Received: 15 December 2022 | Revised: 2 May 2024 | Accepted: 24 May 2024

DOI: 10.1002/bse.3838

RESEARCH ARTICLE

Business Strategy and the Environment WILEY

Positive, negative or graded sustainability labelling? Which is most effective at promoting a shift towards more sustainable product choices?

John Thøgersen[1] | François J. Dessart[2] | Ginevra Marandola[3] | Stefanie Lena Hille[4]

[1]Department of Management, Aarhus University, Aarhus, Denmark

[2]European Commission—Joint Research Centre, Seville, Spain

[3]Ministero dell'Economia e delle Finanze, Rome, Italy

[4]European Commission—DG Connect, Luxemberg, Luxemburg

Correspondence
John Thøgersen, Department of Management, Aarhus University, Fuglesangs Allé 4, 8210 Aarhus V, Denmark.
Email: jbt@mgmt.au.dk

Funding information
European Commission's Joint Research Centre

Abstract

Sustainability labels convey information about different product attributes, such as its environmental impact, lifespan or ethical performance. The labelling can be either positive (only identifying the most sustainable products available on the market), negative (only identifying the least sustainable products available on the market) or graded (comparing the sustainable performance of a product with that of all other products on the market). We assess the relative performance of these three labelling approaches in terms of influencing product choices. A nationally representative sample of 1243 consumers from Germany, Spain and the Czech Republic participated in an incentive-compatible online discrete choice experiment with random allocation to different labelling approaches. Compared with positive and negative labels, graded labels were most effective in guiding consumers towards more sustainable product choices. These findings support policy interventions that convey product sustainability with graded labels.

KEYWORDS

choice experiment, incentive compatible, sustainability label, sustainability labelling, traffic light colours

1 | INTRODUCTION

As emphasized in the UN Sustainable Development Goal #12, sustainable development requires changes (also) on the demand side (e.g. IPCC, 2022). Especially, middle- and high-income consumers need to change their consumption patterns, including a substantial shift towards more sustainable products and services (Creutzig et al., 2022). One of the key impediments blocking such a shift is the lack of clear and credible information that consumers can use to assess the sustainability performance of alternative products and services (Girod et al., 2014; Ölander & Thøgersen, 2014). Labelling is an increasingly popular tool to inform consumers about the sustainability performance of products and services, which is generally liked by consumers (e.g. Gadema & Oglethorpe, 2011; Schuitema et al., 2020) and, when designed well, used by producers and trusted by consumers, is documented effective at influencing their choices (Majer et al., 2022). From a consumer point of view, sustainability labelling serves especially two purposes: (1) to inform about important credence characteristics of the product or service (Fernqvist, 2018), compensating for the information asymmetry between sellers and buyers (Brach et al., 2018), in a summarised, salient and understandable fashion, and (2) to remind consumers of these issues at the

Abbreviations: IPCC, Intergovernmental Panel on Climate Change; UN, United Nations.

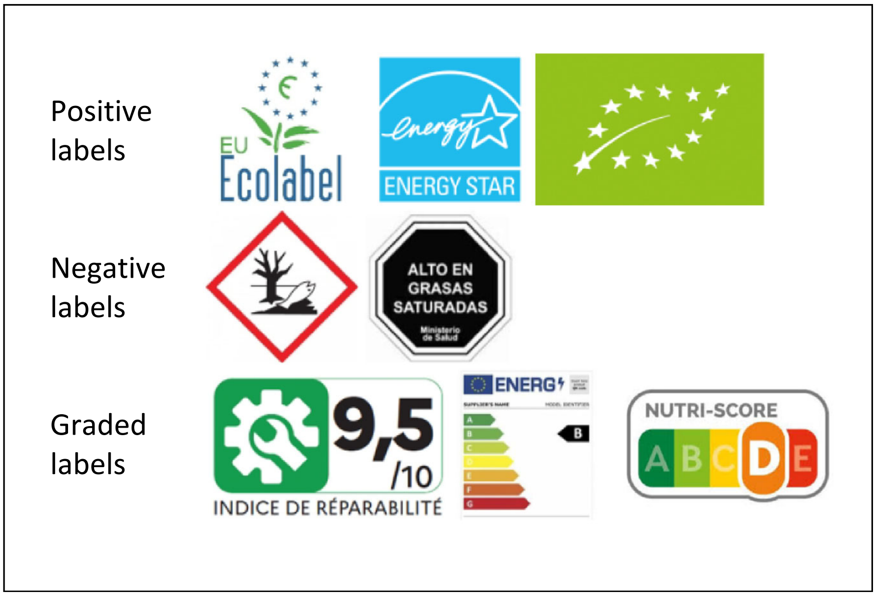


FIGURE 1 Examples of existing positive, negative and graded labels.

point-of-purchase, compensating for consumers' limited attention (Peschel et al., 2019).

Sustainability is a complex, multifaceted construct, and in practice, sustainability labels rarely cover all aspects (Torma & Thøgersen, 2021). Instead, most of them focus on a single, important sustainability attribute or dimension. Some of the more comprehensive are general environmental impact labels based on a product life-cycle assessment, such as the EU Ecolabel[1] (the 'EU Flower') or the Nordic Swan Ecolabel.[2] Other sustainability labels cover a narrower aspect, such as energy efficiency labelling (European Commission, 2019) or the emerging labelling of carbon footprint (Taufique et al., 2022) or product durability (Jacobs & Hörisch, 2022). Yet, others inform about ethical dimensions, such as Fairtrade,[3] certified B-Corps[4] or Cruelty Free.[5]

Existing sustainability labelling schemes follow different approaches to labelling (see examples in Figure 1). The most common is *positive labelling*, which identifies only those products that perform particularly well on a certain sustainability attribute. For instance, an organic food label is only awarded to products whose agricultural production meets the national organic standards. Positive labels are sometimes referred to as 'best-in-class' labels (e.g. Heinzle & Wüstenhagen, 2012). *Negative labelling* identifies only those products that perform particularly poorly or involve a particularly high risk (Machín et al., 2018), such as warning labels informing that a product is hazardous to the environment (Carrero et al., 2021). A third approach is *graded labelling* (also sometimes referred to as comparative labels, e.g. Grankvist et al., 2004), where all covered products are ranked along a rating scale on their relative performance. Examples include the EU energy label and the French Eco-Score label.

[1]https://environment.ec.europa.eu/topics/circular-economy/eu-ecolabel-home_en
[2]https://www.nordic-ecolabel.org
[3]https://www.fairtrade.org.uk
[4]https://www.bcorporation.net/en-us/certification
[5]https://crueltyfreeinternational.org/go-cruelty-free-leaping-bunny

There are many purposes for which a graded label is used in some and a categorical (usually positive) label in other contexts, such as the graded energy labelling in the EU versus the positive Energy Star labelling in the United States. A specific labelling approach may be chosen for different reasons. An important reason why positive labelling is the most frequently used approach, by far, is that few producers will voluntarily label their products as low performers. Hence, for negative and graded labelling to work as intended, they need to be mandatory (Edenbrandt & Nordström, 2023). However, mandatory labelling carries higher implementation costs (Mantilla Herrera et al., 2018) because not only the performance of those products that ask to be assessed to get the label needs to be assessed (Sunstein, 2019), and it provokes political resistance, especially from producers and their organizations (François-Lecompte et al., 2017). Such resistance is more difficult to overcome because of the lack of scientific evidence documenting that one approach is more effective than the other. Hence, there is a need for more research investigating which type of labelling is most effective at producing a shift in consumer choices towards more sustainable products and services. Complicating the matter further, the effectiveness of different labelling types might depend on a range of factors, such as the product, the labelled sustainability attribute or the national context.

On this background, it is the objective of this research to provide such much-needed evidence on the relative effectiveness of the three mentioned labelling approaches (positive, negative and graded) to shift consumers' choices towards more sustainable products and whether (and then how much) it depends on the country, product and/or labelled sustainability dimensions.

We do that by answering the following question: Which of the following three types of labelling approaches is most effective at making consumers choose more sustainable products?

1. Positive labelling, allowing the identification of the best performing products on the market in terms of sustainability;

2. Negative labelling, allowing the identification of the worst performing products on the market in terms of sustainability;
3. Graded labelling, identifying the sustainability performance of a product compared with that of all other products on the market, displaying a traffic light-coloured rating scale with letters from A to E.

Hence, this paper extends previous sustainability labelling research by examining and comparing the impacts of three common labelling approaches (positive vs. negative vs. graded), across different countries, product categories and sustainability dimensions.

2 | PRIOR RESEARCH

According to a recent systematic review of extant research, there is strong evidence, across diverse labelling schemes and research designs, that sustainability labelling indeed affects consumer perception and behaviour in the intended way (Majer et al., 2022). Majer et al. (2022) identified and reviewed 26 empirical journal articles covering labelling of different attributes (e.g. organic, fair trade, circular production) and different product categories (but mostly food and beverages, only four non-food). This independent, systematic review is an excellent basis for identifying gaps in the current state of the art in sustainability labelling research and for positioning this study's incremental contributions, free from potential selection biases.

Notably, none of the studies reviewed by Majer et al. (2022) compare the relative effectiveness of positive, negative and graded labelling. To be sure that this apparent gap in labelling research was not just due to oversight from these authors, we did a targeted search for published research that compare different labelling approaches to sustainability labelling but which was missed by Majer et al. (2022). Our targeted search confirmed that research that compares different sustainability labelling approaches is scarce. We identified seven studies comparing the effectiveness of a graded, traffic light-coloured label to some other way(s) of labelling sustainability performance in terms of influencing consumer intentions or choices (Bengart & Vogt, 2023; Holenweger et al., 2023; Meyerding et al., 2019; Muller et al., 2019; Slapø & Karevold, 2019; Thøgersen & Nielsen, 2016; Vlaeminck et al., 2014). In addition, we identified two studies comparing positive and negative sustainability labelling (Grankvist et al., 2004; Van Dam & De Jonge, 2015).

Among these studies, only one, Slapø and Karevold (2019), compared the effects of positive, negative and graded eco-labelling, for meal choices in a university cafeteria. However, all seven studies comparing a graded label using traffic light colours found that this approach is more effective at influencing sustainable consumer choices than the other approaches it was compared with. Other studies found that traffic light colours work best when combined with a straight alphabetic letter marking, compared with a combination of letters and other signs (Heinzle & Wüstenhagen, 2012; Ölander & Thøgersen, 2014). The two studies comparing positive and negative labelling both found that negative labelling is more effective at influencing sustainable consumer choices than positive (Grankvist et al., 2004; Van Dam & De Jonge, 2015). This is consistent with research on the negativity bias (Petersen et al., 2021; Vaish et al., 2008) and with research finding that negatively framed advertising messages are more effective than positively framed ones in promoting pro-environmental choices (e.g. Amatulli et al., 2019).

