and new questions are developed in collaboration with Skard and Landsvik. Another concern is that any form of contamination of the respondent's answers will reduce the data's reliability (Saunders et al., 2016, p. 442). A type of contamination that may occur is if the respondent does not have sufficient experience or knowledge to answer the question, the participant may attempt to guess the answer, a phenomenon known as an uninformed response. This is particularly likely when the questionnaire is incentivized (Saunders et al., 2016, p. 442), which is the case of the study as respondents are given a 150 NOK gift card after completing the survey (more on this in section 7.4.2). However, uninformed responses are likely avoided, as the questions do not require any previous knowledge or experience, and ask for the participants' subjective opinions. The likelihood of contamination is also low as long as the participants are left to complete the survey by themselves (Saunders et al., 2016, p. 441). A more detailed discussion of the survey can be found in section 7.4.5.

7.3. Sampling

The sample used for the experiment including the survey consisted of 101 people. As the experiment had a within-subjects design, all 101 participants completed both the experiment and the incorporated survey. The participants were actively recruited by the researchers throughout the duration of the experiment at IKEA Åsane (more on the location in section 7.4.1). In the end, the sample consisted of participants that purchase and use household cleaning products (assumed to be every participant that chose to be a part of the study) with an age range of 16 to 95, along with both genders and different levels of education. It was determined that the minimum age to participate in the experiment was 16, due to the notion that participants should have a relationship to purchasing and using household cleaning items, as well as participants receiving the incentive of a 150 NOK gift card. As the study was a field experiment conducted at IKEA Åsane, most participants can be assumed to be living in Bergen and therefore being Norwegian residents and/or citizens. This assumption is further strengthened by the fact that participants were required to be able to speak Norwegian fluently.

Achieving a large enough sample size is crucial to ensure external validity in the study (Saunders et al., 2016, p. 204). Furthermore, a goal of the study was to have a larger sample size than Skard et al. (2021) achieved during their unpublished studies (88 for the questionnaire which was study 1, and 58 for the field experiment which was study 2), as they expressed concern that their sample sizes were too small, particularly since the samples were comprised of students.

The required sample size for the experiment can be calculated based on certain criteria, including test family, sample groups, number of tails, effect size, significance level, and power (AI-Therapy Statistics, 2021a). T-test is the test family used, as discussed in section 7.6.1.2. Despite the study having a within-subjects design, the sample groups are defined as independent groups, as they are divided based on the choice of product (as demonstrated in section 7.6.4). The number of tails is one, since the hypotheses presented have an implied direction (AI-Therapy Statistics, 2021a), as indicated in section 6. For the significance level and power, commonly applied values of 0.05 and 0.8 are used (AI-Therapy Statistics, 2021a). The final variable is effect size, which is the standardized way to report the strength of a relationship between variables (AI-Therapy Statistics, 2021a). Effect size can only be calculated after data has been collected from the participants (Zint, 2006). However, according to Cohen's d, the lowest effect size to be accepted as small is 0.2 (AI-Therapy Statistics, 2021b). With an effect size of 0.2, the calculated required sample size is 620 (AI-Therapy Statistics, 2021a). Assuming a medium effect size (between 0.5-0.8) of for instance 0.6, the calculated required sample size is 72.

There were restrictions to the sample size due to logistical restrictions such as funding and time. The funding received from the Centre for Sustainable Business amounted to approximately 16,000 NOK. It was determined that every participant should receive 150 NOK as an incentive to complete the study (more on this in section 7.4.2), and therefore 15,000 NOK was allocated towards this, meaning the intended sample size would be 15,000 NOK/150 NOK per participant = 100 participants. The final sample size was 101 participants, as one participant declined to receive the 150 NOK gift card. The remaining 1000 NOK were allocated to buy additional equipment for the experiment. Furthermore, for the study to be cross-sectional, it was critical to keep the time frame to a minimum, and it was estimated that getting 100 participants would take approximately 1-3 days at IKEA Åsane. The experiment was indeed time-consuming, with one participant taking about 3-10 minutes

to complete it. In addition, there was not a continuous flow of people willing to participate, and therefore there were stretches of up to 30 minutes where there were no participants despite the experiment being set up. It is worth noting that the sample size of 101 is satisfactory when the assumed effect size is above 0.51 (medium effect size), as the required sample size is 98 (AI-Therapy Statistics, 2021a).

7.4. Experimental Design

The experiment with the incorporated survey was conducted physically at IKEA Åsane in Bergen. The data was therefore collected in the field experiment setting with a laptop that participants used to fill out the questionnaire. The choice of product was completed in the same questionnaire by the researchers before asking the participants to complete the rest of the questionnaire. This process will be discussed in detail in section 7.4.5. To provide clear and simple instructions for participants and given the requirement that participants spoke Norwegian fluently, all instructions and survey questions were presented in Norwegian (Bokmål).

7.4.1. Location and Time

The location of the experiment was IKEA Åsane in Bergen, Norway.

The choice of location for the field experiment began by narrowing down options to large indoor public spaces. The experiment was to be conducted indoors to not be directly affected by the weather (both cold temperatures and rain), and a public space was necessary to gain access to a large, general population. The conclusion was that a shopping mall or a supermarket would be the ideal spot. Multiple shopping malls and supermarkets in Bergen were contacted, and IKEA Åsane was quick to respond with interest. IKEA Åsane was a good fit due to its large size and customer capacity. It has an area of 62,500 m² (Multiconsult, 2012), and approximately 800 000 customer visits on an annual basis, encompassing all ages, genders, and levels of education (Hitland, 2021). Furthermore, there was complementary value to the location of IKEA Åsane and providing incentives for potential participants. IKEA Åsane was willing to supply 100 IKEA gift cards valued at 150 NOK each that were funded by the Centre for Sustainable Business at the Norwegian School of Economics.

IKEA Åsane also recommended both the specific times and places for conducting the experiment. The best option was to set up the experiment just inside the entrance of IKEA,

across from an information desk, as customers would be optimally incentivized to participate in the experiment as they could use the 150 NOK IKEA gift card immediately after completing the experiment. Customers would also not be tired after shopping, which would be the case if the experiment took place by the exit (since the entrance and exit of IKEA are located at different places). Details of the gift card incentive are discussed in section 7.4.2. With the consultation of IKEA Åsane, the majority of the experiment occurred during peak customer time, which is Saturdays from noon till around 4 p.m. The experiment began Saturday October 30, 2021 at noon. After conducting the experiment on Saturday October 30th, Monday November 1st, and Tuesday November 2nd, the intended goal of 100 participants was surpassed.

7.4.2. Equipment

A variety of equipment was used to conduct the experiment, including plastic bottles with stickers of fictional cleaning brands, room dividers, a large sign to draw people in, gift cards, a non-disclosure agreement, printed instructions, records of participants receiving a gift card, tape, a dish brush, a cleaning sponge, rubber gloves, steel wool, hand wipes, tables, and a chair.

The plastic bottles used were cleaning spray bottles with attached stickers of two fictional brands, *Eco-Sera* and *Sera*, which were inspired by the studies of Skard et al. (2020) and Skard et al. (2021). Orkla supplied the stickers of the fictional brands. The stickers of Eco-Sera and Sera share multiple similarities, including the statements that the product is intended to clean ovens and grills, and that it can clean food waste and dirt that has been burnt and stuck. There are a few differences between the two stickers to indicate that Eco-Sera is the sustainable product, whereas Sera is the regular product. The Eco-Sera sticker is the color green and contains the additional statement "100% natural ingredients". The color green was chosen as it has been long associated with nature and has been adopted as a prevalent marketing tool for environmentally conscious consumption (Pancer et al., 2017). The "100% natural ingredients" statement is in relation to the green core attributes aspect studied by Skard et al. (2020). The Sera sticker is the color orange and does not include an additional statement. The color orange was chosen to clearly differentiate from the color green. Orange is also often used by brands to draw attention, as it is a stimulatory color (Lischer, 2021). An example is Tide, a large American producer of cleaning detergents, which uses orange as its main brand color (Tide, 2021). The "100% natural ingredients"

statement and one of the brand names (*Sera*) were used in the study by Skard et al. (2021) and were therefore adopted in this experiment, with the main differences being the product (oven cleaner instead of dish- and kitchen spray), the second brand name (*Eco-Sera* instead of *Aveno*), and colors (green and orange instead of black and white). Pictures of the Eco-Sera and Sera cleaning spray bottles are displayed below as Figure 3.

Figure 3: Cleaning spray bottles used with stickers of the fictional brands, Eco-Sera and Sera.

Room dividers were used to create the feeling of isolation for the participants, thereby reducing or possibly eliminating social signaling. Additionally, room dividers were used to create two separate rooms for the experiment, with the first room being where the participants made the choice between Eco-Sera and Sera, and the second room being where participants were instructed to complete the questionnaire (more on this in sections 7.4.4. and 7.4.5).

A large sign was used to draw attention to potential participants. The sign displayed the logo of the Norwegian School of Economics and a statement in Norwegian which translates to: "Participate in a product test as a part of our master's thesis and receive an IKEA gift card!". The Norwegian School of Economics logo was displayed as the school and brand are well known in both Norway and Bergen, meaning that customers at IKEA would likely recognize it. The statement was used to attract consumers as well, informing them that they would be of

help in a product test, as well as receiving a gift card at IKEA in return. The reason why the experiment was disclosed as a product test for participants is discussed in section 7.4.3. Pictures of the sign are seen below in Figure 4.

Figure 4: The large sign used to advertise the experiment and to draw attention to potential participants.

100 IKEA gift cards valued at 150 NOK each were a crucial element of the experiment to create an incentive for the participants when making a choice between Eco-Sera and Sera. This is similar to the 250 NOK incentive provided by Skard et al. (2021), where they concluded that the monetary incentive led to participants choosing the product that they believed was the most effective when anticipating a cleaning challenge. The limitation of 150 NOK gift cards was due to funding as highlighted in section 7.3. Further discussions of the incentive structure are found in sections 7.4.4 and 7.4.5.

Other equipment used for the experiment included a non-disclosure agreement, printed instructions, records of participants receiving a gift card, tape to hold these documents in place, cleaning tools including a dish brush, a cleaning sponge, rubber gloves, and steel wool to create the feeling for participants that they would be participating in a cleaning challenge (which is discussed in section 7.4.4), a laptop for the questionnaire that participants were asked to complete, hand wipes for participants to clean their hands before and after the

experiment, three tables to place the various items in the different rooms/stations, and a chair for participants to sit in while completing the questionnaire.

7.4.3. Instructions Before Participating

The first step of the field experiment was to attract participants and give them limited but necessary information on what they would be a part of. When entering IKEA Åsane, customers would walk up a flight of stairs or use an escalator to immediately go to the first floor where an information desk would be located. With the experiment being located near this information desk, IKEA customers could clearly observe the researchers and the large sign used to draw attention as illustrated in Figure 4. As mentioned in section 7.4.2, the sign displayed the Norwegian School of Economics logo and stated that customers could participate in a product test as a part of a master thesis while receiving an IKEA gift card. The researchers would also actively recruit participants by approaching IKEA customers and asking a question similar to the statement on the sign.

The reason customers were informed that they would participate in a product test was to make them unaware that they were in fact a part of an experiment. Since participants did not know that they were a part of an experiment, there was a lower likelihood that demand characteristics, which are the clues in an experiment that lead to participants understanding what the researcher is looking for, affected the results (McLeod, 2012). Demand characteristics were also the reason why any suggestion of a focus on sustainability was eliminated, as participants may be influenced to infer that the "right" choice is the sustainable product, and subsequently choose this. If IKEA customers had any questions for the researchers regarding the experiment disguised as a product test, they would be provided with additional details, including that the estimated length of the experiment was approximately five minutes, and that one would receive an IKEA gift card valued between 50 and 150 NOK. If all instructions were followed and the completion of the test was satisfying, one would receive the 150 NOK gift card.

Participants were also instructed to sign a non-disclosure agreement before starting the "product test". The non-disclosure agreement consisted of two main points. The first was that participants would not share information regarding the process of the product test during the next 24 hours to protect the integrity of the test. The reason for this point was to avoid contamination of the experiment by information spreading to other IKEA customers

regarding what would actually occur. The second point was to inform participants that the objective of the test was to collect data and that the participant would remain anonymous. Figure 5 below are pictures of the placement of the non-disclosure agreement, which was to the left of the large advertising sign. The full non-disclosure agreement can be found in the appendix under section 13.1.1.
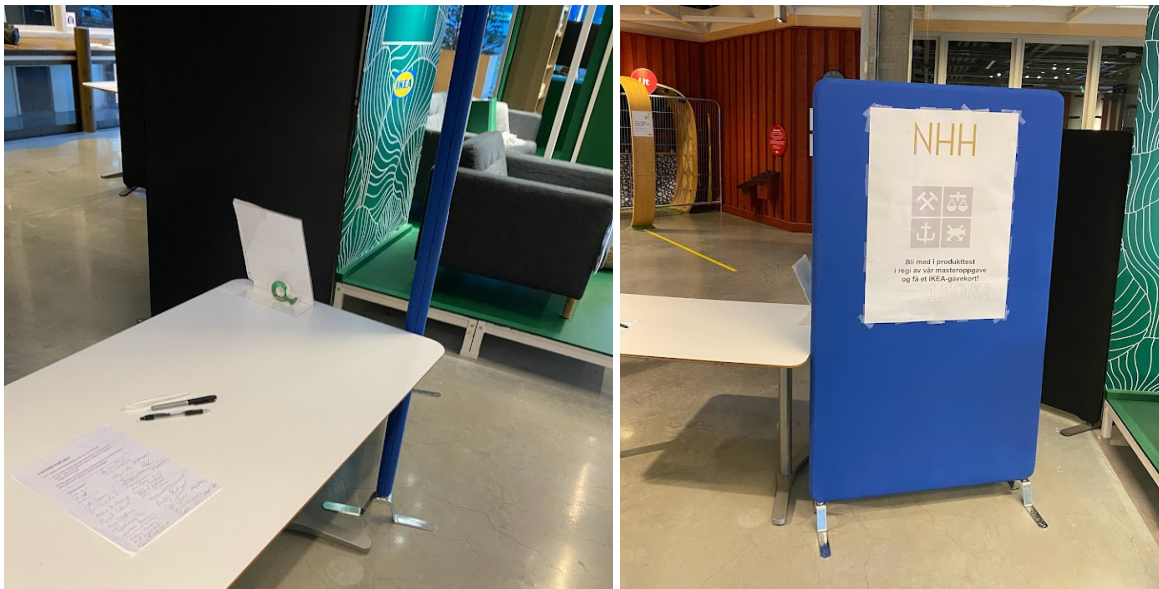
Figure 5: Non-disclosure agreement on the table pictured to the left, the table placed next to the large promotional sign pictured to the right.

7.4.4. Instructions for Choice of Product - Room 1

After signing the non-disclosure agreement, participants were instructed to go to room 1, which was an area isolated by room dividers. Participants were told to follow the instructions placed on the wall in room 1, and after following all of the instructions to move to room number 2 which was past the room divider to the left. Figure 6 below illustrates what the participants observed when entering room 1.

Figure 6: Pictures of room 1 where participants first entered the experiment.

As seen in Figure 6, participants were met with household cleaning items placed on a table, with the most important ones being the fictional sustainable (Eco-Sera) and regular (Sera) oven cleaners, along with a set of instructions on the wall.

The instructions were as follows (when translated to English, the original version can be found in the appendix in section 13.1.2):

READ CAREFULLY - DO NOT ENTER THE NEXT ROOM BEFORE YOU HAVE READ THIS

1. In front of you are two products, SERA and ECO-SERA.
2. You must choose ONE of the products that will be used to clean an oven rack in the next room. If you are able to clean the oven rack properly within 1 minute, you will receive an extra 100 NOK IKEA gift card (a total of 150 NOK). You can also bring other cleaning supplies that you see here.
3. Now, choose EITHER SERA OR ECO-SERA and bring it with you to the next room. In addition, bring a sponge, dish brush, rubber gloves, and/or steel wool!

The key point of the instructions in room 1 is point 2, which informs the participants that there is an incentive and a time limit to complete the task of cleaning the oven rack properly. The gift card incentive is intended to ensure that the consumers choose the product they believe is the most effective. This was the case for Skard et al. (2021), who determined that the sustainability liability effect was found due to the implementation of the monetary incentive. Furthermore, the time limit and cleaning challenge are supposed to create a high-stakes environment that further indicates to the participant that it is advantageous to choose the more effective product.

The room dividers and leaving the participants alone in room 1 were measures taken to eliminate social signaling, as participants were supposed to feel isolated and unobserved when choosing between the sustainable (Eco-Sera) and regular (Sera) oven cleaners. Therefore, participants were not supposed to be influenced to choose a product that would make them seem more generous and fair-minded (DellaVigna, 2009), and in this case would direct them towards the sustainable product (as discussed in section 2.3.6.1). These measures were designed as improvements to the attempts of removing social signaling by Skard et al. (2021). The difference in the IKEA experiment compared to Skard et al.'s (2021) experiment was that the participants were observed and instructed to a lesser degree by the researchers, as the signs with instructions replaced verbal instructions to a larger degree.

