TI 2024-024/I
Tinbergen Institute Discussion Paper

Learning the value of Eco-Labels: The role of information in sustainable decisions

Alejandro Hirmas[1]
Jan B. Engelmann[2]

1 University of Amsterdam and Tinbergen Institute
2 University of Amsterdam and Tinbergen Institute

Learning the value of Eco-Labels:

The role of information in sustainable decisions

Last updated: April 9, 2024

Link to newest version

Alejandro Hirmas[1] and Jan B. Engelmann[1]

[1]University of Amsterdam, Center for Experimental Economics and political Decision making (CREED), and the Tinbergen Institute

Abstract

Sustainability ratings help consumers understand the environmental impact of their purchases. Such ratings have increased the consumers' sustainable choices in the electrodomestics and housing markets. In the particular case of energy labels, sustainable products are also associated with private benefits due to future cost reductions in energy expenditure. These results question the potential effectiveness of sustainability ratings for other products, such as food, where the link between environmental and private benefits is less clear. In two incentivized experiments (N=749), we study how consumers use sustainability ratings when these ratings are dissociated from private benefit, i.e. product quality. Participants chose between two products based on their quality and sustainability, which were presented in separate rating scales, alongside the products' prices. Furthermore, we study how consumers integrate the usage of ratings with other information provided from other sources. Halfway through the experiment, we provide information regarding the underlying value behind the ratings. Using a between-subject design, we modify the information provided and analyze the impact of such information on the participants' subsequent choices. Our findings indicate that even when sustainability ratings are not connected to the products' quality, participants make use of them to decide which products to buy. We also find that participants underreact to new information, and make inefficient choices based on their decisions from before. Moreover, to track the participants' attention and analyze potential heterogeneous usage of the information we use process-tracing methods. We find that participants show highly heterogeneous attention patterns, which are linked to differential weighting of the product's attributes (price, quality, and sustainability) during the decision. While our information treatment has little effect on attention allocation to individual attributes, participants correctly recall the information at the end of the experiment. These results suggest that participants partially neglect new information, and anchor to their initial decision rules formed before the information treatments.

Keywords: Attention, Sustainability ratings, conjoint analysis, information treatments, Mouselab
JEL Codes: D81, D83, D87, D91.

1 Introduction

The consequences of global warming are costly and are already visibly affecting the majority of the world's societies. So far, the efforts to reduce the global levels of CO2 emissions, and the rise in temperatures associated with that, have proven to be largely ineffective (European Commission et al., 2021). Although evidence shows strongly positive attitudes of consumers' towards sustainability, this is not reflected in their actual consumption decisions (ElHaffar et al., 2020). One reason for consumers not to follow through with their sustainable intentions is due to their unawareness, or underestimation, of the climate impact of their food choices due to the general lack of accessible information (Camilleri et al., 2019; Pace and van der Weele, 2020).

Eco-labels and sustainability ratings, such as the eco-score or the EU Energy labels, have been developed to provide consumers with credible and understandable information regarding the sustainability of their choices. Evidence shows that these labels and ratings increase the sustainable choices for several types of purchases, such as buying a house (e.g., Chegut et al., 2020; Khazal and Sønstebø, 2020) or choosing appliances (e.g., Labandeira et al., 2020; European Commission, 2021). Although these results suggest that labels increase the sustainable choices of consumers, there is evidence that these decisions are driven by the private benefits associated with their choices instead of their environmental benefits (Stadelmann and Schubert, 2018).

Specifically, when choosing an electrodomestic product, such as a fridge or a microwave, with a better energy rating, one is not only choosing the sustainable option, but also the option with less energy costs in the future. Therefore, the energy ratings inform consumers about both private (cost reduction) and environmental benefits. The relationship between quality and sustainability for other type of products, such as food or cleaning products, is not necessarily as positive and clear as it is with electrodomestics. Moreover, the food industry has great mitigation potential, since it contributes about 21% to 37% of the total emissions (Ritchie and Roser, 2023). Therefore, understanding the underlying reasons for choosing a sustainable product is fundamental when assessing the impact of a sustainability rating for other types of products. In this study, we investigate the role of sustainability ratings in the consumer's decision, when the sustainability of a product is decoupled from its quality. In this way, we disentangle the personal from the environmental concerns behind choosing a sustainable product, and therefore, can predict better the impact of sustainability ratings for other types of goods, such as food.

Besides the sustainability-quality relationship of a product, consumers have preconceived ideas of what it means for a product to have a certain sustainability rating. Public and social media are loaded with conflicting pieces of information. Consumers are frequently exposed to news and information about sustainability, but also about practices like greenwashing (Delmas and Burbano, 2011; Yokessa and Marette, 2019), which not only affect their environmental concerns but also their beliefs about what the sustainability certifications entail. The second objective of this project is to understand how consumers integrate such types of external information into their subsequent decisions when ratings are involved.

We address these questions in the current study by developing an incentivized experimental task that simulates an online shopping experience. Participants choose between different pairs of products depending on three attributes, namely price, quality, and sustainability. While the products are not real, the attributes of the chosen product will have real consequences for the participants' outcomes: (1) participants receive a bonus payment depending on the quality level of the chosen product and the price they pay for it; (2) additionally, depending on the sustainability level of the chosen product, a donation to plant trees is made on the participant's behalf. Finally, we record how participants pay attention to the product attributes. We expect that individual preferences for the different attributes will be reflected by participants' attentional patterns.

Participants see the product's quality and sustainability via visual rating systems when making their decisions. After several decision trials, we provide participants with precise information about how the different ratings affect their bonus payment and the donations for planting trees. Participants then make the same purchasing decisions in the second half of the experiment. A comparison of pre- and post-information decisions enables us to assess how participants' decisions are affected by precise knowledge about the consequences of their decisions via the rating system.

To test the impact of external information on the usage of the rating systems, we vary the information provided between subjects. In our baseline condition, the information for sustainability and quality is comparable. We contrast our baseline with two treatment conditions, in which some products provide distinctively less benefits regarding their sustainability or quality. Specifically, ratings scales in these conditions deviate from linearity by downgrading the intermediate tier's value to a value that is only slightly larger than the preceding tier. Comparing the different treatment conditions allows us to not only measure the impact of the new information on the usage of ratings, but also to test whether participants differentially adapt to news regarding the sustainability and quality of the product.

Using a random-utility model framework (RUM) we estimate the value attributed to each rating and for each attribute. Specifically, we estimate the decision weights for each attribute and then use these weights to estimate the willingness to pay (WTP) for each rating. Our results show that even in the absence of any link between quality and sustainability, participants use both quality and sustainability ratings for their decisions. The willingness to pay (WTP) increases more steeply from low- to mid-level ratings than from mid- to high-level ratings suggesting a non-linear (concave) increase in WTP as the rating for each attribute increases. We contrast these results to the elicited beliefs regarding the ratings value, and discard that beliefs are driving this concavity. Our results suggest that regardless of the underlying values, participants normalize the value of labels[1].

Based on the baseline behavior of the participants, we assess the effects of the different information treatments on the willingness to pay for the ratings. Our results show that participants under-react to the information provided when compared to the optimal decision. Namely, in both treatment conditions where the marginal sustainability or quality value of some products is reduced to almost nothing (compared to products with a lower rating), participants are still willing to pay for those products. These results suggest that participants are "leaving money on the table" by choosing to buy more expensive products that provide little to no additional benefits. Moreover, we find asymmetric reactions towards information about the underlying values of individual quality and sustainability ratings in our treatment conditions. When considering the sustainability of the products, participants shift from the products whose sustainability rating was downgraded towards products with both lower and higher sustainability ratings. On the other hand, when the quality rating of some products is downgraded, participants shift only towards higher quality products.

Finally, we explore the mechanisms driving these results. Using methods from (Hirmas et al., 2023), we show that participants are highly heterogeneous in the decision weight they allocate towards the different attributes. We find that most of this heterogeneity is driven by differences in the attention paid to the sustainability attribute. Namely, participants who consistently focus more on the sustainability ratings, compared to the rest of the sample, are willing to pay up to 88% more for higher sustainability ratings. We study whether individual characteristics, such as gender, age or their attitudes towards sustainability, help explain the differences in attention. We find no strong correlation between attention and any of the individual measures we used. Therefore, we conclude that attention provides additional information about the decision process that post-questionnaire measures cannot capture.

We also study whether changes in attention after the information treatments can better explain

[1]Evidence in neuroscience supports these types of value normalization (e.g., Bavard and Palminteri, 2022)

the treatment effects. We find that the attention allocated to the different attributes is minimally affected by the information treatments. These results provide evidence that participants are not fully adjusting their decision strategies after the information treatments. At the end of the experiment, we find that participants almost-perfectly recall the information provided by the information treatments. Therefore, they are aware of the information, they can remember it, but they do not incorporate it into their choices. These results provide evidence that participants anchor to their decision strategies and do not fully incorporate new information (Tversky and Kahneman, 1974).

This paper presents a novel paradigm that mitigates the challenges of studying environmental decisions in an incentivized setting, which is suitable for both lab and online experiments. Using bundles of attributes (i.e., combining price, earnings and trees planted) framed as 'products', we can dissociate the private and sustainable benefits associated with a 'product'. Our findings demonstrate how consumers make use of sustainability ratings when the ratings are not linked to private benefits, and how they adapt the usage of these ratings when new information comes to light. These results inform policymakers about the (heterogeneous) consumer attitudes. Importantly, our results suggest that consumers normalize the value driving the ratings and under-adjust when new information comes to light. The results show that participants understand and remember the information, but choose to neglect it.

2 Related Literature

Classic economic theory works on the assumption that consumers use and process all information available to make their decisions. There is substantial evidence from the work on behavioral economics, cognitive psychology, and neuroeconomics that challenges such assumptions. Namely, we find that consumers not only avoid available information but also under-use, or misinterpret, the information that is obtained (for reviews on the topic: Orquin and Mueller Loose, 2013; Mackowiak et al., 2021).

Nowadays, consumers face an increasing flow of information when making decisions. As the information load approaches consumer's cognitive capacity limits, they become more likely to avoid, neglect, and even misapply information that would have been valuable for their decision (Simon, 1971). This trade-off between the availability of information and the costs of attention has been extensively studied in the literature on "rational inattention" (e.g. Sims, 2003), which assumes that agents are aware of the costs of obtaining information. The theory predicts that rationally inattentive agents optimally decide their search rules for information such that the marginal cost of obtaining information equals the expected marginal value of the information obtained.

Parallelly, empirical work in economics has demonstrated that consumers fail to incorporate fundamental information in multiple decision contexts. When information is not easily available, consumers are less likely to use it in their decisions (e.g., Chetty et al., 2009; Allcott and Taubinsky, 2015). Moreover, if information becomes more salient (i.e., more striking compared to the rest), agents are more likely to 'over-weight' this information in their decisions (e.g., Bordalo et al., 2013; Hirmas et al., 2023). The literature in Marketing and cognitive psychology shows that there are multiple factors, not linked to the product's value, that still affect the decision. In the meta-analysis from Orquin et al. (2021), the authors compile evidence that factors in the spatial presentation of the options, such as the position, size, and overall salience of a product, have a strong impact on the consumers' choices.

These different strands of literature provide evidence that consumers' choices are not fully informed and that there are contextual factors, seemingly unrelated to the decision, that consistently influence choice. Therefore, helping consumers during their deliberation process can have a large impact on their choices. A commonly used tool for helping consumers in their decisions are labels and ratings, which aggregate information that otherwise would be too difficult to obtain and process. Consumers are exposed and affected by multiple types of labels and ratings, such as quality ratings when choosing service providers (e.g., hotels, transportation, restaurants, products in online shops), energy efficiency of appliances or houses (i.e., EU energy labels; European Commission, 2021), nutritional value of food (e.g., Jürkenbeck et al., 2022; Barahona et al., 2023; Crosetto et al., 2024); and also sustainability (e.g., EU Eco-label Yokessa and Marette, 2019).

There is an extensive body of literature that studies the impact of sustainability labels and ratings on consumer choices (for reviews on the impact of sustainability ratings, see Yokessa and Marette, 2019; Bastounis et al., 2021; Majer et al., 2022). Nonetheless, most of these studies use hypothetical choices, where participants are primed to think of their purchasing experience and to state their preference for different products (e.g., Staples et al., 2020; Potter et al., 2022). There are significantly fewer studies that either use incentivized settings in the laboratory (e.g., Vecchio and Annunziata, 2015; Engel and Szech, 2020) or implement field studies in cafeterias, canteens, or supermarkets (e.g., Tilling, 2023; Vlaeminck et al., 2014). Hypothetical choices make the design and implementation of the experiments simpler at the cost of having real incentives. On the other hand, incentivized experiments regarding sustainability ratings are specific to their context, since the incentives require real products. Participants make choices based on their preferences for that specific product, which decreases the external validity of the results. Our design lies in between these two types of studies. First, we provide an incentivized choice paradigm, but without having to incorporate any specific product into the design. Moreover, the implementation of our paradigm is rather simple and cost-effective, as opposed to lab experiments with products or field experiments.

Within the experimental literature, several studies that have assessed the impact of labels and ratings in combination with other attributes. Some of these studies compare the impact of different labels/information (e.g., Andor et al., 2019; Sigurdsson et al., 2022; Kolber and Meixner, 2023). Parallelly, some studies test the provision of direct information regarding how sustainable the products are (e.g., Steiner et al., 2017). De Bauw et al. (2021), the authors study the role of a sustainability rating called eco-score in combination with the Nutri-score, a nutritional rating. While the Nutri-score directly informs the participants about the nutritional value of the product, the eco-score informs them about environmentally friendly the product is. In this specific context, the authors find that the combination of both scores still enhances the nutritional value of the participants' choices, but does not affect the sustainability of their choices. These results suggest that when more information is available, and this information is personally relevant and beneficial, participants decrease their levels of attention to the sustainability ratings.

Providing sustainability ratings helps to ensure that consumers have information regarding the environmental impact of their choices. This information will only affect choices if attended and considered important. Thus, the study of the attention allocated to these labels is relevant to understanding their (seemingly heterogeneous) impact on the consumers' choices. In two hypothetical food choice studies (Van Loo et al., 2015, 2021), the authors show that (1) participants pay more attention to nutritional information than the environmental impact of the products and (2) the increased attention to the labels enhances their importance for the decision. Another study shows that informing participants about the underlying information of sustainability labels increases their attention toward the label. Our study further explores the role of attention in consumption choices with sustainability ratings. In our incentivized experiment, we analyze both within- and between-subject variations in attention and choices to assess the channels in which attention connects to the decision process. Moreover, we explore how additional information can affect the decision process, and whether attention can predict part of these effects.

3 Experimental design

3.1 Participants

We recruited participants via the online platform Prolific. As preregistered, we excluded participants who did not pass attention and manipulation checks; or who took too long to finish the task (see Appendix section A for a detailed description of the exclusion criteria). Our final database includes 290 participants (135 females, 2 non-binary, average age = 29.96). Most of the participants belong to the United Kingdom (N=68) and continental Europe (N=167). The experiment lasted 26 minutes on average. Participants received a participation fee of 2.3 pounds and they were told that depending on their choices, they could earn additional bonus payments. Additionally, they were informed that, depending on their choices, trees will be planted on their behalf in a location of their choosing. On average, participants earned 1.91 pounds as a bonus payment and we donated 1.47 trees per participant to the organization One Tree Planted, which will plant trees in locations specified by our participants around the globe (See locations in appendix section B). All procedures were approved by the Ethics Committee of Economics and Business (EBES), University of Amsterdam.

3.2 Materials and Procedure

The experiment was programmed using the web-based software oTree. The front end of the experiment was web-based (i.e., programmed using a combination of HTML, JavaScript, and CSS)[2]. Since our experiment relied heavily on how participants perceive visual stimuli, participation from smartphones and tablets was blocked to ensure that visual stimuli were presented on a sufficiently large screen (i.e., on desktop and laptop computers). Moreover, participants were instructed to keep their Fullscreen mode on and not to switch to other pages. Participants were not able to continue with the experiment if their page was not on Fullscreen. Additionally, we told participants that if they jumped to other pages repeatedly, we reserved the right to exclude them from payment (and analysis).

Participants began agreeing with the informed consent; then they were instructed to set their browser to Fullscreen mode. Following this, they read the instructions and answered several comprehension check questions. After answering all questions correctly, they proceeded to the main task, consisting of two decision blocks, each consisting of 17 rounds of the main task. Between the

[2] code available here.

decision blocks, we elicited the participants' beliefs about how the attributes used in their choices affected their outcomes. Then, we provided clear information clarifying how these attributes affected their outcomes. After completing both decision blocks, participants answered a questionnaire with demographics, comprehension questions, questions regarding their behavior in the experiment, and the Connectedness to Nature scale (Mayer and Frantz, 2004). We also included the question "How do you see yourself: are you generally a person who is prepared to behave sustainably, even when this is inconvenient or costly to you?" (Likert scale 1-5). Three attention checks were included within the questionnaire to ensure that participants were paying attention. These questions prompted participants to give a specific answer (e.g., 'Please answer 2'). Participants who failed to answer two out of the three attention checks were also excluded from payment and further analysis.