Note, however, that there is a bigger, but still relatively limited research in the nutrition and health areas comparing graded, traffic light colour labelling of food products to different other approaches, and this research reports more equivocal results (Kanter et al., 2018). Also in this area, no other approaches have appeared more successful than graded labelling using traffic light colours, in this case as a means to guide consumers towards healthier food choices (Temple, 2020). Graded, traffic light colour labelling, such as the French Nutri-Score label, is generally found to be more effective than nutrition facts alone, like GDA (Crosetto et al., 2016), or positive labelling, such as the Nordic Keyhole label (Ducrot et al., 2016). However, compared with negative labelling, such as the Chilean warning system, graded labelling has been found to be either more (Egnell et al., 2020), equally (Ares et al., 2018; Machín et al., 2018) or less effective (Khandpur et al., 2018).

Going back to the studies reviewed by Majer et al. (2022), all except two investigated the effects of a voluntary, positive label, in most cases already existing labels. The two exceptions investigated the effects of a graded, traffic light-coloured label, in both cases confirming the expected effect on meal choices in a restaurant (Brunner et al., 2018; Osman & Thornton, 2019). Studies that compared the effects of different labelling schemes mostly focused on the effects of the labelling organisation, and they generally found that official, governmental labels are superior to non-governmental (e.g. Ardeshiri et al., 2019; Risius et al., 2017). A few studies compared environmental and ethical labelling, with inconsistent findings regarding their relative impact (e.g. Grunert et al., 2014; Meyerding, 2016). Only two of the reviewed studies investigated whether the behavioural effects of a labelling scheme vary between different products (Banovic et al., 2019; Grunert et al., 2014), which they both confirmed. Grunert et al. (2014) found a stronger labelling effect for chocolate than for coffee, ice cream, soft drinks, cereals and ready meals (but no differences between these other products). Similarly, Banovic et al. (2019) found a stronger labelling effect for fresh fish products compared with smoked or canned fish products (again, no difference between the latter two products). The reviewed studies were carried out in many different countries, but only two (Banovic et al., 2019; Grunert et al., 2014) compared different (European) countries. Grunert et al. (2014) found significant differences in the understanding of different sustainability labels across countries whereas Banovic et al. (2019) found no significant differences between countries regarding the effect of sustainability labelling on choices. In sum, an important insight from the systematic review of extant research by Majer et al. (2022) is the lack of comparative studies of sustainability labelling across countries, products or sustainability attributes (and the few comparative studies report inconsistent results), which limits our ability to generalize results from labelling studies.

3 | HYPOTHESES

Based on previous research demonstrating the power of sustainability labelling to affect consumer behaviour (Majer et al., 2022), we propose and test the following pre-registered hypothesis[6]:

H1. Compared with no labelling, sustainability labelling (whether positive, negative or graded) is effective at making consumers choose more sustainable (e.g. ethical/environmentally friendly/durable) products.

Moreover, based on research on the negativity bias (Petersen et al., 2021; Vaish et al., 2008) and prior empirical research showing that negative labelling has a higher impact on people's choices compared with positive labelling (e.g. Grankvist et al., 2004; Van Dam & De Jonge, 2015), we hypothesize:

H2. Negative labelling is more effective than positive labelling at making consumers choose more sustainable (e.g. ethical/environmentally friendly/durable) products.

Furthermore, based on previous studies (Meyerding et al., 2019; Slapø & Karevold, 2019; Thøgersen & Nielsen, 2016; Vlaeminck et al., 2014), we expect that graded labels using a coloured traffic light rating system are more effective at guiding consumers towards more sustainable choices than only identifying the most (or least) sustainable products on the market via a positive (or negative) labelling approach. Graded labelling using traffic light colours allows the mapping of the sustainability information on an evaluative (good/bad) scale, which reduces information asymmetries by allowing a comparison of the performance of all products on the market (Hille et al., 2018). This in turn reduces cognitive effort and thereby increases the weight such information receives in decision-making (Johnson et al., 2012; Peters et al., 2007). Hence, we hypothesize:

H3. Graded labelling is more effective than positive and negative labelling at making consumers choose more sustainable (e.g. ethical/environmentally friendly/durable) products.

We test the robustness of our findings across national contexts by replicating the same study in three different countries with different languages and, to some extent, labelling traditions (Germany, Spain and the Czech Republic). Likewise, we test the robustness of findings across sustainability attributes by comparing the labelling of three different sustainability attributes: environmental performance, ethical performance and expected durability (in all three countries). Finally, to check the findings' robustness across different products, participants made choices in two different product categories: smartphones and microwave ovens. Because prior research report inconsistent results regarding the stability of labelling effects across countries, sustainability dimensions and product categories, we approach the impact of these factors in an exploratory fashion rather than formulating hypotheses.

4 | METHOD

The study was pre-registered (https://aspredicted.org/3L7_2N2), and ethical approval was obtained by the University of Valencia's Ethics Board. Data were collected by means of an online survey of 1243 consumers from Germany, Spain and the Czech Republic in April–June 2020. Quota sampling ensured that each sample is representative of the country in terms of gender, age and education level. The questionnaire was developed in English and translated into the three national languages. The translations were controlled by native speakers, who compared them with the original English version and settled all uncertainties and ambiguities with the second author. Participant recruitment, data collection and the organization and presentation of the data were handled by *Laboratorio de Investigación en Economía Experimental* (LINEEX) at the University of Valencia. Rather than recruiting from existing participant panels, LINEEX recruited participants for a panel for this study using targeted social media marketing campaigns and a form with filtering questions including age, country of residence and educational level. This procedure was employed in all three countries, in the national language. From this panel, nationally representative samples in terms of gender, age and education level (quota sampling) were drawn for the study. The sample obtained is well representative of each country on sex, age and education level (see Appendix Table A1).

Discrete choice experiments (DCEs) are an increasingly popular method for investigating how variable product attributes impact consumer preferences and choices (Haghani, Bliemer, & Hensher, 2021). The survey included two DCEs: one about smartphones and the other one about microwave ovens, presented in random order. Besides representing two different product categories, these products were chosen for practical reasons,[7] and we expected most consumers to own this type of products. This expectation was confirmed as 90% of our sample report to own a smartphone and 94% a microwave oven. Each participant made 10 choices among smartphones and 10 choices among microwave ovens. In each choice task, participants had to select their preferred product within a set of three alternatives.

Compared with answers on rating scales, preference measures based on choice-based experiments are less influenced by social desirability bias (Auger & Devinney, 2007; Meyerding, 2016) and by response styles from scale usage (Grunert et al., 2015). Also, preference estimates from choice experiments have been found to predict actual purchase behaviour (Grunert et al., 2009; Mueller et al., 2010).

The main weakness of choice experiments is that the choices are hypothetical, in and off themselves having no consequences for the

[6]The hypotheses were slightly, but not substantially, reworded after the pre-registration.

[7]To make the experiment incentive compatible, we needed data on brands' real performance on the selected product attributes (i.e. environmental performance, ethicality and average product lifespan). The budget for sending the real products to some participants also mattered.

participant, creating a risk of hypothetical bias (Haghani, Bliemer, Rose, et al., 2021a). We mitigated this risk by means of two measures. First, in the introduction to the choice experiments, participants were asked to imagine they needed to buy a new mobile phone/microwave oven (see Appendix Table A2) (Haghani, Bliemer, Rose, et al., 2021b). Second, and more importantly, we made the experiments incentive-compatible (Johnston et al., 2017) by informing participants in the beginning of the survey that one in 150 participants would be randomly drawn to receive a real smartphone or microwave oven with the characteristics of the alternative they selected in one of their choices (also randomly selected). If they selected an alternative cheaper than the most expensive alternative, they would receive the price difference as a money transfer. This meant that participants could face real consequences of their choices, therefore reducing the hypothetical nature of the experiment. Prior research has documented that, as long as decisions are incentivized, random lottery payments generally produce results consistent with paying all (Charness et al., 2016; Clot et al., 2018). In the weeks following the data collection, LINEEX randomly drew eight participants and sent them a smartphone or a microwave oven, together with the change if applicable.

The use of DCEs for studying consumer preferences and choices is based on Lancaster's (1966) consumer theory, assuming that consumer product preferences are composed of preferences regarding different attributes of the product. DCEs are used to 'reverse engineer' choices to quantify the impact of changes in attribute levels on choice (Hauber et al., 2016), effectively 'unbundling' consumers' product preferences into its component parts (Green & Srinivasan, 1990). The product alternatives that consumers can choose from typically vary on several attributes. In the present study, the smartphones varied on internal memory, camera resolution, price and brand name (fake), with or without a sustainability labelling attached to the brand name. The microwave ovens varied on power, presence of a digital display, price and brand name (again fake and again with or without an attached sustainability label). Note that behind the fake brand names, there were real brands, which we rated (from A to E, as explained below) on the selected sustainability dimension (i.e. environmental performance, ethicality and average product lifespan) using data on each brands' real performance.

We used fake brand names because the sustainability attributes vary by brand, and because many consumers are brand loyal, especially in the smartphone market, the labelling effect would have been confounded by the brand effect had we used real brand names. So, in our design, we treat the brand name and the sustainability labelling as a unit, always combining a (fake) brand name with the same labelling (which in the no-labelling condition is no label). This means that the expected effect of the brand name as such is 0, so a not significant effect in the control condition, and that all effects of the brand name/sustainability labelling unit can be attributed to the sustainability labelling.

In both cases, there were five different brands/sustainability levels, four different price levels (adapted to the market prices in the country) and the two remaining attributes had two levels each, resulting in $5 \times 4 \times 2 \times 2 = 80$ possible combinations. The number of different sets of three products that can be created from these 80 options is 82,160 ($n! / [m! \times (n - m)!]$, where $n = 80$ and $m = 3$). Therefore, we used SAS JMP to generate an efficient fractional factorial design based on the principles of minimum overlap and level balance and setting the number of choices to 10. Participants were presented with the 10 choice sets in random order.

The behavioural assumption behind DCEs is random utility theory (McFadden, 1974), which assumes that consumers (and, hence, participants) choose the product that maximizes their expected utility, with an amount of uncertainty in consumer judgments, which ads a random element (or random error) to their choices. Hence, it is assumed that consumers' choices reveal their preferences and the trade-offs they make between different attributes and levels. To estimate participants' (stated) preferences from the alternatives' characteristics, we use a mixed conditional multinomial logit model (MNL) with the alternatives' characteristics as explanatory variables (Hauber et al., 2016; McFadden, 1974), estimated by means of Latent Gold Choice 6.0 (LGC) (Vermunt & Magidson, 2016).

