Validating Perceived Sustainable Design Features Using a Novel Collage Approach

Nasreddine El Dehaibi[1]
Department of Mechanical Engineering,
Stanford University,
Stanford, CA 94305
e-mail: ndehaibi@stanford.edu

Ting Liao
Assistant Professor
School of Systems and Enterprises,
Stevens Institute of Technology,
Hoboken, NJ 07030
e-mail: tliao@stevens.edu

Erin F. MacDonald
Assistant Professor
Department of Mechanical Engineering,
Stanford University,
Stanford, CA 94305
e-mail: erinmacd@stanford.edu

Designers are challenged to create sustainable products that resonate with customers, often focusing on engineered sustainability while neglecting perceived sustainability. We previously proposed a method for extracting perceived sustainable features from online reviews using annotations and natural language processing (NLP), testing our method with French press coffee carafes. We identified that perceived sustainability may not always align with engineered sustainability. We now investigate how designers can validate perceived features extracted from online reviews using a relatively new design method of collage placement where participants drag and drop products on a two-by-two axis collage and select features from a dropdown menu. We created collage activities for participants to evaluate French press products on the three aspects of sustainability: social, environmental, and economic, and on how much they like the products. During the activity, participants placed products along the two axes of the collage, sustainability and likeability, and labeled products with descriptive features. We found that participants more often selected our previously extracted features when placing products higher on the sustainability axis, validating that the perceived sustainable features resonate with users. We also measured a low correlation between the two axes of the collage activity, indicating that perceived sustainability and likeability can be measured separately. In addition, we found that product perceptions across sustainability aspects may differ between demographics. Based on these results, we confirm that the collage is an effective tool for validating sustainability perceptions and that features perceived as sustainable from online reviews resonate with customers when thinking of various sustainability aspects.
[DOI: 10.1115/1.4052584]

Keywords: data-driven design, design evaluation, design for humans, design for the environment, design methodology, design theory and methodology, design validation, product design, product development, sustainable design, user-centered design

1 Introduction

With the growth of e-commerce platforms, designers are challenged to create products that resonate with customers so that they stand apart from the competition. When creating sustainable products, designers typically rely on engineered sustainability tools like life cycle analyses to guide their decisions. Perceived sustainability, however, is an often-missed factor that can help designers differentiate their products and influence purchasing decisions [1]. Perceptions of sustainability reflect what customers think is sustainable, which may not always align with engineered sustainability. We showed previously that customers can perceive sustainability differently from engineered sustainability [2]. For example, customers perceive that "natural" materials like stainless steel and glass are what make a coffee carafe sustainable, but in reality, it is the engineered sustainability like the auto-off energy saving feature that has the most benefit to the environment.

This disconnect between perceived and engineered sustainability can lead to misinformed purchasing decisions, such as a customer not purchasing a sustainable product because they perceive it as not sustainable [3]. Perceived sustainability is often overlooked when making sustainable engineering decisions. Following the coffee carafe example above, a carafe that includes an auto-off feature should also include natural materials, instead of being all plastic. This creates an alignment for customers that the product is sustainable. With this alignment, designers can potentially differentiate the product in the market and drive purchases. The benefits of sustainable features are limited if the product does not also have market success, and achieving that success relies on both perceived sustainability as well as engineered sustainability.

Designers can differentiate their products by adding features that resonate with customers compared with other options. For customers to resonate with sustainable products, they must (1) identify them as sustainable and (2) like them. The current approach for designers to differentiate sustainable products is to provide information about engineered sustainability, either through online product descriptions or eco-label packaging [4]. Engineered sustainability information, however, can lead to anxiety or confusion when a person has a limited understanding of it [5]. Eco-labels can also trigger the altruism-sacrifice heuristic, where customers expect to sacrifice performance for sustainability. O'Rourke and Ringer investigated how sustainability information affected 40,000 purchase interactions on GoodGuide.com over a 12-month period [6]. The authors found that engineered sustainability information tends to be significant only for those that directly seek it, which is not enough to influence mainstream customer behavior. Therefore, only providing engineered sustainability information has a limited effect on consumers.

An alternative approach for designers to differentiate sustainable products is to design-in perceived sustainability. Although perceived sustainability features, such as using natural materials, may not contribute toward engineering sustainability goals, such as low energy use, they resonate with customers as "sustainable," and such features have been shown to change purchase intentions [5]. She and MacDonald investigated visible product features that capture sustainability perceptions termed "sustainability

[1]Corresponding author.

Contributed by the Design Theory and Methodology Committee of ASME for publication in the Journal of Mechanical Design. Manuscript received April 23, 2021; final manuscript received September 19, 2021; published online October 21, 2021. Assoc. Editor: Scarlett Miller.