Main Task

The main task consisted of a series of purchasing choices. Participants chose between two fictitious products depending on their attributes, which included price, quality, and sustainability. While the products were not real, the decisions had real outcomes for the participants. Choosing products with a higher price resulted in a lower bonus payment, while products with higher quality yielded a higher bonus. Choosing products with higher sustainability ratings meant that more trees were planted on participants' behalf. At the end of the experiment, one decision was randomly selected to be payout relevant, and participants received the two-part bonus associated with that decision: (1) a bonus payment reflecting the private value of the product (based on the chosen products' quality) minus the price, and (2) a given amount of trees to be planted on their behalf. The value of the products ranged from 60 to 90 quality points (10 points = 0.5 pounds), resulting in a range of 3 to 4.5 pounds for the bonus. The price of the products ranged between 1 to 3 pounds. Therefore, the products never had a cost greater than their monetary value, and participants were aware that the bonuses would never be negative. The products' sustainability ranged between 0 to 30 sustainability points (10 points = 1 tree planted). To personalize the sustainability attribute, participants could choose a specific location around the globe, or choose 'anywhere' if they had no preferences (See Appendix B for more information about the selected locations).

The quality and sustainability values were presented to participants as two distinct three-tier ratings. Quality was presented as a one-to-three 'star' rating, while sustainability would be presented as a one-to-three 'leaf' rating. Participants were informed that a higher tier in the rating represented higher values of the respective attribute and they were informed of the respective interval ranges (quality range = 60-90 points, sustainability range = 0-30 points), but they were not informed about the exact value underlying each tier before the experiment.

Figure 1 shows an example of the information provided to participants during a decision trial. For illustration purposes, Panel A displays all the information available. Of note, participants never observed all the information simultaneously because we hid this information behind grey cells (Panel B) to record how participants paid attention to different attributes while they made their decisions. Participants could reveal the information by moving their cursor over the attribute's row, which made the information of that attribute available as shown in Panel C. This method is similar to standard approaches, such as provided via MouselabWeb (e.g., Willemsen and Johnson, 2011), but differs from it as the information is presented simultaneously for an entire row. When developing the task, we noticed that participants took significantly longer if the information was revealed one by one (about five times more). Since longer studies lead to fatigue and boredom (Miller, 2023), and because our main focus is to analyze the changes in decisions after the information treatments, we opted for revealing the attributes simultaneously, thus facilitating the comparison between products within a given attribute. Importantly, since the scales and dimensions of the attributes were significantly different, we expected comparisons to be more prominent within-attribute relative to between-attribute. Evidence shows that when participants are exposed to more attributes than options, which is our case, then they are more likely to compare products attribute-wise instead of integrating the information product-by-product (Meißner et al., 2020; Jenke et al., 2021). It is pertinent to note that our statistical models require only the time spent looking at an attribute, not at each specific value, eliminating the need for fixation times on individual values. Therefore, we concluded that our *row-wise Mouselab* approach is more fitting for our paradigm.

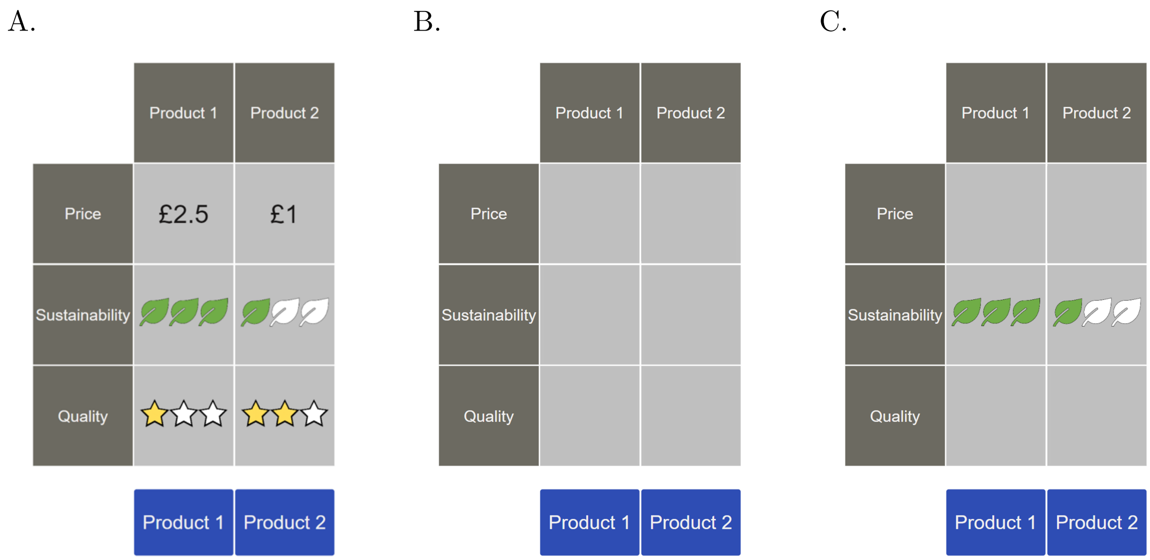


Figure 1: Example of a decision screen

This figure shows an example of the decision screen. Panel A provides all the information that participants could observe on a trial. Participants never observed all the information at the same time. At the beginning of a trial, all the information is covered (Panel B). To reveal information for a specific attribute, participants are required to hover the mouse over any column of that attribute, and the attribute values are revealed simultaneously for both products (Panel C).

The order of the attributes was randomized at the participant level, remaining constant for each participant throughout the experiment. The price was always on the top to increase external validity, as it is most common to have the price first in online webshops. Quality and Sustainability were randomized in the second and third positions. Between every decision, participants clicked on a button to continue to the next trial. This button was positioned in the middle of the vertical axis (y=0.5 screen height), while the position of the horizontal axis was randomized across trials to apper on either the left (x=10% of the screen width) or right side (x=90% of the screen width). This intermediate button ensured that the participant's cursor was never on top of any attribute before the beginning of a trial. By randomizing the horizontal location, we minimize the risk of priming the products on one specific side.

Participants performed 17 trials for each decision block, before and after the information treatment. The attribute values for each trial were pseudo-randomized to maximize the power of our estimations and increase the external validity (see Appendix C for a detailed description). Every participant completed 5 trials for each combination of sustainability ratings (bottom vs. middle, middle vs. top, and bottom vs. top) plus two trials with the same value of sustainability for both products. The product with the highest sustainability was randomized between left and right. We did not present trials where any choice was strictly dominated (i.e., more Quality and Sustainability and lower price for the same product).

Information treatment and belief elicitation

After the first decision block, we elicited the participants' beliefs about the underlying values of the quality and sustainability rating scales. We asked participants to guess how many sustainability points they would obtain for a product of each tier. Similarly, we asked how many quality points they would obtain for each tier of the quality rating[3]. Participants were reminded of the range for both quality (60 to 90) and sustainability (0 to 30) points. They answered in steps of 2.5 points and received 10 pennies extra for each correctly estimated rating.

After eliciting the participants' beliefs about the ratings, we provided clear information regarding the underlying values for both ratings. Each participant observed two graphs, one for sustainability and one for quality. The graphs showed the amount of points they could obtain by choosing products with a specific rating. Figure 2 shows all the possible graphs that could be presented to participants. The information in these graphs allowed participants to know exactly how many points they would obtain for each tier of both quality and sustainability ratings. To assess how sensitive the participants are to the information, we assigned participants to three between-subject treatment conditions, each with different underlying values for either the sustainability or quality rating scale that were communicated via the graphs shown in Figure 2. In the baseline condition, which we refer to as the *linear* treatment, values for *both* sustainability and quality increased linearly, as reflected by the information provided in panels A and B. In the second condition, named *sustainability convex* treatment, values of the sustainability attribute increased significantly only for the top-tier of the sustainability rating, such that the mid-tier of sustainability had a similar value to the bottom tier (panel C). Quality outcomes increased linearly, as described in panel B. Analogously, the third condition, the *quality convex* treatment, displayed a linear value increase for sustainability (panel A) and a convex value increase for quality (panel D) that is equivalent to the convex sustainability increase in panel C.

After the completion of the second block of decisions, we elicited participants' beliefs about the ratings once again. This assessment is a manipulation check confirming whether participants processed the rating scale values provided to them after the first block and whether they remember

[3]See Appendix D for more details about the belief elicitation method.

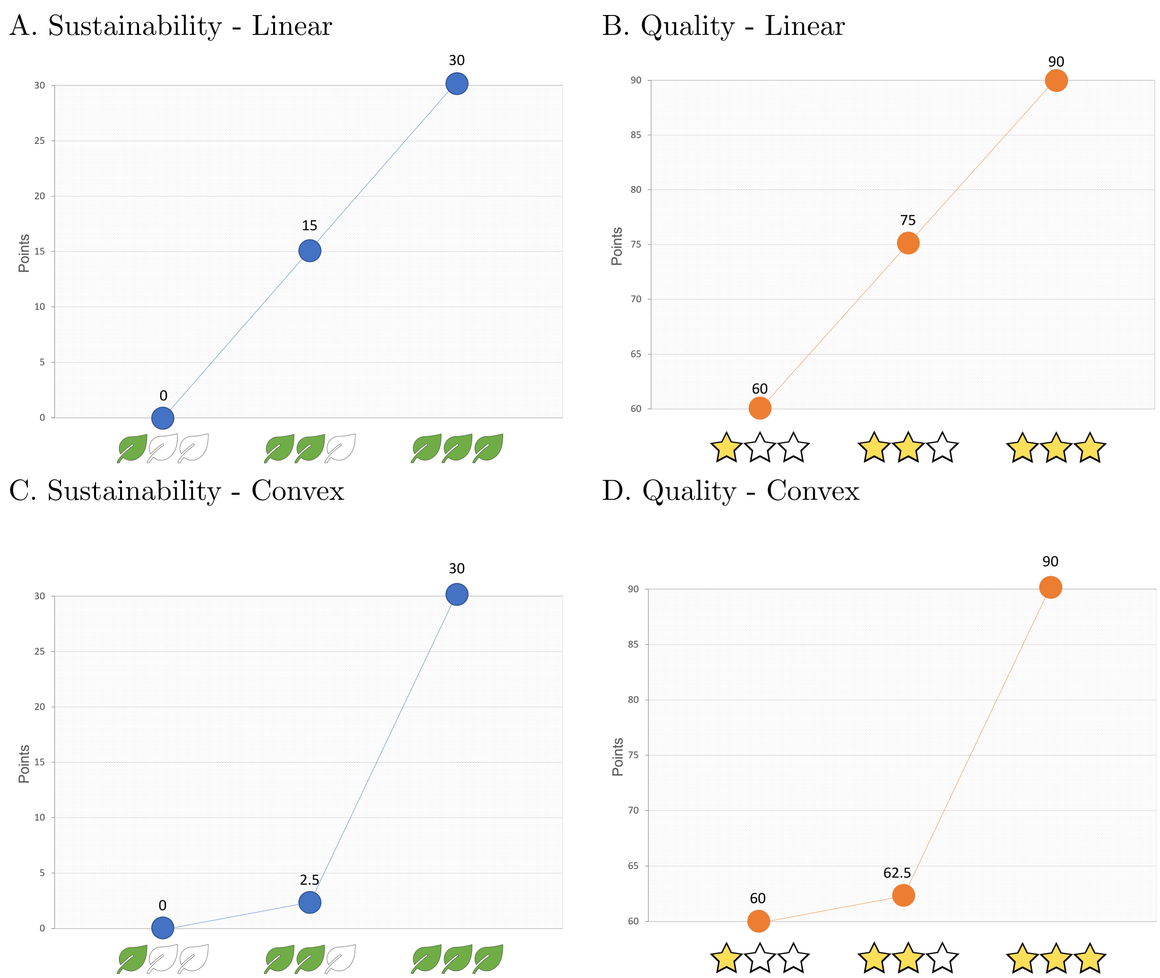


Figure 2

This figure shows values underlying each tier of the sustainability and quality rating scales. At the halfway point of the experiment (after 17 trials), participants are shown different combinations of these rating scales depending on their assigned treatment condition: Participants in the *linear treatment* condition were shown panels A and B, participants in the *sustainability convex* treatment were shown panels C and A, and participants in the *quality convex* treatment were shown panels A and D. This information was provided after the first block of decisions, so at the halfway point of the experiment. The left column shows the graphs for sustainability, while the right column shows the graphs for quality; top panels show linear values, while bottom panels show convex values.

this information after making additional decisions in block two. Following the belief elicitation, we presented participants with three graphs for each rating, a linear, a convex, and a concave graph. We asked participants to choose the type of graph that they saw after the completion of block one. Across all treatments, only four participants (0.013 percent of all participants) failed to remember the correct graph for either sustainability or quality. Next, participants proceeded to fill out the questionnaire.

We conducted a pilot study with equivalent experimental procedures (N = 459), with the exception that we used different information treatment conditions. Specifically, the treatment ratings scales were closer to a linear scale and therefore differed from the linear baseline condition to a lesser extent compared to the current study. Also, ambiguity was added to the exact values of the rating scale as values were surrounded by confidence bounds within which final values could fall, instead of the specific point values used in the current study (see Appendix E). Under these conditions of added ambiguity and difficulty, our manipulation checks showed that more than two-thirds of the participants did not perceive or remember any differences across treatments by the end of the experiment. Of note, in the pilot study, we ran an additional treatment condition, in which no additional information was provided after block 1. Given that the results showed no differences in behavior between the linear and no-information conditions, we decided to omit that condition in the current study.

4 Theoretical Framework

In this section, we describe the empirical methods used in our analysis and state our hypotheses. First, we present the baseline model for representing the participants' choices. Then, we analyze the impact of the information treatments on the baseline behavior. Finally, we incorporate the method from Hirmas et al. (2023) to capture heterogeneity in choice behavior via the participants' attention patterns.

4.1 Decision process

We model the choice between two products using a random utility model (RUM; McFadden, 1973). The RUM assumes a stochastic decision process due to agents' imprecise assessment of the subjective value of the available options due to distractions, time pressure, impulsiveness, attentional biases or misperceptions.

To describe the decision model, consider an agent i belonging to a population J, who makes a series of purchasing decisions (indexed by $t \in \{1, ..., T\}$) between two products. The agent assigns a subjective value to each option, ($V^o_{i,j,t}$ for $o \in \{1, 2\}$). The subjective value is specified as a linear combination of the option's attributes ($x_{k,i,t}$ for $k \in K$). The difference in subjective value for both options, $\Delta V_{i,t} = V^1_{i,j,t} - V^2_{i,j,t}$, is described in equation 1:

$$\Delta V_{i,t} = \alpha_{0,i,t} + \sum_{k \in K} \omega_{k,i,t} \Delta x_{k,t} + \epsilon_{i,t} \tag{1}$$

Where $\omega_{k,i,t}$ represents the weight that agent i allocates to attribute k in trial t, $\alpha_{0,i,t}$ is an ex-ante predisposition for option 1 over 2; and $\epsilon_{i,t}$ is the error term, which follows an independent Gumbel distribution.

In our experiment, each product has three attributes, price, quality, and sustainability. Since the attribute values of quality and sustainability are categorical variables (bottom, middle, and top tier), we allow the increase between tiers to be non-linear. For $k \in \{Q, S\}$, we define the attribute values as follows:

$$\omega_{k,i,t} \Delta x_{k,t} := \begin{cases} \omega^{a,b}_{k,i,t} & \text{if} \quad x^1_{k,t} = a \text{ and } x^2_{k,t} = b \\ -\omega^{a,b}_{k,i,t} & \text{if} \quad x^1_{k,t} = b \text{ and } x^2_{k,t} = a \end{cases} \tag{2}$$

Where $(a, b) \in \{(2, 1), (3, 2)\}$ represent the rating tiers for the different products. Based on our previous experiment (see Appendix E), we predict that in the absence of any additional information, participants expect a linear increase in the underlying values behind the attributes' ratings. Nonetheless, we expect that participants will not value the increase in the ratings linearly (reflecting marginal decreasing returns). Which lead us to our first pre-registered hypothesis:

Hypothesis 1. *The marginal value of choosing a product with a higher attribute rating $k \in \{Q, S\}$ is decreasing.*

$$\omega^{2,1}_{k,i,t} > \omega^{3,2}_{k,i,t} \quad \forall k \in \{Q, S\}$$

One possible reason for a non-linear valuation of the ratings is that participants expect the underlying attribute to increase non-linearly. In our pilot study we found that participants expect the underlying attributes driving the ratings to increase linearly. Our second hypothesis aims to confirm our previously found results:

Hypothesis 2. *Participant's belief's about label values increase linearly for attribute $k \in \{Q, S\}$.*

We test the hypotheses above in a baseline model. Following the common practice in economics, we use as a baseline a random-effects logistic regression, which can be represented in our model by imposing the following constraints:

$$\omega_{k,i,t} := \omega_k \qquad \text{(common weights for all agents and periods)}$$

$$\alpha_{0,i,t} := \alpha_i \qquad \text{(random-intercepts for each participant)}$$

4.2 Information treatments and changes in preferences

The additional information, provided in the middle of the experiment, precisely explains the underlying values driving the quality- and sustainability ratings. To assess the impact of the information on the decision process, we estimate the changes in all decision weights after the new information, conditional on the treatments. We operationalize the effect of the information treatments as follows. For every attribute $k \in K$, the decision weight is:

$$\omega_{k,i,t} \quad := \quad \pi_{k,\tau} + \pi_{k,after}\mathbb{1}\{t > T/2\} + \pi_{k,DiD,\tau}\mathbb{1}\{t > T/2\} \tag{3}$$

Where the $\pi_{k,\tau}$ captures the treatment group effects, $\pi_{k,after}$ captures the before vs after effects, and $\pi_{k,DiD,\tau}$ is the differences-in-differences effect of treatment τ on the decision weight for attribute k.