Another purpose of room 1 was to create the feeling for the participants that they would clean an oven rack. Due to time restrictions and practical reasons, it was determined that it was not necessary for the participants to actually clean an oven rack. Instead, the belief that they would clean it was sufficient, as the choice would be made before actually taking on the perceived task in room 2. Therefore, the critical moment of the first part of the experiment was the participants' choice of Eco-Sera versus Sera based on the belief that they would clean an oven rack.

7.4.5. Instructions after Choice of Product and Survey - Room 2

After making the choice between Eco-Sera and Sera, participants would bring the chosen product along with other cleaning supplies to room 2. Here, the participants would be met by one of the researchers who would explain to them that there was in fact no cleaning challenge, but instead that the participant would receive the full 150 NOK IKEA gift card by

completing a short questionnaire. The participant would then be left alone to complete it. The layout of the second room is illustrated in Figure 7 below.
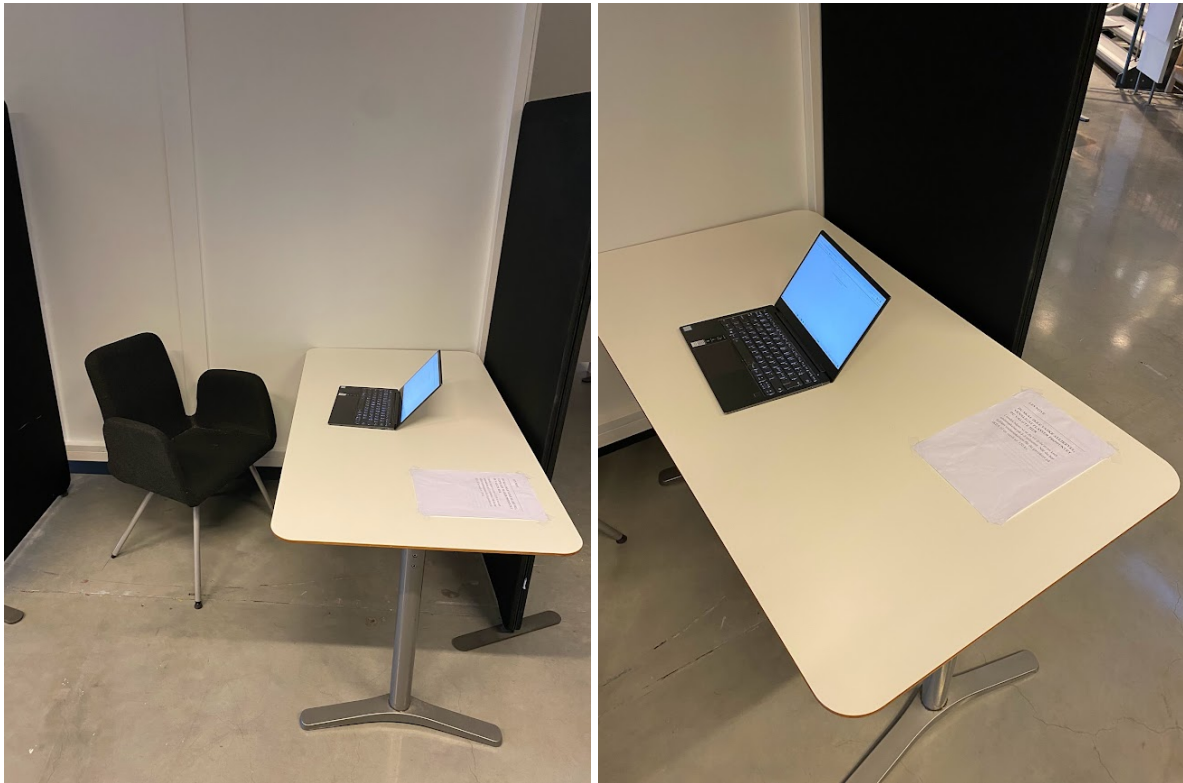
Figure 7: Pictures of room 2 of the experiment.

As observed in Figure 7, room 2 consisted of a piece of paper with a set of instructions, as well as a laptop where participants would complete the questionnaire.

The instructions on the paper were as follows (when translated to English, the original version can be found in the appendix in section 13.1.3):
READ CAREFULLY - YOU ARE NOT SUPPOSED TO CLEAN AFTER ALL, PLEASE PLACE THE PRODUCT YOU CHOSE HERE.
Instead, we hope that you will complete a short questionnaire that will take approximately 2-3 minutes. When you have completed the questionnaire you will receive an IKEA gift card with a value of 150 NOK.

Once again, the room dividers, printed instructions, and leaving the participants alone when completing the questionnaire were all measures taken to make the participants feel unobserved, therefore attempting to eliminate social signaling. As mentioned in section 7.2,

Dillman et al. (2014) note that social signaling may be eliminated when respondents feel unobserved when answering a questionnaire.

The reason for the researchers also explaining the steps of the second room to the participants was mainly for the researchers to record the choice of product as the answer to the first question of the questionnaire. The researchers recorded this instead of the participants to ensure that the observed and recorded choice was correct. After recording the answer to question 1, the laptop was placed on the table for the participant to complete the rest of the survey. Question 1 of the survey is illustrated in Figure 8 below.

Figure 8: Question 1 of the survey, which asks which product the participant chose. The question was answered by the researchers instead of the participants. The participants therefore started the survey by viewing question 2.

The survey used to collect data was developed and completed on Qualtrics, a cloud-based platform for creating and distributing surveys on the internet (Kent, 2021). The full questionnaire can be found in section 13.2 of the appendix. The questionnaire was written in Norwegian (bokmål), but all questions will be translated to English and explained in this section. The reasoning for including the questionnaire is discussed in section 7.2. The concepts of the questionnaire are based on the research question (section 5) and hypotheses (section 6), which in turn are based on existing literature (section 2). Existing literature

includes the unpublished studies by Skard et al. (2021) and Landsvik (2021), where the surveys they used serves as the foundation to the questionnaire in the IKEA study, as intended and approved by Skard and Landsvik.

To avoid leading questions, the questionnaire was structured to ensure that the answers to the dependent variables were not influenced by the answers regarding the mediating variable, moderating variables and different checks. Therefore, the questions regarding choice (dependent variable) were presented first, followed by questions on perceived effectiveness (mediating variable), different checks, and green identity and demographics (moderating variables). The exception is the possible moderating variable of social signaling, which is addressed throughout the questionnaire by comparing self condition and other condition questions.

The survey was also structured to only allow closed questions, meaning that there were limited options for answers for each question. Most of the questions used a variation of the 5-point Likert scale format, which is the most frequently used rating question format in which the respondent is asked how strongly he/she agrees or disagrees with a statement (Saunders et al., 2016, p. 457). The Likert scale is effective at measuring attitudes, and traditionally uses five points (anchors) where the middle point is neutral (Chyung et al., 2017). Agree/disagree anchors were mainly replaced with Eco-Sera/Sera, with the structure of the alternatives being as follows:

Sera is much better | Sera is a little better | Sera and Eco-Sera are equal | Eco-Sera is a little better | Eco-Sera is much better

The same foundational structure for the variation of the 5-point Likert scale questions was used consequently for both consistency and to eliminate confusion among respondents (Dillman et al., 2014). In addition, different versions of closed questions were used where it served as a better fit, including a 7-point Likert scale for the questions regarding the theme of green identity, along with a few questions with only two options as answers (being Sera or Eco-Sera). The reasoning for the alternate use of closed questions will be addressed during the breakdown of the questions later in this section. Furthermore, there was an equal distance between the options, which is critical for the Likert scale to be categorized as an interval scale (more on this in section 7.5.1) (Chyung et al., 2017).

45

As mentioned previously, the first question of the questionnaire was regarding the choice of Sera versus Eco-Sera and was completed by the researchers upon observing the participant's choice of product. The second question (meaning the first one completed by the participants) was reading and agreeing to the statement that the questionnaire is conducted in compliance with the Declaration of Helsinki, which is a set of ethical guidelines for research involving human subjects (World Medical Association, 2021). The statement also included that the study is performed by students at the Norwegian School of Economics, participating is voluntary and anonymous, and by completing the questionnaire one agrees to participate and the answers will be analyzed as part of a study.

Questions 3 and 4 were concerning the dependent variables of the experiment, specifically the choice of sustainable versus regular product, and were as follows:

Question 3: Which product do you believe the majority of other participants will choose?

Sera | Eco-Sera

Question 4: If you were to buy this product in the store, which product would you have chosen?

Sera | Eco-Sera

For both questions, participants were presented with two options. The reason for this (instead of using a Likert scale) was to be able to directly compare the answers to the observed choice of regular versus sustainable products. Question 3 is based on the use of self condition versus other condition (with question 1 being the self condition and question 3 being the other condition, with the goal of eliminating social signaling as a moderating variable) by Luchs et al. (2010), whereas question 4 is inspired by the Skard et al. (2021) questionnaire. For more detailed reasoning of the questions, see sections 6.1.4 and 6.1.5, which explain hypothesis 4 (linked to question 3) and hypothesis 5 (linked to question 4).

Questions 5 and 6 focus on the mediating variable being the perception of product effectiveness. The questions were presented as follows:

Question 5: During tests of these products, which one do you believe was the most effective against burnt food scraps?

Sera is much more effective at removing burnt food scraps | Sera is a little more effective at removing burnt food scraps | The products are equally effective at removing burnt food scraps | Eco-Sera is a little more effective at removing burnt food scraps | Eco-Sera is much more effective at removing burnt food scraps

Question 6: Which product do you think the majority of other participants believe is the most effective one?

Sera is much more effective at removing burnt food scraps | Sera is a little more effective at removing burnt food scraps | The products are equally effective at removing burnt food scraps | Eco-Sera is a little more effective at removing burnt food scraps | Eco-Sera is much more effective at removing burnt food scraps

Here, the variation of the 5-point Likert scale is introduced for both questions. Question 5 is based on a similar question used in the unpublished study by Skard et al. (2021), whereas question 6 presents the other condition version of question 5 (with question 5 being the self condition). For more detail on the reasoning of the questions, see sections 6.1.1 and 6.1.2, with hypothesis 1 being linked to question 5, and hypothesis 2 being linked to question 6.

Question 7, 8, and 9 serve as different checks of certain variables of the experiment. Variations of the 5-point Likert scale were used for these questions:

Question 7: Which of the products do you believe is the most environmentally friendly?

Sera is much more environmentally friendly | Sera is a little more environmentally friendly | The products are equally environmentally friendly | Eco-Sera is a little more environmentally friendly | Eco-Sera is much more environmentally friendly

Question 8: To what degree do you associate the products with strength/efficiency?

Sera is much stronger/more effective | Sera is a little stronger/more effective | The products are equally strong/effective | Eco-Sera is a little stronger/more effective | Eco-Sera is much stronger/more effective

Question 9: Which of the products do you believe has the best design?
Sera has a much better design | Sera has a little better design | The products' designs are
equal | Eco-Sera has a little better design | Eco-Sera has a much better design

Question 7 is a manipulation check for sustainability. A manipulation check is conducted to evaluate the efficacy of an experimental variable, meaning that it confirms that a variable or manipulation affects the participants as intended (APA Dictionary of Psychology, 2021b). This means that question 7 aims to verify that participants recognize Eco-Sera (versus Sera) as the sustainable product. Question 7 is based on previous studies on the sustainability liability effect that have included similar manipulation checks, such as Luchs et al. (2010) and Lin & Chang (2012).

Question 8 serves as a proxy of each product's strength, testing if participants associate the products with strength, and if participants believe that Sera (versus Eco-Sera) is stronger/more effective. The question can also be used as an alternative measurement to the belief of effectiveness as a mediating variable for the choice of product.

The purpose of question 9 is two check whether participants view the aesthetic design of the sustainable versus regular products as equal, which is the intention of the researchers. Luchs et al. (2012) argues that a superior aesthetic design has a disproportionately positive effect on the choice likelihood for sustainable (versus regular) products. Therefore, it is of interest whether consumers view either Sera or Eco-Sera as having superior aesthetic design, and in turn if this could explain the product choice to some degree. This is discussed further in sections 7.6.3.3 and 9.1.1.4.

The remaining questions concern moderating variables that may have had an impact on the choice of product. To begin, question 10 deals with the concept of green identity, and is a set of four statements:

Question 10: Here we want you to share your opinion on various statements.
-   It is important to me that the products I use are not environmentally harmful.
-   I think about how my choices and actions can affect the environment.
-   I would describe myself as environmentally conscious.

- I am willing to do something that is more cumbersome to make a choice that is more environmentally friendly.

All of these statements were accompanied by the same 7-point Likert scale:

Strongly Disagree | Disagree | Slightly Disagree | Neutral | Slightly Agree | Agree | Strongly Agree

The set of green identity questions is taken from the unpublished study by Landsvik (2021) as she recommended and approved the inclusion of it to further investigate the link between the choice of product and the extent to which participants view themselves as environmentally conscious consumers. Landsvik (2021) has based the green identity questions on a scale of green consumption values developed by Haws et al. (2014). The reason for the 7-point Likert scale is to maintain similarity to the green identity questions by Landsvik (2021), since a 7-point Likert scale was used here as well. This topic is further elaborated on in section 6.2.4, and the set of statements correspond with hypothesis 10.

Questions 11, 12, and 13 are included to identify possible links between demographics and choice of product:

Question 11: What is your age?
Answers were submitted on a slider with a lower limit of 16 and an upper limit of 100.

Question 12: What is your gender?
Male | Female | Other/Prefer Not To Say

Question 13: What is your level of education?
Primary/Secondary School | High School | College/University | Prefer Not To Say

Similar demographic questions were also included in the unpublished study by Skard et al. (2021) and are of interest as multiple studies have investigated the relationship between age, gender, education, and sustainable views, as outlined in section 2.2. The relationship between demographics and choice of sustainable versus regular product is further discussed in sections

6.2.1, 6.2.2, and 6.2.3, with question 11 being linked to hypothesis 7, question 12 being linked to hypothesis 8, and question 13 being linked to hypothesis 9.

The final page of the questionnaire thanks the participant for completing the questionnaire, and states that for the participant to receive an IKEA gift card worth 150 NOK, the participant must agree to sign a form that confirms that the gift card was received. The confirmation of participants receiving the gift card was completed at the request of the Centre for Sustainability at the Norwegian School of Economics as a requirement to receive funding. The form can be viewed in section 13.1.4 of the appendix.

7.5. Data Collection

7.5.1. Data Types and Collection
The observed product choice along with the questionnaire completed by participants was collected using Qualtrics, with the data being stored in the Qualtrics database. The data was also cleaned in Qualtrics by removing incomplete answers manually. After the experiment was completed, the data was downloaded as an Excel file where it was checked once more before it was uploaded to Jamovi (more on Jamovi in section 7.6.1).

The data used for analysis consists of four different data types - data describing product choice, data describing factors of choice, data on different checks, and data on individual differences. The data describing product choice consisted of the alternatives Sera (the regular product) and Eco-Sera (the sustainable product) as the answers. These were assigned the numbers 1 (Sera) and 2 (Eco-Sera) for quantitative analysis. The data describing the factors of choice and the different checks consistently use a variation of a 5-point Likert scale (ranging from 1 being Sera - much to 5 being Eco-Sera - much) as detailed in section 7.4.5, and can be defined as interval data since the difference between any two data values of a variable can be stated, but for which the relative difference cannot be stated (Saunders et al., 2016, p. 719). The data on individual differences were assigned values depending on response alternatives. Responses to age (which were inputted using a slider ranging from 16 to 100) were directly analyzed as it is a type of ratio data, as both the difference and relative difference between any two values can be stated (Saunders et al., 2016, p. 725). The response alternatives for gender and education were assigned number values for quantitative analysis. For gender, male was assigned the number 1 and female was assigned the number 2. For

education, primary/secondary school was assigned the number 1, high school was assigned the number 2, and college/university was assigned the number 3. Finally, the concept of green identity used a range of 1-7 as values (ranging from 1 being strongly disagree to 7 being strongly agree), since a 7-point Likert scale was used for its collection (as discussed in section 7.4.5), and thereby analyzed as a type of interval data.

7.5.2. Data Combination and Processing

The dataset from the survey was downloaded as number values and assigned variable names in Jamovi. The assigned variable names were as follows:

Actual Choice - The choice between sustainable and regular products.

Choice Other Condition - A participant's belief of what the majority of other participants would choose.

Hypothetical Choice in Store - Product choice by the participant in a hypothetical store setting.

Efficient Self Condition (Perceived Effectiveness SC) - A participant's own perception of product effectiveness.

Efficient Other Condition (Perceived Effectiveness OC) - A participant's belief of the majority of other participants' perception of product effectiveness.

Sustainability Check - A participant's perception of which product is more sustainable.

Strength/Efficiency Check - A participant's perception of which product is stronger/more effective.