 Copyright © 2021 by ASME

triggers" [7]. The authors found that the triggers led customers to think about sustainability-related criteria as well as prioritize sustainability features in simulated decision scenarios of realistic toaster prototypes. Designers can therefore use perceptions to help lead customers to accurate information about a product.

As more customers rely on online shopping, designers have access to a growing source of customer perceptions in the form of product reviews. These perceptions can offer designers insights and help guide their decisions. A growing body of research is developing methods for designers to tap into perceptions in online reviews. For example, Joung and Kim filtered online reviews for product feature perceptions using latent Dirichlet allocation (LDA) to identify automated keywords and validate their method using Amazon reviews of Android smart phones [8]. They found that their approach with LDA yielded better topic coherence than previous methods. Moreover, Hou et al. captured changes in customer perceptions over time using a rule-based natural language processing (NLP) method to extract features and conjoint analysis to categorize the features, validating the approach using reviews of two generations of a Kindle [9]. Their case study demonstrated how designers can use their method to improve new and existing products. In a previous paper, we developed a method for designers to extract features that are perceived as sustainable using annotations of online reviews and natural language processing, testing the method with Amazon reviews of French press products (see Sec. 2.2) [2]. While we demonstrated that there is a gap between engineered and perceived sustainability, we did not validate if perceived sustainable features extracted from online reviews resonate with customers as sustainable.

In this study, we propose a method to validate the effectiveness of features perceived as sustainable from online reviews to help customers resonate with sustainable products. With this method, sustainable designers can confirm whether features extracted from online reviews resonate with customers as sustainable; thus enabling designers to confidently consider perceived sustainability in their products. The method involves a novel collage approach where participants drag and drop products onto a two-by-two axis and select features from a dropdown menu to describe the products. Specifically, we asked participants to evaluate products and features based on their perceived sustainability and likeability. The collage activity has previously been used in design to gain insights into creating sustainable products that resonate with customers (see Sec. 2.3 for more on collages).

The rest of the paper is organized as follows: Sec. 2 presents a background on customer perceptions in design, Sec. 3 outlines the research propositions and hypotheses, we describe our method in Sec. 4 and Sec. 5 presents the results and analysis, our findings are discussed in Sec. 6, and we make conclusions in Sec. 7.

2 Background

Humans develop perceptions using sensory input information and interpret them based on available information, their thoughts, and prior experiences [10]. Perceptions can differ from person to person and change over time [11]. MacDonald et al. showed how customers develop product perceptions on a case-by-case basis [12]. The authors investigated perceptions of paper towels using a discrete choice survey and found inconsistencies in preferences when provided with crux (complex) attributes versus sentinel (simple) attributes. The findings highlight the designer's role to communicate relevant product information to customers when making purchase decisions.

Due to the subjective nature of perceptions, they may or may not accurately represent the context. Understanding perceptions is therefore critical for designers to communicate information accurately to customers. Below we provide a review on (1) works that investigate customer perceptions in sustainable design, (2) our previous work on developing a method to extract customer perceptions from online product reviews, and (3) the use of collages in the design space to assess customer perceptions.

2.1 Customer Perceptions in Sustainable Design. In this section, we provide a literature overview on understanding customer perceptions in design.

Borin et al. investigated the effects of positive, negative, and no environmental information on consumer perceptions [13]. The authors recruited 329 participants and evaluated products in different categories including apples, bath soap, MP3 headphones, and printed paper. Within each product category, there were five environmental messages ranging from very positive to very negative. The authors found that the positive environmental information did not change customer perceptions or purchase intent compared with having no environmental message; however, participants viewed products with positive environmental information better than those with negative environmental information. Therefore, highlighting negative features that a sustainable product avoids may be more effective than highlighting its positive features.

Maccioni et al. investigated the difference between conscious and unconscious perceptions of sustainable products [14]. They recruited 43 participants to evaluate 20 baseline products and 20 sustainable products in the same categories. The authors measured conscious perceptions with self-assessments and unconscious perceptions with biometric measurements. They found participants did not experience any emotional reactions to eco-design efforts because they could not identify sustainability. The authors also found that baseline products were perceived as more functional

