





Cognitive bias can depend on eco-label design: Evidence from a negative footprint illusion in energy labeled household appliances

Patrik Sörqvist^{a,b,*} , Emil Skog^a 

^a Department of Health, Learning and Technology, Luleå University of Technology, Luleå, Sweden

^b Department of Building Engineering, Energy Systems, and Sustainability Science, University of Gävle, Gävle, Sweden

ARTICLE INFO

Keywords:
 Kilowatt hour
 Energy efficiency
 Scale
 Carbon footprint
 Ecodesign

ABSTRACT

Eco-labeling is the main policy instrument to communicate products' energy efficiency and environmental diagnosticity to consumers. Eco-labels can be graded, such as the European Union's (EU) new eco-design scale (with energy classes A, B, C, D, E, F, G) or the EU's old eco-design scale (with energy classes A+++ , A++ , A+ , A, B, C, D). Whether one scale is superior to the other in its effect on consumers is debated. The new scale tends to be understood as a linear interval scale, whereas the old scale tends to be perceived as skewed. We contribute to this debate by taking the novel approach of investigating how linear and skewed eco-labels elicit cognitive biases in consumers. We found that consumers perceive the energy demand (Experiment 1) and carbon footprint (Experiment 2) of a bundle containing an appliance with an intermediate energy class and an appliance with an excellent energy class to be lower than the energy demand/carbon footprint of the appliance with the intermediate energy class alone—a negative footprint illusion. A distortion of the mental organization of eco-labels (Experiment 2) showed that the magnitude of the negative footprint illusion depends dynamically on how eco-labels are organized in mental space, sometimes resulting in a greater effect with the skewed scale and sometimes a greater effect with the linear scale. The results stress the importance of communicating the energy efficiency of individual eco-label classes, as well as how their relation to other eco-label classes should be understood, to mitigate cognitive biases.

1. Introduction

Eco-labeling is the main policy instrument to communicate products' energy efficiency and environmental diagnosticity to consumers and helps end-users (consumers) to make energy-efficient and sustainable choices (Basiru et al., 2024). These eco-labels can either be positive (e.g., 'energy star', 'organic', 'fair trade') or graded (a product can be given a score that reflects its value relative to other products, e.g., nutrition labels with categorical letters A to E; Thøgersen et al., 2024; Zhong et al., 2025). Within the European Union (EU), for example, electronic products are labeled in accordance with the 'Ecodesign and Energy Labelling Directive (2009/125/EC) and Regulation (2017/1369)'. The eco-design label within EU is a graded scale with seven labels, used to classify household appliances (e.g., refrigerators, freezers, and washing machines) based on energy efficiency. In the scale that has long been in use (henceforth called the *old scale*), energy classes range the color-coded

letters A+++ , A++ , A+ , A, B, C, to D (Fig. 1, bottom row). In 2021, the EU decided to switch to a *new scale* that spans the color-coded range A, B, C, D, E, F, to G (Fig. 1, top row). This new scale is gradually replacing the old scale, although the old scale is presently still in use.

Whether one scale is superior to the other in its effect on consumers is debated (Heinzle and Wüstenhagen, 2012). The new scale seems to increase demand for the most energy efficient appliances (Faure et al., 2021) but also reduce intention to purchase energy efficient products (Beck and Toulouse, 2023). Here, we contribute to this debate by taking the novel approach of investigating how graded labels with different psychological characteristics elicit cognitive biases. We begin by outlining a theoretical framework and building hypotheses before reporting the present empirical work.

Sörqvist was supported by a grant from Stiftelsen Riksbankens Jubileumsfond (P23-0067) and Skog was supported by a grant from Kempestiftelsen (JCSMK23-0179). The study was also supported by a grant from Magnus Bergvalls stiftelse (2025-235).

* Corresponding author. Luleå University of Technology, Laboratorievägen 14, SE-971 87, Luleå, Sweden.

E-mail address: patrik.sorqvist@ltu.se (P. Sörqvist).

<https://doi.org/10.1016/j.enpol.2026.115391>

Received 12 September 2025; Received in revised form 12 May 2026; Accepted 13 May 2026

Available online 23 May 2026

0301-4215/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1.1. Theoretical framework

1.1.1. Perception of graded eco-labels

To understand how and why the new and old labeling scales differ in their tendencies to elicit cognitive biases, we need to consider how graded labels are perceived. Products are seen as having lower environmental friendliness when labeled with energy classes from the new (A to G) scale in comparison with the old (A+++ to D) scale (Stasiuk and Maison, 2022). An in-depth analysis showed that the reason for this is that the A-classes in the old scale (A+++ , A++ , A+ , and A) are all perceived as highly environmentally friendly, with only small differences between them (Skog and Sörqvist, 2026). While the relation between perceived environmental value and energy class is roughly linear in the new scale, with each scale step of about the same size, the function between perceived environmental value and energy class is skewed in the old scale (i.e., larger differences in perceived carbon footprint between low energy classes [e.g., C and D] and smaller differences between high energy classes [e.g., A+++ and A+]). Thus, the mental distance (or perceived difference) between the A-classes in the old scale is relatively small, while the distance between the lower energy classes (B, C and D) is relatively large, similar to the distance between all energy classes in the new scale.

What is the psychological basis of this difference in perception of the old and new energy scale? First, human judgment is shaped by heuristics (i.e., rules of thumb that simplify complexity) such as ‘representativeness’ (Tversky and Kahneman, 1974). For example, stimuli that appear representative of environmentally friendly objects will be judged as more environmentally friendly, even when such representativeness is misleading. Letter ‘A’ often represents ‘high’, and the old scale has several A-classes which inflate the perceived environmental friendliness of the scale (Skog and Sörqvist, 2026). Second, consumers tend to assign different weight to different cues when making judgments (Cox, 1962; Sullivan and Burger, 1987). Regarding EU’s ecodesign, the categorical letters are assigned more weight than colors (Skog and Sörqvist, 2026), which further contributes to the differences in perception of the old and new scales. This results in an important difference in how the intermediate energy classes of the two scales are perceived. The ‘yellow’ energy class (i.e., ‘A’ in the old scale, and ‘D’ in the new scale) has the rank four out of seven, in both scales, but is perceived as more environmentally friendly in the old scale than in the new scale due to the larger weight assigned to categorical letters than color-coding.

Finally, according to range-frequency theory (Parducci, 1965), the range of a scale is a strong determinant of psychological judgment, especially the endpoints. People use the endpoints as references when making judgments, which provide context and meaning to the to-be-estimated stimulus. For example, the energy class ‘D’ is the lowest class in the old scale but the middle class in the new scale. Such contextual knowledge about scale boundaries has a surprisingly small

effect on eco-judgment of energy classes, however, when the scales’ categorical letters begin with the alphabet (Skog and Sörqvist, 2026). It is only when the range of categorical letters is more arbitrary (e.g., ‘K’ to ‘Q’) that knowledge of the best and worst energy classes has a tangible effect on judgment.

In sum, judgments of energy classes appear mainly determined by the labels’ categorical letters, and the main reason why people judge the old and new ecodesign scales differently is the presence of four A-classes in the old scale, which inflate the perceived eco-value and make the top four energy classes in the old scale appear close in mental space. In other words, the new scale is organized more linearly, while the old scale is organized with a skew in mental space.