Aesthetic Design Check - A participants' perception of aesthetic product design.

Green Identity (1-4) - Measurement of green consumption values. Consists of four questions, therefore 1-4.

Age - Age of participant.

Gender - Gender of participant.

Education - Level of completed education of participant.

For *Efficient Self Condition, Efficient Other Condition, Sustainability Check, Strength/Efficiency Check, Aesthetic Design Check, Green Identity (1-4),* and *Age,* the number format was changed from the standard format in Jamovi to the correct data type as stated in 7.5.1. For the remaining variables, the standard format in Jamovi was correct.

Data columns for *Green Identity (1-4), Age, Gender,* and *Education* were transformed using the transform function in Jamovi. *The Green Identity (1-4)* 7-point Likert scales were combined into *Green Identity (Total Average)*, measuring the average of the responses to the four statements for each participant. *Age* was filtered into 10-year intervals, except for the "16-19" and "70+" categories, and called *Age (10-year intervals)*. As there was only one non-binary gender respondent, *Gender* was transformed to *Gender (2 values)*. *Education* was transformed to *Education (2 values)*, meaning the two values were "no college degree" and "college degree", and this was further transformed to *Education (filtered)*, which filtered out those who did not disclose their academic record.

7.6. Data Analysis

This section will cover the methods and thought processes used to understand the sample data. This includes analytical methods, descriptive statistics, different checks, and testing the hypotheses.

7.6.1. Analytical Methods

The data sample was analyzed with Jamovi, a free software for statistical analysis (Jamovi, 2021). The application runs on R-scripts, but users are not required to be familiar with R as they interact directly with the Jamovi interface. With the use of drag and drop applications of variables, users can perform statistical analysis relatively easily. Therefore, users only need to understand the purpose of each test and how to interpret the results. A drawback is that Jamovi is not customizable to the same degree as more advanced statistical software

programs. However, a high degree of customization is not needed for this data analysis. An additional extension, "medmod" (mediator and mediation) was downloaded to assist in the analysis.

7.6.1.1. Data Visualization

Visualization simplifies the process of interpreting quantitative data. Furthermore, data visualization is important in data cleaning, making outliers easier to spot (Unwin, 2020). Graphs display the connections between the choice, thoughts, and beliefs of participants, giving further insight into how the underlying beliefs of a participant coincide with their product choice. Even though data visualization can be helpful to identify trends, it may not serve as evidence to state that there is proof of significant dependency. Data visualization only serves to complement statistical analysis. The visuals will be used to extrapolate and explain the results for deeper insight.

7.6.1.2. Statistical Methods for Different Checks and Hypothesis Tests

Several statistical methods are used to investigate the different checks and hypotheses. An independent t-test measures if there is a significant difference in the mean between groups (Keller, 2018). The defined groups were participants who chose Sera and participants who chose Eco-Sera. A matched pairs test is used with the objective of comparing two populations where the parameter is the difference between the two means (Keller, 2018). Moderating effects may also be measured through indirect relationships of variables. Moderation is measured through Jamovi's moderation test. A linear regression is used to determine the relationship between a variable and a scalar response . Linear regression produces $y = b + ax$, where the slope coefficient represents the expected change between the variables (Keller, 2018). A chi-squared test of association measures the significance of dependence. It is often used to measure if there are significant differences between results (Keller, 2018). It requires a relational metric, in this case an interaction rate. The chi-squared test will be supplemented with a relative risk test (also known as risk ratio), a test quantifying the strength of association between two events. Specifically, it measures the ratio for A to occur when B is present, and the ratio of A happening in the absence of B. The two events are independent if the risk ratio equals 1, the effect is decreasing when the risk ratio < 1, and increasing when the risk ratio > 1 (Cummings, 2009). A p-value is used to measure the statistical confidence of the effect found. P-values are directly related to the statistical power of the analysis, which

in turn are directly influenced by the strength of an effect and the sample size that it was measured in (Walmsley & Brown, 2017). Finally, Cohen's d is an effect size measurement and is used as a supplement to hypothesis testing. As mentioned in section 7.3, effect size is the standardized way to report the strength of a relationship between two variables (AI-Therapy Statistics, 2021a). Large sample sizes can give a significant p-value that does not directly imply a relevant effect. Furthermore, if the effect size is small, a significant p-value may not be of relevance. Cohen's d verifies whether the significant effect is of relevant size (AI-Therapy Statistics, 2021b).

7.6.2. Descriptive Statistics - Demographics of Participants

For the analysis to be precise, a clear overview of the demographics of participants is presented here. The experiment was conducted at IKEA Åsane, implying that most participants will be residents of Bergen. The relevant characteristics for the experiment are age, gender, and education, as discussed in sections 6.2.1-6.2.3. The tables below in sections 7.6.2.1-7.6.2.3 are divided into four columns - the relevant demographic (age group, gender, or level of education), the total number of participants, percentage of the total number of participants, and the corresponding percentage in the Norwegian population aged above 16. The inclusion of the last column is to compare the sample representation with the Norwegian population. This is based on the assumption that the demographics of the Norwegian customer base purchasing household cleaning products have the same demographic attributes as the overall Norwegian population (given an age above 16).

7.6.2.1. Age

Age Group	Total (Participants)	% of Total (Participants)	% of Norwegian Population*
16-19	7	6.9 %	5.7 %
20-29	23	22.8 %	16.0 %
30-39	17	16.8 %	16.8 %
40-49	13	12.9 %	16.3 %
50-59	13	12.9 %	16.2 %
60-69	20	19.8 %	13.4 %
70-79	7	6.9 %	10.2 %
80-89	0	0.0 %	4.3 %
90-99	1	1.0 %	1.0 %

*Table 1: Age distribution of participants. *% of the Norwegian population refers to the Norwegian population above the age of 16 (Statistics Norway, 2021a).*

The age range of participants is from 16 to 95. As observed in Table 1, the age groups 20-29 and 60-69 have the two largest sample pools and are overrepresented in the experiment when comparing percentages to the Norwegian population above the age of 16. Participants in the age groups 40-49, 50-59 and 80-89 are the most underrepresented in the experiment.

7.6.2.2. Gender

Gender	Total (Participants)	% of Total (Participants)	% of Norwegian Population*
Male	42	41.6 %	50.2 %
Female	58	57.4 %	49.8 %
Other	1	1.0 %	0.0 %

*Table 2: Gender distribution of participants. *% of the Norwegian population refers to the Norwegian population above the age of 16 (Statistics Norway, 2021b).*

As observed in Table 2, the gender distribution of the participants is skewed, with female participants being overrepresented and male participants being underrepresented, compared to the benchmark which is the Norwegian population.

7.6.2.3. Education

Education	Total (Participants)	% of Total (Participants)	% of Norwegian Population*
Primary/Secondary School	1	1.0 %	24.8 %
High School	35	34.7 %	39.9 %
College/University	61	60.4 %	35.3 %
Prefer Not To Say	4	4.0 %	0.0 %

*Table 3: Education distribution of participants. *% of the Norwegian population refers to the Norwegian population above the age of 16 (Statistics Norway, 2021c).*

As seen in Table 3, participants in the experiment with college/university education are overrepresented, and participants with only primary/secondary school education are underrepresented. Participants with a high school degree are slightly underrepresented, although being relatively accurately represented compared to primary/secondary and college/university education groups. All in all, the sample data does not accurately resemble the Norwegian population when it comes to the level of education.

7.6.3. Results of Different Checks

7.6.3.1. Manipulation Check - Sustainability

A manipulation check on sustainability was included in the survey to ensure that participants correctly identified Eco-Sera as the more sustainable product. Since Eco-Sera was marketed as sustainable (versus Sera which had no indication of sustainability), participants should have recognized Eco-Sera (versus Sera) as the more sustainable product.

Sustainability Check	All	Chose Sera	Chose Eco-Sera	Independent Samples T-Test Output		One Sample T-Test Output	
Mean	4.26	3.95	4.46	p-value	0.003	p-value	< 0.001
Standard Deviation	0.91	1.15	0.65	Cohen's d	-0.577	Cohen's d	1.38

Table 4 (left), Table 5 (middle), and Table 6 (right): Means and standard deviations of the sustainability check (left), independent samples t-test of sustainability check (middle), and one sample t-test of the sustainability check (right).

As observed in Tables 4, 5, and 6, there is a strong argument for participants perceiving Eco-Sera as the more sustainable product. The mean of all participants is 4.26, with the value 4 corresponding to "Eco-Sera is a little more environmentally friendly" and the value 5 corresponding to "Eco-Sera is much more environmentally friendly" as response alternatives to question 7 of the survey. The mean of 4.26 therefore indicates that the average participant perceived Eco-Sera (versus Sera) as somewhere between a little and much more environmentally friendly. In addition, a p-value of 0.003 from the independent samples t-test indicates a statistically significant difference between the means of participants who chose Sera and Eco-Sera. This demonstrates that participants who chose Eco-Sera viewed Eco-Sera as even more sustainable compared to participants who chose Sera. A Cohen's d of -0.577 implies a medium effect size. Furthermore, a p-value of < 0.001 from the one sample t-test provides support for that Eco-Sera is perceived as the most sustainable product since the mean of 4.26 is statistically significantly different from the neutral value of 3. Cohen's d of 1.38 implies a large effect size. The distribution of responses to question 7 is illustrated in Diagram 1 below.

Diagram 1: Distribution of responses to question 7 of the survey, which is the sustainability check.

7.6.3.2. Check - Strength/Efficiency

A question regarding perceived strength/effectiveness was included to check if participants associated the products with strength/effectiveness, and if participants observed Sera as the stronger/more effective product. If Sera (versus Eco-Sera) was associated to a higher degree with strength, this would underline the notion that the participants viewed the products as products with strength-related attributes. This is due to the confirmation of the perceived trade-off between strength and ethicality attributes for strength-dependent products (Luchs et al., 2010; Luchs et al., 2012). This is discussed further in section 2.3.3.

Strength/Efficiency Check	All	Chose Sera	Chose Eco-Sera	Independent Samples T-Test Output		One Sample T-Test Output	
Mean	2.54	1.82	3.02	p-value	< 0.001	p-value	< 0.001
Standard Deviation	1.14	0.78	1.10	Cohen's d	-1.21	Cohen's d	-0.398

Table 7 (left), Table 8 (middle), and Table 9 (right): Means and standard deviations of the strength/efficiency check (left), independent samples t-test of strength/efficiency check (middle), and one sample t-test of the strength/efficiency check (right).

As indicated in Tables 7, 8, and 9, there is an argument for participants perceiving Sera as the stronger/more efficient product. The mean for all participants is 2.54, with the value 2 corresponding to "Sera is a little stronger/more effective", and the value 3 corresponding to "The products are equally strong/effective" as response alternatives to question 8 of the survey. The mean of 2.54 therefore indicates that the average participant perceived Sera as either a little stronger/more effective compared to Eco-Sera, or Sera as equally

strong/efficient compared to Eco-Sera. Additionally, a p-value of < 0.001 from the independent samples t-test indicates a statistically significant difference between the means of participants who chose Sera versus participants who chose Eco-Sera. In other words, participants who chose Sera viewed Sera as more strong/efficient compared to participants who chose Eco-Sera. A Cohen's d of -1.21 implies a large effect size. Moreover, a p-value of < 0.001 from the one sample t-test provides support for that Sera is perceived as stronger/more efficient since the mean of 2.54 is statistically significantly different from the neutral value of 3. Cohen's d of -0.398 implies a small to medium effect size. The distribution of responses to question 8 is illustrated in Diagram 2 below.
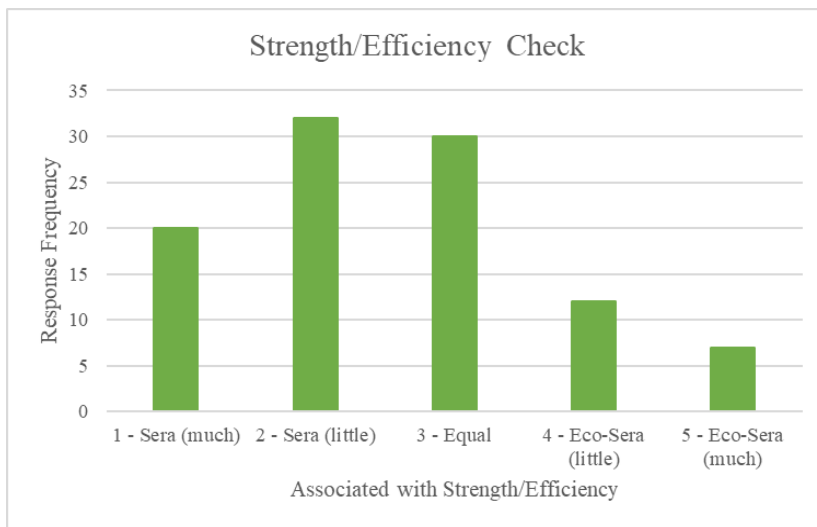
Diagram 2: Distribution of responses to question 8 of the survey, which is the strength/efficiency check.

7.6.3.3. Check - Aesthetic Design Variable

How participants perceive the design of Eco-Sera and Sera may have had an impact on the choice of product. As discussed in section 2.3.6.2, a superior product aesthetic design has a disproportionately positive effect on the choice likelihood for sustainable products (Luchs et al., 2012). In the case of the experiment, the design of the products is as similar as possible (with the exception of the colors and sustainability statement on Eco-Sera), with the goal that participants view the aesthetic design of the products as equal, and therefore it is not a variable that affects choice.

Aesthetic Design Check	All	Chose Sera	Chose Eco-Sera	Independent Samples T-Test Output		One Sample T-Test Output	
Mean	3.81	3.38	4.10	p-value	0.001	p-value	< 0.001
Standard Deviation	1.21	1.46	0.91	Cohen's d	-0.625	Cohen's d	0.673

Table 10 (left), Table 11 (middle), and Table 12 (right): Means and standard deviations of the aesthetic design check (left), independent samples t-test of the aesthetic design check, and one sample t-test of the aesthetic design check (right).

Tables 10, 11, and 12 provide evidence that participants perceive Eco-Sera as the product with superior aesthetic design. The mean for all participants is 3.81, with the value 3 corresponding to "The products' designs are equal", and the value 4 corresponding to "Eco-Sera has a little better design" as response alternatives to question 9 of the survey. The mean of 3.81 therefore indicates that the average participant perceived Eco-Sera as either having a little better design compared to Sera, or Eco-Sera as having an equal design compared to Sera. In addition, a p-value of 0.001 from the independent samples t-test implies a statistically significant difference in means between participants who chose Sera and Eco-Sera. In other words, participants who chose Eco-Sera perceived Eco-Sera (versus Sera) as the product with an even better aesthetic design compared to participants who chose Sera. A Cohen's d of -0.625 implies a medium to large effect size. Furthermore, a p-value of < 0.001 from the one sample t-test provides support for that Eco-Sera is perceived as having a superior aesthetic design since the mean of 3.81 is statistically significantly different from the neutral value of 3. A Cohen's d of 0.673 implies a medium to large effect size. The distribution of responses to question 9 is illustrated in Diagram 3 below.
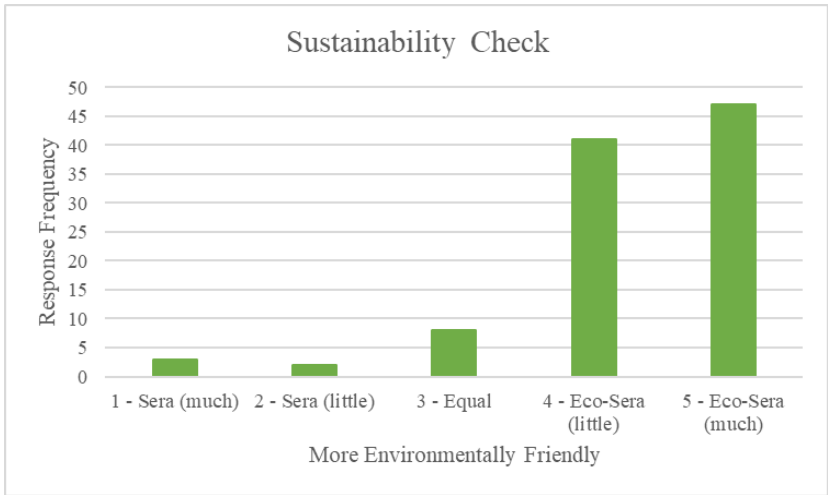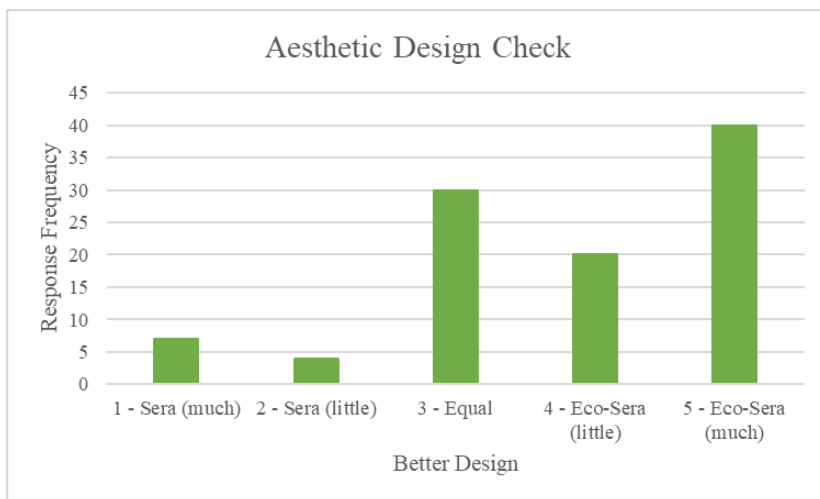

Diagram 3: Distribution of responses to question 9 of the survey, which is the aesthetic design check.