When comparing our baseline linear condition (L) with our two information treatments, Sustainability Convex (SC) and Quality Convex (QC), we expect the following results. First, in the Sustainability Convex condition, the middle tier of sustainability rating is decreased almost to the level of the bottom tier. Therefore, if participants incorporate this information, they will decrease their willingness to pay for switching from the bottom to the middle tier of sustainability (2 vs 1), but will increase their willingness to pay to switch from the middle to the highest (3 vs 2). This leads us to our next hypotheses:

Hypothesis 3a. *The willingness to pay for the middle tier of the sustainability rating decreases* *after the information treatment in the sustainability convex condition (compared to the linear condition)*

$$\pi^{2,1}_{S,DiD,SC} < 0 \quad \pi^{3,2}_{S,DiD,SC} > 0$$

Similarly, in the Quality convex treatment, we expect participants to reduce their willingness to pay for the second tier of quality.

Hypothesis 3b. *The willingness to pay for the middle tier of the quality rating decreases after the information treatment in the quality convex condition (compared to the linear condition)*

$$\pi^{2,1}_{Q,DiD,QC} < 0 \quad \pi^{3,2}_{Q,DiD,QC} > 0$$

In the next section, we enhance our decision model by incorporating differences in the decision weights based on attention (Hirmas et al., 2023). We allow the decision weights, $\omega_{k,i,t}$, and the intercepts, $\alpha_{0,i,t}$, to be contingent on the individual making the decision (i.e., individual differences), but also the context in which the decision is made (i.e., contextual differences). The next subsection elaborates on how we use attentional measures to capture these individual- and contextual-differences in the decision model.

4.3 Decision weights and attention

We refer to attention as the cognitive mechanism by which available information is processed and filtered during decision-making. Attention supports decision making, as effective decision making relies on efficiently acquiring choice-relevant information. Therefore, examining how individuals seek information provides insights into the cognitive processes underlying decisions. This notion is supported by prior theoretical and empirical work. Theories such as rational inattention theory or salience theory, suggest that information that appears more often or is more salient gains greater relevance in the decision process (e.g., Bordalo et al., 2013; Matějka and McKay, 2015). Overall, these theories predict a greater weighting of those attributes that receive more attention in the decision process. These predictions are supported by recent empirical work that has shown that information that is more relevant for the choice is sampled more frequently and for longer (e.g., Orquin and Mueller Loose, 2013; Hirmas et al., 2023). Conversely, when information is salient (e.g., visually striking), one is also more likely to fixate on that information, which in turn is weighted more heavily in the decision (e.g., Alós-Ferrer and Ritschel, 2022; Li and Camerer, 2022).

Given the positive link between attention and decisions, we can expect attentional patterns to vary across participants in the presence of heterogeneous preferences for product attributes, and across choice contexts. Our approach makes use of these attentional differences in heterogeneous decision processes. Specifically, by analyzing systematic differences in attention, we can capture differences in the relative importance of the available information during the decision. Differences in attention can arise due to (1) individual specific factors, such as preferences, personality traits, or individual-specific information (e.g., between-subject treatments), as well as (2) context-specific factors that are not constant across the whole experiment, such as the relative value of the options' attributes. To capture variations in attention due to individual and contextual factors we construct two independent indexes of attention for each attribute.

4.3.1 Individual-average attention

We denote the attention allocated to attribute k by agent i in trial t as $a_{k,i,t}$. We define the individual average of attention, $\bar{a}_{k,i}$, in equation (4). This index represents *how much more participant i attends to attribute k compared to the other participants (measured in standard deviations).*

$$\bar{a}_{k,i} \quad := \quad \frac{\frac{1}{T}\left(\sum_{t=1}^{T} a_{k,i,t}\right) - \bar{a}_k}{sd(a_k)}. \tag{4}$$

The individual-average attention index captures idiosyncratic effects that remain relatively constant for the participant across the experiment. These constant effects reflect individual preferences, personality traits, but are also influenced by between-subject treatment effects. Based on previous work (e.g., Fisher, 2021; Hirmas et al., 2023), we expect these factors to influence both the attention and decision process. Hence, variations in the individual-average attention should be correlated with changes in the decision weights. Using these indexes as moderators for the attribute value on the choice allows us to capture the correlation between decision weights and individual-average attention. Hypothesis 4 aims to confirm our previous results but in the context of purchasing decisions.

Hypothesis 4. *The individual-average attention index ($\bar{a}_{k,i}$)for attribute k correlates positively* *with the decision weight of the attribute in the choice.*

$$\rho(\bar{a}_{x,i}, \omega_{k,i,t}) > 0$$

4.3.2 Trial-wise deviations of attention

The previous index captures stable effects about the individual that affect both attention and the decision at a low frequency (i.e., with relatively little variation throughout the experiment). Now, we turn into contextual effects that change across the experiment and can have additional impacts on both attention and the decision, but at a higher frequency. The trial-wise deviations of attention, $\tilde{a}_{x,i,t}$, in equation (5) represents *how much more participant i attends to attribute k in trial t compared to the participant's average behavior (measured in standard deviations).*

$$\tilde{a}_{x,i,t} \quad := \quad \frac{a_{x,i,t} - \frac{1}{T}\sum_{t=1}^{T} a_{x,i,t}}{sd(a_x)}. \tag{5}$$

Any context-specific factor that affects attention and changes throughout the experiment will be captured by this index. Factors such as experience on the task or variation in attribute values would be captured by this index. If these contextual factors affect both the decision and the attention process, then this index can be used as a proxy to estimate the effect of the contextual factors on the decision. We hypothesize that that this index will be correlated with the decision weights.

Hypothesis 5. *The index of trial-wise deviations in attention for attribute k correlates positively with the decision weight of the attribute.*

$$\rho(\tilde{a}_{k,i}, \omega_{k,i,t}) > 0$$

4.4 Incorporating the attention indexes to the decision process

Now, we further develop the decision model, defined by equations (1) and (3), by allowing the decision weights to depend on our attention indexes defined above. Namely, for every attribute $k \in K$, the decision weight will be:

$$\omega_{k,i,t} \quad := \quad \pi_{k,\tau} + \pi_{k,after}\mathbb{1}\{t > T/2\} + \pi_{k,DiD,\tau}\mathbb{1}\{t > T/2\} + \pi_{k,\bar{a}}\bar{a}_{k,i} + \pi_{k,\tilde{a}}\tilde{a}_{k,i,t} \tag{6}$$

In practical terms, for each attribute we additionally incorporate an interaction effect of the attention indexes with the attribute's value[4]. We can contrast this model with our benchmark with constant slopes, where $\pi_{k,\bar{a}} = \pi_{k,\tilde{a}} = 0$. Note that Hypothesis 4, stating that the individual-average attention positively correlates with the decision weights, implies that $\pi_{k,\bar{a}} > 0$. Similarly, Hypothesis 5, stating that the trial-wise deviations are positively correlated with the decision weights, is represented by $\pi_{k,\tilde{a}} > 0$.

Now that we have established our decision model, we proceed to use our estimations to calculate the willingness to pay for the different attributes, depending on the treatments, the rating comparison and the attention allocated to the attributes by the participants.

4.5 Empirical estimation and Willingness to Pay

In the following section, we show our estimations of the model based on equations (1) and (6). For better readability, we repeat the relevant equations below. As equation (1, revisited) shows, price enters linearly into our decision process. The attributes presented in ratings, quality, and sustainability, are introduced discretely (i.e., signed dummy variable depending on the combination c).

$$\Delta V_{i,t} = \alpha_{0,i,t} + \omega_{P,i,t}\Delta P_{k,t} + \sum_{k\in\{Q,S\}} \sum_{c\in\{(2,1),(3,2)\}} \omega^c_{k,i,t}\Delta x^c_{k,t} + \epsilon_{i,t} \tag{1, revisited}$$

$$\omega_{k,i,t} = \pi_{k,\tau} + \pi_{k,after}\mathbb{1}\{t > T/2\} + \pi_{k,DiD,\tau}\mathbb{1}\{t > T/2\} + \pi_{k,\bar{a}}\bar{a}_{k,i} + \pi_{k,\tilde{a}}\tilde{a}_{k,i,t} \tag{6}$$

Using the estimates of the decision weights, we construct the willingness to pay of participants for the different attributes. Namely, we estimate the ratio between the weight of a given attribute and the weight of the price.

$$WTP^c_{i,t} = \frac{\omega^c_{k,i,t}}{\omega_{P,i,t}} \qquad \begin{array}{c} \forall k \in \{Q,S\} \\ \forall c \in \{(2,1),(3,2)\} \end{array} \tag{7}$$

The ratio in equation (7) represents the value in pounds that participants are willing to pay to increase the attribute k of the chosen product from the low to the middle tier (2 vs 1) or from the

[4]We also incorporate the attention indexes without the interaction. Thus, the attention indexes can also affect the intercept. For simplicity, we do not describe these equations in this section.

middle to high tiers (3 vs 2). We use the WTP estimations to calculate the monetary impact of our treatments and attention on the decisions.

5 Results

In this section, we describe the results of our analysis concerning the participants' sustainable decisions. We use a random-effects logistic regression to estimate the model described by equations (1, revisited) and (6). For a more comprehensive presentation of the results, we separate the analysis based on our hypotheses. The analysis is structured as follows. First, we estimate the decision weights for the different attributes before the information treatments, and calculate the participants' initial willingness to pay for quality and sustainability conditional on the attribute's rating. Subsequently, we analyze the impact of the different information treatments on the decision weights and their willingness to pay for the different attributes. Finally, we investigate the role of attention in the decision process and the predictive power of our attention measures on the decision process.

5.1 Participants' choices and their Willingness to pay (WTP) for Quality and Sustainability

To determine the value attributed to each attribute, we estimate our baseline random-utility model (1, revisited), with the constraint of equal weights for all participants (i.e., $\omega_{k,i,t} = \omega_k$). We focus on the data collected before the information treatments to estimate the pre-information willingness to pay and decision weights for each attribute[5]. Our decision models use two assumptions, (1) there are no significant differences in the treatment groups' behavior before the additional information, and (2) the participants' decisions are transitive. We refer to transitive decisions when the willingness to pay to switch from the first to the third tier for any attribute is equal to the willingness to pay for increasing from the first to the second, plus increasing from the second to the third (i.e., $\omega_k^{2,1} + \omega_k^{3,2} = \omega_k^{3,1}$). We find evidence supporting both assumptions in an unrestricted model (See Appendix section H for further details).

Table 1 shows the estimated decision weights for the different attributes, where we find that all attributes significantly influence choices[6]. Additionally, we find that the marginal value of increasing the attributes from the bottom to the middle tier is larger than the increase from the

[5]Section 5.2 shows that these results are similar once the information after the treatments is incorporated.

[6]Appendix section F shows similar results when analyzing the raw choice data.

middle to the top. Both, the difference for quality ($\omega_Q^{32} - \omega_Q^{21} = -.519$, p$< 0.001$) and sustainability ($\omega_S^{32} - \omega_S^{21} = -.232$, p$= 0.001$) are significant. These results support our hypothesis 1 which states that the marginal value of choosing a product with a higher rating is decreasing.

Table 1: Decision weights before information treatment

ΔP	-1.093***	(0.072)
ΔQ^{21}	1.754***	(0.096)
ΔQ^{32}	1.235***	(0.094)
ΔS^{21}	1.233***	(0.070)
ΔS^{32}	1.001***	(0.068)
Constant	0.0236	(0.034)
Observations	4929	
AIC	4904.0	
BIC	4943.0	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Result 1. *Participants' willingness to pay for increasing the quality and sustainability of the products is concave with respect to the attribute's tier.*

The results above suggest that, on average, the marginal value of increasing the ratings decreases along the rating scale (i.e., concave WTP). Based on the results from our pilot study (see Appendix section E), we expect that the concavity is not driven by the participants beliefs regarding the underlying attributes driving the ratings. Before the information treatment, we elicited participants beliefs for every tier of both, the quality and sustainability scale. Using these beliefs, we test whether the expected marginal increase on the ratings is non-linear. The results of a signed-rank test (Wilcoxon et al., 1970) show that the marginal increase in the participants beliefs regarding quality is not significantly different from a linear function ($[B(Q(3)) - B(Q(2))] - [B(Q(2)) - B(Q(1))] = 0.64$ points, $p = 0.128$). In Appendix Section G, we show the distribution of the beliefs before the information treatments and at the end of the experiment.

On the other hand, participants expect a slightly convex increase in the sustainability ratings ($[B(S(3)) - B(S(2))] - [B(S(2)) - B(S(1))] = 0.497$ points, $p < 0.001$). It must be noted that the underlying values for sustainability range from 0 to 30 sustainability points. Hence, a difference of about half a point is quite small. Since the WTP for the different ratings is concave, and the beliefs increase rather linearly, we conclude that the concavity on the WTP is not driven by their beliefs. These results are similar to our pilot study and therefore partially support our hypothesis 2, that the beliefs before the information treatments approach linearity[7].

Result 2. *Participants believe that the underlying value driving the sustainability ratings increases convexly.*

To be able to quantify the value assigned to each attribute rating, we proceed to estimate the participants' willingness to pay (WTP) for a higher tier rating within a specific attribute category. We calculate the willingness to pay, defined in equation 7, as the ratio between the weight for a specific attribute k given the tier combination a, b, $\omega_k^{a,b}$, and the weight allocated to the price difference, ω_P. This ratio represents the amount in pounds that participants are willing to pay to increase the rating of attribute k from tier b to tier a. Figure 3 shows the estimated willingness to pay for both Quality (orange) and Sustainability (Green) and for each type of comparison (2 vs 1 and 3 vs 2). The participants' willingness to pay to increase the quality of a product from the bottom to the highest tier was 2.74 pounds. Since the value of a product ranged from 3 to 4.5 pounds, participants overpaid on average 1.24 pounds (2.74 - 1.5) for the product's quality. The WTP for increasing the sustainability rating from the bottom to the top tiers was 2.04 pounds. The value of planting a tree was 1 USD (0.8115 GBP / USD based on the current exchange rate) and depending on the participants' choices, zero to three trees would be planted on their behalf. Therefore, participants underpaid, on average, 0.39 pounds (2.04 - 2.43) for the product's sustainability. It must be noted, that participants were not aware of the cost of planting trees.

5.2 The effect of new information

In the previous section, we show that participants are willing to pay for both the quality and sustainability of the products. Now, we analyse whether providing clear information about such ratings has an impact on the willingness to pay for increasing the products' attributes. To this end, we estimate the model represented by equations (1, revisited) and (3). Table 2 shows the estimated treatment effects of the information treatments on the decision weights. The first column shows the estimates of the decision weights using only the data before the information treatments (same results as in Table 1). The second column presents the treatment effects (and baseline decision weights) based on the whole dataset (For the complete regression table see Appendix Table I).

[7] See the results presented in appendix section G

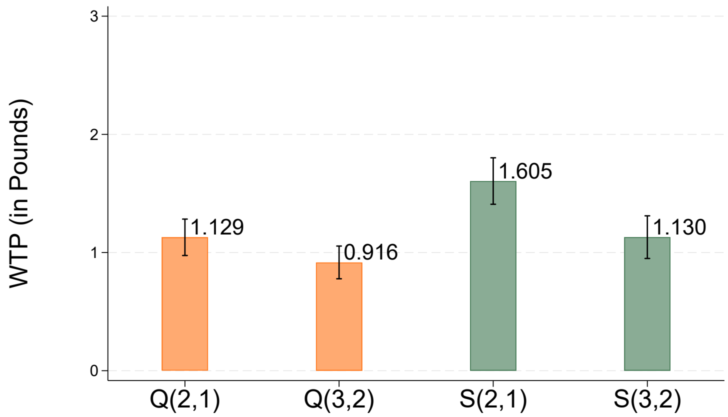


Figure 3: Willingness to pay for the attributes depending on the tier comparison

The Figure above shows the estimated willingness to pay (WTP) for increasing the attribute rating by one tier for each attribute category (Quality in orange, Sustainability in green). The WTP is calculated as in equation 7. The confidence intervals are estimated via the delta-method at a level of 95% confidence.

Table 2: Differences-in-Differences estimates of information effects on decision weights

	(1) Before		(2) After	
ΔP	-1.093***	(0.072)	-0.979***	(0.123)
$\delta^{SC}(P)$			0.159	(0.136)
$\delta^{QC}(P)$			0.082	(0.138)
ΔQ^{21}	1.754***	(0.096)	1.690***	(0.175)
$\delta^{SC}(Q,21)$			0.005	(0.218)
$\delta^{QC}(Q,21)$			-0.218	(0.244)
ΔQ^{32}	1.235***	(0.094)	1.254***	(0.165)
$\delta^{SC}(Q,32)$			-0.096	(0.190)
$\delta^{QC}(Q,32)$			0.696**	(0.249)
ΔS^{21}	1.233***	(0.070)	1.114***	(0.124)
$\delta^{SC}(S,21)$			-0.447**	(0.165)
$\delta^{QC}(S,21)$			0.043	(0.162)
ΔS^{32}	1.001***	(0.068)	0.903***	(0.123)
$\delta^{SC}(S,32)$			0.369*	(0.184)
$\delta^{QC}(S,32)$			0.080	(0.175)
N	4929		9851	
AIC	4903.967		9950.416	
BIC	4942.984		10216.643	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Based on the results of column (2), we find that neither the Quality convex (QC) nor Sustainability convex (SC) treatments have a significant effect on the decision weights for the price ($\pi_{P,DiD,SC} = 0.159, p = 0.242$; $\pi_{P,DiD,QC} = 0.082, p = 0.554$). Hence, any effect on the willingness to pay for quality and sustainability will come through changes in the decision weights for said attributes and not the price[8].