1.1.2. Cognitive bias in judgment of product bundles

Products are classified and assigned eco-labels individually, but they are not always bought in isolation. A typical shopping basket includes a mix of items—often with varying eco-labels. The way people perceive this mix might not necessarily reflect the sum of its individual components (Sörqvist et al., 2020). Similarly, electronic appliances, such as a refrigerator and a freezer, are often bought in pairs; a television might be bought together with a soundbar system; a washing machine might be bought in combination with a tumble dryer; and an oven might be bought in combination with a stove. A focus of the present study is therefore the perceptual shift of one electronic product to a bundle of two products.

Eco-labels can bias perception of an individual item in various ways (Hahnel et al., 2015; Lee et al., 2013; Moreno-Fernández et al., 2025). When products are combined, however, a specific cognitive effect called a *negative footprint illusion* can emerge (Gorissen and Weijters, 2016; Sörqvist et al., 2020): the footprint of an item with high carbon footprint (e.g., beef) and an item with low carbon footprint (e.g., an eco-labeled side dish) is perceived as lower than the footprint of the high carbon footprint item (the beef) alone—as if A + B would be less than A. This cognitive phenomenon applies to many stimulus contexts. For example, the carbon footprint of ‘green’ (low footprint) and ‘yellow’ (high footprint) buildings in combination is perceived as less than the carbon footprint of the ‘yellow’ buildings alone (Andersson et al., 2024; Holmgren, Andersson et al., 2018; Sörqvist et al., 2022); the carbon footprint of ‘organic’ and ‘conventional’ food is perceived as less than the ‘conventional’ food alone (Gorissen and Weijters, 2016; Kusch and Fiebelkorn, 2019); and ‘hybrid’ and ‘petrol’ cars are seen as having a smaller carbon footprint than the ‘petrol’ cars alone (Holmgren et al., 2021). The negative footprint illusion is seemingly robust to individual differences in attitudes (Threadgold et al., 2021) and expertise (Holmgren, Kabanshi et al., 2018) and is quite resistant to debiasing techniques (Holmgren et al., 2021) and surrounding contextual information (Skog et al., 2026). As such, the effect seems to be a basic and general phenomenon with a marked impact on how consumers perceive

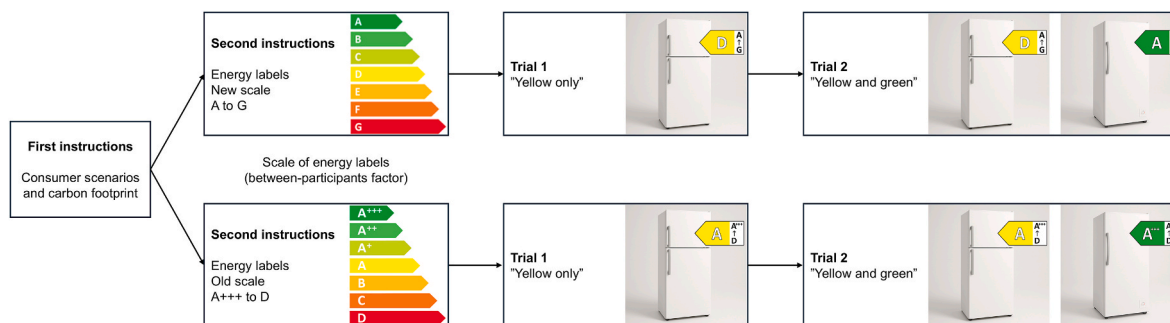


Fig. 1. Design and procedure. The figure represents central features of the experimental design and procedure. Top and bottom rows of boxes show differences across the between-participants factor ‘Scale of energy labels’, where some participants were shown the new scale (A to G) and others the old scale (A+++ to D). The stimulus arrangement of the two trials of the negative footprint illusion task is illustrated. Notice that the white goods contain different energy labels, drawn from the two scales.

the environmental diagnosticity of their purchase decisions.

Furthermore, the effect seems relatively independent on how the task and the judgment scale are framed. For instance, the negative footprint illusion is found when people are asked to rate items on a scale ranging from 'low carbon footprint' to 'high carbon footprint' (Gorissen and Weijters, 2016), when asked to assign a quantity of kilogram CO₂ associated with the construction of buildings (Sörqvist and Holmgren, 2022), and when asked to guess how many trees that would be needed to compensate for a building's CO₂ emissions through a carbon binding process (Holmgren, Kabanshi et al., 2018). The reason for this independence appears to be an attribute substitution process (Kahneman and Frederick, 2001), whereby the observer does not respond to the actual question they are faced with, but instead (subconsciously) replaces the relatively complex task of estimating the items' carbon footprint with the simpler task of estimating the items' environmental friendliness (Sörqvist and Marsh, 2024). This attribute substitution process leads to an averaging bias—a tendency to think in terms of the average rather than the sum of multiple items—which appears to explain why the negative footprint illusion emerges (Andersson et al., 2024; Holmgren, Andersson et al., 2018; Sörqvist and Marsh, 2024). When faced with the task to estimate the carbon footprint of a bundle of 'low footprint' and 'high footprint' items, people tend to average the items' carbon footprint instead of adding them together, as this average corresponds to the combined items' environmental friendliness.

1.2. Hypothesis building and the present study

In the present paper, we address how and why graded eco-labels produce cognitive biases of different magnitudes and test whether the negative footprint illusion—usually studied in the context of carbon footprint judgments (Gorissen and Weijters, 2016; Holmgren, Andersson et al., 2018)—can be extended to judgments of energy demand. Specifically, we compared EU's old (skewed) and new (linear) ecodesign scales, and test a series of hypotheses that build on psychological theory concerning how energy labels are perceived and how people judge item bundles. First, we expected to find a negative footprint illusion in judgment of electronic appliances, where a bundle with a 'yellow and green' item should produce lower 'kilowatt-hour demand' estimates than the 'yellow' item alone. Our hypotheses predicted this bias to emerge with labels from the old scale (Hypothesis 1) and the new scale (Hypothesis 2). Furthermore, we expected to find a larger negative footprint illusion with the new scale (Hypothesis 3) because the perceived difference (or mental distance) between the intermediate and the best energy class should be larger in the new scale ('A' and 'D') and smaller in the old scale ('A+++)' and 'A') (Skog and Sörqvist, 2026).

In Experiment 1, we requested participants to rate the eco-labels from the new and old ecodesign scales, to establish a direct measure of mental distance between energy classes. This rating confirmed that the new ecodesign is perceived as a linear scale in mental space, whereas the old scale is skewed in mental space, compressed in the upper energy classes. Experiment 1 then moved on to test Hypotheses 1-3 by asking participants to first rate the energy demand of an appliance with a 'yellow' energy class (either 'A' or 'D' depending on condition) and then rate the energy demand of a bundle of two appliances, one with a 'yellow' and one with a 'green' (either 'A+++)' or 'A' depending on condition) energy class, respectively. All three hypotheses were supported.