7.6.4. Hypothesis Testing

7.6.4.1. Main Hypotheses Regarding the Sustainability Liability Effect

7.6.4.1.1. Effectiveness Self Condition

H1: Consumers will perceive the regular (versus sustainable) product as more effective.

Efficient Self Condition	All	Chose Sera	Chose Eco-Sera
Mean	2.65	1.85	3.18
Standard Deviation	1.30	0.95	1.23

Independent Samples T-Test Output	
p-value	< 0.001
Cohen's d	-1.18

Table 13 (left) and Table 14 (right): Mean and standard deviation of efficient self condition (left), and independent samples t-test of efficient self condition (right).
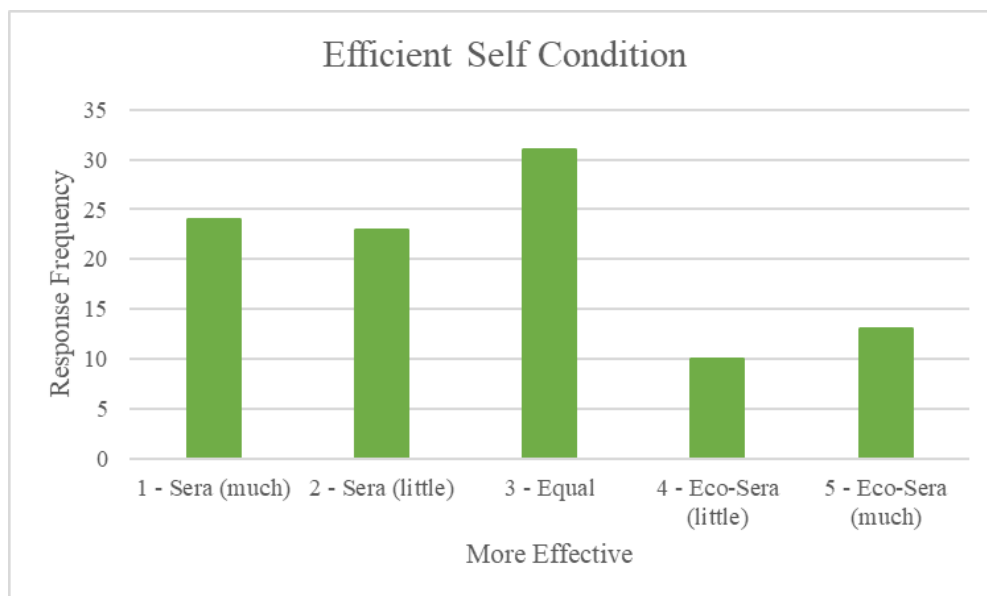
Diagram 4: Distribution of responses to question 5 of the survey, which is the efficient self condition variable.

As observed in Tables 13 and 14, consumers perceive Sera as more effective than Eco-Sera. The mean of all participants being 2.65 indicates that consumers perceive the regular product Sera as between a little more effective and equally effective (versus Eco-Sera). For participants that chose Sera, the mean was 1.85 which translates to a perception of Sera being between much more and little more effective (versus Eco-Sera). For participants that chose Eco-Sera, the mean was 3.18 which translates to a perception of Eco-Sera being between

equally effective or a little more effective (versus Sera). The independent samples t-test investigates if there is a significant difference in perceived efficiency between participants that chose the sustainable and regular product. The p-value of < 0.001 confirms that the perceived efficiency mean for participants who chose Sera is significantly lower than the perceived efficiency mean for participants who chose Eco-Sera. Cohen's d is -1.18, which implies a large effect size. These hypothesis tests confirm that the participants perceive the regular product to be more efficient than the sustainable. Therefore, H1 is supported.

7.6.4.1.2. Effectiveness Other Condition

H2: The effect postulated in H1 will be stronger when participants respond in the other condition (versus self condition).

Efficient Other Condition	All	Chose Sera	Chose Eco-Sera
Mean	2.33	1.85	2.64
Standard Deviation	1.28	1.08	1.32

Paired Samples T-Test Output	
p-value	0.005
Cohen's d	0.265

Table 15 (left) and Table 16 (right): Mean and standard deviation of efficient other condition (left), and paired samples t-test of efficient self condition versus efficient other condition (right).
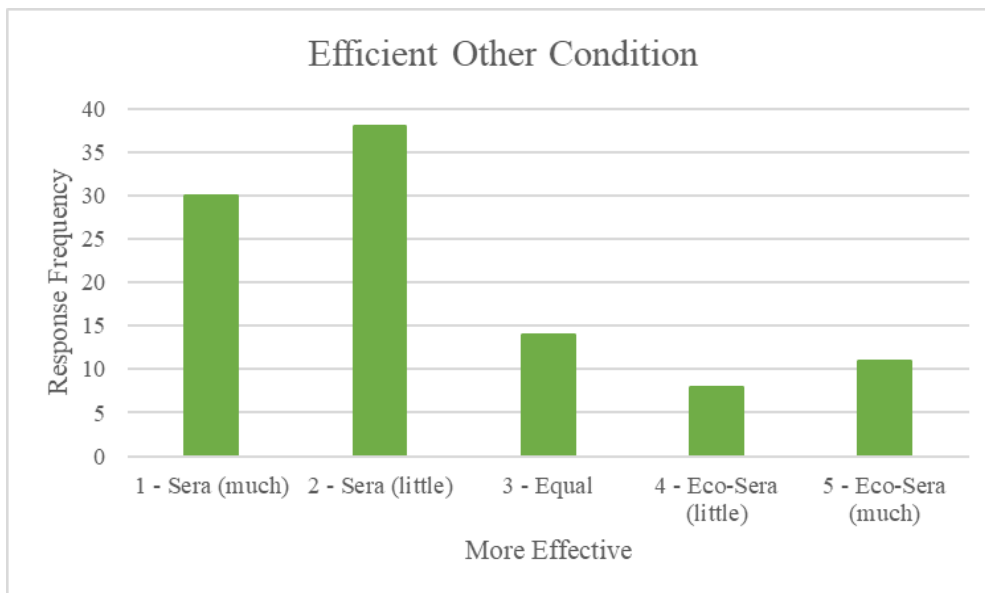
Diagram 5: Distribution of responses to question 6 of the survey, which is the efficient other condition variable.

Tables 15 and 16 show that the mean of the efficient other condition variable is 2.33, which is lower than the efficient self condition mean of 2.65 (Table 13). This suggests that participants view Sera as even more effective (versus Eco-Sera) in the other condition (versus the self condition). The paired samples t-test investigates if the mean for the efficient other condition variable is statistically significantly lower than the mean of the efficient self condition variable. A p-value of 0.005 confirms a statistically significant difference. A Cohen's d of 0.265 suggests a small effect size. In conclusion, H2 is supported.

7.6.4.1.3. Choice of Regular versus Sustainable Product

H3: The choice share of the regular product will be higher than that of the sustainable product.

Results from the survey show that the regular product was chosen by 40 participants, and the sustainable product was chosen by 61 participants, which indicates the opposite of H3.

Binomial Test Output	
p-value	0.023

Table 17: Binomial test checking if the proportion of Sera is statistically significantly larger than 50 %.

Furthermore, as observed in Table 17, the p-value of 0.023 of the binomial test rejects that the proportion of Sera is statistically significantly larger than 50 %. Therefore, H3 is not supported.

7.6.4.1.4. Choice Other Condition

H4: The effect postulated in H3 will be stronger when participants respond in the other condition (versus actual choice).

In the other condition, 38 participants selected Sera and 63 participants selected Eco-Sera. 38 participants selecting Sera in the other condition is lower than the 40 participants selecting Sera during actual choice, which illustrates the opposite of H4 (which argues that the choice share of Sera should further increase compared to actual choice).

Paired Samples T-Test Output	
p-value	0.354
Cohen's d	-0.037

Table 18: Paired samples t-test of actual choice versus choice other condition.

Furthermore, as illustrated in Table 18, a p-value of 0.354 of the paired samples t-test of actual choice versus choice other condition indicates that there is no significant difference in the choice share. A Cohen's d of -0.037 indicates a trivial effect. <u>Therefore, H4 is not supported.</u>

7.6.5.1.5. Hypothetical Choice in Store

H5: The effect postulated in H3 will be weaker when participants respond in the hypothetical choice setting (versus actual choice).

In the hypothetical choice setting, 37 participants selected Sera and 64 participants selected Eco-Sera. 37 participants selecting Sera is lower than the actual choice of 40 (and subsequently 64 participants selecting Eco-Sera is higher than the actual choice of 61), which indicates the same direction of choice as H5.

Paired Samples T-Test Output	
p-value	0.221
Cohen's d	-0.077

Table 19: Paired samples t-test of actual choice versus hypothetical choice in store.

However, as illustrated in Table 19, a p-value of 0.221 of the paired samples t-test of actual choice versus hypothetical choice in store indicates that there is no significant difference in the choice share. Furthermore, Cohen's d of -0.077 indicates a trivial effect. <u>As such, H5 is not supported.</u>

7.6.4.1.5. Link Between Sustainability Liability Effect and Choice of Product

H6: The effect of sustainability (regular versus sustainable product) on choice share (H3) will be mediated by beliefs about product effectiveness: The less a participant associates the sustainable product with efficiency (H1), the lower the choice share of the sustainable product.

A linear regression was performed to investigate the link between the mediating variable of perceived effectiveness (self condition) and dependent variable of product choice. A full mediation test was not performed as the independent variable (being the presented sustainable and regular products) was not manipulated against a pre-intervention (as discussed in section 7.2), and therefore there is no change in the independent variable itself. Furthermore, the support of H1 and H2 (in sections 7.6.4.1.1 and 7.6.4.1.2) established that participants viewed the sustainable (versus) regular product as less effective, and this demonstrates the link between the independent variable and mediating variable.

Linear Regression Output		
Efficient Self Condition on Actual Choice	Estimate	p-value
1-3	-0.515	< 0.001
2-3	-0.459	< 0.001
4-3	0.094	0.541
5-3	0.117	0.401

Table 20: Linear regression output of efficient self condition on actual choice.

As observed in Table 20, 1-3 (with an estimate of -0.515) and 2-3 (with an estimate of -0.459) demonstrate a link between the perceived high efficiency of the regular product and high choice likelihood of the regular product. A p-value of < 0.001 for both 1-3 and 2-3 indicates that these links are statistically significant. Furthermore, since the p-values of 4-3 and 5-3 are 0.541 and 0.401, the estimated link between perceiving the sustainable product as more efficient and an increase in the choice likelihood of the sustainable product is not statistically significant. Moreover, as noted in section 7.6.3.2, there is statistically significant evidence that participants who chose the regular product associated it with strength/efficiency to a higher degree, whereas participants who chose the sustainable product associated the regular product with strength/efficiency to a lower degree. Therefore, H6 is supported.

7.6.5.2. Other Hypotheses

7.6.4.1.6. Age

H7: The effect of sustainability on choice (H3) will be moderated by age: The choice share of the sustainable (versus regular) product will decrease as age increases.

Percentage of Participants that Chose Eco-Sera	
16-19	42.9%
20-29	65.2%
30-39	64.7%
40-49	46.2%
50-59	61.5%
60-69	65.0%
70+	62.5%

Table 21: Breakdown of the percentage of participants that chose the sustainable product based on age.

As observed in Table 21, the percentage of participants that chose Eco-Sera seems to be consistently between 61% and 66% for all age groups, with the exception of 16-19 (42.9%) and 40-49 (46.2%). This is different to the prediction by H7.

The moderation test investigates whether there is a significant relationship between age and choosing the sustainable product.

Moderation Test Output	
Estimate	0.002
p-value	0.502

Table 22: Moderation test output of the moderating effect of age on the interaction between perceived sustainability of the products and actual choice.

The estimate of 0.002 suggests that age has a very small moderating effect on product choice in the direction that as age increases, the choice likelihood of the sustainable product increases. However, a p-value of 0.502 indicates that this effect is not statistically significant. Therefore, H7 is not supported.

7.6.4.1.7. Gender

H8: The effect of sustainability on choice (H3) will be moderated by gender: The choice share of the sustainable (versus regular) product will be higher for females (versus males).

Percentage of Participants that Chose Eco-Sera	
Female	60.3%
Male	61.9%

Table 23: Breakdown of the percentage of participants that chose the sustainable product based on gender.

As illustrated in Table 23, slightly more men chose the sustainable product (61.9%) compared to women (60.3%), which is the opposite as predicted by H8.

Moderation Test Output	
Estimate	0.036
p-value	0.732

Table 24: Moderation test output of the moderating effect of gender on the interaction between perceived sustainability of the products and actual choice.

The moderation test investigates whether there is a significant relationship between gender and choosing the sustainable product. A p-value of 0.732 demonstrates that the moderating effect of gender is not statistically significant.

Chi-squared test: $\chi^2 = 0.0249$

$$df = 1$$

$$p\text{-value} = 0.875$$

Relative risk output: 1.04* (*Eco-Sera and Sera compared).

Furthermore, a chi-squared test of association and a risk ratio test were performed to investigate if the choice is dependent on gender. The risk ratio of 1.04 implies that there is a slightly higher chance that a participant that chooses the sustainable product is a male (versus female). However, a p-value from the chi-squared test of 0.875 suggests that there is no statistically significant dependence between gender and actual choice. Therefore, H8 is not supported.

7.6.4.1.8. Education

H9: The effect of sustainability on choice (H3) will be moderated by education: The choice share of the sustainable (versus regular) product will increase as the level of education increases.

Percentage of Participants that Chose Eco-Sera	
No College Degree	60.0%
College Degree	60.7%

Table 25: Breakdown of the percentage of participants that chose the sustainable product based on level of education.

As illustrated in Table 25, slightly more participants with a college degree chose the sustainable product (60.7%) compared to participants with no college degree (60.0%).

Moderation Test Output	
Estimate	0.100
p-value	0.361

Table 26: Moderation test output of the moderating effect of education on the interaction between perceived sustainability of the products and actual choice.

The moderation test investigates whether there is a significant relationship between the level of education and choosing the sustainable product. The estimate of 0.100 suggests that level of education has a small moderating effect on product choice in the direction that as the level of education increases, the choice likelihood of the sustainable product increases. However, a p-value of 0.361 demonstrates that this effect is not statistically significant.

Chi-squared test: $\chi^2 = 0.0043$

$$df = 1$$

$$\text{p-value} = 0.947$$

Relative risk output: 1.01* (*Eco-Sera and Sera compared).

A chi-squared test of association and a risk ratio test were also performed to investigate if the choice is dependent on education. The risk ratio of 1.01 implies that there is a slightly higher chance that a participant who chooses the sustainable product has a college degree (versus no college degree). However, a p-value from the chi-squared test of 0.947 indicates no statistically significant dependence between education and actual choice. Therefore, H9 is not supported.

7.6.4.1.9. Green Identity

H10: The effect of sustainability on choice (H3) will be moderated by green identity: The choice share of the sustainable (versus regular) product will increase the more participants identify themselves as green.

The moderation test investigates whether there is a significant relation between green identity and choosing the sustainable product. The green identity (1-4) variables were combined into one variable using the average number value of the responses to the four Likert-scale questions by a participant.

Moderation Test Output	
Estimate	0.046
p-value	0.210

Table 27: Moderation test output of the moderating effect of green identity on the interaction between perceived sustainability of the products and actual choice.

The estimate of 0.046 suggests that green identity has a small moderating effect on product choice in the direction that as green identity increases, the choice likelihood of the sustainable product increases. However, a p-value of 0.210 indicates that this effect is not statistically significant. Therefore, H10 is not supported.

7.7. Research Ethics

The main threat to the ethical concern of the experiment is handling the participant's data. If the data is not handled with care it can be used to identify respondents. To ensure the anonymity of participants, the non-disclosure agreement and proof of participants receiving the IKEA gift card were completed separately using pen and paper, while the digital survey was anonymous. Participants were instructed to randomly select where to sign on the paper, and the papers were discarded after sending a confirmation to the Centre for Sustainable Business at NHH (to receive funding as outlined in section 7.4.5). It was also important that the participants consented to the storage and analysis of the data. The confirmation of consent was received through question 2 of the survey as discussed in section 7.4.5. Finally, the data stored in the Qualtrics database will be deleted when it is no longer necessary.