When we consider the SC condition, we find no significant effect on the decision weights for quality ($\pi^{2,1}_{Q,DiD,SC} = 0.005, p = 0.982$; $\pi^{3,2}_{Q,DiD,SC} = -0.0957, p = 0.614$). On the other hand, we find a significant decrease in the decision weights for S(2,1) ($\pi^{2,1}_{S,DiD,SC} = -0.447, p = 0.007$) and an increase for S(3,2) ($\pi^{3,2}_{S,DiD,QC} = 0.369, p = 0.045$). These results are in line with our hypothesis 3a, which predicts that the participants' willingness to pay for the mid-tier will decrease in the SC condition.

Result 3. *In the sustainability convex condition (SC), participants shift from choosing the middle sustainability products to both low and top sustainability products.*

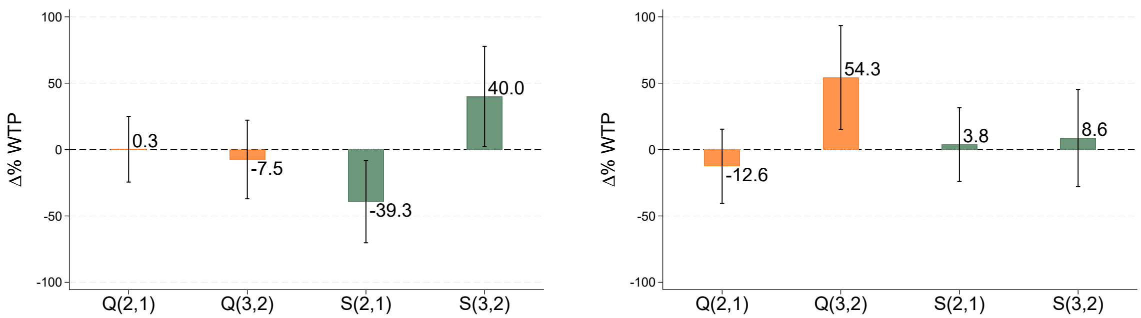
In the QC condition, we find a large and significant increase in the decision weight on Q(3,2) ($\pi^{3,2}_{Q,DiD,QC} = 0.696, p = 0.005$), but no significant decrease in the weights for Q(2,1) ($\pi^{2,1}_{Q,DiD,QC} = -0.218, p = 0.372$). These results partially confirm our hypothesis 3b, stating that participants decrease their willingness to pay for the middle-tier quality in the QC treatment. We find no significant effects on the decision weights for sustainability from the QC treatment ($\pi^{2,1}_{S,DiD,QC} = 0.0432, p = 0.789$; $\pi^{3,2}_{S,DiD,QC} = 0.0797, p = 0.649$).

Result 4. *In the quality convex condition (QC), participants shift from choosing middle quality products to top quality, but not to lower quality products.*

In both, the sustainability and quality convex conditions, the middle tier of the affected attribute has a similar value to the bottom lowest tier. In theory, this means that participants should be willing to pay an amount close to zero to increase the attribute from the bottom to the middle tier (2 vs 1). Figure 4 shows the difference in difference effects relative to the WTP before the information. As we can see in the left panel, the WTP for S(2,1) decreases by 39.29% (p=0.013) and the WTP for S(3,2) increases by 39.99% (p=0.038). These results show that while participants clearly take the information on convex sustainability ratings into account in their decisions, they nonetheless underreact to this new information in the SC condition. Since the added value of the middle tier compared to the bottom tier is practically zero, the willingness to pay for increasing that

[8]The raw choice data analysis in appendix section F shows similar results.

tier should decrease in about 100%. Instead, we find a much lower, yet still significant reduction of 39%. In the Quality convex condition, participants do not significantly decrease the WTP for Q(2,1) (Δ = 12.6%, $p = 0.377$, relative to the expected 100 %), while they increase their WTP for Q(3,2) in 54.33% (p=0.006). These results show not only an under-reaction to the information but also an asymmetric shift in their preferences. Participants do not shift down, but only up in response to negative information about the mid-quality products.



Sustainability Convex

Quality Convex

Figure 4: Treatment effects relative to WTP before information

The Figure above shows the estimated percentage differences in the willingness to pay (WTP) for each attribute (Quality in orange, Sustainability in green) and treatment condition (relative to the linear condition) depending on the rating comparison. The WTP is calculated as in equation 7. The estimated difference above is the ratio between the treatment effect (i.e., $\pi_{k,DiD,\tau}$ from equation 3) and the WTP before the information treatment. This ratio represents the percentage change of the WTP given the treatment. The confidence intervals are estimated via the delta-method at a level of 95% confidence.

5.3 The role of attention in the decision process

Our results demonstrate that participants use the quality and, importantly, the sustainability ratings in their decisions in a choice setting where the private benefits of quality and sustainability are decoupled. Moreover, they react to clear information regarding the rating distributions underlying the attributes' ratings. To shed additional light on the cognitive processes underlying participants' decisions, we analyze their attention patterns. To this end, we perform the following analyses: First, we analyze what the drivers of attention in our experimental setting are, i.e., are individual factors related to participants preferences, or contextual factors related to attribute salience more important in participants' decisions. Then, we analyze if the information treatments have a significant impact on attention, and whether these changes can help predicting differences in the treatment effects across participants.

First, we analyze whether attention is driven by individual and contextual factors; and whether the information treatments have an impact on the attention to the attributes. In the linear regressions reported in Appendix section J we show that attention to the three attributes is partially explained by both individual and contextual factors. Moreover, we also find treatment effects on how attention is allocated during the decisions. Specifically, across conditions, participants attend less to the price after the information treatments. This suggests that the quality and sustainability attributes become more salient after the information treatments since the information provided by the ratings is richer now and focused on these attributes. Additionally, participants attend more to the quality ratings and even less to the price after the quality convex condition. We find no effects on attention after the sustainability convex treatment.

Now that we have established that both individual and contextual factors affect the attention paid to the different attributes, it is important to understand how the differences in attention driven by these factors predict differences in the decision process as well. To do so, we construct the indexes described in section 4.3 and incorporate them into our decision models. Attention therefore becomes a moderator of the decision weights associated with each attribute. As outlined above and in our previous paper (Hirmas et al., 2023), we include two attention indexes in this regression, one reflecting individual attention effects related to goal-driven attention, and one reflecting contextual attention effect, related to stimulus-driven attention. Table 3 presents the moderating effect of the attention indexes on the decision weights from the model defined by equations (1, revisited) and (6). In this table, we only present the baseline decision weights and the moderating effects of attention (see Appendix table K for the complete estimations). Column (1) presents the model with no attention, which is identical to the decision model in the previous section (Table 2, column 2). The second and third columns present the models which sequentially introduce the effect of individual-average and the trial-wise deviations of attention.

Now, we proceed to test the link between the attention indices, individual-average and trial-wise deviations of attention, with the decision weights. Namely, we test our hypotheses 4 and 5, which respectively state that individual-average and trial-wise deviations of attention positively correlate with the decision weights. The joint tests show that individual average attention is a significant predictor of choice (H0: $\pi_{k,\bar{a}} = 0 \ \forall k$, $\chi^2(5) = 102.333, p < 0.0001$), supporting our hypothesis 4, which states that individual-average attention is positively correlated with the decision weights. We also find that all the moderator effects of the individual attention on the choice are significant (except for ΔQ^{21}), suggesting that this effect is independent of the attribute.

Table 3: Effects of attention indexes on decision weights

	(1) No Attention		(2) Individual Effects		(3) Individual + Contextual Effects	
ΔP	-0.979***	(0.123)	-1.141***	(0.128)	-1.143***	(0.129)
$\Delta P \times \bar{a}(P)$			-0.324***	(0.057)	-0.324***	(0.058)
$\Delta P \times \tilde{a}(P)$					-0.041	(0.027)
ΔQ^{21}	1.690***	(0.175)	1.766***	(0.180)	1.769***	(0.182)
$\Delta Q^{21} \times \bar{a}(Q)$			-0.043	(0.079)	-0.041	(0.079)
$\Delta Q^{21} \times \tilde{a}(Q)$					0.120**	(0.046)
ΔQ^{32}	1.254***	(0.165)	1.391***	(0.166)	1.401***	(0.167)
$\Delta Q^{32} \times \bar{a}(Q)$			0.182*	(0.082)	0.181*	(0.082)
$\Delta Q^{32} \times \tilde{a}(Q)$					0.010	(0.046)
ΔS^{21}	1.114***	(0.124)	1.166***	(0.120)	1.180***	(0.119)
$\Delta S^{21} \times \bar{a}(S)$			0.373***	(0.057)	0.373***	(0.057)
$\Delta S^{21} \times \tilde{a}(S)$					0.018	(0.042)
ΔS^{32}	0.903***	(0.123)	0.951***	(0.123)	0.954***	(0.124)
$\Delta S^{32} \times \bar{a}(S)$			0.356***	(0.055)	0.362***	(0.055)
$\Delta S^{32} \times \tilde{a}(S)$					0.089	(0.047)
Observations	9851		9851		9851	
AIC	9950.416		9399.544		9381.839	
BIC	10216.643		9716.138		9748.801	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Result 5. *Individual-average attention to an attribute is positively correlated with the weight of said attribute in the decision.*

When analyzing the impact of the trial-wise deviations in attention over the decision weights, we find a significant joint effect (H0: $\pi_{k,\tilde{a}} = 0 \ \forall k$, $\chi^2(5) = 24.63, p = 0.0002$), although most of the individual parameters for the trial-wise deviations in attention are not significant (except for ΔQ^{21}). Moreover, model comparisons based on BIC show that the model including both indexes of attention describes the data better than a model with no attention, but the marginal contribution of the trial-wise deviations does not compensate for the cost of more parameters. Thus, we conclude that if there is an effect of trial-wise deviations of attention on choice, this effect is weak and specific to certain attributes. Our results therefore only partially confirm our hypothesis 5. Based on these results and the fact that the best-fitting model (according to BIC) includes only the individual effects, we continue the rest of our analysis based on the model described in column (2).

Result 6. *Trial-wise deviations in attention to an attribute have a weak effect on the decision weights, and only for specific attributes.*

To further assess the economic significance of the captured individual differences in the decision weights, we estimate the participants' willingness to pay (WTP) for quality and sustainability conditional on their attention to the attributes (compared to the rest of the sample). To do so, we measure the average attention for the different attributes for participants that are on the lowest 10th percentile of the individual average attention, participants around the median, and participants above the 90th percentile. It is important to note, that the WTP of an attribute is the ratio between the weight of the attribute and the weight of the price. Hence, if for example participants, that allocate low levels of attention to sustainability, focus mostly on the price, the negative effects of attention on the willingness to pay will be enhanced (compared to the case where the focus shifts to quality).

Figure 5 shows the willingness to pay (WTP) for Sustainability (in green) and Quality (in orange) across the 10th, 50th, and 90th percentiles of attention distribution toward the respective attributes. When comparing the WTP for elevating the sustainability rating from low to middle tier, participants with low attention (10th percentile) to sustainability exhibit a 64.66% decrease in their WTP compared to the median participant ($\chi^2 = 36.24, p < 0.001$). Conversely, those with high attention levels to sustainability demonstrate a 54% increase in WTP for the same upgrade ($\chi^2 = 19.95, p < 0.001$). Similar patterns emerge when assessing the WTP for increasing the sustainability rating from middle to top tier, with participants paying 71.56% less at the low attention (10th percentile, $\chi^2 = 40.61, p < 0.001$) end and 88.48% more at the high attention end (90th percentile, $\chi^2 = 23.43, p < 0.001$), both compared to median.

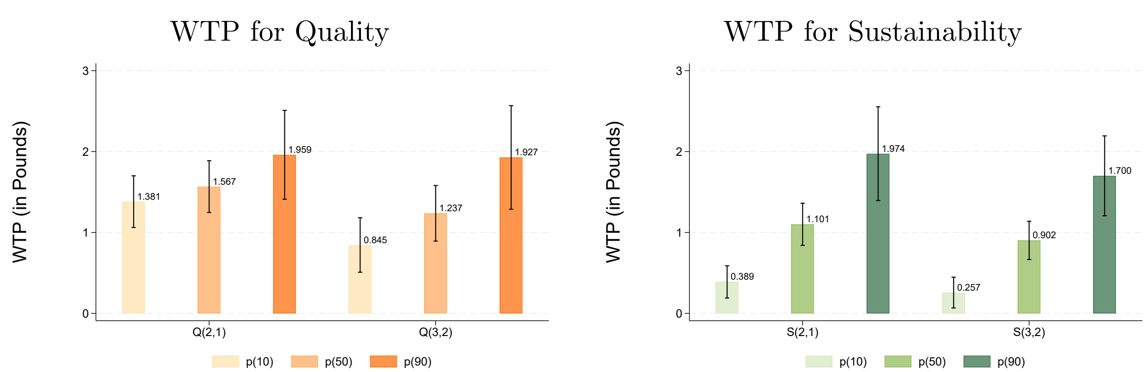


Figure 5: Willingness to pay for attributes depending on level of attributes and levels of attention

The Figures above show the estimated willingness to pay (WTP) for each attribute (Quality in orange, Sustainability in green) depending on the attention allocated to the attributes. The WTP is calculated as in equation 7 using the decision weights from equation 6. In the left (right) panel, the bars present the estimated WTP for the participants belonging to the 10th, 50th and 90th percentile of attention to quality (sustainability). The confidence intervals are estimated via the delta-method at a level of 95% confidence.

We find similar, but smaller effects for the quality ratings. When comparing low-to-middle tier decisions, participants at the low end of attention towards quality are willing to pay 11.88% percent less ($\chi^2 = 3.78, p = 0.0517$), while participants at the high end are willing to pay 25% more ($\chi^2 = 5.03, p = 0.025$). When comparing middle-to-top-tier decisions, participants are willing to pay 31.7% less ($\chi^2 = 12.74, p < 0.001$) and 55.8% more ($\chi^2 = 10.41, p = 0.001$) for the low- and high-end participants. The lower differences compared to sustainability suggest that participants that are inattentive to quality, are usually also inattentive to price. Thus, the attentional effects due to lower attention to quality get counteracted by the lower attention to price.

6 Discussion

In this study, we investigate participants' preferences in a decision context that provides a trade off between sustainability versus their own private benefits. Our experimental design simulates an online shopping experience, where participants choose between constructed products based on their quality and sustainability ratings, and the products' prices. In our setup, sustainability is decoupled from any private benefits (i.e., higher sustainability does not bring any personal profit). Thus, if participants are willing to pay for sustainability, it cannot be due to any personal benefits associated to more sustainability. Our results show that participants consider both the level of quality and sustainability when making a decision. While they display a larger willingness to pay (WTP) for quality than sustainability on the average, it is important to note that sustainability, which has no direct private benefits and was costly for consumers, still generated positive decision weights.

In our experiment, we study the information-gathering process of participants while they make their choices. Our experimental design allowed us to test whether participants fully integrate any additional information regarding the rating values underlying the attributes. We find that if participants receive additional and clear information regarding the ratings, they react to the information and adapt their WTP for the different tiers of the rating scale. While information treatments led to significant changes in behavior in the expected direction (Figure 4), participants fall short of fully adapting their decisions to the new information. This result is surprising, particularly because participants correctly update their beliefs about the underlying ratings for each attribute. These results are consistent with anchoring behavior (Tversky and Kahneman, 1974), which lead to sub-optimal updating in behavior.

Finally, we also analyze the role of attention in the decision process and study if different attentional patterns are linked to preferences towards specpfic ratings. We implement the methods used in (Hirmas et al., 2023) to use individual- and contextual differences in attention to predict heterogeneous patterns in behavior. Our results show that participants are quite heterogeneous in how much attention they allocate to the different attributes, and in turn, these differences are correlated with the weight the attribute plays in the decision. These individual differences in attention are not correlated with any of the demographics we measured (i.e., gender, age, attitudes towards sustainability and connectedness to nature. See Appendix J). We do find a position effect on the attention of the attributes. Namely, if either quality or sustainability are presented in the middle instead of at the bottom, participants will allocate more attention to that attribute. In appendix section M, we show that even when controlling for other moderating effects, such as the position of the attribute, the individual-average index of attention still has a strong and significant descriptive power on the participants decision weights. Therefore, we show that the individual-average attention provides additional insights that other commonly used measures cannot.

We also find that the participants' attention is driven by contextual factors such as the attribute values presented on the specific trial. Our tests show that these effects also significantly predict differences in the decision weights. Namely, if a specific participant allocated more attention to an attribute in a specific trial, this attribute will be weighted more in that specific trial. It is important to note that these effects are milder in scale and significance. Hence, in alignment with previous results (Hirmas et al., 2023), we find stronger individual compared to contextual effects of attention.