Experiment 2 tested whether the main findings from Experiment 1 generalize to carbon footprint judgments—the standard dimension of judgment in the negative footprint illusion paradigm. The results pattern was like that of Experiment 1 and provided further support for Hypotheses 1-3. In a different condition, Experiment 2 also explored the consequences of distorting the organization of eco-labels in mental space by guiding perceptions of the 'yellow' energy class. This revealed that the magnitude of the negative footprint illusion depends on how people understand energy classes. The 'yellow' energy class of the old scale (i.e., 'A') was perceived as quite low in energy demand/carbon footprint

when participants made judgments without guidance, resulting in a smaller negative footprint illusion when a 'green' energy class was added to the set. However, when participants were told that the 'yellow' energy class should be seen as having an intermediate carbon footprint, adding an appliance with a 'green' ('A+++)' energy class produced a larger negative footprint illusion.

2. Experiment 1

2.1. Method

2.1.1. Participants

Participants were recruited via Prolific (www.prolific.com) in this internet-based study. The sample size was determined with an a-priori power analysis in G*Power (Faul et al., 2007). The targeted effect size was estimated based on Sörqvist et al. (2022), who in a methodologically similar experiment found that stimulus arrangement can modulate the negative footprint illusion with a Cohen's $d = .32$ (Cohen's d was estimated using a student's t -test based on a re-analysis of Sörqvist et al.'s data). With a one-tailed hypothesis, $\alpha = .05$, and power ($1 - \beta$ error probability) = .80, the required sample size was $N = 122$ in each scale group. With two scale groups, the final sample comprised 244 participants (133 female, 110 male, and 1 who preferred not to say, Mean age = 45.36 years, $SD = 13.54$). The study was conducted in accordance with the Declaration of Helsinki and conformed to the guidelines of the Swedish Research Ethics Authority (Dnr, 2024-05795-01). All participants resided in the UK and had English as their first language. They reported normal or corrected-to-normal vision and no color-blindness. Participants were compensated with £.9 for a total study time of approximately 6 min, amounting to a per-hour rate of £9.

2.1.2. Materials

The experiment was created in PsychoPy (version 2024.2.2, Peirce et al., 2019) and made available online on Pavlovia. The experiment ran in the participants' web-browsers on their own desktop computers in an uncontrolled environment.

In this experiment, participants conducted two tasks: a 'class arrow' eco-label rating task and a negative footprint illusion task. In the class arrow eco-label rating task, the seven eco-labels from the ecodesign scales (made available by the EU on the file exchange tool CIRCABC) were presented, one-by-one, on the computer screen (see Fig. 1 for examples of the yellow and the green class arrow eco-label). Half of the participants saw labels from the old scale, the other half saw labels from the new scale. The task was conducted in three blocks. In one block, participants rated each eco-label on perceived 'energy efficiency'. The seven eco-labels were presented in random order, and ratings were made by clicking a slider located at the bottom of the screen. The slider contained eleven response categories. In addition, the left- and right-most slider response options stated 'Not energy efficient' and 'Very energy efficient', respectively. In a second block, the same class arrow eco-labels were presented in new random order, but here participants rated each eco-label on perceived 'environmental friendliness'. Rating responses were given on a similar slider, but the left- and right-most response options stated 'Not environmentally friendly' and 'Very environmentally friendly'. In a third block, the class arrow eco-labels were rated along perceived 'sustainability', with the left- and right-most slider response options stating 'Not sustainable' and 'Very sustainable'. The order of the three blocks was randomized anew for each participant. A composite score (the average of the three ratings of each eco-label, respectively) was used as the dependent variable in analysis of the rating task.

In the second task (the negative footprint illusion task), participants were presented with two images of white goods and two energy labels (Fig. 1). An image of a refrigerator and an image of a freezer were created with the assistance of a web-based artificial intelligence service.

These appliances were paired with class arrow energy labels. The yellow labels, drawn from each scale, were superimposed within the image of the refrigerator, and the green labels were superimposed within the image of the freezer (see the rightmost boxes in Fig. 1). The size of the stimulus images (refrigerator and freezer) was scaled to 40% of the participant's monitor's height (1:1 aspect ratio). Text was displayed above the refrigerator stating 'Refrigerator', and text above the freezer said 'Freezer'. The first trial displayed a 'yellow only' stimulus arrangement—the refrigerator with a yellow energy label—shown in the middle of the screen. Participants were instructed to rate the energy demand of the refrigerator (kilo-watt hours [kWh] per year). kWh ratings were made by clicking a slider located at the bottom of the screen. The slider contained nine categories that were labeled with numbers 1–9. In addition, the left- and right-most slider response options stated 'Many kWh' and 'Few kWh', respectively. The second trial displayed a 'yellow and green' stimulus arrangement, with the refrigerator (with a yellow label) on the left and the freezer (with a green label) on the right (Fig. 1; 'Trial 2'). Participants were instructed to rate the energy demand of both products (kilo-watt hours [kWh] per year). Our design with one 'yellow only' trial that will be compared to one 'yellow and green' trial follows a typical design style within the negative footprint illusion paradigm (e.g., Andersson et al., 2024), where a trial with a 'conventional' or 'medium' stimulus (yellow eco-label) is compared to a trial which contains both a conventional and an 'eco-friendly' stimulus (yellow and green eco-label). As both trials contain the yellow stimulus, the design focuses on what happens when a 'green addition' is introduced.

2.1.3. Design and procedure

A between-subjects design was used wherein half of all participants were shown eco-labels from the new scale (A to G), and the other half were shown eco-labels from the old scale (A+++ to D). The class arrow rating task was subjected to a 2 (Scale of energy labels: A to G, A+++ to D [between-participants factor]) \times 7 (energy class [within-subjects factor]) analysis. The negative footprint illusion task was subjected to a 2 (Scale of energy labels: A to G, A+++ to D [between-participants factor]) \times 2 (Trial: yellow only, yellow and green [within-participants factor]) analysis.

Participants provided informed consent prior to beginning the experiment. Participants were then presented with an instruction screen which explained how energy labels represent different levels of energy efficiency. This was followed by the class arrow eco-label rating task. After making the ratings in three blocks, participants were presented with a new instruction screen. Here, participants were told that they would evaluate energy demand (kilo-watt hours) in two separate consumer scenarios involving electronic products. Participants were also told that a lower energy demand is relatively better for the environment, and a higher energy demand is relatively worse for the environment. Participants then completed the two trials of the negative footprint illusion task. Participants were debriefed upon completion.

2.1.4. Data availability and transparency

We declare our hypotheses, methods, analyses (statistical tests conducted in Jamovi, version 2.3.28.0), results, and that neither experiment was preregistered. The data supporting this article are publicly available via the Open Science Framework (<https://doi.org/10.17605/OSF.IO/W6QY3>).