In addition to the choice experiment, participants were randomly allocated to 10 experimental conditions. One of the 10 conditions was the control group whose members were not exposed to any sustainability label when making their choices. In the other nine conditions, participants were exposed to a sustainability label, which differed across conditions on two factors with three levels each, that is, in a 3×3 full-factorial design. The first factor consists of three different types of labelling: (1) a positive label (a green check mark label, which identified the 20% products with the best sustainability performance), (2) a negative label (a red cross label, which identified the 20% products with the worst sustainability performance) or (3) a graded label (an A to E label, which ranked all products on their relative sustainability performance). The second factor consists of three different sustainability dimensions along which the product was characterized: (1) its environmental performance, (2) its lifespan performance or (3) its ethical performance. How the sustainability dimensions and the labelling type were presented to participants is shown in Appendix Table A2. Examples of the choice sets presented to participants are shown in Appendix Table A3.

Supplementary analyses were done on the mean sustainability scores of choices and on the 'market shares' of the best and worst performers when using different labelling approaches.

5 | RESULTS

5.1 | The impacts of labelling and other product attributes on product choices

As we will explain in more detail later, separate MNL analyses revealed that the effects of sustainability labelling are robust across the covered countries and sustainability dimensions. We therefore pooled the country samples and the three groups being exposed to labelling of different sustainability dimensions (environmental label, lifespan label or ethical label) for the main analysis, estimating the effects of the different labelling approaches. When using the control condition as benchmark, it is possible to estimate the effect of sustainability labelling as such, in addition to the differential effects of the three labelling approaches. Because the only attribute that differed between conditions was the type of label used, whereas the other attributes were invariant, the effects of all attributes except the labelling were fixed to be identical across conditions. Multigroup MNL analysis results for the two products with a comparison between the three sustainability labelling conditions—positive, negative and graded, defined as 'known classes'—are reported in Table 1.

TABLE 1 Multinomial logit model estimates for the choice of smartphones and microwave ovens with comparison of the effectiveness of labelling approaches ($n = 1243$).

Variables	Control (no label) B	Positive B	Negative B	Graded B	Wald	Wald(=)	Mean	Std. dev.
Smartphones	$R^2 = 0.17$	$R^2 = 0.20$	$R^2 = 0.18$	$R^2 = 0.30$				
Label[a] (relative importance)	(0.09)	(0.30)	(0.33)	(0.47)	2348.059***	1253.806***		
A	0.10	0.84	0.22	1.26			0.70	0.45
B	−0.17	−0.42	0.36	0.73			0.19	0.47
C	0.13	0.02	0.41	0.20			0.21	0.15
D	−0.04	−0.18	0.04	−0.78			−0.28	0.34
E	−0.02	−0.26	−1.04	−1.42			−0.82	0.53
Memory					775.090***			
Low	−0.33	−0.33	−0.33	−0.33			−0.33	
High	0.33	0.33	0.33	0.33			0.33	
Camera					1072.092***			
Low	−0.47	−0.47	−0.47	−0.47			−0.47	
High	0.47	0.47	0.47	0.47			0.47	
Price					1885.750***			
Lowest	0.71	0.71	0.71	0.71			0.71	
Low	0.30	0.30	0.30	0.30			0.30	
High	−0.32	−0.32	−0.32	−0.32			−0.32	
Highest	−0.69	−0.69	−0.69	−0.69			−0.69	
Microwave ovens	$R^2 = 0.18$	$R^2 = 0.22$	$R^2 = 0.20$	$R^2 = 0.35$				
Label[a] (relative importance)	(0.11)	(0.28)	(0.33)	(0.46)	2540.944***	1326.181***		
A	0.15	0.91	0.18	1.33			0.73	0.49
B	−0.17	−0.37	0.40	0.81			0.23	0.48
C	0.19	0.02	0.45	0.20			0.22	0.17
D	0.03	−0.21	0.11	−0.90			−0.29	0.41
E	−0.20	−0.36	−1.14	−1.44			−0.90	0.49
Display					884.291***			
Low	−0.36	−0.36	−0.36	−0.36			−0.36	
High	0.36	0.36	0.36	0.36			0.36	
Power					835.459***			
Low	−0.43	−0.43	−0.43	−0.43			−0.43	
High	0.43	0.43	0.43	0.43			0.43	
Price					2151.282***			
Lowest	0.78	0.78	0.78	0.78			0.78	
Low	0.34	0.34	0.34	0.34			0.34	
High	−0.24	−0.24	−0.24	−0.24			−0.24	
Highest	−0.89	−0.89	−0.89	−0.89			−0.89	

[a]The labelling attribute contains both fictitious brand names and label information. In all cases, A–E refers to the brands that are labelled as such in the graded labelling case. LatentGold uses effects coding for nominal variables per default.
***$p < .001$.

The analysis confirms that all included product attributes for both products significantly influence consumers' choices and in the expected direction. Participants generally prefer a lower price and superior performance on all attributes (e.g. bigger memory). This is no surprise but confirms that participants generally took the choices seriously.

Table 1 shows that the brand/labelling information becomes more important for consumer choices when sustainability labelling information is added and that it influences choices in the way proposed by Hypothesis 1. Further, the Wald(=) test shows that the different labelling approaches led to a significantly different impact on consumer choices for both products. Supplementary pairwise Wald tests revealed that the effects of the brand/labelling information on choices are significantly different (p < .001) between all four labelling conditions. As expected, the importance of the labelling/brand information is lowest, and small, in the control condition, where there was no sustainability labelling, only fictitious brand names. Adding a positive sustainability label identifying the best performers substantially increases the importance of this attribute, and especially, the preference for the product carrying the positive label increases substantially. Confirming prior research (e.g. Grankvist et al., 2004; Van Dam & De Jonge, 2015) and Hypothesis 2, negative labelling has a slightly stronger effect on choices than positive labelling and especially the preference for the brand carrying the negative label decreases substantially.

However, as predicted by Hypothesis 3, the effect of the labelling is strongest when consumers are exposed to graded sustainability labels with information about the relative performance of all brands, compared with only labelling the 20% best or worst performing brands. In this case, preferences for the brands follow the order of their sustainability grade. Paired comparisons show that the utility profiles for the brands carrying different labels differ significantly between all four conditions (p < .001, Wald test). The MNL analysis does not include an overall test of how the different labelling conditions or types influence the sustainability of choices, so we carried out supplementary analyses to investigate the relative sustainability of choices under the different conditions.

5.2 | Differences in average sustainability scores across labelling conditions

To further investigate the effects of the different labelling approaches, we did a rough calculation of participants' average sustainability scores over the 10 choices within a product category. As outlined in our pre-registration, we transformed the A to E scale, with each letter representing 20% of the variation in sustainability performance, to a five-point scale from E = 1 to A = 5 (for further explanation, see Box 1). We also analysed the market shares (i.e. share of choices) of the most and least sustainable products.

As shown in Figure 2, in the control (no label) condition, the average sustainability score of participants' choices was close to the midpoint of the scale, as one should expect, both for smartphones and microwave ovens. However, adding a sustainability label

BOX 1 Supplementary analyses

1. Average sustainability scores: the average score of points assigned according to the sustainability performance of the product that participants chose in each choice task (see table below), calculated across the 10 choice tasks

Sustainability performance	Label (if shown)	Points
Best	A B C D E (A highlighted) or ✓	5
Second-to-best	A B C D E (B highlighted)	4
Middle range	A B C D E (C highlighted)	3
Second-to-worst	A B C D E (D highlighted)	2
Worst	A B C D E (E highlighted) or ✗	1

2. Market shares of the most (least) sustainable product: the number of times a participant chose a product with the highest (lowest) sustainability performance, divided by the number of choice tasks for the product category.

3. Types of tests: because the distributions of these outcome variables are significantly different from the normal distribution (Kolmogorow-Smirnow test p values < .05), we use non-parametric Kruskal-Wallis tests for pairwise comparisons.

significantly increased the average sustainability score (p's < .001; non-parametric Kruskal-Wallis tests) as predicted (H1). The graded labelling produces the highest average sustainability score, which is significantly higher than in the positive and negative labelling conditions (all p's < .001), also as predicted (H3).

5.3 | Market share of the most sustainable product

The market shares of products characterized by different sustainability level, calculated (as pre-registered) as the share of choices of each of the five brands that are assigned labels from A to E in the graded labelling condition, are reported in Figure 3, again separately for smartphones and microwave ovens. The bottom (dark green) bars represent the market share of the product with the best sustainability performance.

In the control condition, where participants did not see any sustainability labels, the most sustainable smartphone has an average market share close to the expected 20% (given that there are five levels). When identifying the most sustainable smartphone by means of a positive label, its average market share rises substantially to 31.6%. This is significantly more than in the control (i.e. no label) condition (p < .001; non-parametric Kruskal-Wallis test). In the graded label condition, the most sustainable smartphone was identified with

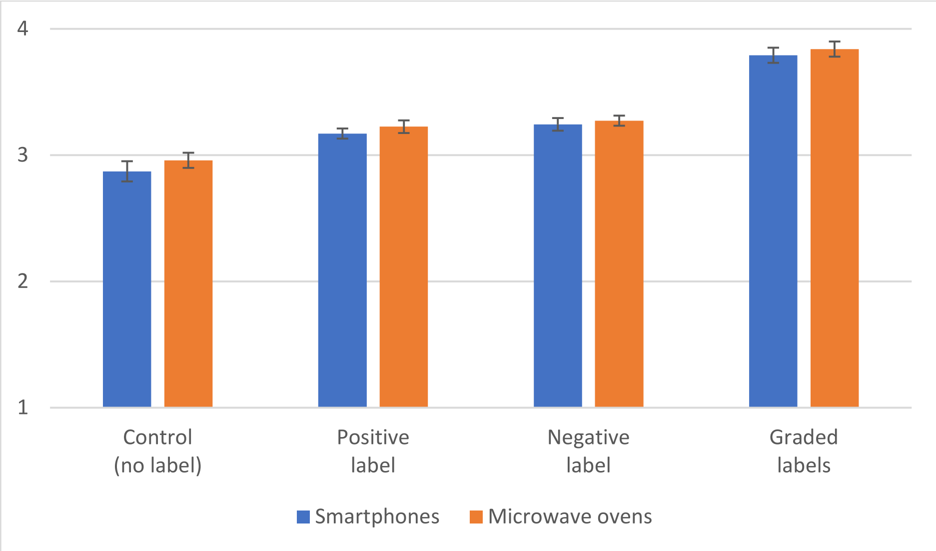


FIGURE 2 Average sustainability scores ($n = 1243$). All differences in average sustainability scores between control and labelling conditions and between categorical and graded labelling are significant ($p < .001$, non-parametric Kruskal-Wallis tests). Whiskers = 95% confidence intervals.