Finally, we found that the information treatments do not cause significant changes in attention. In our pre-registration, we aimed to look for alternative ways that attention can help predict heterogeneous effects of the information. In appendix section L, we show results from alternative model specifications that allow attention to moderate the effect of the information treatments. We find that incorporating the changes in attention from before and after do not improve the model fit of our estimations. Moreover, we test whether the individual average attention also moderates the treatment effects (e.g., participants that attend to sustainability more, also react more to the new information). Our results show no significant moderating effects on the treatments either. Since almost all participants perfectly recall the information provided by the treatment conditions, we cannot attribute the lack of changes in attention to misunderstanding or forgetting the new information. Hence, we conclude that although participants are aware of the new information, they do not incorporate it fully into their decisions, which is consistent with anchoring behavior

(Tversky and Kahneman, 1974).

Following prior work on process tracing, we chose the proportion of time spent looking at an attribute as our proxy for attention. The proportion of time spent looking at an attribute is a recommended measure for capturing relative differences in the importance of the attributes Rahal and Fiedler (2019). Other measures commonly used are the absolute dwell-time (i.e., time spent looking at an attribute) or number of fixations on an attribute (i.e., how many times an attribute was looked at). In appendix section N, we estimate our decision model using these different attention measures, finding similar results. Thus, our results are robust to the measure of attention used.

7 Conclusion

In this study, we studied participants preferences for constructed products with real consequences using an experimental design that uses a simulation of online purchasing behavior and decouples the personal benefits to participants from the sustainability of the product. Our findings show that participants consider both quality and sustainability when choosing the desired products. We find that even though participants are unaware of how the ratings reflect on the underlying attributes, they expect them to be linear, but assign a decreasing value to getting products with higher labels. Using the attention data, we find strong heterogeneity in the participants preferences for sustainability (relative to quality).

When we provide additional information regarding the ratings, participants change their preferences accordingly, though in a smaller proportion than what would be optimal for them. This means that given the information that they get, they could act differently and obtain higher benefits (based on their decisions before this information) and achieve better outcomes. This evidence suggests that any amendments or news about a new sustainability rating, such as the adaptations to the recent changes to the EU energy ratings, will not be as effective as the initial implementation of the rating itself. Therefore, if a new sustainability rating is created, we suggest to consider that (1) consumers will likely anchor to their previous behavior and require time to update to new rating systems, and (2) consumer demand can have decreasing marginal returns on purchasing products with higher ratings. While we find that consumers updated their beliefs at the end of the experiment, this update is not fully reflected in their decisions. With the current data we cannot speak how long it will take the consumers to integrate their beliefs into their decisions, which is an interesting avenue for future research. Moreover, based on the two previous recommendations, additional studies on the supply side, that use our results as a benchmark for the demand, can provide interesting insights on how to lead firms into making their products more sustainable.

8 Acknowledgments

We would like to thank 'A Sustainable Future' (ASF) from the University of Amsterdam for funding this project. We would also like to thank Suvi Vaisanen for her work during her internship in this project. Thank you Natalie Lee, Diana Garcia for all the useful comments and help.

References

Allcott, H. and Taubinsky, D. (2015). Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market. *American Economic Review*, 105(8):2501–38.

Alós-Ferrer, C. and Ritschel, A. (2022). Attention and salience in preference reversals. *Experimental Economics*, 25(3):1024–1051.

Andor, M. A., Gerster, A., and Götte, L. (2019). How effective is the European Union energy label? Evidence from a real-stakes experiment. *Environmental Research Letters*, 14(4):044001.

Barahona, N., Otero, C., and Otero, S. (2023). Equilibrium Effects of Food Labeling Policies. *Econometrica*, 91(3):839–868.

Bastounis, A., Buckell, J., Hartmann-Boyce, J., Cook, B., King, S., Potter, C., Bianchi, F., Rayner, M., and Jebb, S. A. (2021). The impact of environmental sustainability labels on willingness-to-pay for foods: A systematic review and meta-analysis of discrete choice experiments. *Nutrients*, 13(8):2677.

Bavard, S. and Palminteri, S. (2022). The functional form of value normalization in human reinforcement learning. Preprint, Cold Spring Harbor Laboratory.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2013). Salience and consumer choice. *Journal of Political Economy*, 121(5):803–843.

Camilleri, A. R., Larrick, R. P., Hossain, S., and Patino-Echeverri, D. (2019). Consumers underestimate the emissions associated with food but are aided by labels. *Nature Climate Change*, 9(1):53–58.

Chegut, A., Eichholtz, P., Holtermans, R., and Palacios, J. (2020). Energy Efficiency Information and Valuation Practices in Rental Housing. *The Journal of Real Estate Finance and Economics*, 60(1-2):181–204.

Chetty, R., Looney, A., and Kroft, K. (2009). Salience and taxation: Theory and evidence. *American Economic Review*, 99(4):1145–77.

Crosetto, P., Muller, L., and Ruffieux, B. (2024). Label or taxes: Why not both? Testing nutritional mixed policies in the lab. Technical report.

De Bauw, M., Matthys, C., Poppe, V., Franssens, S., and Vranken, L. (2021). A combined Nutri-Score and 'Eco-Score'approach for more nutritious and more environmentally friendly food choices? Evidence from a consumer experiment in Belgium. *Food quality and preference*, 93:104276.

Delmas, M. A. and Burbano, V. C. (2011). The Drivers of Greenwashing. *California Management Review*, 54(1):64–87.

ElHaffar, G., Durif, F., and Dubé, L. (2020). Towards closing the attitude-intention-behavior gap in green consumption: A narrative review of the literature and an overview of future research directions. *Journal of Cleaner Production*, 275:122556.

Engel, J. and Szech, N. (2020). A little good is good enough: Ethical consumption, cheap excuses, and moral self-licensing. *PloS one*, 15(1):e0227036.

European Commission, D.-G. f. E. (2021). *Ecodesign Impact Accounting Annual Report 2020: Overview and Status Report*. Publications Office of the European Union, LU.

European Commission, J. R. C., Olivier, J. G. J., Guizzardi, D., Schaaf, E., Solazzo, E., Crippa, M., Vignati, E., Banja, M., Muntean, M., Grassi, G., Monforti-Ferrario, F., and Rossi, S. (2021). *GHG Emissions of All World: 2021 Report*. Publications Office of the European Union, LU.

Fisher, G. (2021). Intertemporal Choices Are Causally Influenced by Fluctuations in Visual Attention. *Management Science*.

Hirmas, A., Engelmann, J. B., and van der Weele, J. (2023). Individual and Contextual Effects of Attention in Risky Choice.

Jenke, L., Bansak, K., Hainmueller, J., and Hangartner, D. (2021). Using Eye-Tracking to Understand Decision-Making in Conjoint Experiments. *Political Analysis*, 29(1):75–101.

Jürkenbeck, K., Mehlhose, C., and Zühlsdorf, A. (2022). The influence of the Nutri-Score on the perceived healthiness of foods labelled with a nutrition claim of sugar. *Plos one*, 17(8):e0272220.

Khazal, A. and Sønstebø, O. J. (2020). Valuation of energy performance certificates in the rental market–Professionals vs. nonprofessionals. *Energy Policy*, 147:111830.

Kolber, A. and Meixner, O. (2023). Effects of Multi-Level Eco-Labels on the Product Evaluation of Meat and Meat Alternatives—A Discrete Choice Experiment. *Foods*, 12(15):2941.

Labandeira, X., Labeaga, J. M., Linares, P., and López-Otero, X. (2020). The impacts of energy efficiency policies: Meta-analysis. *Energy Policy*, 147:111790.

Li, X. and Camerer, C. F. (2022). Predictable effects of visual salience in experimental decisions and games. *The Quarterly Journal of Economics*, 137(3):1849–1900.

Mackowiak, B., Matejka, F., and Wiederholt, M. (2021). Rational inattention: A review. Working Paper 2570, ECB Working Paper.

Majer, J. M., Henscher, H. A., Reuber, P., Fischer-Kreer, D., and Fischer, D. (2022). The effects of visual sustainability labels on consumer perception and behavior: A systematic review of the empirical literature. *Sustainable Production and Consumption*, 33:1–14.

Matějka, F. and McKay, A. (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–98.

Mayer, F. S. and Frantz, C. M. (2004). The connectedness to nature scale: A measure of individuals' feeling in community with nature. *Journal of Environmental Psychology*, 24(4):503–515.

McFadden, D. (1973). *Conditional Logit Analysis of Qualitative Choice Behavior*. Institute of Urban and Regional Development, University of California.

Meißner, M., Oppewal, H., and Huber, J. (2020). Surprising adaptivity to set size changes in multi-attribute repeated choice tasks. *Journal of Business Research*, 111:163–175.

Miller, J. (2023). How Many Participants? How Many Trials? Maximizing the Power of Reaction Time Studies. *Behavior Research Methods*.

Orquin, J. L., Lahm, E. S., and Stojić, H. (2021). The visual environment and attention in decision making. *Psychological Bulletin*, 147(6):597.

Orquin, J. L. and Mueller Loose, S. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica*, 144(1):190–206.

Pace, D. and van der Weele, J. J. (2020). Curbing Carbon: An Experiment on Uncertainty and Information About CO2 Emissions. *SSRN Electronic Journal*.

Potter, C., Pechey, R., Clark, M., Frie, K., Bateman, P. A., Cook, B., Stewart, C., Piernas, C., Lynch, J., and Rayner, M. (2022). Effects of environmental impact labels on the sustainability of food purchases: Two randomised controlled trials in an experimental online supermarket. *PLoS One*, 17(11):e0272800.

Rahal, R.-M. and Fiedler, S. (2019). Understanding cognitive and affective mechanisms in social psychology through eye-tracking. *Journal of Experimental Social Psychology*, 85:103842.

Ritchie, H. and Roser, M. (2023). How much of global greenhouse gas emissions come from food? *Our World in Data*.

Sigurdsson, V., Larsen, N. M., Pálsdóttir, R. G., Folwarczny, M., Menon, R. V., and Fagerstrøm, A. (2022). Increasing the effectiveness of ecological food signaling: Comparing sustainability tags with eco-labels. *Journal of Business Research*, 139:1099–1110.

Simon, H. A. (1971). Designing organizations for an information-rich world. *Computers, communications, and the public interest*, 72:37.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, page 26.

Stadelmann, M. and Schubert, R. (2018). How Do Different Designs of Energy Labels Influence Purchases of Household Appliances? A Field Study in Switzerland. *Ecological Economics*, 144:112–123.

Staples, A. J., Reeling, C. J., Widmar, N. J. O., and Lusk, J. L. (2020). Consumer willingness to pay for sustainability attributes in beer: A choice experiment using eco-labels. *Agribusiness*, 36(4):591–612.

Steiner, B. E., Peschel, A. O., and Grebitus, C. (2017). Multi-product category choices labeled for ecological footprints: Exploring psychographics and evolved psychological biases for characterizing latent consumer classes. *Ecological Economics*, 140:251–264.

Tilling, A. S. (2023). Changing consumption behavior with carbon labels: Causal evidence on behavioral channels and effectiveness. Technical report, Mimeo, University of Bonn.

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157):1124–1131.

Van Loo, E. J., Caputo, V., Nayga Jr, R. M., Seo, H.-S., Zhang, B., and Verbeke, W. (2015). Sustainability labels on coffee: Consumer preferences, willingness-to-pay and visual attention to attributes. *Ecological Economics*, 118:215–225.

Van Loo, E. J., Grebitus, C., and Verbeke, W. (2021). Effects of nutrition and sustainability claims on attention and choice: An eye-tracking study in the context of a choice experiment using granola bar concepts. *Food Quality and Preference*, 90:104100.

Vecchio, R. and Annunziata, A. (2015). Willingness-to-pay for sustainability-labelled chocolate: An experimental auction approach. *Journal of Cleaner Production*, 86:335–342.

Vlaeminck, P., Jiang, T., and Vranken, L. (2014). Food labeling and eco-friendly consumption: Experimental evidence from a Belgian supermarket. *Ecological Economics*, 108:180–190.

Wilcoxon, F., Katti, S. K., and Wilcox, R. A. (1970). Critical values and probability levels for the Wilcoxon rank sum test and the Wilcoxon signed rank test. *Selected tables in mathematical statistics*, 1:171–259.

Willemsen, M. C. and Johnson, E. J. (2011). Visiting the decision factory: Observing cognition with MouselabWEB and other information acquisition methods. *A handbook of process tracing methods for decision research*, pages 21–42.

Yokessa, M. and Marette, S. (2019). A Review of Eco-labels and their Economic Impact. *International Review of Environmental and Resource Economics*, 13(1-2):119–163.

Appendix

A Exclusion Criteria

We recruited 328 participants from the online platform Prolific. As preregistered, we excluded from analysis any participant that did not pass the attention checks. The attention checks were three questions within the questionnaire, where we prompted participants to give a specific answer. Participants with two or more mistakes were excluded from the analysis and exempt from payment (N=1). Additionally, we check how long participants spent on each page during the instructions. We also excluded participants that skipped the instructions and tried to force the answers for the comprehension questions. When participants signed the informed consent, we told them that they needed to keep the experiment on Full-screen and that they should not change tabs to go to other pages during the length of the experiment. We excluded from the analysis all participants that spent more than 30 seconds on other pages or that would change tabs more than three times . Since we cannot verify that participants are actually paying attention to the task (e.g., they could still check their phones, N=3), we excluded any participant that had an average response three standard deviations above the sample average. We excluded from the analyses all participants that skipped the instructions and answered the control questions by trial-and-error (i.e. tried all possible combinations until they got them all right, N=4), In total, 12 participants were excluded due to the aforementioned criteria.

In our experiment, participants were asked to state their beliefs about the underlying values behind the quality and sustainability ratings. In the instructions, we explicitly explained to them that a higher rating meant higher underlying values for the represented attribute. We excluded any participant that showed decreasing beliefs about any of the ratings (i.e. beliefs that a higher rating yields lower values, N=18). Additionally, for some participants the system did not constrain the range of values that participants could state for the beliefs. This was due to software compability (problems with the browser). We excluded all participants whose beliefs were out of range (N=8). In total, 26 participants were excluded due decreasing or out-of-range beliefs.

The final sample consists on 290 participants. In order to see that our results were not driven due to the exclusion of these participants, we ran the analysis with the whole sample. We do not report these results, as they yield very similar results. All our scripts and databases for the analysis are available here.

B Locations of planted trees

Location of planted trees

The map above shows the amount of trees planted in different locations. Participants could either choose to plant the trees in a specific location, or they could choose the default option of planting them anywhere. The organisation 'One Tree planted' offers an option to plant trees where they are needed the most at the moment. We used this option for all participants that chose the default of planting the trees "anywhere". The amount of trees planted in our two studies adds up to 1.153 trees across the world. We thank Diana Garcia for providing us with this graph.

C Pseudo-randomization of attribute values

We ensure that participants observe sufficient trials for every relevant combination of sustainability ratings. Specifically, every product pair will offer choice options that differ in sustainability ratings (i.e., 1 vs. 3 leaves, 2 vs. 3 leaves and 1 vs. 2 leaves) across multiple trials. Sustainability pairs will be offered with varying combinations of price and quality. Across the three sustainability pairs, however, the presented price-quality combinations will be the same for each participant. We show which type of combinations below. To describe these price-quality combinations, we refer to the products as the sustainable option (one with higher sustainability rating) and the competing option. Each sustainability pair will be presented in combination with the following sets of prices and quality ratings:

1. The price and quality of the competing product are higher.

2. The price of the sustainable product is higher and the quality of the competing product is higher.

3. The sustainable product has higher price and quality

4. Both products have the same price, but the competing product has higher quality

5. Both products have the same quality, but the competing product is more expensive.

For comparing quality and sustainability ratings, we also add two trials where both products have equal levels of sustainability and the price and quality of one of the products is higher. These trials will be shown to participants in randomized order and the values of price and quality are randomized within the aforementioned restrictions.

D Belief elicitation

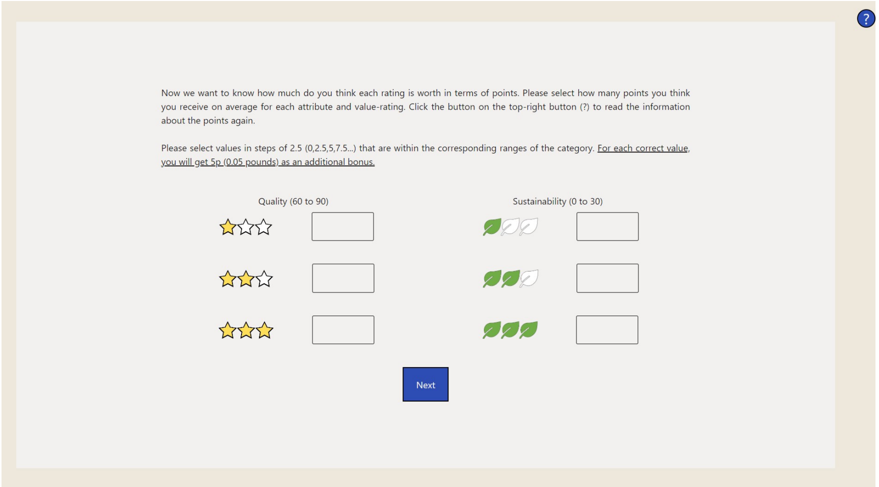


Figure A1: Belief elicitation screen

The Figure shows the decision screen for eliciting beliefs about the ratings. After each block of decisions, participants were asked to state their beliefs about the underlying values for each rating level. Participants were reminded of the ranges for each attribute, but also could revisit the general information about each attribute by pressing the '?' button on the top-right of the screen. They were informed that they would receive 0.1 pounds per correct belief. To make participants' decision easier and increase the chance that the beliefs were correct, we restricted the participants responses in steps of 2.5 points (e.g., 0, 2.5, 5,...). The attribute presented on the left is the attribute that is presented above in the purchasing decisions screen.