2.2. Results and discussion

2.2.1. Class arrow eco-label rating task

The full results from the class arrow eco-label rating task are presented in Fig. 2. The eco-labels from the new scale were discriminated in a linear fashion, with even spaces between energy classes. In contrast, the ratings of the eco-labels from the old scale formed a skewed response function. The mental distance between higher energy classes in the old scale was relatively small, whereas the mental distance between lower energy classes was relatively large. This was confirmed by a repeated measures 2 (Scale: new, old) \times 7 (energy class) analysis of variance (ANOVA) with Greenhouse-Geisser correction. The analysis revealed a significant main effect of scale, $F(1, 242) = 392, p < .001, \eta^2_p = .618$, a significant main effect of energy class, $F(3.04, 734.89) = 3601, p < .001, \eta^2_p = .937$, and a significant interaction between the factors, $F(3.04, 734.89) = 119, p < .001, \eta^2_p = .330$. Thus, the findings from this task confirm that the energy classes in the old scale and the energy classes in the new scale are organized differently in mental space, linearly in the

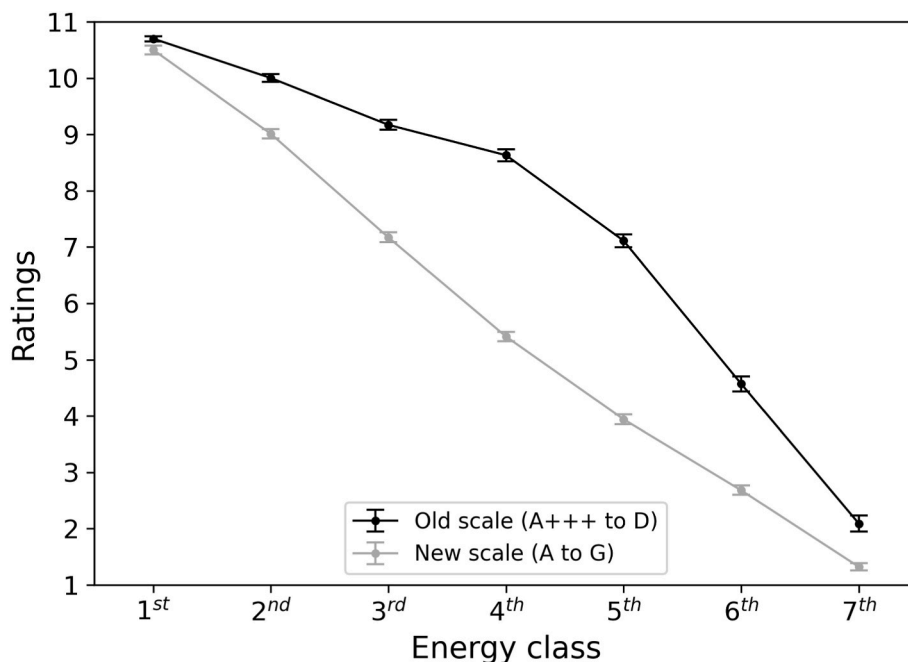


Fig. 2. Results of the class arrow eco-label rating task in Experiment 1. Means represent composite scores from three measures (ratings of energy efficiency, environmental friendliness and sustainability). Error bars represent the standard errors of the means.

new scale and skewed in the old scale.

2.2.2. Negative footprint illusion task

We expected to discover a negative footprint illusion in both scales (Hypotheses 1 and 2), and that it would be stronger for the new scale (Hypothesis 3). As can be seen in Fig. 3, these hypotheses were all supported. For this analysis, we used a repeated measures 2 (Trial: yellow only, yellow and green) \times 2 (Scale: new, old) analysis of variance (ANOVA). The analysis revealed a significant main effect of scale, $F(1, 242) = 72.0, p < .001, \eta^2_p = .229$, a significant main effect of trial, $F(1, 242) = 69.9, p < .001, \eta^2_p = .224$, and a significant interaction between the factors, $F(1, 242) = 13.3, p < .001, \eta^2_p = .052$. Pair-wise comparisons confirmed that the negative footprint illusion was present for both scales (new A to G scale: $W(121) = 616, p < .001$, rank-biserial correlation = .746; old A+++ to D scale: $W(121) = 951, p < .001$, rank-biserial correlation = .441). Also, energy classes in the new scale were given a higher kWh demand rating overall, as initially anticipated and aligning with previous research (e.g., Skog and Sörqvist, 2026).

3. Experiment 2

Experiment 1 confirmed that the new ecodesign scale is linearly perceived and the old scale is perceived with a skew, consistent with past research (Skog and Sörqvist, 2026). This provides a manipulation check of the key construct of psychological distance between energy classes. Experiment 1 also established, for the first time, a negative footprint illusion in judgment of energy demand (i.e., kilowatt-hours). The negative footprint illusion is, however, usually studied in relation to carbon footprint estimates (see, e.g., Andersson et al., 2024; Gorissen and Weijters, 2016; Kusch and Fiebelkorn, 2019). One purpose of Experiment 2 was to conduct a conceptual replication of the results from Experiment 1 and test if they generalize to carbon footprint judgments. This is important because it might reveal potential differences or similarities between judgmental dimensions. Furthermore, perceived carbon footprint has a strong influence on product demand (Andrade and Vieites, 2025; Baca and Reshidi, 2025) and broadly applies to all eco-labeled products (including energy labels), thus increasing the relevance of the results beyond energy labels.

Another purpose of Experiment 2 was to investigate what will happen if the organization of energy classes in mental space is distorted,

by guiding participants to perceive the 'yellow' energy classes of the two scales as having the same intermediate level of sustainability. As seen in Experiment 1, people spontaneously perceive the yellow 'A' class in the old scale as low and the yellow 'D' class in the new scale as high in energy demand, but what will happen to the negative footprint illusion if these perceptions are aligned? While the mental distance (or perceived difference) is small between 'A' and 'A+++' in the old scale when people make unguided interpretations of energy classes (Fig. 2), the mental distance between 'A' and 'A+++' might appear larger if people are told to consider that energy class 'A' has an intermediate carbon footprint. This guidance might make the negative footprint illusion larger in the old scale and reveal new dynamics on how energy classes are interpreted and influenced by e.g., consumer guidance and communication.

3.1. Method

3.1.1. Participants

Participants were recruited via Prolific (www.prolific.com) in this internet-based study and were randomly assigned to two sub-experiments. In the first sub-experiment (140 female, 104 male, Mean age = 43.9 years, $SD = 12.6$), participants were not instructed what carbon footprint to associate with the 'yellow' energy class (henceforth called the "unguided sub-experiment"). In the second sub-experiment (133 female, 110 male, 1 did not reveal their sex, Mean age = 46.1 years, $SD = 14.1$), participants were told to associate the 'yellow' energy class with an intermediate carbon footprint (henceforth called the "guided 'yellow is intermediate' sub-experiment"). Two participants were replaced and their data excluded from analysis due to them not completing the full experiment. This resulted in a final sample of 488 participants. The sample size was based on the same power analysis as Experiment 1. The study was conducted in accordance with the Declaration of Helsinki and conformed to the guidelines of the Swedish Research Ethics Authority (Dnr, 2024-05795-01). All participants resided in the UK and had English as their first language. They reported normal or corrected-to-normal vision and no color-blindness. Participants were compensated with £.3 for a total study time of approximately 2 min, amounting to a per-hour rate of £9.