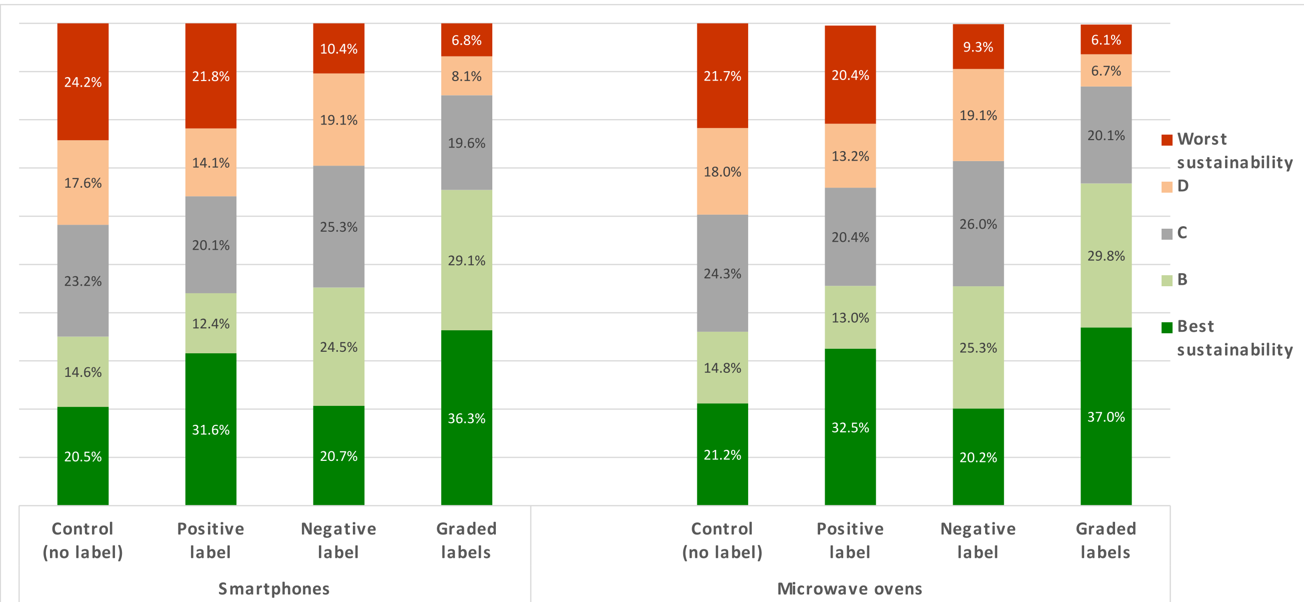


FIGURE 3 Average market shares ($n = 1243$). Differences in shares between control and labelling conditions and between categorical and graded labelling are significant ($p < .001$, non-parametric Kruskal-Wallis tests).

a dark green colour and an 'A' label on a scale from A to E. Here, the market share of the most sustainable smartphone rose even further to 36.3%. This is significantly more than in the control condition ($p < .001$), and also significantly more than in the positive label condition ($p < .001$). In the negative label condition, the most sustainable smartphone was not labelled, so it is no surprise that its market share does not differ from the control (i.e. no label) condition ($p = .77$).

The market share of the microwave ovens with the best sustainability performance follows the exact same pattern—we find the same significant versus non-significant differences between the different types of labels. In summary, these analyses show that graded labels are most effective at guiding consumer choices towards the most sustainable product, both for smartphones and for microwave ovens.

5.4 | Market share of the least sustainable product

In Figure 3, the top (dark red) bars represent the average market share of the product with the worst sustainability performance. For smartphones, this market share is 24.2% in the control condition (i.e. no label on any of the options). Adding a negative label on the least sustainable smartphone significantly ($p < .001$) and substantially reduced its market share to 10.4%, which is less than half of the control condition. In the graded label condition, the market share of the least sustainable smartphone is 6.8%, which is significantly lower than in the control condition ($p < .001$), and also significantly lower ($p < .001$) than in the negative label condition. In the positive label condition, participants did not know which smartphone was the least sustainable. Still, in this condition, the least sustainable smartphone had a slightly, yet significantly lower average market share (21.8%), than in the control condition ($p = .003$).

Again, we find the same pattern in terms of significant versus non-significant differences for microwave ovens. In summary, graded labels are also the most effective at guiding consumer choices away from the least sustainable product, both for smartphones and for microwave ovens.

5.5 | Robustness across products

Separate MNL analyses for the two products are reported in Appendix Table A4. These and the next two robustness checks only included the experiment groups, that is, participants who were exposed to sustainability labels on at least some products, but with no differentiation between labelling approaches. It appears from Appendix Table A4 that the sustainability labelling was equally important for consumer choices within the two product groups, relative to other included product attributes. The sustainability labelling was equally important for choices as the price and substantially more important than the two functional attributes. Not surprising, participants appear to prefer a mobile phone and a microwave oven with superior functional attributes and the lowest price. It also appears that they prefer a more sustainable product. Most importantly, the effects of the labelling appear to be robust across the two products.

5.6 | Robustness across countries

Multigroup MNL analyses for the two products with a comparison between the three countries, with countries defined as 'known classes', are reported in Appendix Table A5. It appears from this analysis that the sensitivity to functional attributes and the price differs significantly between countries, the German participants in general being less influenced by these attributes than participants from the other two countries. However, most important for our purpose, the effects of labelling on choices are not significantly different across countries (at the conventional level, $p < .05$). Hence, the effects of the labelling are robust across the three countries.

5.7 | Robustness across sustainability dimensions

Multigroup MNL analyses for the two products with a comparison between the three sustainability dimensions—environmental, ethical and lifespan ranking—defined as known classes, are reported in Appendix Table A6. Note that the other attributes are invariant across conditions, so it is not surprising that their effects on choices are not significantly different. However, despite the different sustainability dimensions, the effects of the sustainability labelling are also not significantly different (again, at the conventional level, $p < .05$). Hence, the effects of the labelling are also robust across the three sustainability dimensions.

6 | DISCUSSION

6.1 | Summary of results

This study confirms all the hypotheses that it set out to test. It confirms Hypothesis 1 and, using incentive compatible settings, bolsters research finding that sustainability labelling significantly influences consumer choices (Majer et al., 2022). It further shows that this result is robust across labelled sustainability dimensions and (European) countries. Hypothesis 2 is not fully confirmed. Negative labelling is more effective at decreasing preferences for the 20% least sustainable products than positive labelling is at increasing preferences for the 20% most sustainable products. However, the positive and negative labels have a similar effect on the overall sustainability of choices as measured by the sustainability score. Finally, this study confirms Hypothesis 3 that a graded label is more effective compared with positive and negative labels at promoting more sustainable buying decisions and thus shifting consumer choices towards more sustainable alternatives. We even find that a dark green 'A' label is more effective at promoting purchases of the most sustainable alternative than a positive label on the same alternative, and that a dark red 'E' label is more effective than a negative label at shying consumers away from the least sustainable alternative.

6.2 | Limitations

In the reported experiments, except for the control group, sustainability labelling was treated as de facto 'mandatory' and participants were implicitly made aware of this. This means, for instance, that participants had reasons to assume that a product that did not display a positive sustainability label did not belong to the category of the most sustainable products. Similarly, a product without negative label did not belong to the category of the least sustainable products. In practice, categorical (especially positive) labelling is often voluntary, which means that consumers cannot know for sure the performance of non-labelled products. It is possible that ambiguity about non-labelled products, which better reflects real-world choices, would have affected outcomes of the study, but it is difficult to predict in which direction.

This experiment was not designed to assess the absolute effect of sustainability labelling, but rather the effectiveness of each of the tested approaches relative to the others (positive, negative and graded). The experimental setting creates higher focus on the sustainability labelling than in real life, where consumers may also use other brands and attributes not included in this experiment to evaluate products, and where it is easier to miss a label. Therefore, the absolute effect of sustainability labelling may be smaller in real life than in experiments such as this.

A possible bias in the opposite direction might have been created by the way the DCE was made incentive compatible, which might have created an incentive to always choose the cheapest product to get the maximum amount of money if one's number was drawn. This might have inflated the importance of the price and, by implications, attenuated the importance of labelling and other product characteristics. Speaking against such a bias is the fact that under all circumstances, the value of the product would be much higher than the received change and that any bias in this direction was clearly not strong enough to lead to (falsely) rejecting our hypotheses.

Also, although DCEs are less influenced by desirability and instrument biases than rating scales, participants might still have exaggerated their willingness to buy sustainable products.

Moreover, in the study design, there was intentionally no correlation between the other attributes (e.g. the price) and labelling. In practice, sustainability-labelled options might carry a premium price, or might entail compromising some other, important quality, which might influence choices, especially in the graded conditions, where compromise does not necessarily mean giving up sustainability completely.

We cannot be sure how far our findings regarding the higher effectiveness of graded labels generalize. The fact that graded labels were most effective in all three studied countries and for both products makes us cautiously confident that our results can be generalized to many other countries and products, but future research should replicate our study in different national contexts and for other types of products. It is especially uncertain whether our findings can be extended to fast-moving consumer goods, where choices might be strongly influenced by habit and might not involve deliberate decision-making.

6.3 | Conclusions and implications

The positive effect of sustainability labelling on consumer choices has been attributed to consumers valuing sustainability attributes in products (Thøgersen & Nielsen, 2016) and needing labelling information to be able to take 'credence attributes' into account when making choices (Gorton et al., 2021). The superior performance of graded, traffic light-coloured labels has been attributed to several causes. The most parsimonious explanation is that, when using the graded labelling approach, consumers receive relevant information about a wider range of alternatives (Ducrot et al., 2016). Graded labelling implies that all products are labelled whereas only a minority of products are labelled when using either positive or negative labelling, which is therefore less powerful in reducing information asymmetries. It is an important caveat, however, that more information is not always better when it comes to labelling information. For example, studies find no positive effect on consumer choices of increasing the amount of information in graded, traffic light-coloured Nutri-Score labelling by extending it from a simple unidimensional to a multidimensional labelling (Ducrot et al., 2016; Egnell et al., 2020). Hence, it appears that more information is only better to a certain point (Bogliacino et al., 2023).

A second explanation for the effectiveness of graded labelling is that it allows the mapping of information about the environmental performance of a product on an evaluative rating scale ranging from good to bad (Hille et al., 2018). The use of categories with clear end points, such as grades, leads to a reduction of cognitive effort by simplifying information evaluation (Johnson et al., 2012; Peters et al., 2009). The less cognitive effort needed, the more the information weighs during the decision-making process (Peters et al., 2007).