E Design of previous study

In a previous version of this study, we ran a similar experiment analyzing the role of ratings on the decision process. In that experiment, we aimed to see whether different underlying information regarding sustainability ratings will influence choice. In that study, we had four between-subject treatments pertaining the information regarding sustainability. This experiment is identical to the one presented in section 3.2, with the exception of the information presented in the middle of the experiment.

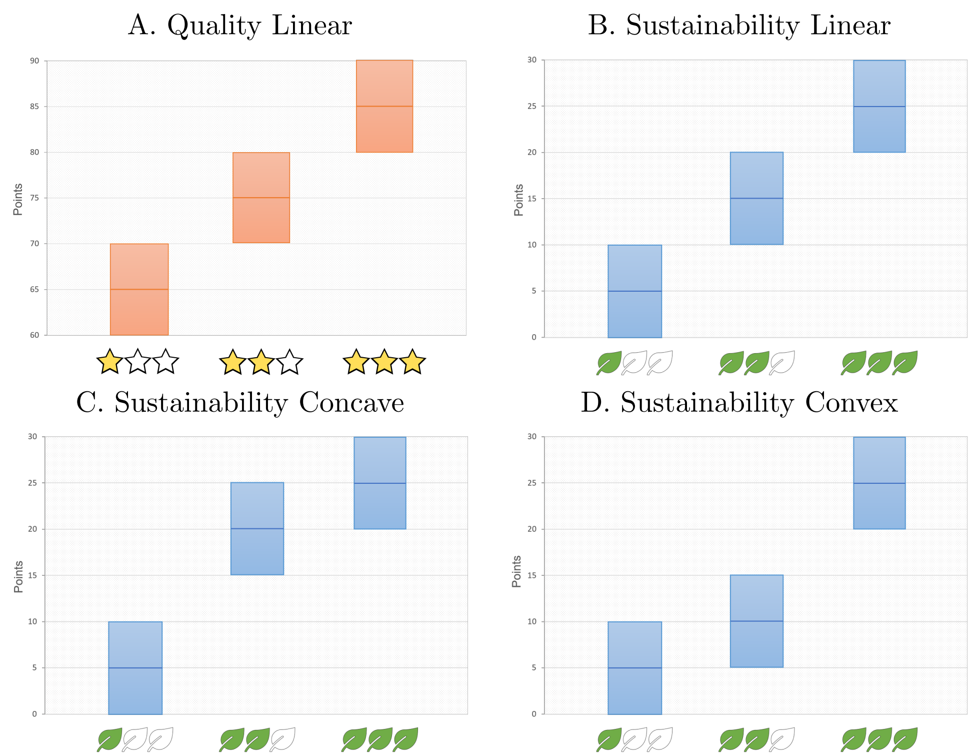


Figure A2: Information treatments for previous study

The figure above describes the possible information provided to participants. The different treatments were (1) a no-information treatment, where no information was provided in the middle of the experiment, (2) a linear treatment, where participants observed the information presented in panels A and B, and then (3) and (4) where instead of panel B, participants observed panels C and D respectively.

Another relevant difference of the information provided here compared to the experiment in our main task is that the underlying values driving the ratings had a range of possible values instead of being just one specific value. Moreover, if we compare the information contained in panels, B, C and D, we can see that the difference across treatments is much more nuanced compared to our treatments in section 3.2.

When analyzing our confirmation checks, we found that 21.6% of participants could not identify the treatment they were in (compared to 98.6% in our current experiment). Thus, we concluded that the information treatments in that experiment were too mild for people to internalize.

F Raw choice data

In this section, we inspect the raw choice data to test whether in the context of our experiment, in which personal benefits are decoupled from sustainability, participants respond to the sustainability attribute. To this end, we first analyze the aggregate behavior of the participants regarding their purchasing choices, focusing on decisions made before the information treatment to avoid distortion by the enhanced salience of the attributes post treatment. Namely, we test the impact of the different attributes on the decisions. Table A1 summarizes the purchasing decisions of participants based on these three attributes before the information treatments. Panel A shows the percentage of choices favoring the most sustainable product based on the sustainability ratings. Panel B shows the percentage of choices where the chosen product had the highest quality rating. Finally, the third panel presents the percentage of decisions where the chosen price had the lowest price.

Table A1: Percentage of choices given attribute differences

Panel A:		Difference in Sustainability			
		2vs1	3vs2	3vs1	Total
Highest Sustainability chosen		49.7%	45.7%*	65.4%***	53.6%***

Panel B:		Difference in Quality			
		2vs1	3vs2	3vs1	Total
Highest Quality chosen		65.7%***	54.6%*	82.7%***	67.8%***

Panel C:	Difference in Price				
	0.5	1.0	1.5	2.0	Total
Lowest Price chosen	40.9%***	51.8%	60.1%***	63.1%***	51.2%

The estimated percentages above are calculated based on the decisions before the information treatments. For each panel, observations where both products have the same level of the relevant attribute are omitted. Significance level for rejecting null-hypothesis that decisions are made by chance (i.e. proportion equal to 50%) with Bonferroni correction for multiple hypotheses ($\alpha/3$).

* $p < 0.05$; ** $p < 0.01$, *** $p < 0.001$

The overall results in Table A1 show that participants are sensitive to each attribute, including the sustainability attribute. If participants were indifferent between choice options, we would expect these proportions to be around 50%. Moreover, if participants were purely self-interested, we would expect them to respond to quality and price attributes only. We find that participants are more likely to choose the higher quality product when the product difference is higher (3 vs 1) and they are more likely to choose the cheapest product when the price difference is larger than 1. Importantly, participants are more likely to make sustainable decisions when the difference between the products' sustainability is large (3 vs 1).

So far, we have determined how participants use the quality and sustainability ratings in the absence of further information regarding the underlying value of each attribute rating. In this section, we investigate how the decision process changes when participants are presented with new information regarding the attribute values. To this end, we first analyze the raw data regarding the purchasing choices and compare the different treatments with the average behavior before the information.

Table A2 provides an overview of the changes in behavior after the information treatments. In Panel A, we examine the shift in the proportion of sustainable choices after the different information treatments are provided (compared with the behavior before). The proportions are conditional on the comparison of sustainability levels between the two products (columns). In the Sustainability convex (SC) treatment, following the expected direction, participants chose the sustainable option less frequently when comparing between the lowest and the middle tier (p=0.031). Conversely, they chose more the sustainable option when comparing between the middle and the highest tier (p=0.014). As expected, there are no large differences when comparing the bottom to the highest tier (p=0.969).

A. Sustainable choices			
	2vs1	3vs2	3vs1
Before	0.497	0.457	0.654
Linear	0.52	0.468	0.639
SC	0.442	0.521*	0.655
QC	0.535	0.489	0.68

B. Highest quality choices		
2vs1	3vs2	3vs1
0.658	0.546	0.827
0.603	0.565	0.826
0.722*	0.574	0.765*
0.592*	0.618*	0.808

C. Lowest price choices			
0.5	1	1.5	2
0.409	0.518	0.601	0.631
0.437	0.509	0.581	0.559
0.357	0.479	0.601	0.616
0.321**	0.481	0.575	0.627

Table A2: Proportion of choices before and after treatments depending on attribute values.

The tables above show the proportion of sustainable (Panel A), highest quality (Panel B) and lowest price choices (Panel C) for each treatment condition (Linear, SC and QC) after the information appears. As a reference, the first row of each table presents the proportion of choices before all condition treatments. The columns describe the relevant attribute's comparison (Sustainability for A, Quality for B and Price for C). Significance level for rejecting the null-hypothesis that the proportion of choices is equal to the proportion before the information treatment (p values are adjusted using a Bonferroni correction for multiple hypotheses).
* $p < 0.05/3$; ** $p < 0.01/3$, *** $p < 0.001/3$
SC: Sustainability convex, QC: Quality convex

In the sustainability convex condition (SC), the marginal benefit between choosing a product from the bottom versus the middle tier of sustainability is close to zero. Hence, we expected even a stronger reaction from the treatment. On the other hand, We find no significant differences in the proportion of sustainable choices between the linear and quality convex treatment (QC). This result was expected since the information regarding the sustainability ratings is equivalent in both treatment conditions.

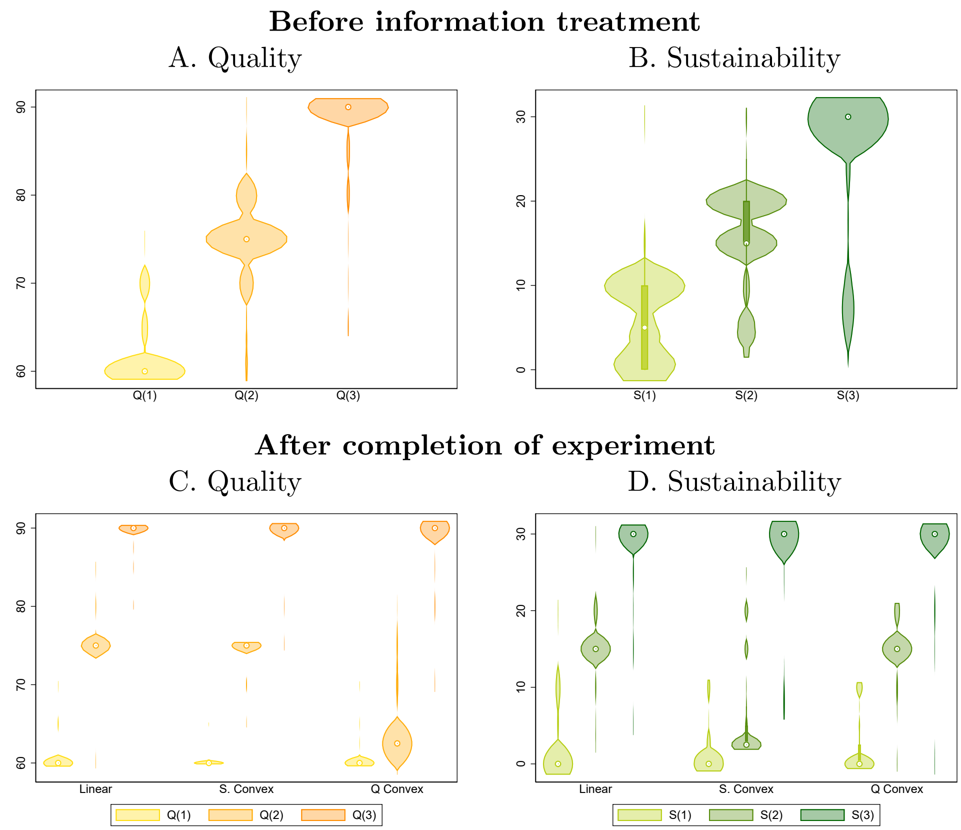
Panel B shows the proportion of choices, where the highest quality product was chosen, conditional on the treatments and levels of quality. We find an analogous effect regarding quality-driven choices. When comparing products from the lowest- and middle-quality tiers in the QC treatment, participants' choices for the highest-quality product were less than before the information treatment (p=0.016). When comparing middle to high tier, participants chose more the highest quality option than before (p=0.01).

Additionally, we find a somewhat striking result. Participants are more likely to choose the highest quality option in the SC treatment when comparing both low- to middle-tier products (p=0.011) and low-to-high (p=0.004). This effect is most likely driven by the fact that the intermediate sustainable option becomes less appealing due to the SC condition.

In Panel C, we can see the proportion of choices where the lowest price was chosen depending on the price difference. We find a decrease in the choices of the lowest price when the price difference is 0.5 (SC: p=0.044; QC: p=0.001). These effects suggest that participants become more sensitive to the labels (relative to the price) after the information treatments.

G Beliefs about ratings

The figures below show the distribution of the participants' beliefs about the quality and sustainability ratings before the information treatments (panels A and B), and the beliefs at the end of the experiment for each information treatment (panels C and D). The violin plots in panel A show the distribution of beliefs for sustainability for each tier of the rating. Panel B displays the beliefs about each tier of the quality ratings.



As we can see, the participants' beliefs about the underlying values driving both ratings are fairly linear. In the case of sustainability, the beliefs seem to be bi-modal, where some participants expected the ratings to cover the whole range (0 to 30 points), while others expected it to start from 10 sustainability points.

It is clear from panels C and D, that participants understand the treatments quite clearly. This can be seen by comparing the beliefs regarding the middle tier (Q(2) and S(2)) for the different treatments. Most participants perfectly recall the values of the information treatments at the end of the experiment.

H Transitivity of decisions

The table below present three representations of the decision model using the data before the information treatment. In column (1), we use the difference in ratings as a linear function. Column (2) allows for all possible non-linear combinations of parameters. Finally, column (3) incorporates non-linearities on the ratings, but assumes that the preferences are transitives (i.e. value of going from bottom to top is equal to the sum of the values going from bottom to middle and from middle to top).

	(1) Linear		(2) Unconstrained		(3) Constrained	
ΔQ	1.485***	(0.069)				
ΔQ^{21}			1.732***	(0.106)	1.754***	(0.096)
ΔQ^{32}			1.218***	(0.100)	1.235***	(0.094)
ΔQ^{31}			3.031***	(0.166)		
ΔS	1.116***	(0.059)				
ΔS^{21}			1.266***	(0.087)	1.233***	(0.070)
ΔS^{32}			1.033***	(0.086)	1.001***	(0.068)
ΔS^{31}			2.204***	(0.111)		
ΔP	-1.083***	(0.072)	-1.097***	(0.072)	-1.093***	(0.072)
Constant	0.027	(0.034)	0.023	(0.034)	0.024	(0.034)
var(_cons[id])	0.000	(0.000)	0.000	(0.000)	0.000*	(0.000)
Observations	4929		4929		4929	
AIC	4936.629		4906.754		4903.967	
BIC	4962.641		4958.777		4942.984	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

When comparing the models by their Bayesian information criteria (BIC), the best fitting model is the constrained non-linear model presented in column (3). Moreover, we use the models in column (2) and (3) to test our assumptions. The results of the different linear tests are provided below:

Hypothesis	Model	Test	χ^2	p-value
Transitiveness (Q)	(2)	$\Delta Q^{21} + \Delta Q^{32} = \Delta Q^{31}$	0.22	0.6399
Transitiveness (S)	(2)	$\Delta S^{21} + \Delta S^{32} = \Delta S^{31}$	0.81	0.3692
Non-linearity (Q)	(3)	$\Delta Q^{21} = \Delta Q^{32}$	16.25	0.0001
Non-linearity (S)	(3)	$\Delta S^{21} = \Delta S^{32}$	10.24	0.0014

I Estimation tables of decision model

	(1) Before Treatment		(2) All Data		(3) All Data + Ind. Attention	
idec						
ΔP	-1.093***	(-15.14)	-0.979***	(-7.95)	-1.141***	(-8.91)
After × ΔP			0.0963	(1.24)	0.106	(1.26)
SC × ΔP			-0.282	(-1.54)	-0.168	(-0.87)
QC × ΔP			-0.125	(-0.72)	-0.0468	(-0.25)
After × SC × ΔP			0.159	(1.17)	0.166	(1.14)
After × QC × ΔP			0.0815	(0.59)	0.0867	(0.58)
$\Delta P \times \bar{a}(P)$					-0.324***	(-5.65)
ΔQ^{21}	1.754***	(18.33)	1.690***	(9.67)	1.766***	(9.82)
After × ΔQ^{21}			-0.209	(-1.43)	-0.228	(-1.45)
SC × ΔQ^{21}			0.298	(1.20)	0.361	(1.42)
QC × ΔQ^{21}			-0.0189	(-0.08)	0.0367	(0.15)
After × SC × ΔQ^{21}			0.00482	(0.02)	0.0222	(0.10)
After × QC × ΔQ^{21}			-0.218	(-0.89)	-0.227	(-0.87)
$\Delta Q^{21} \times \bar{a}(Q)$					-0.0429	(-0.54)
ΔQ^{32}	1.235***	(13.20)	1.254***	(7.58)	1.391***	(8.38)
After × ΔQ^{32}			-0.115	(-0.92)	-0.123	(-0.93)
SC × ΔQ^{32}			0.108	(0.47)	-0.0260	(-0.11)
QC × ΔQ^{32}			-0.0872	(-0.37)	-0.129	(-0.54)
After × SC × ΔQ^{32}			-0.0957	(-0.50)	-0.103	(-0.51)
After × QC × ΔQ^{32}			0.696**	(2.79)	0.749**	(2.80)
$\Delta Q^{32} \times \bar{a}(Q)$					0.182*	(2.23)
ΔS^{21}	1.233***	(17.65)	1.114***	(9.00)	1.166***	(9.73)
After × ΔS^{21}			-0.00613	(-0.06)	-0.00658	(-0.06)
SC × ΔS^{21}			0.220	(1.29)	0.307	(1.82)
QC × ΔS^{21}			0.224	(1.27)	0.231	(1.31)
After × SC × ΔS^{21}			-0.447**	(-2.71)	-0.464**	(-2.69)
After × QC × ΔS^{21}			0.0432	(0.27)	0.0442	(0.26)
$\Delta S^{21} \times \bar{a}(S)$					0.373***	(6.58)
ΔS^{32}	1.001***	(14.64)	0.903***	(7.36)	0.951***	(7.72)
After × ΔS^{32}			-0.0873	(-0.79)	-0.0920	(-0.79)
SC × ΔS^{32}			0.161	(0.94)	0.235	(1.38)
QC × ΔS^{32}			0.195	(1.15)	0.188	(1.09)
After × SC × ΔS^{32}			0.369*	(2.01)	0.392*	(2.03)
After × QC × ΔS^{32}			0.0797	(0.46)	0.0819	(0.44)
$\Delta S^{32} \times \bar{a}(S)$					0.356***	(6.51)
After			-0.0377	(-0.62)	-0.0402	(-0.62)
SC			-0.0140	(-0.17)	0.00803	(0.09)
QC			0.0236	(0.28)	0.0488	(0.57)
After × SC			-0.0246	(-0.27)	-0.0248	(-0.25)
After × QC			0.129	(1.39)	0.140	(1.41)
$\bar{a}(P)$					-0.0334	(-0.95)
$\bar{a}(Q)$					-0.0603	(-1.95)
$\bar{a}(S)$					0	(.)
Constant	0.0236	(0.69)	0.0233	(0.41)	0.0284	(0.48)
Ind. Random Effects	2.33e-34*	(2.28)	0.0822***	(3.64)	0.0721**	(3.26)
Observations	4929		9851		9851	
AIC	4904.0		9950.4		9399.5	
BIC	4943.0		10216.6		9716.1	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

J Treatment effects on attention

In order to estimate which factors affect attention, we estimate a linear regression with random effects for the attention variables of each attribute. The results are presented in appendix table J. We separate the factors into three categories: (1) individual factors, which represent characteristics that are fixed throughout the experiment, (2) contextual factors, which include any element that changes over the different trials and (3) treatment effects. The dependent variable is the standardized measure of attention.