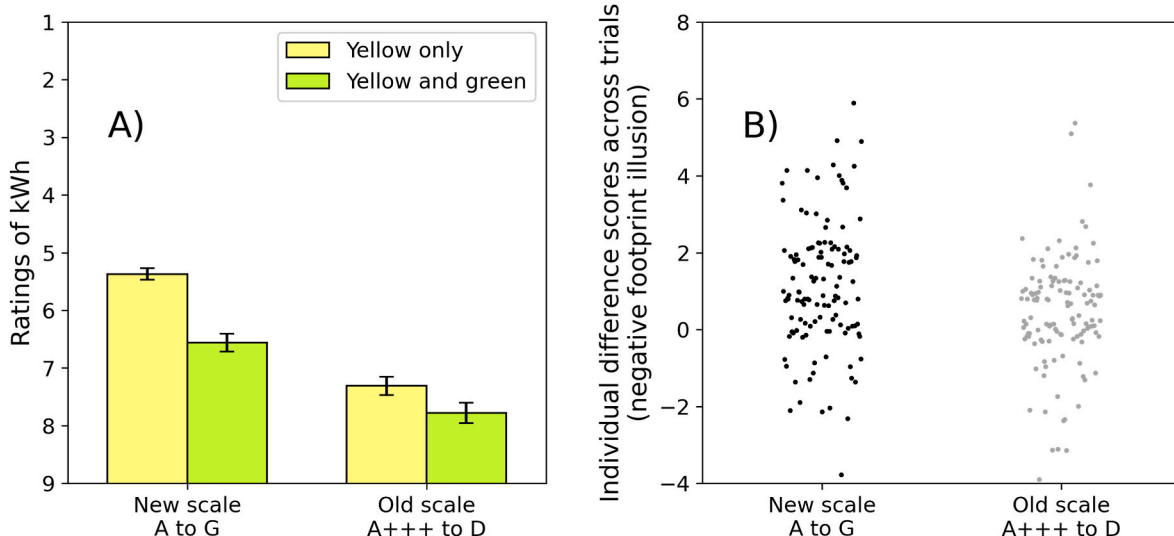


Fig. 3. Results of the negative footprint illusion task in Experiment 1. Left panel (A) shows estimates of sets with either one 'yellow' appliance or two 'yellow + green' appliances (error bars represent the standard errors of the means). The y-axis is flipped to help interpretation, as higher values represent less energy demand. Right panel (B) shows difference scores across trials ('yellow only' subtracted by 'yellow and green'; with data jitter) for each participant, revealing the size of the negative footprint illusion (>0 indicates that a green addition reduced the perceived energy demand). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

3.1.2. Materials, design and procedure

The present study consists of two sub-experiments—the unguided sub-experiment and the guided ‘yellow is intermediate’ sub-experiment—that shared largely similar methods. Both sub-experiments were mixed 2 (Trial: yellow only, yellow and green [within-participants factor]) × 2 (Scale of energy labels: A to G, A+++ to D [between-participants factor]) designs. Key features of the design and the experimental procedure are summarized in Fig. 1.

The materials and procedure for the unguided sub-experiment were identical to that of the negative footprint illusion task of Experiment 1 (see Fig. 1), but with the exception that participants made carbon footprint estimates instead of kilowatt hour estimates. Carbon footprint ratings were made by clicking a slider located at the bottom of the screen. The slider contained nine categories that were labeled with numbers 1-9. In addition, the left- and right-most slider response options stated ‘Low carbon footprint’ and ‘High carbon footprint’, respectively. In the unguided sub-experiment, participants were not given any specific instructions on where to place their responses on the slider. But in the guided ‘yellow is intermediate’ sub-experiment, participants were

given additional instructions on how to think about the first trial stimulus and on how to respond to it. Here, participants were told: ‘In this scenario, please select ‘5’ as your response. Please consider this refrigerator to have a medium carbon footprint. Your response (‘5’) will then represent the middle of the carbon footprint scale.’ Before the second trial in the guided sub-experiment, participants were told to respond freely and that there were now no specific instructions on how to respond. These instructions constituted the only difference across the two sub-experiments.

Participants provided informed consent prior to beginning the experiment. Participants were initially told that they would evaluate environmental impact (carbon footprint) in two separate consumer scenarios. Here, participants were also told that a lower carbon footprint is relatively better for the environment, and a higher carbon footprint is relatively worse for the environment. The second instruction screen told participants about energy efficiency, it showed a scale of energy classes (Fig. 1), and it stated what the most and least energy efficient classes were. Participants were shown the full range of the scale since this can help to organize meaningful ratings (e.g., Skog and Sörqvist, 2026).