A third explanation for the superiority of graded labelling using traffic light colours emphasizes the salience and symbolic value of colours (Hille et al., 2018). The use of colours can influence the salience of provided information, which may in turn influence attention and decision-making (Jarvenpaa, 1990; Lurie & Mason, 2007; Sunstein, 2014). Moreover, research suggests that colours affect people's evaluations in an automatic way, without conscious awareness (Elliot & Maier, 2014). For example, the colour red is often unconsciously negatively associated with concepts such as risk, danger, red traffic lights, a schoolteacher's red pen to correct mistakes or warning signals (Michalek et al., 2015; Selinger & Whyte, 2011). In contrast, the colour green is often associated with positive content, including success (Moller et al., 2009). For example, an experimental study found that, compared with numerical information on real-time energy consumption, the same information communicated by means of traffic light-coloured ambient lighting (glowing red during high levels and green during low levels) was easier to process and led to lower energy consumption (Maan et al., 2011). The same effect was not found when using other colours, less strongly associated with energy consumption (Lu et al., 2014). In the present study, colours were used consistently across the three labelling approaches (i.e. a green check mark in the positive label condition, a red cross label in the negative label condition and a traffic light-coloured rating scale in the graded label condition). Still, the combined use of all the traffic light colours (and colour nuances in between) might have contributed to the identified superior effectiveness of the graded labelling approach.

The reported findings provide important input to upcoming sustainability labelling policies regarding the most effective way of assisting consumers' shift to a more sustainable consumption pattern. For instance, the EU Green Deal[8] acknowledges the importance of consumption policies to reach the goal of no net emissions of greenhouse gases by 2050. Also, in the EU Circular Economy Action Plan,[9] it is explicitly stated that the European Commission will make proposals so that consumers receive information regarding the sustainability of products. Among other things, the EU plans to introduce labelling on product environmental footprint[10] and product reparability.[11] Graded, traffic light-coloured sustainability labelling has been field tested in

[8]European Commission (2019). EU Green Deal (https://eur-lex.europa.eu/resource.html%3Furi%3Dcellar:b828d165-1c22-11ea-8c1f-01aa75ed71a1.0002.02/DOC_1%26format%3DPDF).

[9]European Commission (2020). EU Circular Economy Action Plan (https://eur-lex.europa.eu/resource.html?uri=cellar:9903b325-6388-11ea-b735-01aa75ed71a1.0017.02/DOC_1&format=PDF).

[10]https://ec.europa.eu/environment/eussd/smgp/ef_methods.htm

[11]https://commission.europa.eu/law/law-topic/consumer-protection-law/consumer-contract-law/rules-promoting-repair-goods_en

restaurants (Brunner et al., 2018; Osman & Thornton, 2019; Slapø & Karevold, 2019; Spaargaren et al., 2013) and grocery stores (Muller et al., 2019; Vlaeminck et al., 2014) with promising results and are currently being tested in the form of an 'eco-score' label in supermarket chains in France and Belgium.[12] The present research lends further support for the use of a graded, traffic light-coloured approach to sustainability labelling. However, there are still open questions that need to be answered to develop optimal sustainability labelling. There are strong reasons to believe that, for maximum impact, sustainability labelling should preferably be mandatory. However, it is still an open question how large an effect can be achieved by means of voluntary labelling implemented by a supermarket chain, which is currently being tested in different countries. Prior research suggests that this will depend on consumer trust in the supermarket chain and the certifying organization (Darnall et al., 2018; Gorton et al., 2021) as well as the size of exposure to labelled products in the supermarket (Gadema & Oglethorpe, 2011; Thøgersen et al., 2010). Future research should investigate the importance of these and other contingencies for labelling effectiveness in practice (Taufique et al., 2022). Another important question is what is the most effective reference category for a graded sustainability label? In electronic products, it seems obvious that the reference category should be the product category, like mobile phones or microwave ovens. However, for example regarding foods, it is hotly debated whether the reference category should be all foods or a narrower category, such as vegetables or meat products (French Agency for Ecological Transition, 2021). The former approach has, for example, been chosen for the French eco-score label,[13] but the latter approach discriminates better between close substitutes and has been found to influence consumer choices more in a simulated online shopping context (Suchier et al., 2023). Hence, future research should determine optimal reference categories for sustainability labelling in different sectors, considering both functionality and costs (Taufique et al., 2022).

These open questions should be answered, preferably by means of field tests and some of them before full-scale implementation. Based on our research, we urge governments, businesses and other stakeholders to speed up their testing and implementation of effective sustainability labelling. This research strongly supports the use of graded, traffic light-coloured sustainability labelling, which clearly outperform other labelling types at making the sustainable choice the easy choice, for businesses and their customers.

ACKNOWLEDGEMENTS

The data collection for this study was financed by the European Commission's Joint Research Centre. The views expressed in this paper are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

[12]https://en.wikipedia.org/wiki/Eco-score#cite_note-auto1-1

[13]https://docs.score-environnemental.com/v/en/

ORCID

John Thøgersen https://orcid.org/0000-0003-0613-8062

François J. Dessart https://orcid.org/0000-0001-6783-5684

REFERENCES

Amatulli, C., De Angelis, M., Peluso, A. M., Soscia, I., & Guido, G. (2019). The effect of negative message framing on green consumption: An investigation of the role of shame. *Journal of Business Ethics*, *157*, 1111–1132. https://doi.org/10.1007/s10551-017-3644-x

Ardeshiri, A., Sampson, S., & Swait, J. (2019). Seasonality effects on consumers' preferences over quality attributes of different beef products. *Meat Science*, *157*, 107868. https://doi.org/10.1016/j.meatsci.2019.06.004

Ares, G., Varela, F., Machin, L., Antúnez, L., Giménez, A., Curutchet, M. R., & Aschemann-Witzel, J. (2018). Comparative performance of three interpretative front-of-pack nutrition labelling schemes: Insights for policy making. *Food Quality and Preference*, *68*, 215–225. https://doi.org/10.1016/j.foodqual.2018.03.007

Auger, P., & Devinney, T. (2007). Do what consumers say matter? The misalignment of preferences with unconstrained ethical intentions. *Journal of Business Ethics*, *76*, 361–383. https://doi.org/10.1007/s10551-006-9287-y

Banovic, M., Reinders, M. J., Claret, A., Guerrero, L., & Krystallis, A. (2019). A cross-cultural perspective on impact of health and nutrition claims, country-of-origin and eco-label on consumer choice of new aquaculture products. *Food Research International*, *123*, 36–47. https://doi.org/10.1016/j.foodres.2019.04.031

Bengart, P., & Vogt, B. (2023). Effects and interactions of labels' color scheme and the individual difference variable lay rationalism on pro-environmental choices. *Journal of Environmental Psychology*, *87*, 101998. https://doi.org/10.1016/j.jenvp.2023.101998

Bogliacino, F., Charris, R., Codagnone, C., Folkvord, F., Gaskell, G., Gómez, C., Liva, G., & Montealegre, F. (2023). Less is more: Information overload in the labelling of fish and aquaculture products. *Food Policy*, *116*, 102435. https://doi.org/10.1016/j.foodpol.2023.102435

Brach, S., Walsh, G., & Shaw, D. (2018). Sustainable consumption and third-party certification labels: Consumers' perceptions and reactions. *European Management Journal*, *36*, 254–265. https://doi.org/10.1016/j.emj.2017.03.005

Brunner, F., Kurz, V., Bryngelsson, D., & Hedenus, F. (2018). Carbon label at a university restaurant—label implementation and evaluation. *Ecological Economics*, *146*, 658–667. https://doi.org/10.1016/j.ecolecon.2017.12.012

Carrero, I., Valor, C., Díaz, E., & Labajo, V. (2021). Designed to be noticed: A reconceptualization of carbon food labels as warning labels. *Sustainability*, *13*(3), 1581. Retrieved from https://www.mdpi.com/2071-1050/13/3/1581

Charness, G., Gneezy, U., & Halladay, B. (2016). Experimental methods: Pay one or pay all. *Journal of Economic Behavior & Organization*, *131*, 141–150. https://doi.org/10.1016/j.jebo.2016.08.010

Clot, S., Grolleau, G., & Ibanez, L. (2018). Shall we pay all? An experimental test of Random Incentivized Systems. *Journal of Behavioral and Experimental Economics*, *73*, 93–98. https://doi.org/10.1016/j.socec.2018.01.004

Creutzig, F., Niamir, L., Bai, X., Callaghan, M., Cullen, J., Díaz-José, J., Figueroa, M., Grubler, A., Lamb, W. F., Leip, A., Masanet, E., Mata, É., Mattauch, L., Minx, J. C., Mirasgedis, S., Mulugetta, Y., Nugroho, S. B., Pathak, M., Perkins, P., … Ürge-Vorsatz, D. (2022). Demand-side solutions to climate change mitigation consistent with high levels of well-being. *Nature Climate Change*, *12*(1), 36–46. https://doi.org/10.1038/s41558-021-01219-y

Crosetto, P., Muller, L., & Ruffieux, B. (2016). Helping consumers with a front-of-pack label: Numbers or colors?: Experimental comparison between Guideline Daily Amount and Traffic Light in a diet-building exercise. *Journal of Economic Psychology*, *55*, 30–50. https://doi.org/10.1016/j.joep.2016.03.006

Darnall, N., Ji, H., & Vázquez-Brust, D. A. (2018). Third-party certification, sponsorship, and consumers' ecolabel use. *Journal of Business Ethics*, *150*, 953–969. https://doi.org/10.1007/s10551-016-3138-2

Ducrot, P., Julia, C., Méjean, C., Kesse-Guyot, E., Touvier, M., Fezeu, L. K., Hercberg, S., & Péneau, S. (2016). Impact of different front-of-pack nutrition labels on consumer purchasing intentions: A randomized controlled trial. *American Journal of Preventive Medicine*, *50*, 627–636. https://doi.org/10.1016/j.amepre.2015.10.020

Edenbrandt, A. K., & Nordström, J. (2023). The future of carbon labeling—factors to consider. *Agricultural and Resource Economics Review*, *52*, 151–167. https://doi.org/10.1017/age.2022.29

Egnell, M., Galan, P., Farpour-Lambert, N. J., Talati, Z., Pettigrew, S., Hercberg, S., & Julia, C. (2020). Compared to other front-of-pack nutrition labels, the Nutri-Score emerged as the most efficient to inform Swiss consumers on the nutritional quality of food products. *PLoS ONE*, *15*(2), e0228179. Retrieved from. https://doi.org/10.1371/journal.pone.0228179

Elliot, A. J., & Maier, M. A. (2014). Color Psychology: Effects of Perceiving Color on Psychological Functioning in Humans. *Annual Review of Psychology*, *65*, 95–120. https://doi.org/10.1146/annurev-psych-010213-115035

European Commission. (2019). Europeans' attitudes on EU energy policy. Retrieved from Luxembourg.