8. Discussion of Results

In this section, the relevant outcomes of the hypotheses are synthesized into the main findings of the study. For clarity reasons, the study is often referred to as the IKEA study.

8.1. The Sustainability Liability Effect is Present in Choice

The key part of the sustainability liability effect is people's perception of the regular product to be more efficient than the sustainable. Among 101 participants, 47 responded that they perceived the regular (versus sustainable) product to be much more efficient or a little more efficient, as seen in Diagram 4 (section 7.6.4.1.1). Only 23 participants perceived the sustainable product as a little or much more efficient, and therefore the effect is quite convincing. When answering in the other condition, there was an increased perception that the regular product was more effective. 68 out of 101 respondents believed other participants would perceive the regular (versus sustainable) product as a little or much more efficient, as observed in Diagram 5 (section 7.6.4.1.2). The results supporting H1 and H2 give important information, as they verify that consumers are biased in their perception of sustainable products, believing sustainable products to be less effective than regular products.

H6 was supported, meaning that when participants viewed the sustainable product as less effective, participants were less likely to choose the sustainable product. When combining this with the support for H1 and H2 (which found that participants viewed the sustainable product as less effective), the sustainability liability effect is found to be present when it comes to product choice, leading to participants choosing the regular (versus sustainable) product.

These results validate and add to findings from earlier research. The study adds to the findings of Luchs et al. (2010), as it further builds on study 5 which investigates how the sustainability liability effect is present in a field experiment setting, specifically the use of hand sanitizers in a cafeteria. For the IKEA study, the addition is product choice rather than product use. Furthermore, consistency is achieved with study 5 of Luchs et al. (2010) as products with strength-related attributes are investigated in both cases. In addition, the support of H1 in the IKEA study is consistent with Lin & Chang's (2012) support of their H1, which finds that green products are perceived as less effective than regular products. Results are also consistent with study 1 of Luchs et al. (2012), although the Luchs et al. (2012) study

was slightly different as it investigated a trade-off between sustainability and functionality. This meant that one product had superior sustainability attributes and the other had superior functionality attributes, rather than the IKEA study which only differentiated products on the sustainability attribute. The findings of the IKEA study are also in line with Newman et al. (2014), but once again the Newman et al. (2014) study was slightly different as it investigated the intentionality of sustainable messaging and purchasing decision rather than a presented sustainable versus non-sustainable product and product choice. Moreover, the results of the IKEA study add to the findings of Pancer et al. (2017). Pancer et al. (2017) find the sustainability liability effect when a product displays the color-green or a certified eco-label in the absence of supporting environmental cues, and the IKEA study finds the sustainability liability effect to still be present despite the color-green being supported by an environmental cue (being the "100% natural ingredients" statement). Finally, the findings of the IKEA study add to the results of Skard et al. (2020) and Skard et al. (2021). Skard et al. (2020) found the sustainability liability effect to be present for product preference and stated that future research should validate their findings in a field experiment setting on purchasing decision. The IKEA study does validate the findings in a field experiment setting, although on product choice rather than purchasing decision. The goal of the IKEA study was to serve as a new and improved study to Skard et al. (2021) by improving sampling and reducing social signaling. Results are consistent with Skard et al. (2021), and sampling is improved through sample size and representation as detailed in section 10.1. However, it is hard to determine whether the reduction or removal of social signaling was achieved, as discussed in section 9.1.1.3.

8.2. The Continued Positive Attitude Towards Sustainable Products May Eventually Outweigh the Sustainability Liability Effect

H3 was rejected, meaning that the study found the majority of participants to prefer the sustainable product. This is in contrast to the unpublished study by Skard et. al. (2021), who found that the majority of participants preferred the regular product when incentivized by money. This may suggest that the continued green trend of consumers viewing sustainable products increasingly positively may eventually outweigh the sustainability liability effect, as discussed in sections 2.1 and 4.2.

However, the participants in the Skard et al.'s (2021) study had a larger incentive (250 NOK) compared to the participants in the IKEA study (150 NOK). In addition, the incentive is a much larger sum compared to earnings for the student sample of Skard et al. (2021) compared to working adults. The different effect of motivation between the two experiments is unknown and hard to measure. One may also ask if the results from this study are comparable to Skard et. al.'s (2021) unpublished report, due to differences in location, sampling, incentive, and products among others, as discussed in section 4.2 and 10.1.

Skard et al. (2021) neutralized the effect of age as the sample pool consisted exclusively of students, unlike the IKEA study. However, age did not have an effect on choice in the IKEA study either. This might be due to the urban environment as discussed in section 9.1.1.2.

9. Validity, Reliability, and Research Limitations

Quantitative research is vulnerable to external factors. Validity and reliability are measures of quality of the research project, and research limitations are possible threats to validity and reliability. Validity is to what degree one can draw valid conclusions on behalf of the experiment, and can be divided into categories including external, internal, construct, and content validity. (Saunders et al., 2016, p. 730).

9.1. Validity

External validity is defined as the extent to which the research results from a study are generalizable (Saunders et al., 2016, p. 716). With external validity in mind, the field experiment was designed to be as similar to a real-life setting as possible, and participants were sampled from the entrance of IKEA Åsane in an attempt to get a participant profile similar to the Norwegian population. However, as further detailed in sections 9.1.1.1 and 9.1.1.2, there were threats to external validity due to limitations in the field experiment environment and sample representation. Internal validity is defined as the extent to which the findings can be attributed to interventions rather than flaws in the research design (Saunders et al., 2016, p. 718). A confounding variable, which is defined as an extraneous variable that can potentially undermine the inferences drawn between the independent and dependent variables (Saunders et al., 2016, p. 713), may threaten internal validity. In the study, skepticism of greenwashing could be a confounding variable. As discussed in section 2.3.5, skepticism of greenwashing may lead to consumers not choosing a product with sustainable

claims on it. In the experiment, participants may have been repelled by the environmental claim on the sustainable product. Greenwashing could therefore serve as an alternative explanation to the relationship between indication of sustainability and product choice. Furthermore, social signaling, aesthetic product design, belief of cleaning, and monetary incentive are all confounding variables that may have threatened internal validity, as detailed in sections 9.1.1.3-9.1.1.6. Another threat to the internal validity of the experiment is the variation of the within-subjects design used, as discussed in section 7.2. Since the experiment does not include a pre-intervention observation or measurement, internal validity is threatened as it is harder to detect confounding variables that may influence the dependent variable.

In relation to the survey, both content validity and construct validity should be addressed. Content validity is the extent to which the questions of the survey provide adequate coverage of the investigative questions (Saunders et al., 2016, p. 450). To ensure content validity, research was defined through reviewed literature and the questions in the survey were based on previous questionnaires by Skard et al. (2021) and Landsvik (2021), as well as carefully developed in collaboration with Skard and Landsvik (as detailed in section 7.2). Construct validity is the extent to which measurement questions actually measure the constructs intended to be measured (Saunders et al., 2016, p. 713). For the questionnaire, simple and clear language was used to solidify construct validity. There were incidents where construct validity may have been threatened due to inconsistent information, as highlighted in section 9.1.1.7.

9.1.1. Potential Threats to Validity (Limitations)

9.1.1.1. Field Experiment Environment

Although the field experiment was designed to be as realistic as possible, there were still elements that would deviate from a real-life setting and therefore may have been a threat to external validity, as results produced may not have been fully representative of the actual choice of product taken by a consumer in a store. Elements include that there was no purchasing decision and that the choice was taken in an isolated setting. Since participants were not faced with the purchasing decision between the sustainable and regular product, but rather a decision of choice, participants may not have felt the proper incentives to choose the "superior" product in their mind. In fact, when there is no financial risk (instead the

experiment had a financial reward being the monetary incentive), people have a lower motivation to exercise effort in a decision-making process (Hoyer et al., 2018). Furthermore, the choice was taken in an isolated setting with the goal of fully removing social signaling. However, in a store, social signaling may be a factor to choice as other people may observe a consumer's decision. The degree of observability affecting social signaling may vary depending on several factors, including the number of people in the store and the placement of products. Therefore, to keep social signaling constant, the aim was to remove it altogether during the experiment even if this is not fully representative of the real world.

9.1.1.2. Sample Representation

As discussed in section 7.6.2, the sample representation of the experiment is disproportionate in relation to the Norwegian population and is therefore a threat to external validity. When it comes to age, the sample has an overrepresentation of a few groups, with the most prominent one being the age groups of 20-29 and 60-69. The age group of 20-29 may have been overrepresented since the people of these ages had less spending power compared to other age groups and were therefore more willing to participate in the experiment when learning about the monetary incentive. The age group of 60-69 may have been overrepresented as the restaurant at IKEA is a popular destination for retirees (Hitland, 2021). Looking at gender, there is an overrepresentation of women in the sample, as almost 60% of participants were women. The overrepresentation of women likely occurred during physically recruiting participants at IKEA. Through anecdotal evidence by the researchers, when people were randomly approached, women were seemingly more open and willing to participate. Finally, there was also an inaccurate representation of the level of education in the sample, with an overrepresentation of participants with a college or university degree. This may be attributed to the field experiment being conducted in the city of Bergen. Multiple studies have found evidence that there is increased access to higher education in urbanized areas, and as a result, a higher percentage of the urban population has a college/university degree compared to rural areas (van Maarseveen, 2021; Florida, 2018; Beamer & Steinbaum, 2019).

9.1.1.3. Social Signaling

As discussed extensively throughout the paper, one of the goals of the experiment was to remove the variable of social signaling. The concern was that if participants were observed, social signaling would skew their choice towards the sustainable product (DellaVigna, 2009),

and therefore be a threat to internal validity as it disguised the true impact of the possible sustainability liability effect on product choice. As such, there was an attempt to isolate the participants during the choice of product. However, there were potential disruptions to this feeling of isolation that may have induced social signaling to some level, including lack of privacy, technical issues, and customers in groups.

9.1.1.3.1. Lack of Privacy

Although room dividers were used to create the feeling of isolation for participants, the room dividers did not entirely block out participants from the rest of the store. "Room 1" and "room 2" were instead partially blocked by the room dividers, with these partially blocked areas being hard but not impossible to spot by IKEA customers not participating in the experiment. The researchers conducted checks of observability before the experiment and concluded that it would be difficult to observe participants as an IKEA customer unless an intentional effort was made to observe what was taking place. Furthermore, in conversations with participants after the experiment, the general consensus was that they did not feel observed when making the choice of product.

9.1.1.3.2. Technical Issues

There were a few cases where elderly participants struggled with navigating the questionnaire on the laptop. Before beginning the questionnaire, participants were instructed to let the researchers know about any technical difficulties. If this occurred, the researchers would guide the participants through the questionnaire to the degree that was necessary. As such, some of these participants' responses may have been influenced by social signaling as they were observed by the researchers during parts of the questionnaire.

9.1.1.3.3. Groups

Participants were at times recruited in groups such as friends or couples. The reason for this was that there was a high likelihood that if one person accepted to be a participant, the rest of the group would follow. However, there may have been cases of indirect social signaling, as when participants were faced with a choice, they may have felt the need to signal their type to the rest of their group knowing that the results would likely be discussed at some point. As such, there is a possibility that participants who were a part of a group had a higher choice likelihood of the sustainable product due to indirect social signaling. However, since this was

a case of indirect social signaling rather than a case of social signaling due to observability, the effect of people being in groups is likely smaller than the other factors that may have caused social signaling.

In addition, some participants had children that could not be left unattended. In those rare cases, participants were allowed to bring the child during the experiment. This may have caused the parent to be affected by social signaling if he or she wants to be a good role model for the child by choosing the product viewed as more ethical.

9.1.1.4. Aesthetic Product Design

Ideally, the variable of aesthetic product design would not have factored into the choice likelihood of the sustainable (versus regular) product. However, as indicated in section 7.6.3.3, participants recognized the sustainable product as having the superior aesthetic design. There was evidence that participants who chose the sustainable product perceived it as having a higher degree of superior aesthetic design (versus the regular product), whereas participants who chose the regular product viewed the sustainable product as having a superior aesthetic design (versus the regular product) to a lesser degree. As discussed in section 2.3.6.2, this may have had a disproportionate effect on choice likelihood in the direction of the sustainable product (Luchs et al., 2012). Therefore, the aesthetic design of the products is a confounding variable that could threaten internal validity.

9.1.1.5. Belief of Cleaning

Participants may have anticipated that they would not actually clean an oven rack, and therefore not make a choice based on the product they perceived to be the most effective. This would threaten internal validity as other factors would have played into the decision-making process instead. However, there was anecdotal evidence that participants did believe that they were supposed to clean when making the choice of product, as they expressed this to the researchers during conversations after the experiment, and researchers observed almost all participants bringing extra cleaning supplies with them to room 2. In some cases, participants were even wearing the rubber gloves as they entered room 2.

9.1.1.6. Incentive

There is a possibility that participants did not fully feel an incentive to clean the oven rack as effectively as possible, thereby threatening internal validity. As discussed in section 8.2, Skard et al. (2021) were able to introduce a higher monetary incentive of 250 NOK compared to the IKEA experiment of 150 NOK, and observed stronger evidence of the sustainability liability effect. Although the results may not be directly comparable for reasons highlighted in section 8.2, there is still a concern that the monetary incentive of the IKEA experiment was too low. For instance, since the initial information given was that the participant would receive 50 NOK regardless of completing the challenge or not, people with spending power may not have felt the need to exert a significant amount of effort to receive the extra 100 NOK, and therefore may have not made a choice based on the perception on which product was more effective. Additionally, Norway has one of the leading HDI's in the world (United Nations Human Development Programme, 2020), and the monetary incentive might have been stronger among participants if the relative value was larger compared to their income or wealth. However, it should be noted through anecdotal evidence that quite a few of the participants were significantly incentivized, as they would approach the researchers and ask if they could participate in the "product test" since they wanted to win the IKEA gift card valued at 150 NOK.

9.1.1.7. Inconsistent Information

Inconsistent information was given to participants throughout the day due to human error, which may have threatened construct validity. The researchers often needed to convince IKEA consumers to participate, and the information shared varied during each interaction, which may have influenced how participants interpreted the instructions. Furthermore, the instructions in room 1 were slightly inconsistent. The original instructions were quickly altered during the beginnings of the experiment, with the minor changes including instructing the participants that they had a maximum of 1 minute (instead of 2-3 minutes) to "clean the oven rack" (as the first couple of participants abandoned the experiment in room 1 since they believed it would take too much time/effort), and an additional sentence in point 3 clarifying that participants should bring other cleaning supplies in addition to the chosen Sera or Eco-Sera product (as some participants were confused about this). Although these updates took place during the early stages of the experiment (with less than 5 participants having

undergone the experiment), it may have caused a different effect in the view of incentive and subsequently the product choice between the first 5 and final 96 participants.

9.2. Reliability

Reliability is the extent to which a data collection technique will yield consistent findings if the study is replicated (Saunders et al., 2016, p. 726). Check questions were implemented as part of the survey to ensure that participants had a consistent belief of the attributes regarding sustainability, perceived efficiency/strength and aesthetic design. The attributes measured in each question were the variables *sustainability check, strength/efficiency check,* and *aesthetic design check*.

Each check question had other survey questions that measured similar effects and were bundled together with these related questions to make three constructs. A factorial analysis ensures the internal reliability of each construct, measured by Cronbach's Alpha (Saunders et al., 2016, p. 451). The "Effectiveness" (0.822) and "Green Identity" (0.897) constructs had satisfactory Cronbach's Alpha values. However, the "Perceived Sustainability" construct, consisting of the two variables *sustainability check* and *aesthetic design check* had a Cronbach's Alpha of 0.599, which is below the satisfactory threshold of 0.8. This casts doubt upon the internal reliability of the sustainability construct.

The online questionnaire with cloud storage reduces the chance for human error. A thorough description and documentation of the experiment, data gathering, and handling of data achieves a high level of reliability as described throughout section 7.

10. Theoretical Implications, Practical Implications, and Future Research

10.1. Theoretical Implications

The sustainability liability effect was found to be present during product choice, meaning that there is reason to believe that sustainable products with strength-related attributes suffer in competition with regular products. When looking at the big picture, this study identifies that there is a problem of disconnect between consumers and producers when it comes to sustainable products with strength-related attributes, which in turn may lead to lower

sustainable sales and overuse of sustainable products as outlined in section 2.3.4, ultimately leading to a less sustainable society in a time where environmental issues such as climate change and waste pollution have become a generational problems that must be addressed immediately.

These findings are consistent with and add to previous research, as discussed in section 8.1. In particular, the goal of the study to validate the findings of Skard et al. (2021) in an improved field experiment setting was achieved through improved sampling. The study included consumers of household cleaning items aged between 16 and 95, of both genders, and with varying levels of education. This an improvement compared to Skard et al. (2021), where the study was limited to smaller sample sizes of college students, and thereby a smaller age and education range. However, there is one main inconsistency with the Skard et al. (2021) study, which is that they found that 40% of participants selected the sustainable product (versus 60% in this study). As discussed in section 8.2, this inconsistency may be attributed to multiple reasons, including differences in monetary incentives, sampling, locations, products, as well as the one-year difference between when the studies took place.