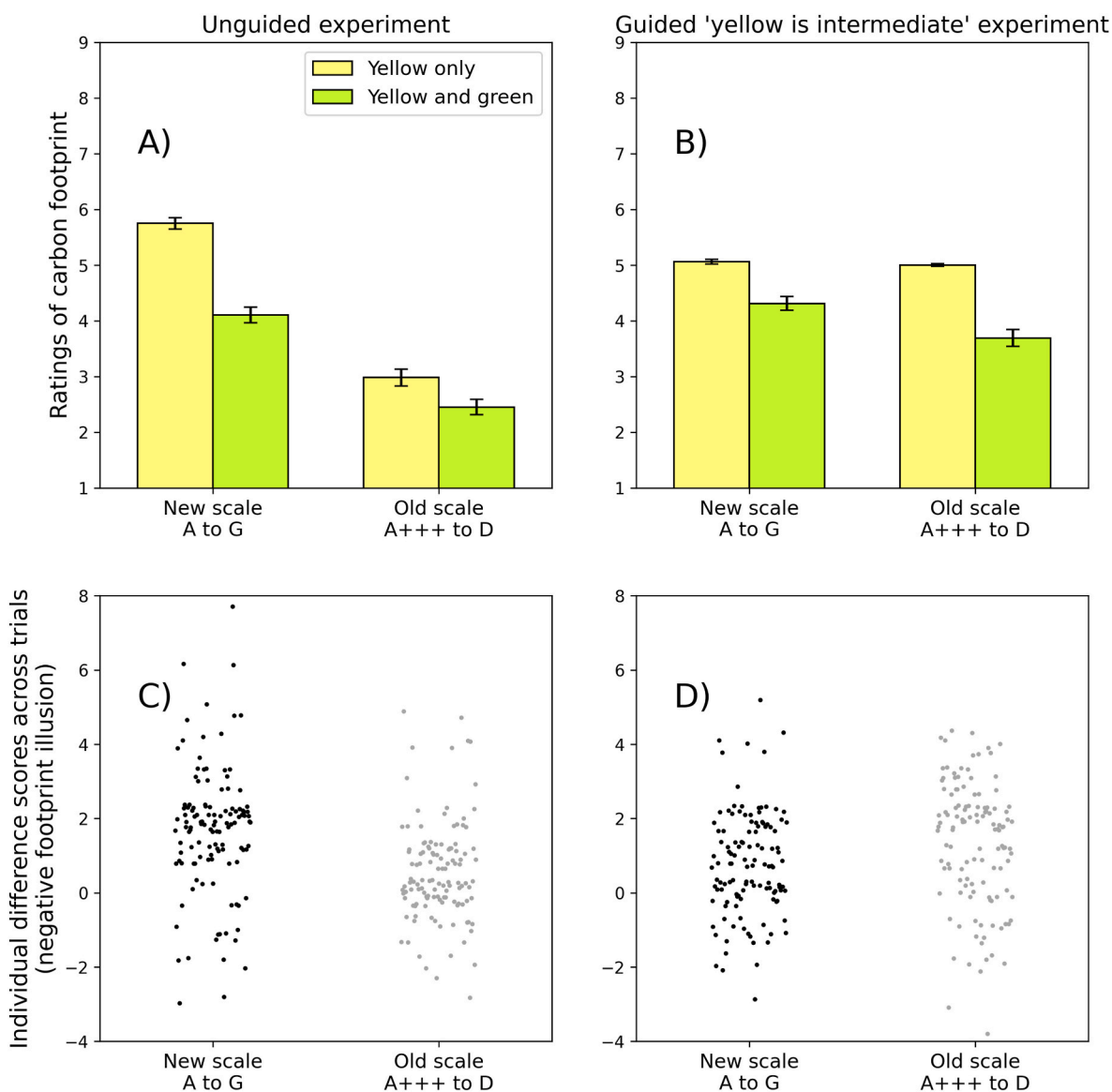


Fig. 4. Full results of Experiment 2. Top panels show the main results of both sub-experiments (error bars represent the standard errors of the means). Bottom panels show difference scores across trials (‘yellow only’ subtracted by ‘yellow and green’; with data jitter) for each participant, revealing the size of the negative footprint illusion (>0 indicates that a green addition reduced the perceived carbon footprint). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

3.2. Results and discussion

Each participant provided responses to two trials, one 'yellow only' trial and one 'yellow and green' trial. Fig. 4 displays these data in panels A and B for the unguided and guided 'yellow is intermediate' sub-experiments, respectively. The difference scores across these two trials reveals the negative footprint illusion, and the individual difference scores are shown in panels C and D in Fig. 4, for each sub-experiment, respectively. Our two sub-experiments were conducted sequentially. The results of the first sub-experiment inspired the second, and we thus present the results in sequence, followed by a joint discussion of both sub-experiments.

3.2.1. Results of the unguided sub-experiment

As in Experiment 1, we expected to discover a negative footprint illusion in both scales (Hypotheses 1 and 2), and that it would be stronger for the new scale (Hypothesis 3). For this analysis, we used a repeated measures 2 (Trial: yellow only, yellow and green) \times 2 (Scale: new, old) analysis of variance (ANOVA). A main effect of trial, $F(1, 242) = 120.2, p < .001, \eta^2_p = .332$, revealed a negative footprint illusion where adding a green item reduced carbon footprint ratings (Fig. 4A and C). Pair-wise comparisons confirmed that the negative footprint illusion was present for both scales (new A to G scale: $W(121) = 5897, p < .001$, rank-biserial correlation = .831; old A+++ to D scale: $W(121) = 2142, p < .001$, rank-biserial correlation = .503).

This negative footprint illusion was different in magnitude across the two scales, as confirmed by a significant interaction across trial and scale, $F(1, 242) = 32.3, p < .001, \eta^2_p = .118$. There was a smaller (negative footprint illusion) effect of adding a green item in the context of the old scale than the new scale (Figs. 4A and 4C).

There was also a main effect of scale (see Fig. 4A), $F(1, 242) = 186.0, p < .001, \eta^2_p = .435$. Energy classes in the new scale were given a higher carbon footprint rating overall. This motivated a test of the difference between the scales when the 'yellow' energy classes' carbon footprints were perceived the same.

3.2.2. Results of the guided 'yellow is intermediate' sub-experiment

Here, participants were asked in the first trial to think of the yellow-labeled product as having a medium carbon footprint, and they were asked to give a '5' response, representing the middle of the response scale. We used the same 2 \times 2 ANOVA as for the first subset of the study and found a main effect of trial, $F(1, 242) = 108.24, p < .001, \eta^2_p = .309$, replicating the expected negative footprint illusion (Figs. 4B and 4D), for both scales (new A to G scale: $W(121) = 3078, p < .001$, rank-biserial correlation = .646; old A+++ to D scale: $W(121) = 5060, p < .001$, rank-biserial correlation = .751).

The ANOVA also revealed a significant main effect of scale, $F(1, 242) = 10.7, p = .001, \eta^2_p = .042$, and an interaction between scale and trial, $F(1, 242) = 8.18, p = .005, \eta^2_p = .033$. The interaction was reversed compared to the first 'unguided' sub-experiment. The negative footprint illusion was greater with the old scale than the new scale, whereas the opposite was true when perception of the 'yellow' energy class was unaligned, in the unguided sub-experiment (Fig. 4). The significance of the reversal across the three factors was confirmed by a 2 \times 2 \times 2 ANOVA, where the key three-way interaction was significant, $F(1, 484) = 36.56, p < .001, \eta^2_p = .070$. The size of the negative footprint illusion changed across the sub-experiments for both scales, as revealed by trial \times sub-experiment interactions for each scale individually analyzed (new A to G scale: $F[1, 242] = 20.5, p < .001, \eta^2_p = .078$; old A+++ to D scale: $F[1, 242] = 16.1, p < .001, \eta^2_p = .062$).

4. General discussion

The negative footprint illusion was always of the same kind and found for EU's new (linear) and old (skewed) ecodesign, confirming Hypothesis 1 and 2. Both scales elicited cognitive bias. These findings

align with the growing body of studies on the negative footprint illusion (e.g., Andersson et al., 2024; Gorissen and Weijters, 2016; Holmgren, Kabanshi et al., 2018; Kusch and Fiebelkorn, 2019) and demonstrate for the first time that energy labeled appliances can produce the effect, both for judgments of energy demand (Experiment 1) and carbon footprint (Experiment 2). Furthermore, the bias was larger in magnitude in the new (linear) scale, confirming Hypothesis 3. However, the size of the negative footprint illusion was instead larger in the old scale when participants' interpretation of the energy classes were manipulated in the guided sub-experiment of Experiment 2 (showing that Hypothesis 3 only holds in some conditions). To understand why the negative footprint illusion differs in the two energy scales, we must consider how eco-labels are organized in mental space.

The eco-labels in the new scale are linearly organized, with even steps between energy classes (Fig. 2). The eco-labels in the old scale, in contrast, are organized with a skew. Here, the perceived distance between the top energy classes is compressed. In view of the statistical magnitude of these effects, the difference in mental organization is very large. We argue this is the mechanism responsible for the larger negative footprint illusion for the new scale (in comparison with the old scale) seen in energy demand estimates (Fig. 3) and in carbon footprint estimates (Fig. 4). While the negative footprint illusion is highly statistically reliable in both scales, it should be noted that the difference in magnitude was quite small.

Our mental space framework can also help explain why the magnitude of the negative footprint illusion varies when the eco-label organization is disrupted (as in the guided sub-experiment of Experiment 2). The negative footprint illusion was smaller in the old (skewed) scale (A+++ to D) with an unguided interpretation of the 'yellow' energy class (Figs. 3 and 4A, right side) compared to when participants were told to interpret the 'yellow' energy class as having a carbon footprint of '5' (Fig. 4B, right side). In the unguided experiments (Figs. 3 and 4A), participants gave the yellow 'A' stimulus a low kilowatt hour and carbon footprint rating. Here, adding a green 'A+++' stimulus to the set had a smaller effect on ratings, because the perceived difference (or mental distance) between the 'A' and the 'A+++' energy classes is small (Fig. 2). When participants' interpretation of the 'yellow' energy class was instead guided to a comparably high starting point (Fig. 4B), the green 'A+++' addition acted as a very strong stimulus, producing a larger shift in perceived carbon footprint, and thus a larger negative footprint illusion. The magnitude of the cognitive bias seems therefore to depend dynamically on the mental representation of the single 'yellow' stimulus and the mental distance between the yellow energy class and the green energy class. This distance (from a yellow 'A' to a green 'A+++') is small when consumers can interpret the energy classes freely, but large when the mental representation of the 'yellow' item is guided to a relatively high value.