Fernqvist, F. (2018). Credence. In G. Ares & P. Varela (Eds.), *Methods in Consumer Research*, Volume *1*. New Approaches to Classic Methods (pp. 531–555). Woodhead Publishing. https://doi.org/10.1016/B978-0-08-102089-0.00020-0

François-Lecompte, A., Bertrandias, L., & Bernard, Y. (2017). The environmental labelling rollout of consumer goods by public authorities: Analysis of and lessons learned from the French case. *Journal of Cleaner Production*, *161*, 688–697. https://doi.org/10.1016/j.jclepro.2017.05.179

French Agency for Ecological Transition. (2021). *Environmental labelling for food products*. French Government. Retrieved from https://expertises.ademe.fr/sites/default/files/assets/documents/environnemental-labelling-food-products-government-report-parliament.pdf

Gadema, Z., & Oglethorpe, D. (2011). The use and usefulness of carbon labelling food: A policy perspective from a survey of UK supermarket shoppers. *Food Policy*, *36*, 815–822. https://doi.org/10.1016/j.foodpol.2011.08.001

Girod, B., van Vuuren, D. P., & Hertwich, E. G. (2014). Climate policy through changing consumption choices: Options and obstacles for reducing greenhouse gas emissions. *Global Environmental Change*, *25*, 5–15. https://doi.org/10.1016/j.gloenvcha.2014.01.004

Gorton, M., Tocco, B., Yeh, C.-H., & Hartmann, M. (2021). What determines consumers' use of eco-labels? Taking a close look at label trust. *Ecological Economics*, *189*, 107173. https://doi.org/10.1016/j.ecolecon.2021.107173

Grankvist, G., Dahlstrand, U., & Biel, A. (2004). The impact of environmental labeling on consumer preference: Negative versus positive labels. *Journal of Consumer Policy*, *27*, 213–230. https://doi.org/10.1023/B:COPO.0000028167.54739.94

Green, P. E., & Srinivasan, V. (1990). Conjoint analysis in marketing: New developments with implications for research and practice. *Journal of Marketing*, *54*(October), 3–19. https://doi.org/10.1177/002224299005400402

Grunert, K. G., Hieke, S., & Wills, J. (2014). Sustainability labels on food products: Consumer motivation, understanding and use. *Food Policy*, *44*, 177–189. https://doi.org/10.1016/j.foodpol.2013.12.001

Grunert, K. G., Juhl, H. J., Esbjerg, L., Jensen, B. B., Bech-Larsen, T., Brunsø, K., & Madsen, C. Ø. (2009). Comparing methods for measuring consumer willingness to pay for a basic and an improved ready made soup product. *Food Quality and Preference*, *20*, 607–619. https://doi.org/10.1016/j.foodqual.2009.07.006

Grunert, K. G., Loose, S. M., Zhou, Y., & Tinggaard, S. (2015). Extrinsic and intrinsic quality cues in Chinese consumers' purchase of pork ribs. *Food Quality and Preference*, *42*, 37–47. https://doi.org/10.1016/j.foodqual.2015.01.001

Haghani, M., Bliemer, M. C. J., & Hensher, D. A. (2021). The landscape of econometric discrete choice modelling research. *Journal of Choice Modelling*, *40*, 100303. https://doi.org/10.1016/j.jocm.2021.100303

Haghani, M., Bliemer, M. C. J., Rose, J. M., Oppewal, H., & Lancsar, E. (2021a). Hypothetical bias in stated choice experiments: Part I. Macro-scale analysis of literature and integrative synthesis of empirical evidence from applied economics, experimental psychology and neuroimaging. *Journal of Choice Modelling*, *41*, 100309. https://doi.org/10.1016/j.jocm.2021.100309

Haghani, M., Bliemer, M. C. J., Rose, J. M., Oppewal, H., & Lancsar, E. (2021b). Hypothetical bias in stated choice experiments: Part II. Conceptualisation of external validity, sources and explanations of bias and effectiveness of mitigation methods. *Journal of Choice Modelling*, *41*, 100322. https://doi.org/10.1016/j.jocm.2021.100322

Hauber, A. B., González, J. M., Groothuis-Oudshoorn, C. G. M., Prior, T., Marshall, D. A., Cunningham, C., IJzerman, M. J., & Bridges, J. F. P. (2016). Statistical methods for the analysis of discrete choice experiments: A report of the ISPOR conjoint analysis good research practices task force. *Value in Health*, *19*, 300–315. https://doi.org/10.1016/j.jval.2016.04.004

Heinzle, S. L., & Wüstenhagen, R. (2012). Dynamic adjustment of eco-labeling schemes and consumer choice—the revision of the EU energy label as a missed opportunity? *Business Strategy and the Environment*, *21*, 60–70. Retrieved from. https://doi.org/10.1002/bse.722

Hille, S. L., Geiger, C., Loock, M., & Peloza, J. (2018). Best in class or simply the best? The impact of absolute versus relative ecolabeling approaches. *Journal of Public Policy & Marketing*, *37*, 5–22. https://doi.org/10.1509/jppm.15.030

Holenweger, G., Stöckli, S., & Brügger, A. (2023). Carbon footprint labels involving traffic lights foster sustainable food choices. *Food Quality and Preference*, *106*, 104813. https://doi.org/10.1016/j.foodqual.2023.104813

IPCC. (2022). In P. R. Shukla, J. Skea, R. Slade, A. Al Khourdajie, R. van Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz, & J. Malley (Eds.), *Climate change 2022: Mitigation of climate change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

Jacobs, K., & Hörisch, J. (2022). The importance of product lifetime labelling for purchase decisions: Strategic implications for corporate sustainability based on a conjoint analysis in Germany. *Business Strategy and the Environment*, *31*, 1275–1291. https://doi.org/10.1002/bse.2954

Jarvenpaa, S. L. (1990). Graphic displays in decision making—the visual salience effect. *Journal of Behavioral Decision Making*, *3*, 247–262. https://doi.org/10.1002/bdm.3960030403

Johnson, E. J., Shu, S., Dellaert, B., Fox, C., Goldstein, D., Häubl, G., Larrick, R. P., Payne, J. W., Peters, E., Schkade, D., Wansink, B., & Weber, E. (2012). Beyond nudges: Tools of a choice architecture. *Marketing Letters*, *23*, 487–504. https://doi.org/10.1007/s11002-012-9186-1

Johnston, R. J., Boyle, K. J., Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T. A., Hanemann, W. M., Hanley, N., Ryan, M., Scarpa, R., Tourangeau, R., & Vossler, C. A. (2017). Contemporary guidance for stated preference studies. *Journal of the Association of Environmental and Resource Economists*, *4*, 319–405. https://doi.org/10.1086/691697

Kanter, R., Vanderlee, L., & Vandevijvere, S. (2018). Front-of-package nutrition labelling policy: Global progress and future directions. *Public Health Nutrition*, *21*, 1399–1408. https://doi.org/10.1017/s1368980018000010

Khandpur, N., de Morais Sato, P., Mais, L. A., Bortoletto Martins, A., Spinillo, C. G., Garcia, M. T., Urquizar Rojas, C., & Jaime, P. (2018). Are front-of-package warning labels more effective at communicating nutrition information than traffic-light labels? A randomized controlled experiment in a Brazilian sample. *Nutrients*, *10*(6), 688. https://doi.org/10.3390/nu10060688

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, *74*, 132–157. https://doi.org/10.1086/259131

Lu, S., Ham, J., & Midden, C. J. H. (2014). Using Ambient Lighting in Persuasive Communication: The Role of Pre-existing Color Associations. Paper presented at the Persuasive Technology, Cham. https://doi.org/10.1007/978-3-319-07127-5_15

Lurie, N. H., & Mason, C. H. (2007). Visual representation: Implications for decision making. *Journal of Marketing*, *71*(1), 160–177. https://doi.org/10.1509/jmkg.71.1.160

Maan, S., Merkus, B., Ham, J., & Midden, C. (2011). Making it not too obvious: The effect of ambient light feedback on space heating energy consumption. *Energy Efficiency*, *4*, 175–183. https://doi.org/10.1007/s12053-010-9102-6

Machín, L., Aschemann-Witzel, J., Curutchet, M. R., Giménez, A., & Ares, G. (2018). Does front-of-pack nutrition information improve consumer ability to make healthful choices? Performance of warnings and the traffic light system in a simulated shopping experiment. *Appetite*, *121*, 55–62. https://doi.org/10.1016/j.appet.2017.10.037

Majer, J. M., Henscher, H. A., Reuber, P., Fischer-Kreer, D., & 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. https://doi.org/10.1016/j.spc.2022.06.012

Mantilla Herrera, A. M., Crino, M., Erskine, H. E., Sacks, G., Ananthapavan, J., Mhurchu, C. N., & Lee, Y. Y. (2018). Cost-effectiveness of product reformulation in response to the Health Star Rating Food Labelling System in Australia. *Nutrients*, *10*(5), 614. https://doi.org/10.3390/nu10050614

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In I. Zarembka (Ed.), *Frontiers in econometrics* (pp. 105–142). Academic Press.

Meyerding, S. G. H. (2016). Consumer preferences for food labels on tomatoes in Germany—a comparison of a quasi-experiment and two stated preference approaches. *Appetite*, *103*, 105–112. https://doi.org/10.1016/j.appet.2016.03.025

Meyerding, S. G. H., Schaffmann, A.-L., & Lehberger, M. (2019). Consumer preferences for different designs of carbon footprint labelling on tomatoes in Germany—does design matter? *Sustainability*, *11*(6), 1587. https://doi.org/10.3390/su11061587

Michalek, G., Meran, G., Schwarze, R., & Yildiz, Ö. (2015). *Nudging as a new "soft" tool in environmental policy—an analysis based on insights from cognitive and social psychology*. Retrieved from Frankfurt (Oder): https://www.europa-uni.de/de/forschung/institut/recap15/downloads/recap15_DP021.pdf

Moller, A. C., Elliot, A. J., & Maier, M. A. (2009). Basic hue-meaning associations. *Emotion*, *9*, 898–902. Retrieved from. https://doi.org/10.1037/a0017811

Mueller, S., Osidacz, P., Francis, I. L., & Lockshin, L. (2010). Combining discrete choice and informed sensory testing in a two-stage process: Can it predict wine market share? *Food Quality and Preference*, *21*, 741–754. https://doi.org/10.1016/j.foodqual.2010.06.008

Muller, L., Lacroix, A., & Ruffieux, B. (2019). Environmental labelling and consumption changes: A food choice experiment. *Environmental and Resource Economics*, *73*, 871–897. https://doi.org/10.1007/s10640-019-00328-9

Ölander, F., & Thøgersen, J. (2014). Informing versus nudging in environmental policy. *Journal of Consumer Policy*, *37*, 341–356. https://doi.org/10.1007/s10603-014-9256-2

Osman, M., & Thornton, K. (2019). Traffic light labelling of meals to promote sustainable consumption and healthy eating. *Appetite*, *138*, 60–71. https://doi.org/10.1016/j.appet.2019.03.015