There are boundary conditions that must be considered when it comes to the presence of the sustainability liability effect. Boundary conditions in research can be defined as the "who", "where", and "when" in relation to a theory (Busse et al., 2016). To begin, the "who" is limited to consumers of products with strength-related attributes possibly being subject to a monetary incentive. Both this study and the study of Skard et al. (2021) used monetary incentives to construct a setting where the consumer would view it as advantageous to select the product that is seemingly more effective. In a real-world setting, consumers will not be directly rewarded by purchasing what is perceived as the more effective product, and therefore perceived effectiveness may not be as significant of a factor when purchasing a household cleaning product. A question that may arise is if the sustainability liability effect impacts the actual purchasing decision of products with strength-dependent attributes, particularly when there is no monetary incentive. The "where" can be related to how the sustainability liability effect impacts choice in a public versus private setting. The study showed that the sustainability liability effect is present in an isolated setting. This may be extended to other areas of shopping where consumers are relatively isolated such as online shopping. Household cleaning brands including Klar and Tide have adapted to sell products on their websites (Klar 2021; Tide 2021). With social signaling likely being removed, there is

reason to believe that the sustainability liability effect could have a larger impact when it comes to online shopping. The "when" is also a notable boundary condition for the sustainability liability effect, as the continued trend of environmental awareness may outweigh the sustainability liability effect in the years to come, as discussed in sections 2.1, 4.2, and 8.2.

10.2. Practical Implications

There are practical implications due to the presence of the sustainability liability effect in the choice of sustainable (versus regular) products.

For sustainable marketers, research and subsequent implementation of behavioral interventions and nudges to counter the sustainability liability effect should be considered. Previous research on possible behavioral interventions to mitigate the sustainability liability effect on choice is plentiful and discussed in section 2.3.6. A suggested behavioral intervention is to market the product as sustainable with a strength/effectiveness guarantee (Luchs et al., 2010; Lin & Chang, 2012; Luchs et al., 2012; Skard et al., 2020). This is also encouraged as part of the 4'C framework by Papadas and Avlonitis (2014), which defines four pillars of environmental business that are useful in developing a company focused on sustainability and that needs to stimulate consumption. The *communication* pillar emphasizes reassurance of the performance of the sustainable products, as well as making sure environmental claims are clear and simple to avoid greenwashing skepticism (as discussed in sections 2.3.5 and 9.1). Performance claims can be further researched and implemented by placing a strength/effectiveness guarantee statement on the sustainable product. A proposed nudge is taking advantage of social signaling by increasing observability when the choice is made (Luchs et al., 2010). This is consistent with the *social influence* part of the SHIFT framework by White et al. (2019), which argues that increasing observability stimulates sustainable consumption due to social desirability. Increasing observability can be achieved by placing the competing sustainable and regular products on shelves located in a relatively well-observed spot of a store. Improving the aesthetic design of sustainable products is a behavioral intervention that also may counter the sustainability liability effect (Luchs et al., 2012). Additionally, using prediction requests in advertisements (Bodur et al., 2015) should be further researched as a nudge that could be implemented by sustainable marketers. Skard et al. (2021) also indicate that they will investigate several behavioral interventions, including

how spillover and licensing effects may affect choice. This can be studied by asking participants to carry out a small sustainable or unsustainable act prior to the actual choice (Skard et al., 2021).

There are also aspects that sustainable suppliers should consider. To begin, if no action is taken to counter the sustainability liability effect, household cleaning products that do not focus on sustainability may continue to maintain the competitive advantage as discussed in section 2.3.4. On the other hand, as the trend of environmental awareness and goals continue to spread and strengthen, the attitude-behavior gap (Skard et al., 2021) may be reduced, resulting in consumers switching to sustainable household cleaning products despite the sustainability liability effect being a factor in decision-making, as discussed in sections 2.1, 4.2, and 8.2. This trend may also cause social signaling to become even stronger in directing decisions towards sustainable products even if they are observed as less effective.

If action is taken to further research and implement behavioral interventions to negate the sustainability liability effect, the cost of R&D will likely fall on the supplier. However, although there is an increase in cost associated with combating the sustainability liability effect, possible behavioral interventions may allow sustainable suppliers of cleaning products to gain a competitive advantage, and therefore be a profitable investment in the long term (Porter & Kramer, 2011). Furthermore, the R&D department should be expanded to increase innovation that spurs green new product introductions. Olsen et al. (2014) find that there is a positive and significant influence of green new product introductions on brand attitude, meaning that the more green products introduced by a company, the more positive attitude the brand will enjoy from consumers, increasing the likelihood of consumers purchasing the brand's products (also discussed in section 2.3.6.2).

10.3. Future Research

Due to the limitations and implications discussed, future research on the sustainability liability effect is necessary.

To begin, an aspect that should be further investigated is how the sustainability liability effect impacts product choice in a store setting. Although the study was conducted in a field experiment setting with the goal of eliminating social signaling, it is not certain that social

signaling was removed, as discussed in section 9.1.1.3. A different approach could be to include social signaling as a variable by for instance recording the choice made by participants in a grocery store setting. This would also be even more realistic in relation to the actual choice a consumer would make when purchasing a sustainable or regular product.

Future research could also include better sample representation and even larger sample sizes, for instance by conducting multiple experiments in both rural and urban areas in Norway or other countries. This addresses the disproportionate sample representation discussed in section 9.1.1.2. In addition, a larger sample size is required if the assumed effect size between certain variables is between small and medium, which is less than this study's assumed effect size of more than 0.51 (medium effect size) as discussed in section 7.3. There were a few cases where Cohen's d was lower than 0.51, the most notable one being in relation to H2 (section 7.6.4.1.2). Therefore, for future research, a lower assumed effect size than 0.51 may be assumed, in turn requiring a larger sample size.

Future research should also investigate how the sustainability liability effect impacts product choice for online shopping, as this trend has become more prominent and accelerated by the COVID-19 pandemic. Research on this matter also has the potential of fully removing social signaling since consumers are likely not observed by others when purchasing products online.

Finally, future research may take place after the COVID-19 pandemic to investigate how the sustainability liability effect may impact product choice without a pandemic being a factor, and to investigate if the seemingly positive consumer trend towards green products discussed in sections 2.1, 4.2, and 8.2 eventually outweigh the sustainability effect.

11. Conclusion

Previous research has expressed concern that consumers may view sustainable products associated with strength and efficiency attributes, such as household cleaning products, to be viewed as less efficient in competition with regular products. This is defined as the sustainability liability effect. The question of whether this directly affects the product choice remains ambiguous, but an unpublished study by Skard et al. (2021) seems to find evidence that the sustainability liability effect is present in the decision-making process for a sustainable versus regular household cleaning product.

Due to weaknesses and limitations in the unpublished report by Skard et al. (2021), a new and improved field experiment was conducted at IKEA in Bergen to further investigate the sustainability liability effect. This study found evidence that the sustainability liability effect was present in the choice of sustainable versus regular household cleaning product, meaning that participants were more likely to view the regular (versus sustainable) product as more effective, and participants who viewed the regular (versus sustainable) product as more effective were more likely to choose the regular (versus sustainable) product.

The presence of the sustainability liability effect should have implications for marketers, producers, and other stakeholders in relation to sustainable products with strength-related attributes. Suggestions include increased investment in R&D by sustainable producers and marketers to investigate and implement behavioral interventions that could counter the sustainability liability effect, including strength guarantee statements and product placement in stores. Finally, future research on the sustainability liability effect should be conducted on aspects including a store setting, rural and urban areas, online shopping, and in a time where the world is not affected by a pandemic.

12. References

Agarwal, S. & Kasliwal, N. (2017). The Effect of Age as a Moderator on Green Purchase
　　　Behavior in Hotel Industry. *IPE Journal of Management, 7*(2), 53-63.
　　　https://www.proquest.com/abiglobal/docview/2234981815/abstract/C0C6958CC3084
　　　6BFPQ/2?accountid=37265.

AI-Therapy Statistics. (2021a). *Sample size calculator.* Retrieved October 15, 2021, from
　　　https://www.ai-therapy.com/psychology-statistics/sample-size-calculator.

AI-Therapy Statistics. (2021b). *Effect size.* Retrieved October 15, 2021, from
　　　https://www.ai-therapy.com/psychology-statistics/effect-size-calculator#sizes.

APA Dictionary of Psychology. (2021a). *Field Experiment.* Retrieved November 9, 2021,
　　　from https://dictionary.apa.org/field-experiment.

APA Dictionary of Psychology. (2021b). *Manipulation Check.* Retrieved November 14, 2021,
　　　from https://dictionary.apa.org/manipulation-check.

Balin, B. (2021). Is Education Compulsory for Environmental Quality? An Empirical Study
　　　on EKC and Education Nexus. *Business and Economics Research Journal, 12*(1),
　　　1-15. http://dx.doi.org/10.20409/berj.2021.307.

Barr, S. et al. (2005). The household energy gap: examining the divide between habitual- and
　　　purchase-related conservation behaviours. *Energy Policy, 33*(11), 1425-1444.
　　　https://doi.org/10.1016/j.enpol.2003.12.016.

Bocken, N. et al. (2016). Product design and business model strategies for a circular
　　　economy. *Journal of Industrial and Production Engineering, 33*(5), 308-320.
　　　https://doi.org/10.1080/21681015.2016.1172124.

Bodur, O. et al. (2015). Will you purchase environmentally friendly products? Using
　　　prediction requests to increase choice of sustainable products. *Journal of Business
　　　Ethics, 129*(1), 59-75. https://doi.org/10.1007/s10551-014-2143-6.

Beamer, L. & Steinbaum, M. (2019). *Unequal and Uneven: The Geography of Higher
　　　Education Access.* Phenomenal World. Retrieved November 14, 2021, from
　　　https://www.phenomenalworld.org/analysis/geography-of-higher-ed/.

Brough, A. et al. (2016). Is Eco-Friendly Unmanly? The Green-Feminine Stereotype and Its
　　　Effect on Sustainable Consumption. *Journal of Consumer Research, 43*(4), 567-582.
　　　https://doi.org/10.1093/jcr/ucw044.

Brown, L. & Robertson, O. (2011). An Analysis of Energy Conservation Among U.S. Age
Cohorts. *Academy of Banking Studies Journal, 10*(2), 45-53.
https://www.proquest.com/abiglobal/docview/886554170/abstract/98764E00205449C
9PQ/1?accountid=37265.

Busse, C. (2016). Boundary Conditions: What They Are, How to Explore Them, Why We
Need Them, and When to Consider Them. *Organizational Research Methods, 20*(4),
574-609. https://doi.org/10.1177%2F1094428116641191.

Cai, Z. et al. (2017). Eco-label credibility and retailer effects on green product purchasing
intentions. *Forest Policy and Economics, 80,* 200-208.
https://doi.org/10.1016/j.forpol.2017.04.001.

Centre for Sustainable Business. (2021). *About CSB.* Norwegian School of Economics.
Retrieved October 20, 2021, from
https://www.nhh.no/en/research-centres/csb/about-csb/.

Chen, Y. & Chang, C. (2012). Enhance green purchase intentions: The roles of green
perceived value, green perceived risk, and green trust. *Management Decision, 50*(3),
502-520. https://doi.org/10.1108/00251741211216250.

Chyung, S., et al. (2017). Evidence-Based Survey Design: The Use of a Midpoint on the
Likert Scale. *Performance Improvement, 56*(10), 15-23.
http://dx.doi.org/10.1002/pfi.21727.

Cotte, J. & Trudel, R. (2009). Does It Pay To Be Good?. *MIT Sloan Management Review,
50*(2), 61-69.
https://www.researchgate.net/publication/253234585_Does_it_Pay_to_Be_Good.

Cottrell, S. (2003). Influence of Sociodemographics and Environmental Attitudes on General
Responsible Environmental Behavior Among Recreational Boaters. *Environment and
Behavior, 35*(3), 347-375. https://doi.org/10.1177%2F0013916503035003003.

Cummings, P. (2009). The Relative Merits of Risk Ratios and Odds Ratios. *Arch Pediatr
Adolesc Med, 163*(5), 438-445.
https://www.researchgate.net/publication/24401019_The_Relative_Merits_of_Risk_R
atios_and_Odds_Ratios.

Dacanay, J. et al. (2011). Pollution mitigation and abatement with productivity: A global look
(Second of Two Parts). *Policies and Issues in Perspective, 18*, 1-5.
https://www.researchgate.net/publication/230648606_Pollution_mitigation_and_abate
ment_with_productivity_A_global_look_Second_of_Two_Parts.

Darnall, N. & Aragon-Correa, J. (2014). Can Ecolabels Influence Firms' Sustainability Strategy and Stakeholder Behavior?. *Organization & Environment, 27*(4), 319-27. https://doi.org/10.1177%2F1086026614562963.

Davidson, D. & Freudenburg, W. (1996). Gender and Environmental Risk Concerns: A Review and Analysis of Available Research. *Environment and Behavior, 28*(3), 302-339. https://doi.org/10.1177%2F0013916596283003.

DellaVigna, S. (2009). Psychology and Economics: Evidence from the Field. *Journal of Economic Literature, 47*(2), 315-372. https://eml.berkeley.edu/~sdellavi/wp/01-DellaVigna-4721.pdf.

Delmas, M. et al. (2013). Choosing the Right Eco-Label for Your Product. *MIT Sloan Management Review, 54*(4), 10-12. https://www.researchgate.net/publication/297274981_Choosing_the_Right_Eco-Label _for_Your_Product.

Deloitte. (2021). *Shifting sands: Are consumers still embracing sustainability? Changes and key findings in sustainability and consumer behaviour in 2021.* Retrieved September 20, 2021, from https://www2.deloitte.com/uk/en/pages/consumer-business/articles/sustainable-consu mer.html.

Dietz, T. et al. (2002). Gender, Values, and Environmentalism. *Social Science Quarterly, 83*(1), 353-364. https://www.jstor.org/stable/42956291.

Dillman, D. et al. (2014). *Internet, phone, mail, and mixed mode surveys: The tailored design method*. (4th Edition). John Wiley & Sons Inc.

Earth Overshoot Day. (2021). *Past Earth Overshoot Days*. Retrieved September 29, 2021, from https://www.overshootday.org/newsroom/past-earth-overshoot-days/.

Ecolabel Index. (2021). *Home.* Retrieved September 23, 2021, from http://www.ecolabelindex.com/?fbclid=IwAR3f-dzk36EKQc5mkEK7-9pSeHyBIbW C-tFwmxRQEu3QBIRbkPl-pyhM8uc.

Ekström, M. (2021a). *Behavioral Economics: Lecture 12: Behavioral Public Policy* [PowerPoint slides]. Norwegian School of Economics.

Ekström, M. (2021b). *Behavioral Economics: Lecture 10: Social Signaling.* [PowerPoint slides]. Norwegian School of Economics.

EPA. (2021). *International Treaties and Cooperation about the Protection of the Stratospheric Ozone Layer.* Retrieved September 15, 2021, from

https://www.epa.gov/ozone-layer-protection/international-treaties-and-cooperation-about-protection-stratospheric-ozone.

Fawehinmi, O. et al. (2020). Assessing the green behaviour of academics: The role of green human resource management and environmental knowledge. International Journal of Manpower, 41(7), 879-900. http://dx.doi.org/10.1108/IJM-07-2019-0347.

Florida, R. (2018). *The Talent Dividend in Urban and Rural Areas.* Bloomberg. Retrieved November 14, 2021, from https://www.bloomberg.com/news/articles/2018-10-31/the-place-of-college-grads-in-the-urban-rural-divide.

Fortune Business Insights. (2021). *Household Cleaning Products Market Size, Share & COVID-19 Impact Analysis, By Product Type (Laundry Detergents, Surface Cleaners, Dishwashing Products, Toilet Cleaners, and Others), and Regional Forecast, 2021-2028.* Retrieved September 28, 2021, from https://www.fortunebusinessinsights.com/household-cleaning-products-market-103286.

Gildea, R. (2001). Consumer Survey Confirms Corporate Social Actions Affects Buying Decisions. *Public Relations Quarterly, 39*(Winter), 20-21.

Gray, S. et al. (2019). Will Millennials save the world? The effect of age and generational differences on environmental concern. *Journal of Environmental Management, 242*, 394-402. https://doi.org/10.1016/j.jenvman.2019.04.071.

Gulbrandsen, C. (2017). *Dette er den mest miljøvennlige måten å vaske huset på.* Aftenposten. Retrieved September 30, 2021, from https://www.aftenposten.no/bolig/i/y3lQoa/dette-er-den-mest-miljoevennlige-maaten-aa-vaske-huset-paa.

Hahn, T. et al. (2019). Toxic Effects of Solvent Exposure. *Handbook of Solvents, 2*(3), 1347-1454. https://doi.org/10.1016/B978-1-927885-41-3.50009-7.