The results show that there is high variability in attention at an individual level, captured by the random effects, but these differences are not correlated with any of the individual measures we elicited from participants, such as gender, age, sustainability attitudes and connectedness to nature. The position of the attributes Quality and Sustainability was randomized at an individual level. Here we see that there is a significant effect of the position on the attention for quality and sustainability.

	(1) Price		(2) Quality		(3) Sustainability	
Individual Factors						
Sustainability First	0.008	(0.062)	-0.657***	(0.056)	0.653***	(0.059)
Female	0.017	(0.062)	0.019	(0.057)	-0.000	(0.060)
Age	-0.001	(0.031)	-0.000	(0.028)	0.011	(0.030)
Sust. Attitude.	-0.023	(0.034)	-0.028	(0.031)	0.046	(0.033)
CNS	-0.016	(0.035)	-0.009	(0.032)	0.030	(0.033)
Contextual Factors						
S(2vs1)	-0.141***	(0.034)	-0.029	(0.033)		
S(3vs2)	-0.096**	(0.034)	-0.070*	(0.033)	-0.021	(0.022)
S(3vs1)	-0.085*	(0.034)	-0.058	(0.033)	-0.043*	(0.022)
Q(2vs1)	-0.145***	(0.029)			-0.099***	(0.027)
Q(3vs2)	-0.090**	(0.029)	-0.018	(0.026)	-0.142***	(0.027)
Q(3vs1)	-0.058*	(0.029)	-0.015	(0.027)	-0.157***	(0.027)
$\|\Delta P\|$	-0.002	(0.021)	-0.095***	(0.015)	-0.096***	(0.015)
Treatment Effects						
After	-0.082*	(0.034)	0.006	(0.033)	0.015	(0.030)
SC	0.107	(0.078)	0.046	(0.072)	-0.181*	(0.075)
QC	0.102	(0.081)	-0.087	(0.074)	-0.038	(0.077)
After × SC	-0.005	(0.047)	0.010	(0.046)	0.035	(0.043)
After × QC	-0.103*	(0.049)	0.115*	(0.047)	0.047	(0.044)
Constant	0.222*	(0.113)	1.131***	(0.103)	-0.673***	(0.106)
Random Effects						
Individual	-0.707***	(0.046)	-0.802***	(0.047)	-0.737***	(0.046)
Constant	-0.141***	(0.008)	-0.173***	(0.008)	-0.199***	(0.008)
Observations	7848		7843		8694	
AIC	20750.426		20205.350		21952.843	
BIC	20889.786		20337.730		22087.181	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We find that the attention to price and sustainability are affected by the attribute values presented in the trial. For quality, the attention is mostly correlated with the differences in price between products. When comparing by treatment effects, we find that quality is attended more after the Quality convex information treatment.

K Models with attention

	(1)		(2)		(3)	
idec						
ΔP	-0.979***	(-7.95)	-1.141***	(-8.91)	-1.143***	(-8.88)
After $\times$ ΔP	0.0963	(1.24)	0.106	(1.26)	0.0958	(1.15)
SC $\times$ ΔP	-0.282	(-1.54)	-0.168	(-0.87)	-0.165	(-0.85)
QC $\times$ ΔP	-0.125	(-0.72)	-0.0468	(-0.25)	-0.0353	(-0.19)
After $\times$ SC $\times$ ΔP	0.159	(1.17)	0.166	(1.14)	0.178	(1.24)
After $\times$ QC $\times$ ΔP	0.0815	(0.59)	0.0867	(0.58)	0.0815	(0.54)
$\Delta P \times \bar{a}(P)$			-0.324***	(-5.65)	-0.324***	(-5.58)
$\Delta P \times \tilde{a}(P)$					-0.0409	(-1.50)
ΔQ^{21}	1.690***	(9.67)	1.766***	(9.82)	1.769***	(9.72)
After $\times$ ΔQ^{21}	-0.209	(-1.43)	-0.228	(-1.45)	-0.229	(-1.45)
SC $\times$ ΔQ^{21}	0.298	(1.20)	0.361	(1.42)	0.354	(1.39)
QC $\times$ ΔQ^{21}	-0.0189	(-0.08)	0.0367	(0.15)	0.0352	(0.15)
After $\times$ SC $\times$ ΔQ^{21}	0.00482	(0.02)	0.0222	(0.10)	0.0163	(0.07)
After $\times$ QC $\times$ ΔQ^{21}	-0.218	(-0.89)	-0.227	(-0.87)	-0.248	(-0.96)
$\Delta Q^{21} \times \bar{a}(Q)$			-0.0429	(-0.54)	-0.0411	(-0.52)
$\Delta Q^{21} \times \tilde{a}(Q)$					0.120**	(2.59)
ΔQ^{32}	1.254***	(7.58)	1.391***	(8.38)	1.401***	(8.40)
After $\times$ ΔQ^{32}	-0.115	(-0.92)	-0.123	(-0.93)	-0.121	(-0.92)
SC $\times$ ΔQ^{32}	0.108	(0.47)	-0.0260	(-0.11)	-0.0214	(-0.09)
QC $\times$ ΔQ^{32}	-0.0872	(-0.37)	-0.129	(-0.54)	-0.146	(-0.60)
After $\times$ SC $\times$ ΔQ^{32}	-0.0957	(-0.50)	-0.103	(-0.51)	-0.114	(-0.58)
After $\times$ QC $\times$ ΔQ^{32}	0.696**	(2.79)	0.749**	(2.80)	0.755**	(2.82)
$\Delta Q^{32} \times \bar{a}(Q)$			0.182*	(2.23)	0.181*	(2.20)
$\Delta Q^{32} \times \tilde{a}(Q)$					0.0101	(0.22)
ΔS^{21}	1.114***	(9.00)	1.166***	(9.73)	1.180***	(9.92)
After $\times$ ΔS^{21}	-0.00613	(-0.06)	-0.00658	(-0.06)	-0.00506	(-0.05)
SC $\times$ ΔS^{21}	0.220	(1.29)	0.307	(1.82)	0.306	(1.83)
QC $\times$ ΔS^{21}	0.224	(1.27)	0.231	(1.31)	0.221	(1.26)
After $\times$ SC $\times$ ΔS^{21}	-0.447**	(-2.71)	-0.464**	(-2.69)	-0.471**	(-2.75)
After $\times$ QC $\times$ ΔS^{21}	0.0432	(0.27)	0.0442	(0.26)	0.0400	(0.23)
$\Delta S^{21} \times \bar{a}(S)$			0.373***	(6.58)	0.373***	(6.55)
$\Delta S^{21} \times \tilde{a}(S)$					0.0184	(0.43)
ΔS^{32}	0.903***	(7.36)	0.951***	(7.72)	0.954***	(7.71)
After $\times$ ΔS^{32}	-0.0873	(-0.79)	-0.0920	(-0.79)	-0.0867	(-0.73)
SC $\times$ ΔS^{32}	0.161	(0.94)	0.235	(1.38)	0.239	(1.39)
QC $\times$ ΔS^{32}	0.195	(1.15)	0.188	(1.09)	0.180	(1.05)
After $\times$ SC $\times$ ΔS^{32}	0.369*	(2.01)	0.392*	(2.03)	0.384*	(1.97)
After $\times$ QC $\times$ ΔS^{32}	0.0797	(0.46)	0.0819	(0.44)	0.0787	(0.42)
$\Delta S^{32} \times \bar{a}(S)$			0.356***	(6.51)	0.362***	(6.60)
$\Delta S^{32} \times \tilde{a}(S)$					0.0892	(1.91)
After	-0.0377	(-0.62)	-0.0402	(-0.62)	-0.0450	(-0.68)
SC	-0.0140	(-0.17)	0.00803	(0.09)	0.00712	(0.08)
QC	0.0236	(0.28)	0.0488	(0.57)	0.0464	(0.54)
After $\times$ SC	-0.0246	(-0.27)	-0.0248	(-0.25)	-0.0172	(-0.17)
After $\times$ QC	0.129	(1.39)	0.140	(1.41)	0.149	(1.48)
$\bar{a}(P)$			-0.0334	(-0.95)	-0.0346	(-1.01)
$\bar{a}(Q)$			-0.0603	(-1.95)	-0.0565	(-1.80)
$\bar{a}(S)$			0	(.)	0	(.)
$\tilde{a}(P)$					0.0536	(1.64)
$\tilde{a}(Q)$					0.0515	(1.58)
$\tilde{a}(S)$					0	(.)
Constant	0.0233	(0.41)	0.0284	(0.48)	0.0305	(0.51)
/						
var(Constant[id])	0.0822***	(3.64)	0.0721**	(3.26)	0.0708**	(3.16)
Observations	9851		9851		9851	
AIC	9950.4		9399.5		9381.8	
BIC	10216.6		9716.1		9748.8	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

L Alternative models with attention

	(1)		(2)		(3)	
idec						
ΔP	-0.979***	(-7.95)	-1.141***	(-8.91)	-1.146***	(-8.44)
After × ΔP	0.0963	(1.24)	0.106	(1.26)	0.0640	(0.64)
SC × ΔP	-0.282	(-1.54)	-0.168	(-0.87)	-0.150	(-0.77)
QC × ΔP	-0.125	(-0.72)	-0.0468	(-0.25)	-0.0564	(-0.30)
After × SC × ΔP	0.159	(1.17)	0.166	(1.14)	0.209	(1.33)
After × QC × ΔP	0.0815	(0.59)	0.0867	(0.58)	0.123	(0.74)
$\Delta P \times \bar{a}(P)$			-0.324***	(-5.65)	-0.304*	(-2.18)
After × SC × $\Delta P \times \bar{a}(P)$					0.0814	(0.46)
After × QC × $\Delta P \times \bar{a}(P)$					-0.154	(-0.96)
ΔQ^{21}	1.690***	(9.67)	1.766***	(9.82)	1.779***	(10.15)
After × ΔQ^{21}	-0.209	(-1.43)	-0.228	(-1.45)	-0.211	(-1.32)
SC × ΔQ^{21}	0.298	(1.20)	0.361	(1.42)	0.343	(1.38)
QC × ΔQ^{21}	-0.0189	(-0.08)	0.0367	(0.15)	0.0891	(0.38)
After × SC × ΔQ^{21}	0.00482	(0.02)	0.0222	(0.10)	-0.00734	(-0.03)
After × QC × ΔQ^{21}	-0.218	(-0.89)	-0.227	(-0.87)	-0.308	(-1.13)
$\Delta Q^{21} \times \bar{a}(Q)$			-0.0429	(-0.54)	0.238	(1.44)
After × SC × $\Delta Q^{21} \times \bar{a}(Q)$					-0.428	(-1.73)
After × QC × $\Delta Q^{21} \times \bar{a}(Q)$					-0.435	(-1.83)
ΔQ^{32}	1.254***	(7.58)	1.391***	(8.38)	1.411***	(8.36)
After × ΔQ^{32}	-0.115	(-0.92)	-0.123	(-0.93)	-0.118	(-0.88)
SC × ΔQ^{32}	0.108	(0.47)	-0.0260	(-0.11)	-0.0450	(-0.19)
QC × ΔQ^{32}	-0.0872	(-0.37)	-0.129	(-0.54)	-0.127	(-0.52)
After × SC × ΔQ^{32}	-0.0957	(-0.50)	-0.103	(-0.51)	-0.121	(-0.60)
After × QC × ΔQ^{32}	0.696**	(2.79)	0.749**	(2.80)	0.775**	(2.78)
$\Delta Q^{32} \times \bar{a}(Q)$			0.182*	(2.23)	0.0825	(0.46)
After × SC × $\Delta Q^{32} \times \bar{a}(Q)$					0.0444	(0.19)
After × QC × $\Delta Q^{32} \times \bar{a}(Q)$					0.233	(1.00)
ΔS^{21}	1.114***	(9.00)	1.166***	(9.73)	1.167***	(9.48)
After × ΔS^{21}	-0.00613	(-0.06)	-0.00658	(-0.06)	0.00865	(0.08)
SC × ΔS^{21}	0.220	(1.29)	0.307	(1.82)	0.299	(1.74)
QC × ΔS^{21}	0.224	(1.27)	0.231	(1.31)	0.260	(1.46)
After × SC × ΔS^{21}	-0.447**	(-2.71)	-0.464**	(-2.69)	-0.506**	(-2.83)
After × QC × ΔS^{21}	0.0432	(0.27)	0.0442	(0.26)	0.0232	(0.13)
$\Delta S^{21} \times \bar{a}(S)$			0.373***	(6.58)	0.451***	(4.04)
After × SC × $\Delta S^{21} \times \bar{a}(S)$					-0.205	(-1.30)
After × QC × $\Delta S^{21} \times \bar{a}(S)$					-0.0693	(-0.46)
ΔS^{32}	0.903***	(7.36)	0.951***	(7.72)	0.977***	(7.92)
After × ΔS^{32}	-0.0873	(-0.79)	-0.0920	(-0.79)	-0.0908	(-0.78)
SC × ΔS^{32}	0.161	(0.94)	0.235	(1.38)	0.207	(1.20)
QC × ΔS^{32}	0.195	(1.15)	0.188	(1.09)	0.164	(0.94)
After × SC × ΔS^{32}	0.369*	(2.01)	0.392*	(2.03)	0.389	(1.90)
After × QC × ΔS^{32}	0.0797	(0.46)	0.0819	(0.44)	0.106	(0.56)
$\Delta S^{32} \times \bar{a}(S)$			0.356***	(6.51)	0.229	(1.94)
After × SC × $\Delta S^{32} \times \bar{a}(S)$					0.221	(1.23)
After × QC × $\Delta S^{32} \times \bar{a}(S)$					0.0818	(0.51)
After	-0.0377	(-0.62)	-0.0402	(-0.62)	-0.0417	(-0.64)
SC	-0.0140	(-0.17)	0.00803	(0.09)	0.00912	(0.10)
QC	0.0236	(0.28)	0.0488	(0.57)	0.0478	(0.55)
After × SC	-0.0246	(-0.27)	-0.0248	(-0.25)	-0.0231	(-0.24)
After × QC	0.129	(1.39)	0.140	(1.41)	0.150	(1.49)
$\bar{a}(P)$			-0.0334	(-0.95)	-0.0346	(-0.97)
$\bar{a}(Q)$			-0.0603	(-1.95)	-0.0627*	(-1.96)
$\bar{a}(S)$			0	(.)	0	(.)
Constant	0.0233	(0.41)	0.0284	(0.48)	0.0294	(0.49)
/						
var(Constant[id])	0.0822***	(3.64)	0.0721**	(3.26)	0.0714***	(3.31)
Observations	9851		9851		9851	
AIC	9950.4		9399.5		9405.6	
BIC	10216.6		9716.1		9902.1	

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

M Effects of attention after controlling for other moderators

The table below presents the moderating effects of attention on the decision weights after controlling for other individual factors that could be measured. In this table we only present the moderating effect of the individual factors on the decision weights. Column (1) presents the full decision model presented in column (3) from appendix table I. The other columns present the introduction of the position of attributes, demographics and the sustainability scales. Although some of these factors predict differences in the decision weights, the correlation between the individual-average index of attention and the decision weights remains fairly similar.