Peschel, A. O., Orquin, J. L., & Mueller Loose, S. (2019). Increasing consumers' attention capture and food choice through bottom-up effects. *Appetite*, *132*, 1–7. https://doi.org/10.1016/j.appet.2018.09.015

Peters, E., Dieckmann, N. F., Västfjäll, D., Mertz, C., Slovic, P., & Hibbard, J. H. (2009). Bringing meaning to numbers: The impact of evaluative categories on decisions. *Journal of Experimental Psychology: Applied*, *15*, 213–227. Retrieved from. https://doi.org/10.1037/a0016978

Peters, E., Hibbard, J., Slovic, P., & Dieckmann, N. (2007). Numeracy skill and the communication, comprehension, and use of risk-benefit information. *Health Affairs*, *26*(3), 741–748. https://doi.org/10.1377/hlthaff.26.3.741

Petersen, L., Hörisch, J., & Jacobs, K. (2021). Worse is worse and better doesn't matter?: The effects of favorable and unfavorable environmental information on consumers' willingness to pay. *Journal of Industrial Ecology*, *25*, 1338–1356. https://doi.org/10.1111/jiec.13141

Risius, A., Janssen, M., & Hamm, U. (2017). Consumer preferences for sustainable aquaculture products: Evidence from in-depth interviews, think aloud protocols and choice experiments. *Appetite*, *113*, 246–254. https://doi.org/10.1016/j.appet.2017.02.021

Schuitema, G., Aravena, C., & Denny, E. (2020). The psychology of energy efficiency labels: Trust, involvement, and attitudes towards energy performance certificates in Ireland. *Energy Research & Social Science*, *59*, 101301. https://doi.org/10.1016/j.erss.2019.101301

Selinger, E., & Whyte, K. (2011). Is there a right way to nudge? The practice and ethics of choice architecture. *Sociology Compass*, *5*, 923–935. https://doi.org/10.1111/j.1751-9020.2011.00413.x

Slapø, H. B., & Karevold, K. I. (2019). Simple eco-labels to nudge customers toward the most environmentally friendly warm dishes: An empirical study in a cafeteria setting. *Frontiers in Sustainable Food Systems*, *3*, 40. https://doi.org/10.3389/fsufs.2019.00040

Spaargaren, G., van Koppen, C. S. A., Janssen, A. M., Hendriksen, A., & Kolfschoten, C. J. (2013). Consumer responses to the carbon labelling of food: A real life experiment in a canteen practice. *Sociologia Ruralis*, *53*(4), 432–453. https://doi.org/10.1111/soru.12009

Suchier, J., Demarque, C., Waroquier, L., Girandola, F., Hilton, D., & Muller, L. (2023). Effects of intra- and inter-category traffic-light carbon labels and the presence of a social norm cue on food purchases. *Journal of Consumer Behaviour*, *22*(3), 597–617. https://doi.org/10.1002/cb.2139

Sunstein, C. R. (2014). Nudging: A very short guide. *Journal of Consumer Policy*, *37*, 583–588. https://doi.org/10.1007/s10603-014-9273-1

Sunstein, C. R. (2019). Ruining popcorn? The welfare effects of information. *Journal of Risk and Uncertainty*, *58*(2/3), 121–142. https://doi.org/10.1007/s11166-019-09300-w

Taufique, K. M. R., Nielsen, K. S., Dietz, T., Shwom, R., Stern, P. C., & Vandenbergh, M. P. (2022). Revisiting the promise of carbon labelling. *Nature Climate Change*, *12*, 132–140. https://doi.org/10.1038/s41558-021-01271-8

Temple, N. J. (2020). Front-of-package food labels: A narrative review. *Appetite*, *144*, 104485. https://doi.org/10.1016/j.appet.2019.104485

Thøgersen, J., Haugaard, P., & Olesen, A. (2010). Consumer responses to ecolabels. *European Journal of Marketing*, *44*, 1787–1810. https://doi.org/10.1108/03090561011079882

Thøgersen, J., & Nielsen, K. S. (2016). A better carbon footprint label. *Journal of Cleaner Production*, *125*, 86–94. https://doi.org/10.1016/j.jclepro.2016.03.098

Torma, G., & Thøgersen, J. (2021). A systematic literature review on meta sustainability labeling—what do we (not) know? *Journal of Cleaner Production*, *293*, 126194. https://doi.org/10.1016/j.jclepro.2021.126194

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development.

Psychological Bulletin, *134*, 383–403. https://doi.org/10.1037/0033-2909.134.3.383

Van Dam, Y. K., & De Jonge, J. (2015). The positive side of negative labelling. *Journal of Consumer Policy*, *38*, 19–38. https://doi.org/10.1007/s10603-014-9274-0

Vermunt, J. K., & Magidson, J. (2016). *Technical guide for Latent GOLD 5.1: Basic, advanced, and syntax*. Statistical Innovations.

Vlaeminck, P., Jiang, T., & Vranken, L. (2014). Food labeling and eco-friendly consumption: Experimental evidence from a Belgian supermarket. *Ecological Economics*, *108*, 180–190. https://doi.org/10.1016/j.ecolecon.2014.10.019

How to cite this article: Thøgersen, J., Dessart, F. J., Marandola, G., & Hille, S. L. (2024). Positive, negative or graded sustainability labelling? Which is most effective at promoting a shift towards more sustainable product choices? *Business Strategy and the Environment*, 1–19. https://doi.org/10.1002/bse.3838

APPENDIX A

TABLE A1 Sample description and comparison with Census Data of the European Statistical System.

	Czech Republic		Germany		Spain	
Variables	**Sample (%)**	**Population (%)**	**Sample (%)**	**Population (%)**	**Sample (%)**	**Population (%)**
Sex						
Male	47.9	48.2	48.5	48.3	49.4	48.8
Female	52.1	51.9	51.5	51.7	50.6	51.3
Age						
18–29	24.8	21.3	17.0	19.7	16.9	16.8
30–49	33.6	34.9	34.0	33.0	40.7	40.3
≥ 50	41.7	43.8	49.0	47.4	42.3	43.0
Educational level						
Primary	15.0	18.8	21.0	23.4	25.8	26.0
Secondary	67.4	66.4	53.0	51.7	47.9	47.6
Tertiary	17.6	14.8	26.00	24.9	26.3	26.4

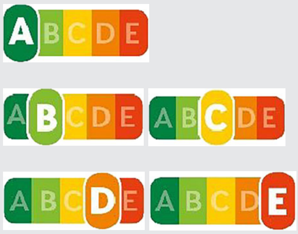
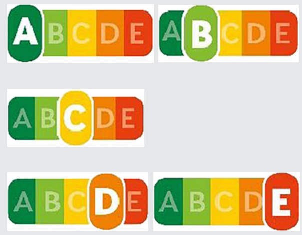
Source: https://ec.europa.eu/CensusHub2/selectHyperCube?clearSession=true.

TABLE A2 How the choice experiment, sustainability dimensions and labelling types were presented to participants.

General introduction	Imagine your mobile phone/microwave oven breaks down. You have looked into the possibility of having it repaired, but it is impossible to do so. You thus need to buy a new mobile phone/microwave oven. If you currently don't own a mobile phone/microwave oven, imagine you need to buy one. You have a budget of maximum €399/€159 to spend. You do not need to spend the entire budget—only if you would in real life. Simply choose the phone/microwave oven that you would choose in real life if you were only proposed the models shown on the screen. On the following screens, you will be shown different smartphones/microwave ovens. All smartphones/microwave ovens proposed are new [and unlocked].[a] These smartphones/microwave ovens vary on the following characteristics: • Brand (in this study, the real name of the brand is not shown to you in order not to influence your choice) • Precision of the main camera (expressed in number of megapixels)/presence of a digital display (allows you to select the exact duration) • Internal memory (expressed in gigabits)/maximum power (expressed in watt) • Price • Environmental/ethical/lifespan ranking[b]		
Further explanation, the labelling conditions	Environmental ranking:	Ethical ranking:	Lifespan ranking:
Common explanation	An independent organisation ranked smartphone/microwave oven brands based on how environmentally friendly these smartphones/ microwave ovens are produced. To do so, the organisation examined the **chemicals used** in the smartphones/microwave ovens and the use of **recycled and recyclable materials.** For instance, if smartphone/ microwave ovens brands use **hazardous toxic chemicals**, they may pollute the environment during mining and disassembling. Conversely, smartphones/ microwave ovens are more environmentally friendly when they are made out of **recycled or recyclable materials**. The independent organisation ranked all smartphone/microwave oven brands according to these environmental criteria.	An independent organisation ranked smartphone/microwave oven brands based on how ethically these smartphones/microwave ovens are produced. To do so, the organisation examined **where and how smartphone/ microwave oven brands source the metals** they put in the smartphones/ microwave ovens, and at **how they treat the workers who assemble them**. For instance, if smartphone/ microwave ovens brands source metals from **countries with armed conflicts**, they may indirectly finance armed groups or support money laundering. Conversely, smartphone/ microwave oven brands may have a policy for **improving their workers' rights, working hours and safety**. The independent organisation ranked all smartphone/microwave oven brands according to these ethical criteria.	An independent organisation ranked smartphone/microwave oven brands based on how long these smartphones/microwave ovens last. To do so, the organisation examined **how long consumers report being able to use their smartphone/ microwave ovens normally** without the need for excessive maintenance or repair. For instance, if the **battery or the operating system/transformer or the magnetron stops working** properly, consumers cannot normally use their smartphone/ microwave oven. [The same happens if the screen breaks easily if one drops their smartphone.[a]] Conversely, consumers can normally use their smartphone/microwave oven when it is **solid, water resistant,** and when it is still **possible to update the operating system**/ when its buttons and door are **solid and does not easily get damaged**. The independent organisation ranked all smartphone/microwave oven brands according to these lifespan criteria.
Positive	In your shopping tasks, **only the top 20% smartphone/microwave brands with the best environmental ranking will be identified as such.**	In your shopping tasks, **only the top 20% smartphone/microwave brands with the best ethical ranking will be identified as such.**	In your shopping tasks, **only the top 20% smartphone/microwave brands with the best lifespan ranking will be identified as such.**
Negative	In your shopping tasks, **only the bottom 20% smartphone/ microwave brands with the worst environmental ranking will be identified as such.**	In your shopping tasks, **only the bottom 20% smartphone/ microwave brands with the worst ethical ranking will be identified as such.**	In your shopping tasks, **only the bottom 20% smartphone/ microwave brands with the worst lifespan ranking will be identified as such.**
Graded	Smartphone/microwave oven brands with an 'A' ranking are among the top 20% most environmentally friendly brands.	Smartphone/microwave oven brands with an 'A' ranking are among the top 20% most ethical brands. Smartphone/microwave	Smartphone/microwave oven brands with an 'A' ranking are among the top 20% with the longest product lifespan. Smartphone/

(Continues)

TABLE A2 (Continued)

	Smartphone/microwave oven brands with an 'E' ranking are among the bottom 20% least environmentally friendly brands. Smartphone/microwave brands with a 'B', 'C' or 'D' ranking are between these two extremes. In your shopping tasks, **the environmental ranking of all smartphone/microwave brands will be shown as such:** 	oven brands with an 'E' ranking are among the bottom 20% least ethical brands. Smartphone/microwave brands with a 'B', 'C' or 'D' ranking are between these two extremes. In your shopping tasks, **the ethical ranking of all smartphone/ microwave brands will be shown as such:** 	microwave oven brands with an 'E' ranking are among the bottom 20% with the shortest product lifespan. Smartphone/microwave brands with a 'B', 'C' or 'D' ranking are between these two extremes. In your shopping tasks, **the lifespan ranking of all smartphone/ microwave brands will be shown as such:** 

[a]Smartphones only.
[b]Left out in the control condition.