Conversely, with the new (linear) scale (A to G), the negative footprint illusion was larger in the unguided experiments (Figs. 3 and 4A, left side) compared to the guided experiment (Fig. 4B, left side). This finding highlights the importance of clearly communicating the energy efficiency and environmental diagnosticity of individual energy classes, as well as how their relation to other energy classes should be understood, to mitigate cognitive biases with the current scale of energy labels. Our experimental manipulation of guidance reveals that people's interpretations of energy classes vary dynamically with how they evaluate different steps in the scales. And this depends on which scale is used.

4.1. Policy implications

Our results contribute to the debate about the effects of the EU's new and old energy scale on consumer behavior (Beck and Toulouse, 2023; Faure et al., 2021; Heinzle and Wüstenhagen, 2012; Skog and Sörqvist, 2026; Stasiuk and Maison, 2022). First, the fact that there is a smaller mental distance between the middle and the top of the old scale,

compared to the new scale, suggests that the old scale carries a problematic feature where it conveys a misinformed sense that products belonging to a middle energy class have relatively low energy demands/carbon footprints. Our data also show that the new scale avoids this problem, as the middle energy class was not considered to have a low energy demand/carbon footprint. These results support the EU's policy direction to discontinue the old scale in favor of the new.

More generally, our results also speak broadly to eco-label design. The present results show that categorical letters gain more weight than color-coding in perception of eco-labels. Designers of eco-labels should thus take this into consideration. Categorical letters (A to G), color-coding (green to red) and special symbols ('+' signs) can all be used to distinguish steps in graded labels. Our results suggest that graded labels should have different categorical letters for each scale step, even if (or instead of) other features such as '+' signs and color-coding are used to distinguish between steps. This is because identical categorical letters might make the steps appear more psychologically similar than intended. We can also speculate on the consequences of creating hypothetically non-linear scales with different categorical letters for each scale step. One example could be a scale with letters 'S', 'A', 'B', 'C', 'D', where 'S' would represent the best (superior) class. With these categorical letters, the mental distance between 'S' and 'A' might not be the same as the difference between the other scale steps, as it does not map onto the alphabetical order structure. Future research could address the psychological consequences of such eco-label scale designs.

Another policy implication of the present study concerns communication strategies. While cognitive bias is present for graded scales with linear and skewed classes, the magnitude of it depends dynamically on scale characteristics and how information about classes is conveyed. People come to form different evaluations of energy labeled products depending on the energy scale and their own understanding of the different steps in the scale. For example, the negative footprint illusion is attenuated for the new scale when people are instructed to consider the midpoint of the scale as having a medium carbon footprint. Methods to help consumers organize their understanding of energy classes can therefore mitigate bias when people judge the new energy scale. This finding highlights the importance of providing consumers with appropriate reference points when marketing appliances with energy labels, especially when products are put up for sale in bundles. The new scale is increasingly implemented on the market, and we see opportunities to mitigate bias when consumers are more informed about the scale and what different energy classes signify. One communication strategy could be to complement energy labels of individual appliances with the addition of eco-labels for bundled products, that clearly communicate the combined products' summative (not average) carbon footprint and energy demand. The potential advantage of such complementary information was not tested in the present study and is therefore a target for future research.

4.2. Limitations

We note that a floor effect with the old scale could have contributed to the smaller negative footprint illusion in the unguided experiments. With a low rating for the yellow stimulus, there is less room to lower ratings in the second trial. This could exaggerate the difference between the two scales in Experiment 1 and the unguided sub-experiment of Experiment 2.

We also note a potential confound in the guided sub-experiment of Experiment 2. The guided condition not only changes participants' interpretation of the yellow label, but also explicitly fixes the response on Trial 1 (see Fig. 1) at '5' (the middle of the response scale). This means that the experimental guidance manipulation combines interpretive guidance with a fixed baseline response, making it difficult to determine whether the now larger negative footprint illusion in the old scale is driven by a change in mental organization of energy classes, or by a baseline-fixing, anchoring, or a demand effect. Future research can

aim to disentangle these alternatives.

Furthermore, we also note that the abstract and ambiguous task of estimating kilowatt hours/carbon footprint on a scale ranging from 'few/low' to 'many/high' can be questioned. Future research should address whether the negative footprint illusion is also found when the task is made more quantitative and more concrete. As the negative footprint illusion can be found with quantitative judgment scales (Sörqvist and Holmgren, 2022), we would expect a pattern of results like the present study. For example, we would expect people to estimate that fewer batteries are needed to run a set of 'yellow and green' battery-dependent appliances together, in comparison with the number of batteries needed to run the 'yellow' battery-dependent appliances alone, as long as attribute substitution drives judgments (cf. Sörqvist and Marsh, 2024).

We also acknowledge that the external validity of the present study can be questioned. First, only one type of bundle was tested, and with only two electronic appliances. The negative footprint illusion has been found across many types of products and objects (e.g., food, buildings, vehicles, and now also in electronic appliances), but whether the results revealed here generalize to other electronic appliances remain unexplored and should not be taken for granted before being tested empirically. Furthermore, while the stimulus arrangement resembles the way products are displayed in online web shops, in real-world appliance purchases, consumers typically consider multiple attributes simultaneously (e.g., price, brand, size, and usage frequency), whereas the experiments deliberately isolated energy labels. As associations with brand and price, for example, might influence judgments, these factors might both strengthen and limit cognitive biases in unknown ways. Future research should test the observed cognitive biases in more realistic purchasing environments, taking interactions with other factors such as price and brand into account.

4.3. Conclusion

People organize graded labels in mental space. This organization has consequences for cognitive biases, depending on the mental distance between classes. The organization can also be disrupted by, for example, information pertaining how the midpoint of the scale should be understood. Cognitive bias was present for both the old (skewed) and the new (linear) ecodesign energy scale, but mitigated in some conditions, suggesting that communication strategies can alter how graded labels are interpreted.