	(1)		(2)		(3)		(4)		(5)	
ΔP	-1.141***	(0.128)	-1.239***	(0.143)	-1.378***	(0.154)	-1.266***	(0.146)	-1.230***	(0.146)
$\Delta P \times \bar{a}(P)$	-0.324***	(0.057)	-0.285***	(0.060)	-0.292***	(0.062)	-0.285***	(0.063)	-0.289***	(0.060)
S. First $\times \Delta P$			0.198	(0.129)	0.176	(0.126)	0.205	(0.130)	0.197	(0.131)
$\Delta P \times$ Female					0.339**	(0.128)				
$\Delta P \times$ Age					0.019	(0.066)				
$\Delta P \times$ CNS							0.170*	(0.066)		
$\Delta P \times$ Sust. Attitude.									0.139*	(0.064)
ΔQ^{21}	1.766***	(0.180)	1.632***	(0.194)	1.713***	(0.199)	1.628***	(0.192)	1.615***	(0.192)
$\Delta Q^{21} \times \bar{a}(Q)$	-0.043	(0.079)	0.157	(0.098)	0.164	(0.099)	0.169	(0.094)	0.158	(0.095)
S. First $\times \Delta Q^{21}$			0.335	(0.203)	0.347	(0.205)	0.341	(0.200)	0.324	(0.197)
$\Delta Q^{21} \times$ Female					-0.150	(0.173)				
$\Delta Q^{21} \times$ Age					-0.027	(0.074)				
$\Delta Q^{21} \times$ CNS							-0.204*	(0.090)		
$\Delta Q^{21} \times$ Sust. Attitude.									-0.235**	(0.079)
ΔQ^{32}	1.391***	(0.166)	1.309***	(0.178)	1.351***	(0.209)	1.334***	(0.181)	1.309***	(0.174)
$\Delta Q^{32} \times \bar{a}(Q)$	0.182*	(0.082)	0.317**	(0.098)	0.326**	(0.101)	0.306**	(0.098)	0.302**	(0.094)
S. First $\times \Delta Q^{32}$			0.173	(0.192)	0.209	(0.196)	0.140	(0.193)	0.157	(0.191)
$\Delta Q^{32} \times$ Female					-0.147	(0.170)				
$\Delta Q^{32} \times$ Age					-0.086	(0.082)				
$\Delta Q^{32} \times$ CNS							-0.037	(0.083)		
$\Delta Q^{32} \times$ Sust. Attitude.									-0.055	(0.077)
ΔS^{21}	1.166***	(0.120)	1.421***	(0.135)	1.393***	(0.143)	1.433***	(0.135)	1.426***	(0.136)
$\Delta S^{21} \times \bar{a}(S)$	0.373***	(0.057)	0.528***	(0.066)	0.535***	(0.066)	0.515***	(0.068)	0.516***	(0.067)
S. First $\times \Delta S^{21}$			-0.572***	(0.138)	-0.581***	(0.137)	-0.558***	(0.139)	-0.553***	(0.137)
$\Delta S^{21} \times$ Female					0.078	(0.124)				
$\Delta S^{21} \times$ Age					-0.046	(0.051)				
$\Delta S^{21} \times$ CNS							0.095	(0.061)		
$\Delta S^{21} \times$ Sust. Attitude.									0.084	(0.056)
ΔS^{32}	0.951***	(0.123)	1.185***	(0.132)	1.206***	(0.149)	1.194***	(0.131)	1.187***	(0.132)
$\Delta S^{32} \times \bar{a}(S)$	0.356***	(0.055)	0.498***	(0.063)	0.501***	(0.063)	0.491***	(0.064)	0.490***	(0.063)
S. First $\times \Delta S^{32}$			-0.532***	(0.137)	-0.529***	(0.138)	-0.521***	(0.139)	-0.515***	(0.138)
$\Delta S^{32} \times$ Female					-0.040	(0.120)				
$\Delta S^{32} \times$ Age					-0.030	(0.048)				
$\Delta S^{32} \times$ CNS							0.057	(0.061)		
$\Delta S^{32} \times$ Sust. Attitude.									0.039	(0.052)
N	9851.000		9851.000		9851.000		9851.000		9851.000	
aic	9399.544		9237.835		9208.354		9155.646		9161.426	
bic	9716.138		9597.601		9654.464		9558.584		9564.365	
Position	No		Yes		Yes		Yes		Yes	
Demographics	No		No		Yes		No		No	
CNS	No		No		No		Yes		No	
Attitudes	No		No		No		No		Yes	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

N Analysis with other attention measures

The table below shows our estimated models when using alternative attention measures. Column (1) and (2) display the decision models presented in column (2) and (3) from appendix table I respectively. The first column corresponds to the full model with no attention, while the second uses the proportion of time spent looking at the attribute as the attention measure. The columns (3) and (4) show the attention model when using the absolute dwell time and number of fixations as attention measures. As we can see, the effects of attention are still significant and quite similar. Moreover, all the attention models display more descriptive power compared to the model with no attention.

	(1) No Attention		(2) Prop(DT)		(3) DT		(4) N	
ΔP	-0.979***	(0.123)	-1.141***	(0.128)	-1.110***	(0.122)	-1.016***	(0.116)
After × ΔP	0.096	(0.078)	0.106	(0.084)	0.100	(0.080)	0.097	(0.079)
SC × ΔP	-0.282	(0.183)	-0.168	(0.193)	-0.206	(0.180)	-0.277	(0.176)
QC × ΔP	-0.125	(0.175)	-0.047	(0.184)	-0.025	(0.172)	-0.108	(0.170)
After × SC × ΔP	0.159	(0.136)	0.166	(0.146)	0.159	(0.139)	0.160	(0.138)
After × QC × ΔP	0.082	(0.138)	0.087	(0.150)	0.080	(0.142)	0.083	(0.141)
$\Delta P \times \bar{a}(P)$			-0.324***	(0.057)	-0.359***	(0.075)	-0.264***	(0.062)
ΔQ^{21}	1.690***	(0.175)	1.766***	(0.180)	1.751***	(0.175)	1.703***	(0.175)
After × ΔQ^{21}	-0.209	(0.146)	-0.228	(0.157)	-0.218	(0.151)	-0.211	(0.147)
SC × ΔQ^{21}	0.298	(0.249)	0.361	(0.253)	0.298	(0.248)	0.303	(0.250)
QC × ΔQ^{21}	-0.019	(0.230)	0.037	(0.240)	-0.024	(0.231)	-0.003	(0.229)
After × SC × ΔQ^{21}	0.005	(0.218)	0.022	(0.229)	0.019	(0.222)	0.006	(0.220)
After × QC × ΔQ^{21}	-0.218	(0.244)	-0.227	(0.262)	-0.217	(0.252)	-0.217	(0.246)
$\Delta Q^{21} \times \bar{a}(Q)$			-0.043	(0.079)	0.150	(0.098)	0.022	(0.083)
ΔQ^{32}	1.254***	(0.165)	1.391***	(0.166)	1.336***	(0.164)	1.283***	(0.169)
After × ΔQ^{32}	-0.115	(0.125)	-0.123	(0.133)	-0.119	(0.130)	-0.117	(0.127)
SC × ΔQ^{32}	0.108	(0.228)	-0.026	(0.229)	0.060	(0.226)	0.115	(0.230)
QC × ΔQ^{32}	-0.087	(0.239)	-0.129	(0.241)	-0.147	(0.237)	-0.120	(0.241)
After × SC × ΔQ^{32}	-0.096	(0.190)	-0.103	(0.200)	-0.100	(0.196)	-0.097	(0.192)
After × QC × ΔQ^{32}	0.696**	(0.249)	0.749**	(0.268)	0.717**	(0.257)	0.707**	(0.254)
$\Delta Q^{32} \times \bar{a}(Q)$			0.182*	(0.082)	0.127	(0.086)	0.001	(0.074)
ΔS^{21}	1.114***	(0.124)	1.166***	(0.120)	1.197***	(0.121)	1.140***	(0.123)
After × ΔS^{21}	-0.006	(0.096)	-0.007	(0.102)	-0.006	(0.099)	-0.006	(0.098)
SC × ΔS^{21}	0.220	(0.170)	0.307	(0.168)	0.217	(0.165)	0.215	(0.171)
QC × ΔS^{21}	0.224	(0.176)	0.231	(0.176)	0.137	(0.172)	0.215	(0.174)
After × SC × ΔS^{21}	-0.447**	(0.165)	-0.464**	(0.173)	-0.452**	(0.168)	-0.450**	(0.167)
After × QC × ΔS^{21}	0.043	(0.162)	0.044	(0.172)	0.046	(0.165)	0.046	(0.163)
$\Delta S^{21} \times \bar{a}(S)$			0.373***	(0.057)	0.322***	(0.071)	0.149*	(0.061)
ΔS^{32}	0.903***	(0.123)	0.951***	(0.123)	0.982***	(0.121)	0.930***	(0.123)
After × ΔS^{32}	-0.087	(0.111)	-0.092	(0.117)	-0.090	(0.114)	-0.089	(0.112)
SC × ΔS^{32}	0.161	(0.171)	0.235	(0.171)	0.152	(0.169)	0.157	(0.173)
QC × ΔS^{32}	0.195	(0.170)	0.188	(0.172)	0.116	(0.169)	0.185	(0.171)
After × SC × ΔS^{32}	0.369*	(0.184)	0.392*	(0.193)	0.383*	(0.188)	0.374*	(0.186)
After × QC × ΔS^{32}	0.080	(0.175)	0.082	(0.187)	0.085	(0.180)	0.082	(0.178)
$\Delta S^{32} \times \bar{a}(S)$			0.356***	(0.055)	0.281***	(0.066)	0.100	(0.058)
After	-0.038	(0.060)	-0.040	(0.064)	-0.039	(0.062)	-0.038	(0.061)
SC	-0.014	(0.084)	0.008	(0.091)	0.010	(0.087)	-0.004	(0.087)
QC	0.024	(0.085)	0.049	(0.086)	0.037	(0.084)	0.033	(0.084)
After × SC	-0.025	(0.092)	-0.025	(0.098)	-0.026	(0.095)	-0.027	(0.093)
After × QC	0.129	(0.093)	0.140	(0.099)	0.134	(0.095)	0.130	(0.094)
$\bar{a}(P)$			-0.033	(0.035)	-0.028	(0.033)	-0.030	(0.034)
$\bar{a}(Q)$			-0.060	(0.031)	-0.098**	(0.034)	-0.072*	(0.028)
$\bar{a}(S)$			0.000	(.)	0.066*	(0.034)	-0.002	(0.034)
Constant	0.023	(0.056)	0.028	(0.060)	0.019	(0.057)	0.017	(0.057)
Var(random effects)	0.082***	(0.023)	0.072**	(0.022)	0.072***	(0.022)	0.081***	(0.023)
Observations	9851		9851		9851		9851	
AIC	9950.416		9399.544		9706.908		9851.018	
BIC	10216.643		9716.138		10030.697		10174.807	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

O Informed consent and instructions

Introduction

Welcome!

- You are participating in a study conducted by: UNIVERSITY OF AMSTERDAM
- Your decisions and information provided will be recorded **anonymously** (Your response in this experiment cannot be traced back to you.).
- Your data will only be used for **academic research**.
- You can choose at any point during the experiment to stop your participation and your data will be discarded.
- If you choose to participate, you will have to make multiple decisions that can lead to **additional payments and donations to charity**.
- The study will last about **15-20 minutes**.

Informed Consent

To participate in this experiment you need to:

- Be **18 or older**
- Be fluent in **English**.
- Focus only on the experiment (It will take around 15-20 minutes).
- Use your **computer** . The experiment does not work on phones or tablets.
- Set this window into **Fullscreen** mode (if you do not know how to do it, we will explain)

We reserve the right to exclude you from payment if we detect that you are not paying attention (jumping between other windows or setting off the Fullscreen too many times) or you do not comply with any of the above mentioned requirements.

☐ I agree to the Terms and Conditions above and to participate in this experiment.

Next

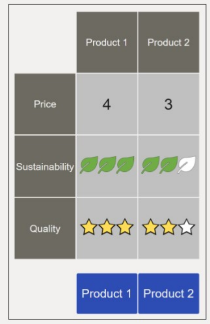
The Experiment

- In this study you will make purchasing decisions.
- The products that you will observe are **not real**.
- There is no physical product, but your decisions will have **real consequences.**
- Depending on your decisions you can obtain an additional **bonus payment of up to £3** and help the environment by **planting trees in a location of your choosing.**
- The trees will be planted by: ONETREEPLANTED
- There are many locations where the trees could be planted.
- Would you like to choose where to plant the trees? Anywhere

Your Decisions

In this experiment you will:

- Make multiple purchasing decisions (**37 rounds**).
- In each round, you will see **2 products** and their characteristics.
- These products are not real, but their benefits are.
- The next slide explains the benefits and characteristics of the products.



The product characteristics

- **Price:** This is the cost you need to pay for the product. The price is in **£**.
  Higher Price → lower bonus payments.
- **Quality:** 1-3 star () rating. Products with more stars () are more likely to have higher quality.
  Higher quality → higher bonus payments.
- **Eco-Label:** 1-3 leaf () rating. Products with more leaves () are more likely to be more sustainable.
  Higher sustainability → more trees planted.

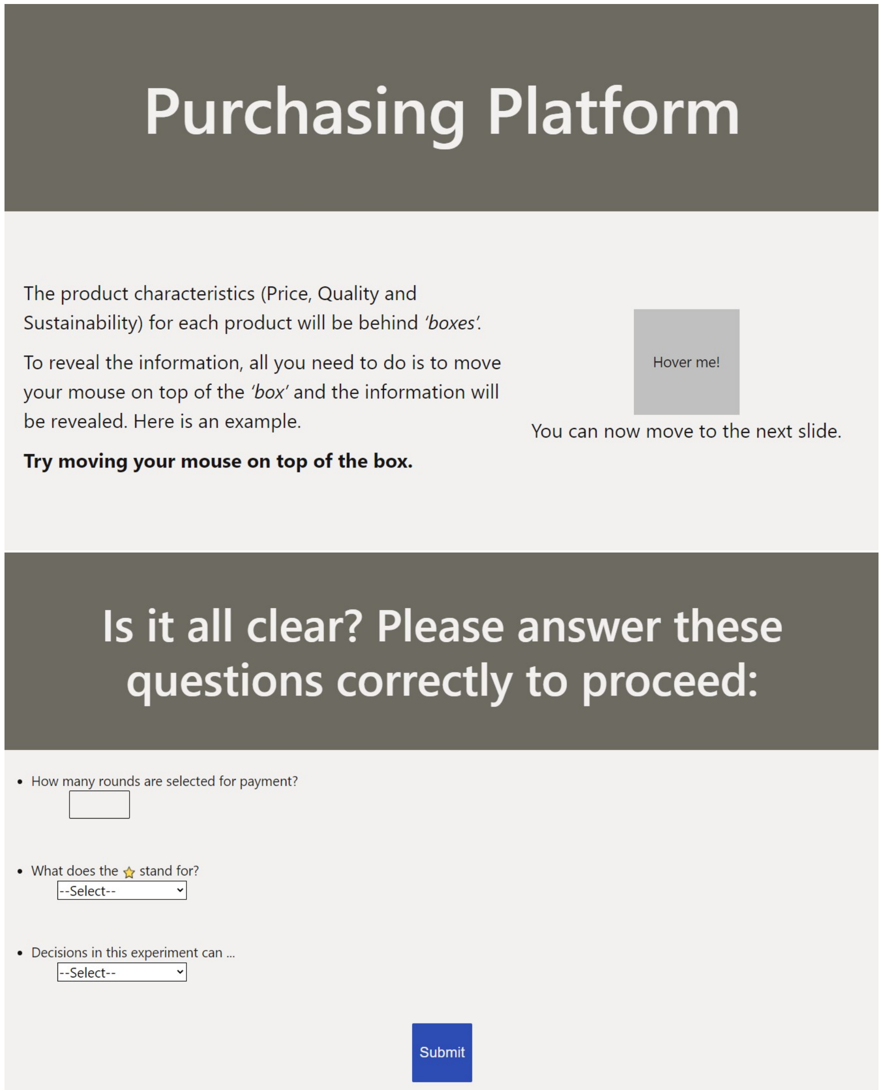
Value of Labels

Quality:

- Quality is measured in Quality points (10 points = 0.5 pound).
- Quality points go from 60 to 90. This means that the worst possible product (in terms of quality) is worth 3 pounds and the best one is worth 4.5.
- If a product has more stars , it means that you get more Quality points on average

Sustainability:

- Sustainability is measured in Sustainability points (10 points = 1 tree planted).
- Sustainability points go from 0 to 30. This means that the worst possible product (in terms of sustainability) will plant 0 trees and the best will plant 3 trees.
- If a product has more leaves , it means that you get more Sustainability points on average
- The total amount of points donated to the area you selected will be rounded up, so no point will be lost. For example, 102 points will mean 11 trees planted (instead of 10).



Purchasing Platform

The product characteristics (Price, Quality and Sustainability) for each product will be behind *'boxes'*.

To reveal the information, all you need to do is to move your mouse on top of the *'box'* and the information will be revealed. Here is an example.

Try moving your mouse on top of the box.

Hover me!

You can now move to the next slide.

Is it all clear? Please answer these questions correctly to proceed:

- How many rounds are selected for payment?
- What does the ☆ stand for?
  --Select--
- Decisions in this experiment can ...
  --Select--

Submit