TABLE A3 Examples of smartphone choice tasks used in the experiment, for participants who were randomly allocated to the lifespan sustainability labelling.

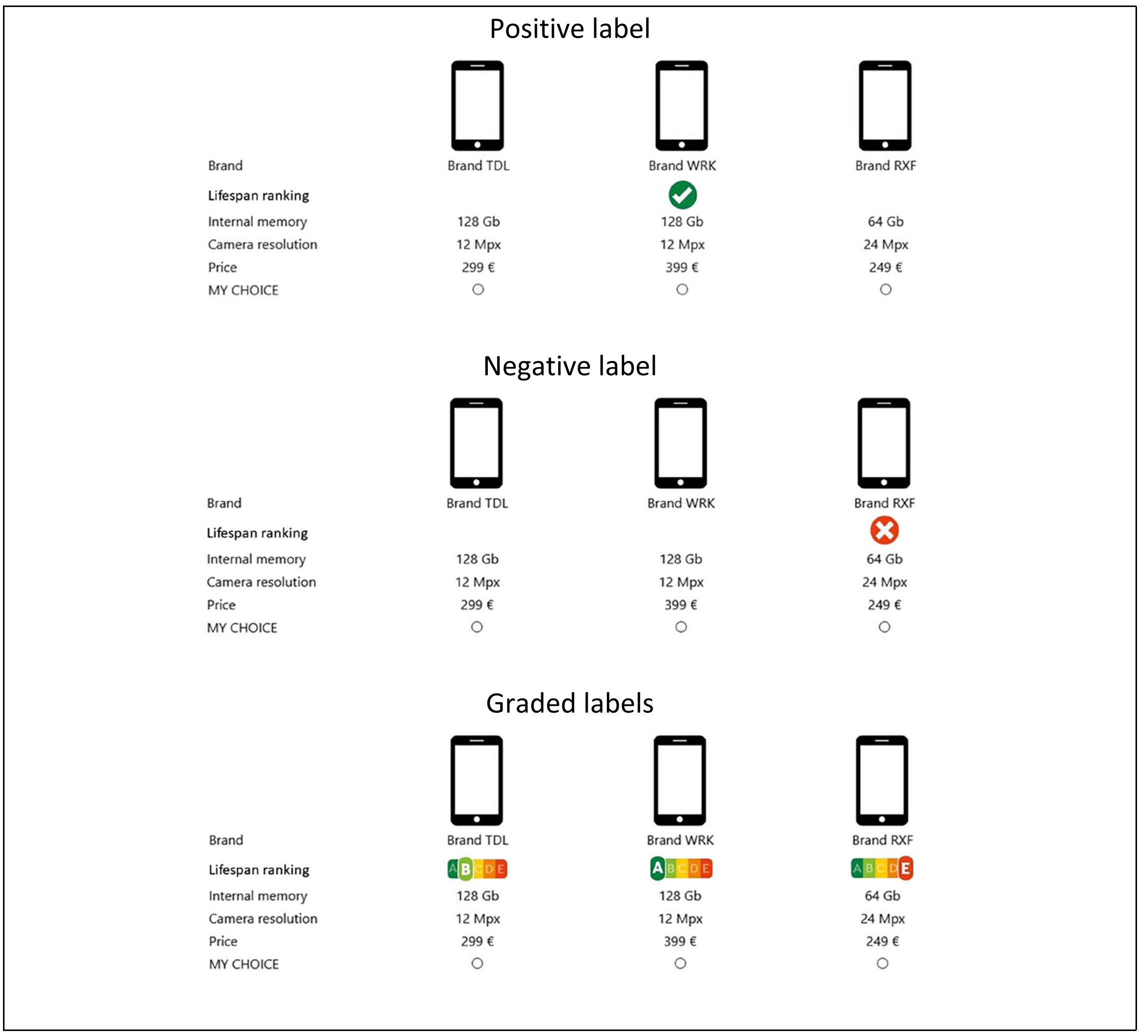


TABLE A4 Multinomial logit model estimates for the choice of smartphones and microwave ovens, experiment groups only ($n = 1120$).

	Smartphones			Microwave ovens		
	$R^2 = 0.17$			$R^2 = 0.19$		
Variables	*B*	Wald	Importance[b]	*B*	Wald	Importance[b]
Label[c]		1516.212***	0.34		1694.782***	0.34
A	0.71			0.74		
B	0.22			0.26		
C	0.18			0.19		
D	−0.29			−0.29		
E	−0.82			−0.89		
Memory/display[a]		724.156***	0.14		837.919***	0.15
Low	−0.32			−0.36		
High	0.32			0.36		
Camera/power[a]		932.427***	0.20		729.806***	0.17
Low	−0.46			−0.42		
High	0.46			0.42		
Price		1662.794***	0.31		1875.386***	0.34
Lowest	0.68			0.75		
Low	0.29			0.33		
High	−0.26			−0.18		
Highest	−0.71			−0.89		

[a]Memory size and camera quality for smartphones, digital display and max power for microwave ovens.
[b]The relative importance of product attributes in the choice experiment.
[c]The labelling attribute contains both fictitious brand names and label information.
****p* < .001.

TABLE A5 Multinomial logit model estimates for the choice of smartphones and microwave ovens, country comparisons ($n = 1120$).

Variables	Czech *B*	Germany *B*	Spain *B*	Wald	Wald(=)	Mean	Std. dev.
Smartphones	$R^2 = 0.22$	$R^2 = 0.12$	$R^2 = 0.20$				
Label[a]				1545.892***	13.745		
A	0.80	0.70	0.68			0.73	0.05
B	0.21	0.21	0.26			0.23	0.02
C	0.23	0.11	0.17			0.17	0.05
D	−0.34	−0.26	−0.24			−0.28	0.05
E	−0.89	−0.76	−0.88			−0.85	0.06
Memory				791.638***	84.855***		
Low	−0.35	−0.20	−0.47			−0.34	0.11
High	0.35	0.20	0.47			0.34	0.11
Camera				950.302***	30.535***		
Low	−0.53	−0.37	−0.55			−0.48	0.08
High	0.53	0.37	0.55			0.48	0.08
Price				1703.171***	222.762***		
Lowest	0.91	0.37	0.82			0.71	0.23
Low	0.33	0.17	0.44			0.31	0.11
High	−0.27	−0.19	−0.31			−0.26	0.05
Highest	−0.97	−0.35	−0.95			−0.76	0.28
Microwave ovens	$R^2 = 0.22$	$R^2 = 0.16$	$R^2 = 0.19$				
Label[a]				1718.544***	15.322		
A	0.71	0.74	0.79			0.75	0.03
B	0.30	0.20	0.29			0.26	0.04
C	0.20	0.20	0.14			0.18	0.03
D	−0.24	−0.35	−0.27			−0.29	0.04
E	−0.97	−0.80	−0.94			−0.91	0.07
Display				876.676***	52.734***		
Low	−0.44	−0.24	−0.42			−0.37	0.09
High	0.44	0.24	0.42			0.37	0.09
Power				752.192***	50.331***		
Low	−0.49	−0.28	−0.52			−0.43	0.11
High	0.49	0.28	0.52			0.43	0.11
Price				1884.014***	81.741***		
Lowest	0.89	0.66	0.73			0.76	0.10
Low	0.48	0.17	0.36			0.34	0.13
High	−0.26	−0.14	−0.13			−0.18	0.06
Highest	−1.11	−0.68	−0.96			−0.92	0.18

[a]The labelling attribute contains both fictitious brand names and label information.
***$p < .001$.

TABLE A6 Multinomial logit model estimates for the choice of smartphones and microwave ovens, sustainability dimension comparisons ($n = 1120$).

Variables	Ethical *B*	Lifespan *B*	Environmental *B*	Wald	Wald(=)	Mean	Std. dev.
Smartphones	$R^2 = 0.17$	$R^2 = 0.17$	$R^2 = 0.17$				
Label[a]				1518.634***	4.677		
A	0.73	0.75	0.67			0.71	0.03
B	0.21	0.25	0.20			0.22	0.02
C	0.17	0.17	0.19			0.18	0.01
D	−0.28	−0.29	−0.28			−0.29	0.01
E	−0.83	−0.88	−0.77			−0.83	0.04
Memory				724.597***	0.536		
Low	−0.33	−0.31	−0.33			−0.32	0.01
High	0.33	0.31	0.33			0.32	0.01
Camera				932.778***	0.844		
Low	−0.47	−0.47	−0.44			−0.46	0.01
High	0.47	0.47	0.44			0.46	0.01
Price				1665.260***	3.971		
Lowest	0.67	0.66	0.71			0.68	0.02
Low	0.32	0.27	0.30			0.29	0.02
High	−0.25	−0.23	−0.30			−0.26	0.03
Highest	−0.74	−0.70	−0.71			−0.71	0.02
Microwave ovens	$R^2 = 0.19$	$R^2 = 0.19$	$R^2 = 0.19$				
Label[a]				1698.790***	5.569		
A	0.69	0.79	0.74			0.74	0.04
B	0.24	0.23	0.31			0.26	0.03
C	0.20	0.21	0.16			0.19	0.02
D	−0.27	−0.31	−0.29			−0.29	0.02
E	−0.86	−0.91	−0.91			−0.89	0.02
Display				840.760***	4.520		
Low	−0.39	−0.37	−0.32			−0.36	0.03
High	0.39	0.37	0.32			0.36	0.03
Power				730.806***	1.875		
Low	−0.41	−0.45	−0.40			−0.42	0.02
High	0.41	0.45	0.40			0.42	0.02
Price				1878.574***	5.185		
Lowest	0.73	0.74	0.78			0.75	0.02
Low	0.35	0.35	0.28			0.33	0.03
High	−0.16	−0.22	−0.15			−0.18	0.03
Highest	−0.91	−0.86	−0.91			−0.89	0.02

[a]The labelling attribute contains both fictitious brand names and label information.
***$p < .001$.