CRedit authorship contribution statement

Patrik Sörqvist: Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing. **Emil Skog:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

We have shared the link to our data in the manuscript

References

- Andersson, H., Holmgren, M., Sörqvist, P., Threadgold, E., Beaman, C.P., Ball, L.J., Marsh, J.E., 2024. The negative footprint illusion is exacerbated by the numerosity

- of environment-friendly additions: unveiling the underpinning mechanisms. *J. Cognit. Psychol.* 36, 295–307. <https://doi.org/10.1080/20445911.2024.2313568>.
- Andrade, E.B., Vieites, Y., 2025. Obstacles and opportunities for sustainable consumption: a comprehensive conceptual model, literature review, and research agenda. *J. Consum. Psychol.* 35, 637–662. <https://doi.org/10.1002/jcpy.70003>.
- Baca, G., Reshidi, N., 2025. Green branding and consumer behavior, unveiling the impact of environmental marketing strategies on purchase decisions. *Bus. Strat. Environ.* 34, 3701–3713. <https://doi.org/10.1002/bse.4172>.
- Basiru, I., Xu, Y., Arkorful, V.E., Lugu, B.K., Ibrahim, A.H., 2024. Energy efficiency labels and urban residents' intention to purchase energy-efficient household appliances: an empirical study based on the theory of planned behavior. *Clean Technol. Environ. Policy* 27, 3433–3448. <https://doi.org/10.1007/s10098-024-03070-z>.
- Beck, M., Toulouze, N.O., 2023. Assessing the impact of energy labels on attitude and behavioral intention: an empirical investigation. *J. Clean. Prod.* 415, 137751. <https://doi.org/10.1016/j.jclepro.2023.137751>.
- Cox, D.F., 1962. The measurement of information value: a study in consumer decision-making. In: Decker, W.S. (Ed.), *Emerging Concepts in Marketing*. American Marketing Association, Chicago, IL, pp. 413–421.
- Faul, F., Erdfelder, E., Lang, A.-G., Buchner, A., 2007. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. <https://doi.org/10.3758/BF03193146>.
- Faure, C., Guetlein, M.C., Schleich, J., 2021. Effects of rescaling the EU energy label on household preferences for top-rated appliances. *Energy Policy* 156, 112439. <https://doi.org/10.1016/j.enpol.2021.112439>.
- Gorissen, K., Weijters, B., 2016. The negative footprint illusion: perceptual bias in sustainable food consumption. *J. Environ. Psychol.* 45, 50–65. <https://doi.org/10.1016/j.jenvp.2015.11.009>.
- Hahnel, U.J.J., Arnold, O., Waschto, M., Korcaj, L., Hillmann, K., et al., 2015. The power of putting a label on it: green labels weigh heavier than contradicting product information for consumers' purchase decisions and post-purchase behavior. *Front. Psychol.* 6, 1392. <https://doi.org/10.3389/fpsyg.2015.01392>.
- Heinzele, 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? *Bus. Strat. Environ.* 21 (1), 60–70. <https://doi.org/10.1002/bse.722>.
- Holmgren, M., Andersson, H., Ball, L.J., Marsh, J.E., 2021. Can the negative footprint illusion be eliminated by summative priming? *J. Cognit. Psychol.* 33, 337–356. <https://doi.org/10.1080/20445911.2021.1903012>.
- Holmgren, M., Andersson, H., Sörqvist, P., 2018a. Averaging bias in environmental impact estimates: evidence from the negative footprint illusion. *J. Environ. Psychol.* 55, 48–52. <https://doi.org/10.1016/j.jenvp.2017.12.005>.
- Holmgren, M., Kabanshi, A., Marsh, J.E., Sörqvist, P., 2018b. When A+B < A: cognitive bias in experts' judgment of environmental impact. *Front. Psychol.* 9 (823), 1–6. <https://doi.org/10.3389/fpsyg.2018.00823>.
- Kahneman, D., Frederick, S., 2001. Representativeness revisited: attribute substitution in intuitive judgment. In: Gilovich, T., Griffin, D.W., Kahneman, D. (Eds.), *Heuristics and Biases: the Psychology of Intuitive Judgment*. Cambridge University Press, pp. 49–81.
- Kusch, S., Fiebelkorn, F., 2019. Environmental impact judgments of meat, vegetarian, and insect burgers: unifying the negative footprint illusion and quantity insensitivity. *Food Qual. Prefer.* 78, 103731. <https://doi.org/10.1016/j.foodqual.2019.103731>.
- Lee, W.-C.J., Shimizu, M., Kniffin, K.M., Wansink, B., 2013. You taste what you see: do organic labels bias taste perceptions? *Food Qual. Prefer.* 29, 33–39. <https://doi.org/10.1016/j.foodqual.2013.01.010>.
- Moreno-Fernández, M.M., Blanco, F., Matute, H., 2025. Eco-friendly labeling biases judgments of environmental impact. *J. Environ. Psychol.* 106, 102715. <https://doi.org/10.1016/j.jenvp.2025.102715>.
- Parducci, A., 1965. Category judgments: a range-frequency model. *Psychol. Rev.* 72, 407–418. <https://psycnet.apa.org/doi/10.1037/h0022602>.
- Peirce, J., Gray, J.R., Simpson, S., MacAskill, M., Höchenberger, R., et al., 2019. PsychoPy2: experiments in behavior made easy. *Behav. Res. Methods* 51, 195–203. <https://doi.org/10.3758/s13428-018-01193-y>.
- Skog, E., Marsh, J.E., Sörqvist, P., 2026. Carbon footprint judgments in context: does surrounding context produce assimilation or contrast? *J. Cognit. Psychol.* (in press).
- Skog, E., Sörqvist, P., 2026. Consumers' understanding of energy labels: perception of eco-design with scale range and color-coding. *Psychol. Market.* 43, 934–952. <https://doi.org/10.1002/mar.70095>.
- Sörqvist, P., Colding, J., Marsh, J.E., 2020. Psychological obstacles to the efficacy of environmental footprint tools. *Environ. Res. Lett.* 15, 091001. <https://doi.org/10.1088/1748-9326/ab9968>.
- Sörqvist, P., Volna, I., Zhao, J., Marsh, J.E., 2022. Irregular stimulus distribution increases the negative footprint illusion. *Scand. J. Psychol.* 63, 530–535. <https://doi.org/10.1111/sjop.12829>.
- Sörqvist, P., Holmgren, M., 2022. The negative footprint illusion in environmental impact estimates: methodological considerations. *Front. Psychol.* 13, 990056. <https://doi.org/10.3389/fpsyg.2022.990056>.
- Sörqvist, P., Marsh, J.E., 2024. Conceptual and methodological considerations to the negative footprint illusion: a reply to Gorissen et al. (2024). *J. Cognit. Psychol.* 36, 954–963. <https://doi.org/10.1080/20445911.2024.2412030>.
- Stasiuk, K., Maison, D., 2022. The influence of new and old energy labels on consumer judgements and decisions about household appliances. *Energies* 15 (4), 1260. <https://doi.org/10.3390/en15041260>.
- Sullivan, G.L., Burger, K.J., 1987. An investigation of the determinants of cue utilization. *Psychol. Market.* 4, 63–74. <https://doi.org/10.1002/mar.4220040107>.
- 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? *Bus. Strat. Environ.* 33 (7), 6795–6813. <https://doi.org/10.1002/bse.3838>.
- Threadgold, E., Marsh, J.E., Holmgren, M., Andersson, H., Nelson, M., Ball, L.J., 2021. Biased estimates of environmental impact in the negative footprint illusion: the nature of individual variation. *Front. Psychol.* 12, 648328. <https://doi.org/10.3389/fpsyg.2021.648328>.
- Tversky, A., Kahneman, D., 1974. Judgment under uncertainty: heuristics and biases. *Science* 185, 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>.
- Zhong, K., Yan, Z., Sun, H., Huang, R., 2025. Overcoming the challenges of "green enough": exploring the dark side of incremental green options. *Psychol. Market.* 42, 892–906. <https://doi.org/10.1002/mar.22155>.