Haugland, S. & Nysveen, H. (2021a). *Sustainable Marketing, Lecture 2: Antecedents and influences of sustainable marketing: Part 1* [PowerPoint slides]. Norwegian School of Economics.

Haugland, S. & Nysveen, H. (2021b). *Sustainable Marketing, Lecture 4: Some interesting (and some surprising) results* [PowerPoint slides]. Norwegian School of Economics.

Haws, K. et al. (2014). Seeing the world through GREEN-tinted glasses: Green consumption values and responses to environmentally friendly products. *Journal of Consumer Psychology, 24*(3), 336-354. https://doi.org/10.1016/j.jcps.2013.11.002.

Hitland, S. (2021). Personal communication [Personal interview].

Holden, R. & Passey, J. (2009). Social desirability. *Handbook of individual differences in social behavior,* 441-454. https://psycnet.apa.org/record/2009-12071-030.

Hoyer, W. et al. (2018). *Consumer Behavior*. (7th Edition). Cengage Learning.

IPCC. (2021). *Climate change widespread, rapid, and intensifying - IPCC*. Retrieved September 15, 2021, from https://www.ipcc.ch/2021/08/09/ar6-wg1-20210809-pr/.

Jamovi (2021). About. Retrieved November 24, 2021, from https://www.jamovi.org/about.html.

Jørgensen, S. & Pedersen, L. (2018). *RESTART Sustainable Business Model Innovation*. (1st Edition). Palgrave Studies.

Keller, G. (2018). *Statistics for Management and Economics*. (11th Edition). Cengage Learning.

Kent. (2021). *Statistical & Qualitative Data Analysis Software: About Qualtrics*. Retrieved November 13, 2021, from https://libguides.library.kent.edu/statconsulting/qualtrics.

Kim, E. et al. (2016). Environment-friendly industrial cleaning agents incorporating plant-oil-based raw materials such as chlorofluorocarbon (CFC) alternatives. *Chemical Engineering Journal, 290*, 103-115. https://doi.org/10.1016/j.cej.2016.01.023.

KIWI. (2021). *Hvor mange gode grunner til å bytte vaskemidler trenger vi?*. Retrieved September 30, 2021, from https://kiwi.no/tema/hus-og-hjem/renhold/hvor-mange-gode-grunner-til-a-bytte-vaske midler-trenger-vi/.

Klar. (2021). *Klar - hvem er vi?*. Retrieved October 2, 2021, from https://klardag.no/om-oss/.

Kooi, M. et al. (2016). The effect of particle properties on the depth profile of buoyant plastics in the ocean. *Scientific Reports, 6*, 1-10. https://doi.org/10.1038/srep33882.

Kronthal-Sacco, R. & Wheland, T. (2021). *Sustainable Market Share Index*. NYU STERN Center for Sustainable Business. Retrieved September 22, 2021, from https://www.stern.nyu.edu/sites/default/files/assets/documents/Final%202021%20CS B%20Practice%20Forum-%207.14.21.pdf.

Kumar, S. et al. (2012). Consumer Awareness and Perception Towards Green Products: A Study of Youngsters in India. *International Journal of Marketing & Business Communication, 1*(4), 35-43. https://www.proquest.com/abiglobal/docview/1478016636/abstract/C0C6958CC3084 6BFPQ/1?accountid=37265.

Landsvik, H. (2021). *Green consumption values and product choice*. Unpublished manuscript.

Lebreton, L. et al. (2018). Evidence that the Great Pacific Garbage Patch is rapidly accumulating plastic. *Scientific Reports, 8*, 1-15. https://doi.org/10.1038/s41598-018-22939-w.

Lee, J. & Holden, S. (1999). Understanding the Determinants of Environmentally Conscious Behavior. *Psychology and Marketing, 16*(5), 373-392. http://dx.doi.org/10.1002/(SICI)1520-6793(199908)16:5%3C373::AID-MAR1%3E3.0.CO;2-S.

Levin, G. (1990). Consumers Turning Green: JWT Survey. *Advertising Age, 61*(12), 3-7.

Lijo, J. & Narayanamurthy, G. (2015). Converging sustainability definitions: industry independent dimensions. *World Journal of Science, Technology and Sustainable Development, 12*(3), 206-232. http://dx.doi.org/10.1108/WJSTSD-04-2015-0017.

Lin Y. & Chang, C. (2012). Double Standard: The Role of Environmental Consciousness in Green Product Usage. *Journal of Marketing, 76*(5), 125-134. http://dx.doi.org/10.1509/jm.11.0264.

Lischer, B. (2021). *Brand Color Psychology: The Power of Color in Branding.* Ignyte. Retrieved November 11, 2021, from https://www.ignytebrands.com/the-psychology-of-color-in-branding/.

Luchs, M. et al. (2010). The Sustainability Liability: Potential Negative Effects of Ethicality on Product Preference. *Journal of Marketing, 74*(5), 18-31. http://dx.doi.org/10.1509/jmkg.74.5.18.

Luchs, M. et al. (2012). Product Choice and the Importance of Aesthetic Design Given the Emotion-laden Trade-off between Sustainability and Functional Performance. *The Journal of Product Innovation Management, 29*(6). 903-915. https://www.proquest.com/scholarly-journals/product-choice-importance-aesthetic-design-given/docview/1112820748/se-2?accountid=37265.

Lüdeke-Freund, et al. (2019). A Review and Typology of Circular Economy Business Model Patterns. *Journal of Industrial Energy, 23*(1), 26-61. https://doi.org/10.1111/jiec.12763.

McLeod, S. (2012). *Experimental Method.* SimplyPsychology. Retrieved November 11, 2021, from https://www.simplypsychology.org/experimental-method.html.

Meyer, A. (2015). Does education increase pro-environmental behavior? Evidence from Europe. *Ecological Economics, 116*, 108-121. https://doi.org/10.1016/j.ecolecon.2015.04.018.

Mishra, D. et al. (1998). Information asymmetry and levels of agency relationships. J. Market. *Journal of Marketing Research, 35*(3). https://doi.org/10.2307/3152028.

Multiconsult. (2012). *IKEA Åsane*. Retrieved November 10, 2021, from https://www.multiconsult.no/prosjekter/ikea-asane/.

Newman, G. et al. (2014). When Going Green Backfires: How Firm Intentions Shape the Evaluation of Socially Beneficial Product Enhancements. *Journal of Consumer Research, 41*(3), 823-839. https://www.proquest.com/scholarly-journals/when-going-green-backfires-how-firm-intentions/docview/1562191795/se-2.

Nidumolu, R. et al. (2009). Why Sustainability Is Now the Key Driver of Innovation. *Harvard Business Review,* 57-64. https://hbr.org/2009/09/why-sustainability-is-now-the-key-driver-of-innovation.

Oates, C. et al. (2008). Marketing sustainability: Use of information sources and degrees of voluntary simplicity. *Journal of Marketing Communications, 14*(5), 351-365. https://doi.org/10.1080/13527260701869148.

Olsen, M. (2014). Green Claims and Message Frames: How Green New Products Change Brand Attitude. *Journal of Marketing, 78*(5), 119-137. https://doi.org/10.1509%2Fjm.13.0387.

Pancer, E. et al. (2017). Isolated Environmental Cues and Product Efficacy Penalties: The Color Green and Eco-labels. *Journal of Business Ethics, 143*(1), 159-177. https://www.proquest.com/abiglobal/docview/1907230596/8F3DDAE161BC446APQ/1?accountid=37265.

Papadas, K. & Avlonitis, G. (2014). The 4 Cs of environmental business: Introducing a new conceptual framework. *Social Business, 4*(4), 345-360. https://doi.org/10.1362/204440814X14185703122928.

Patagonia. (2021a). *Materials.* Retrieved November 25, 2021, from https://www.patagonia.com/materials/.

Patagonia. (2021b). *Women's Web Specials.* Retrieved November 25, 2021, from https://www.patagonia.com/shop/web-specials-womens?start=0&sz=24#tile-9.

Patagonia. (2021c). *Factories, Farms and Mills.* Retrieved November 25, 2021, from https://www.patagonia.com/factories-farms-mills/.

Pelau, C. & Catalina, A. (2018). Econometric Model for Measuring the Impact of the
Education Level of the Population on the Recycling Rate in a Circular Economy.
Amfiteatru Economic, 20(48), 340-355. http://dx.doi.org/10.24818/EA/2018/48/340.

Porter, M. & Kramer, M. (2011). Creating Shared Value. *Harvard Business Review, 89*,
62-77. https://hbr.org/2011/01/the-big-idea-creating-shared-value.

Saunders, M. et al. (2016). *Research Methods for Business Students.* (7th Edition). Pearson.

Sarti, S. et al. (2018). Market segmentation of consumers based on their actual sustainability
and health-related purchases. *Journal of Cleaner Production, 192*, 270-280.
https://doi.org/10.1016/j.jclepro.2018.04.188.

Shahrin, R. et al. (2017). "Green Eco-Label" or "Greenwashing"? Building Awareness About
Environmental Claims of Marketers. *Advanced Science Letters, 23*(4), 3205-3208.
https://doi.org/10.1166/asl.2017.7713.

Skard, S. et al. (2020). When is Sustainability a Liability, and When Is It an Asset? Quality
Inferences for Core and Peripheral Attributes. *Journal of Business Ethics, 173,*
109-132. https://link.springer.com/article/10.1007/s10551-019-04415-1.

Skard, S. et al. (2021). *Attitude-behavior gaps and the sustainability liability: Experimental
evidence from green products - Extended abstract to the HEC SnO Research Day.*
Unpublished manuscript. 1-9.

Smithers. (2021). Sustainable cleaning products market to surge to $110 billion in 2025.
Retrieved September 27, 2021, from
https://www.smithers.com/resources/2021/feb/sustainable-cleaning-market-surge-110-
billion.

Sparebank1 Østlandet. (2018). *Forventningsundersøkelse - Forbrukere på Østlandet*
[PowerPoint].
https://docplayer.me/139306879-Forventningsundersokelse-forbrukere-pa-ostlandet.ht
ml.

Statistics Norway. (2021a). Population and population changes. Retrieved November 18,
2021, from https://www.ssb.no/en/befolkning/folketall/statistikk/befolkning.

Statistics Norway. (2021b). 07459: Population, by sex, age, contents, year and region: 0 The
Whole Country. Retrieved November 18, 2021, from
https://www.ssb.no/en/statbank/table/07459/tableViewLayout1/.

Statistics Norway. (2021c). Educational attainment of the population. Retrieved November
18, 2021, from

https://www.ssb.no/en/utdanning/utdanningsniva/statistikk/befolkningens-utdanningsniva.

Stockholm Resilience Centre. (2015). *What is resilience? An introduction to social-ecological research.* Retrieved September 21, 2021, from https://www.stockholmresilience.org/download/18.2f48c3c31429b6ad0a61cde/1459560221338/SRC_whatisresilience__sida.pdf.

Thaler, R. & Sunstein, C. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. (1st Edition). Yale University Press.

The Investopedia Team. (2021). *Environmental, Social, & Governance (ESG) Criteria*. Investopedia. Retrieved November 20, 2021, from https://www.investopedia.com/terms/e/environmental-social-and-governance-esg-criteria.asp.

The Ocean Clean Up. (2021). The Great Pacific Garbage Patch. Retrieved September 18, 2021, from https://theoceancleanup.com/great-pacific-garbage-patch/.

Thøgersen, J. & Ölander, F. (2006). To what degree are environmentally beneficial choices reflective of a general conservation stance?. *Environment and Behavior, 38*(4), 550-569. https://doi.org/10.1177%2F0013916505283832.

Tide. (2021). *Shop Tide Products.* Retrieved November 11, 2021, from https://tide.com/en-us/shop

United Nations. (2021a). *Sustainability.* Academic Impact. Retrieved September 15, 2021, from https://www.un.org/en/academic-impact/sustainability.

United Nations. (2021b). *The 17 Goals.* Department of Economic and Social Affairs. Retrieved September 15, 2021, from https://sdgs.un.org/goals.

United Nations Human Development Programme (2020). *Latest Human Development Index Ranking.* UN HDR 2020. Retrieved November 24, 2021, from http://hdr.undp.org/en/content/latest-human-development-index-ranking.

Unwin, A. (2020). *Why Is Data Visualization Important? What is Important in Data Visualization?*. Harvard Data Science Review. Retrieved November 20, 2021, from https://hdsr.mitpress.mit.edu/pub/zok97i7p/release/3.

van Maarseveen, R. (2021). The urban-rural education gap: do cities indeed make us smarter?. *Journal of Economic Geography, 21*(5), 683-714. https://doi.org/10.1093/jeg/lbab020.

Walmsley, A. & Brown, M. (2017). *What is Power?*. Statistics Teacher. Retrieved December 7, 2021, from https://www.statisticsteacher.org/2017/09/15/what-is-power/.

White, K. (2019). How to SHIFT Consumer Behaviors to be More Sustainable: A Literature Review and Guiding Framework. *Journal of Marketing, 83*(3), 22-49. https://doi.org/10.1177%2F0022242919825649.

Whitmarsh, L. (2010). Green identity, green living? The role of pro-environmental self-identity in determining consistency across diverse pro-environmental behaviours. *Journal of Environmental Psychology, 30*(3), 305-314. https://doi.org/10.1016/j.jenvp.2010.01.003.

WHO. (2021). *Ultraviolet radiation*. Retrieved September 18, 2021, from https://www.who.int/health-topics/ultraviolet-radiation#tab=tab_1.

World Medical Association. (2021). *WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects*. Retrieved October 15, 2021, from https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for -medical-research-involving-human-subjects/.

WWF (2021). What do sea turtles eat? Unfortunately, plastic bags. Retrieved September 17, 2021, from https://www.worldwildlife.org/stories/what-do-sea-turtles-eat-unfortunately-plastic-ba gs.

Zelezny, L. et al. (2000). New Ways of Thinking about Environmentalism: Elaborating on Gender Differences in Environmentalism. *Journal of Social Issues, 56*(3), 443-457. http://dx.doi.org/10.1111/0022-4537.00177.

Zint, M. (2006). *Power Analysis, Statistical Significance, & Effect Size.* My Environmental Education Evaluation Resource Assistant. Retrieved October 15, 2021, from https://meera.snre.umich.edu/power-analysis-statistical-significance-effect-size.

13. Appendix

13.1. Printed Documents Used During the Experiment (in Norwegian)

13.1.1. Non-Disclosure Agreement

TAUSHETSPLIKT

Ved å signere dette dokumentet er du enig i det følgende:

1. Du kan ikke snakke eller dele informasjon om prosessen til andre i løpet av de neste 24 timene. Dette er for å beskytte integriteten til testen.

Når det gjelder vår behandling, er denne testen kun til innsamling av data og du forblir anonym i denne innsamlingen.

Signatur deltaker:

13.1.2. Instructions Room 1

LES NØYE

IKKE GÅ INN I NESTE ROM FØR DU HAR LEST DETTE

----

1. Foran deg ser du to produkter, SERA og ECO-SERA.
2. Du skal velge ETT av produktene som skal brukes til å vaske en ovnsrist helt ren i det neste rommet. Dersom du får til å vaske ovnsristen rent nok på 1 minutt, får du 100 kr ekstra i gavekort (totalt 150 kr). Du kan også ta med deg andre rengjøringsartikler du ser her slik du ønsker.
3. Velg nå ENTEN SERA ELLER ECO-SERA og ta det med deg videre til det neste rommet. I tillegg ta med svamp, kost, gummihansker, og/eller stålull!

It is worth noting that this is the updated version of the instructions used in room 1. The original instructions were altered during the start of the experiment, with the minor changes being that the participants were instructed that they had a maximum of 1 minute (instead of 2-3 minutes) to "clean" the oven rack, as well as an additional sentence in point 3 that clarifies to participants that they should bring other cleaning supplies in addition to the chosen product. For reasoning and limitations due to these updates, see section 9.1.1.7.

13.1.3. Instructions Room 2

LES NØYE

DU SKAL IKKE VASKE ALLIKEVEL, VENNLIGST PLASSER PRODUKTET DU VALGTE HER

I stedet håper vi at du vil fylle ut en kort undersøkelse på 2-3 minutter. Når du har fullført undersøkelsen får du gavekort på IKEA til en verdi av 150 kr.

13.1.4. Record of Gift Card Recipients

MOTTAK AV GAVEKORT

Jeg bekrefter at jeg har mottatt et gavekort på IKEA til en verdi av 150 kr.

Navn	Signatur

13.2. Questionnaire (in Norwegian)

Question 1:

Hvilket produkt valgte du?

Sera

Eco-Sera

Question 2:

Denne undersøkelsen er utført i henhold til de etiske prinsippene i Helsinki-erklæringen. Studien er utført av studenter ved Norges Handelshøyskole. Deltakelse er frivillig og anonymt. Ved å fullføre undersøkelsen samtykker du til å delta og svarene dine vil bli brukt til analyse i vår masteroppgave.

○ Jeg samtykker

○ Jeg ønsker å trekke meg

Question 3:

Hvilket produkt tror du de fleste ANDRE som deltar i denne testen velger?

Sera

Eco-Sera

Question 4:

Hvis du skulle kjøpt dette produktet i butikken, hvilket ville du ha valgt?

Sera
○

Eco-Sera
○

Question 5:

I tester av disse produktene, hvilket tror du viste seg å være mest effektivt mot fastbrente matrester?

Sera er mye mer effektiv til å fjerne fastbrente matrester	Sera er litt mer effektiv til å fjerne fastbrente matrester	Produktene er like effektive til å fjerne fastbrente matrester	Eco-Sera er litt mer effektiv til å fjerne fastbrente matrester	Eco-Sera er mye mer effektiv til å fjerne fastbrente matrester
○	○	○	○	○

Question 6:

Hvilket produkt tror du de fleste andre mener er mest effektivt?

Sera er mye mer effektiv til å fjerne fastbrente matrester	Sera er litt mer effektiv til å fjerne fastbrente matrester	Produktene er like effektive til å fjerne fastbrente matrester	Eco-Sera er litt mer effektiv til å fjerne fastbrente matrester	Eco-Sera er mye mer effektiv til å fjerne fastbrente matrester
○	○	○	○	○

Question 7:

Hvilket av produktene tror du er mest miljøvennlig?

Sera er mye mer miljøvennlig	Sera er litt mer miljøvennlig	Produktene er like miljøvennlige	Eco-Sera er litt mer miljøvennlig	Eco-Sera er mye mer miljøvennlig
○	○	○	○	○

Question 8:

I hvilken grad forbinder du produktene med styrke/effektivitet?

Sera er mye sterkere/mer effektivt	Sera er litt sterkere/mer effektivt	Produktene er like effektive	Eco-Sera er litt sterkere/mer effektivt	Eco-Sera er mye sterkere/mer effektivt
○	○	○	○	○

Question 9:

Hvilket av produktene synes du er finest?

Sera er mye finere	Sera er litt finere	Produktene er like fine	Eco-Sera er litt finere	Eco-Sera er mye finere
○	○	○	○	○

Question 10:

Her vil vi gjerne at du skal oppgi din mening om forskjellige påstander.

	1 Sterkt uenig	2 Uenig	3 Litt Uenig	4 Nøytral	5 Litt enig	6 Enig	7 Sterkt enig
Det er viktig for meg at produktene jeg bruker ikke er skadelige for miljøet.	○	○	○	○	○	○	○
Jeg tenker på hvordan mine valg og handlinger kan påvirke miljøet	○	○	○	○	○	○	○
Jeg vil beskrive meg selv som miljøvennlig.	○	○	○	○	○	○	○
Jeg er villig til gjøre noe som er mer tungvint for å ta et mer miljøvennlig valg.	○	○	○	○	○	○	○

Question 11:

Så noen spørsmål om deg:

Hva er alderen din?

—

16   100

Question 12:

Hvilket kjønn er du?

○ Mann

○ Kvinne

○ Annet/Ønsker ikke å svare

Question 13:

Hvilken utdanning har du?

○ Grunnskole

○ Videregående skole

○ Høyskole/universitet

○ Ønsker ikke å svare

Question 14:

Tusen takk for at du har deltatt i denne undersøkelsen.

Du vil nå motta et gavekort på kr 150 på IKEA. For å motta gavekortet må du signere på at du har mottatt gavekortet.

○ Ja, jeg ønsker gavekort på 150 kr.

○ Nei, jeg ønsker ikke å motta gavekort på 150 kr.

13.3. Jamovi Scripts

13.3.1. Different Checks

13.3.1.1. Manipulation Check - Sustainability

Independent Samples T-Test		Statistic	df	p	Mean difference	SE difference		Effect Size
Sustainability Check	Student's t	−2.84 ª	99.0	0.003	−0.509	0.179	Cohen's d	−0.577
Note. Hₐ Sera < Eco-Sera								

Table 28: Independent samples t-test of the sustainability check.

One Sample T-Test		Statistic	df	p	Mean difference		Effect Size
Sustainability Check	Student's t	13.8	100	< .001	1.26	Cohen's d	1.38
Note. Hₐ population mean > 3							

Table 29: One sample t-test of the sustainability check.

13.3.1.2. Check - Strength/Efficiency

Independent Samples T-Test		Statistic	df	p	Mean difference	SE difference		Effect Size
Strength/Efficiency Check	Student's t	−5.92	99.0	< .001	−1.19	0.201	Cohen's d	−1.21
Note. H$_a$ Sera < Eco-Sera								

Table 30: Independent samples t-test of the strength/efficiency check.

One Sample T-Test		Statistic	df	p	Mean difference		Effect Size
Strength/Efficiency Check	Student's t	−4.00	100	< .001	−0.455	Cohen's d	−0.398
Note. H$_a$ population mean < 3							

Table 31: One sample t-test of the strength/efficiency check.

13.3.1.3. Check - Aesthetic Design Variable

Independent Samples T-Test		Statistic	df	p	Mean difference	SE difference		Effect Size
Aesthetic Check	Student's t	−3.07 ª	99.0	0.001	−0.723	0.236	Cohen's d	−0.625
Note. H$_a$ Sera < Eco-Sera								

Table 32: Independent t-test of the aesthetic design check.

One Sample T-Test		Statistic	df	p	Mean difference		Effect Size
Aesthetic Check	Student's t	6.77	100	< .001	0.812	Cohen's d	0.673
Note. H$_a$ population mean > 3							

Table 33: One sample t-test of the aesthetic design check.

13.3.2. Hypotheses Tests

13.3.2.1. Main Hypotheses Regarding the Sustainability Liability Effect

13.3.2.1.1. Effectiveness Self Condition

Independent Samples T-Test		Statistic	df	p	Mean difference	SE difference		Effect Size
Perceived Effectiveness SC	Student's t	−5.79	99.0	< .001	−1.33	0.230	Cohen's d	−1.18

Table 34: Independent samples t-test of efficient self condition.

13.3.2.1.2. Effectiveness Other Condition

Paired Samples T-Test			statistic	df	p	Mean difference	SE difference		Effect Size
Perceived Effectiveness SC	Perceived Effectiveness OC	Student's t	2.66	100	0.005	0.327	0.123	Cohen's d	0.265

Note. Hₐ Measure 1 > Measure 2

Table 35: Paired samples t-test of efficient self condition versus efficient other condition.

13.3.2.1.3. Choice of Regular versus Sustainable Product

Binomial Test	Level	Count	Total	Proportion	p	95% Confidence Interval Lower	Upper
AC Eco-Sera		61	101	0.604	0.023	0.517	1.00
Sera		40	101	0.396	0.986	0.314	1.00

Note. Hₐ is proportion > 0.5

Table 36: Binomial test checking if the proportion of Sera is statistically significantly larger than 50 %.

13.3.2.1.4. Choice Other Condition

Contingency Tables		AC Sera	Eco-Sera	Total
OC				
Sera	Observed	25	13	38
	Expected	15.0	23.0	38.0
Eco-Sera	Observed	15	48	63
	Expected	25.0	38.0	63.0
Total	Observed	40	61	101
	Expected	40.0	61.0	101.0

Table 37: Contingency tables for choice other condition compared to actual choice.

Paired Samples T-Test			statistic	df	p	Mean difference	SE difference		Effect Size
AC	OC	Student's t	−0.376	100	0.354	−0.0198	0.0526	Cohen's d	−0.0374

Note. Hₐ Measure 1 < Measure 2

Table 38: Paired samples t-test of actual choice versus choice other condition.

13.3.2.1.5. Hypothetical Choice in Store

Paired Samples T-Test			statistic	df	p	Mean difference	SE difference			Effect Size
AC	STORE	Student's t	−0.773	100	0.221	−0.0297	0.0384	Cohen's d		−0.0769

Note. H₀ Measure 1 < Measure 2

Table 39: Paired samples t-test of actual choice versus hypothetical choice in store.

13.3.2.1.6. Link Between Sustainability Liability Effect and Choice of Product

Model Fit Measures

Model	R	R²
1	0.550	0.303

Model Coefficients - AC

Predictor	Estimate	SE	t	p
Intercept ᵃ	1.8065	0.0752	24.016	< .001
Perceived Effectiveness SC:				
1 − 3	−0.5148	0.1139	−4.521	< .001
2 − 3	−0.4586	0.1153	−3.979	< .001
4 − 3	0.0935	0.1523	0.614	0.541
5 − 3	0.1166	0.1384	0.843	0.401

ᵃ Represents reference level

Table 40: Linear regression output of efficient self condition on actual choice.

13.3.2.2. Other Hypotheses

13.3.2.2.1. Age

Contingency Tables		AC		
Age - (10-year intervals)		Sera	Eco-Sera	Total
16-19	Observed	4	3	7
	% within row	57.1 %	42.9 %	100.0 %
20-29	Observed	8	15	23
	% within row	34.8 %	65.2 %	100.0 %
30-39	Observed	6	11	17
	% within row	35.3 %	64.7 %	100.0 %
40-49	Observed	7	6	13
	% within row	53.8 %	46.2 %	100.0 %
50-59	Observed	5	8	13
	% within row	38.5 %	61.5 %	100.0 %
60-69	Observed	7	13	20
	% within row	35.0 %	65.0 %	100.0 %
70+	Observed	3	5	8
	% within row	37.5 %	62.5 %	100.0 %
Total	Observed	40	61	101
	% within row	39.6 %	60.4 %	100.0 %

Table 41: Breakdown of the percentage of participants that chose the sustainable product based on age.

Moderation Estimates			95% Confidence Interval			
	Estimate	SE	Lower	Upper	Z	p
Sustainability Check	0.13951	0.05319	0.03525	0.24377	2.623	0.009
Age	0.00123	0.00254	−0.00375	0.00622	0.485	0.628
Sustainability Check ✳ Age	0.00204	0.00304	−0.00392	0.00801	0.672	0.502

Table 42: Moderation test output of the moderating effect of age on the interaction between perceived sustainability of the products and actual choice.

13.3.2.2.2. Gender

Contingency Tables		Gender (2 values)		
AC		Male	Female	Total
Sera	Observed	16	23	39
	% within column	38.1 %	39.7 %	39.0 %
Eco-Sera	Observed	26	35	61
	% within column	61.9 %	60.3 %	61.0 %
Total	Observed	42	58	100
	% within column	100.0 %	100.0 %	100.0 %

Table 43: Breakdown of participants that chose the sustainable product based on gender.

Moderation Estimates			95% Confidence Interval			
	Estimate	SE	Lower	Upper	Z	p
Sustainability Check	0.1637	0.0534	0.0591	0.268	3.068	0.002
Gender (2 values)	−0.0578	0.0944	−0.2428	0.127	−0.612	0.541
Sustainability Check ✻ Gender (2 values)	0.0358	0.1044	−0.1688	0.240	0.343	0.732

Table 44: Moderation test output of the moderating effect of gender on the interaction between perceived sustainability of the products and actual choice.

χ² Tests	Value	df	p
χ²	0.0249	1	0.875
N	100		

Table 45: Chi-squared test of actual choice and gender (2 values).

Comparative Measures		95% Confidence Intervals	
	Value	Lower	Upper
Odds ratio	1.07	0.473	2.41
Relative risk	1.04 ᵃ	0.646	1.67
ᵃ rows compared			

Table 46: Relative risk output of actual choice and gender (2 values) (right).

13.3.2.2.3. Education

Contingency Tables		Education (Filtered)		
AC		College Degree	No College Degree	Total
Sera	Observed	24	16	40
	% within column	39.3 %	40.0 %	39.6 %
Eco-Sera	Observed	37	24	61
	% within column	60.7 %	60.0 %	60.4 %
Total	Observed	61	40	101
	% within column	100.0 %	100.0 %	100.0 %

Table 47: Breakdown of participants that chose the sustainable product based on level of education.

Moderation Estimates	Estimate	SE	95% Confidence Interval Lower	95% Confidence Interval Upper	Z	p
Sustainability Check	0.1464	0.0521	0.0444	0.248	2.812	0.005
Education	0.0531	0.0938	−0.1307	0.237	0.566	0.571
Sustainability Check ✳ Education	0.0999	0.1094	−0.1146	0.314	0.913	0.361

Table 48: Moderation test output of the moderating effect of education on the interaction between perceived sustainability of the products and actual choice.

χ^2 Tests	Value	df	p
χ^2	0.00434	1	0.947
N	101		

Table 49: Chi-squared test of actual choice and education (filtered).

Comparative Measures	Value	95% Confidence Intervals Lower	95% Confidence Intervals Upper
Relative risk	1.01 ᵃ	0.731	1.40
ᵃ rows compared			

Table 50: Relative risk output of actual choice and education (filtered).

13.3.2.2.4. Green Identity

Moderation Estimates			95% Confidence Interval			
	Estimate	SE	Lower	Upper	Z	p
Sustainability Check	0.1383	0.0500	0.0403	0.236	2.77	0.006
GreenIDTotalAverage	0.1103	0.0357	0.0403	0.180	3.09	0.002
Sustainability Check ✳ GreenIDTotalAverage	0.0457	0.0365	−0.0257	0.117	1.25	0.210

Table 51: Moderation test output of the moderating effect of green identity on the interaction between perceived sustainability of the products and actual choice.

13.4. Factor Analysis

13.4.1. Reliability Analysis - Effectiveness

Scale Reliability Statistics	
	Cronbach's α
scale	0.822

Table 52: Scale reliability statistics output of the "effectiveness" construct.

Item Reliability Statistics				if item dropped
	mean	sd	item-rest correlation	Cronbach's α
Perceived Effectiveness SC	2.65	1.30	0.727	0.703
Perceived Effectiveness OC	2.33	1.28	0.582	0.851
Strength/Efficiency Check	2.54	1.14	0.736	0.704

Table 53: Item reliability statistics output of the "effectiveness" construct.

13.4.2. Reliability Analysis - Green Identity

Scale Reliability Statistics	
	Cronbach's α
scale	0.897

Table 54: Scale reliability statistics output of the "green identity" construct.

Item Reliability Statistics

	mean	sd	item-rest correlation	if item dropped Cronbach's α
Green ID1	5.63	1.35	0.829	0.847
Green ID2	5.48	1.38	0.815	0.852
Green ID3	5.02	1.39	0.725	0.883
Green ID4	5.06	1.57	0.730	0.886

Table 55: Item reliability statistics output of the "green identity" construct.

13.4.3. Reliability Analysis - Sustainability

Scale Reliability Statistics

	Cronbach's α
scale	0.599

Table 56: Scale reliability statistics output of the "sustainability" construct.

Item Reliability Statistics

	mean	sd	item-rest correlation	if item dropped Cronbach's α
Sustainability Check	4.26	0.913	0.444	0.336
Aesthetic Check	3.81	1.206	0.444	0.587

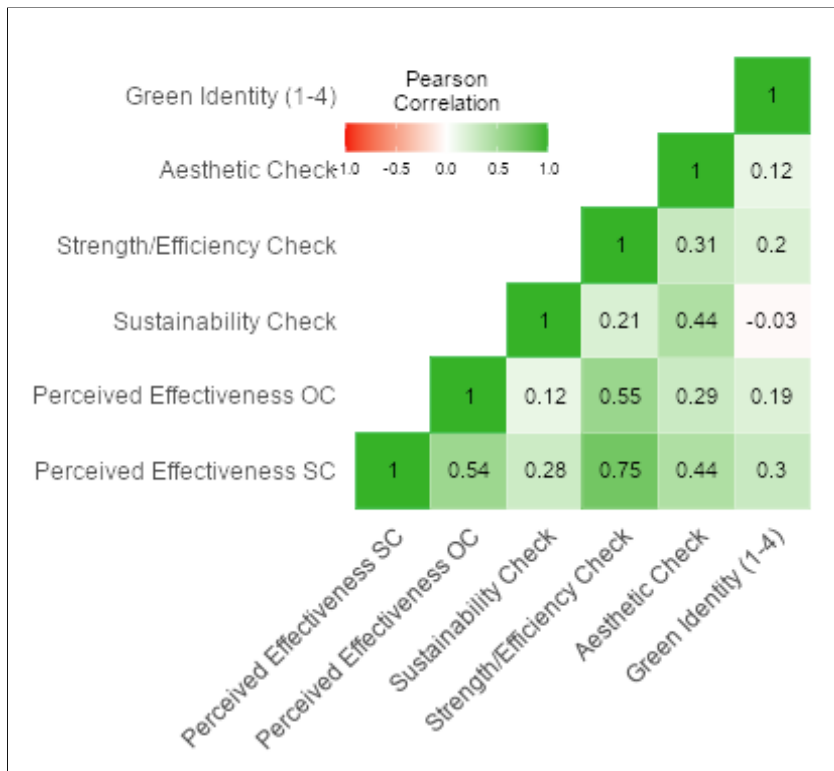*Table 57: Item reliability statistics output of the "sustainability" construct.*

13.5. Correlation Heatmap

Figure 9: Correlation matrix of the measured variables.