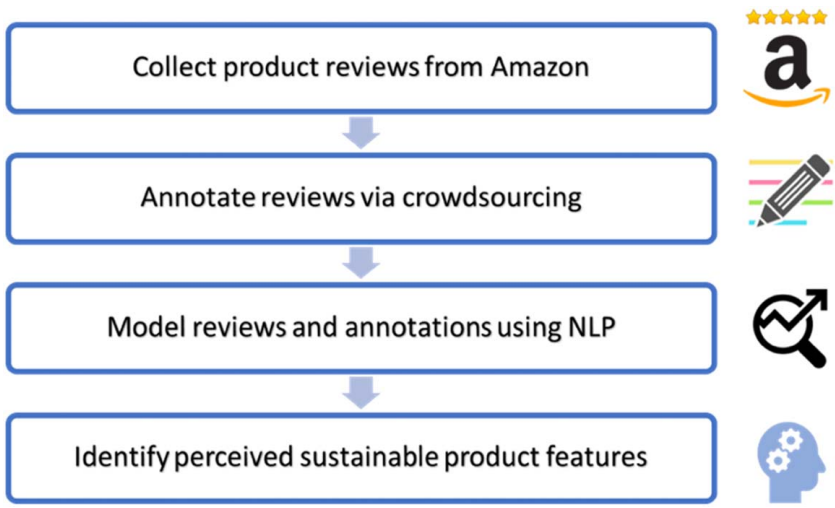

Fig. 1 Extracting customer perceptions method flow

and reliable, while only participants showing high interest in sustainability perceived sustainable products as more innovative.

Steenis et al. investigated the role of product packaging on perceived sustainability [15]. The authors recruited 249 participants and tested their perceptions of soup products with different packaging material and graphics. They found that packaging is not a strong contributor to sustainability perceptions of a product, but that it is a strong contributor to perceptions of product quality and taste. The authors recommend that sustainable packaging can be effective if it enhances perceptions of quality and taste.

Catlin et al. explored the differences between user perceptions of social and environmental sustainability [16]. They recruited 422 participants and asked them to pick between two chocolate bars, one was promoted as socially sustainable while the other was promoted as environmentally sustainable. The participants were then asked to explain their choices. The authors found that participants perceive social sustainability more with affective, short-term, and local considerations while they perceived environmental sustainability with analytical, long-term, and global considerations. Since consumers tend to focus on short term needs over long term, the authors suggest that social sustainability is more likely to resonate with consumers over environmental sustainability.

2.2 Extracting Feature Perceptions From Online Reviews. In this section, we provide a background on identifying perceptions of product features from online reviews. Rai was one of the first in the design space to develop automated methods for extracting value from online reviews [17]. The author used a part-of-speech tagger with a term-document matrix to identify salient features and validated the method using reviews of a camcorder. Several works since then have developed methods using machine learning models to derive design value from reviews, but a gap remained in identifying the differences between perceived and engineered features.

Motivated by this gap, we developed a four-step method in a previous paper to help designers extract product features perceived as sustainable from online reviews (Fig. 1) [2]. We provide details on this method here as we build off it in this paper.

The method involved crowdsourcing annotations of online reviews to build a natural language processing algorithm and then extracting features perceived as sustainable from the parameters of the model. To test this method, we recruited 900 Amazon Mechanical Turk (MTurk) respondents to annotate 1474 reviews of French presses by highlighting phrases based on what they perceive is sustainable and indicating the sentiment in the phrases. We collected annotations for each of the three sustainability aspects: social, environmental, and economic. Using a machine learning NLP algorithm, we identified the most salient features perceived as sustainable from the highlighted phrases that drove positive and negative sentiment for each aspect. See Table 1 for most salient positive features and Table 2 for most salient negative features extracted from this approach. Note that these features represent how users perceived sustainability based on the annotations and may not actually contribute to engineered sustainability. Our method captures features in the form of text directly from reviews, enabling designers to identify what influences sustainability perceptions.

Table 1 Positive features of French presses perceived as sustainable

Social aspects	Environmental aspects	Economic aspects
Easy to use	Well made	Easy to clean
Love it	Easy to use	Great quality
Nice gift	Strong glass	Want more than one
Good for my family	Easy to clean	Reasonable price
Perfect for two	Solid design	Works great
Use with my spouse	Will last	Worth the price
Take to work	Stainless steel	Good customer service
Easy to clean	No plastic	Great value
High quality	Metal frame	Best price
Works great	Sturdy	Hard to beat

Table 2 Negative features of French presses perceived as sustainable

Social aspects	Environmental aspects	Economic aspects
Difficult to use	Too much plastic	Advertised falsely
Looks flimsy	Glass is too thin	Looks cheap
Difficult to wash	Falls apart easily	Waste of money
Glass breaks easily	Glass breaks easily	Glass shatters easily
Sharp corners	Difficult to take apart	Poor design
Metal rusts	Too fragile	Poor customer service
Falls over easily	Plunger leaks	Do not like this brand
Handle hurts	Handle is plastic	Would not buy this
Fragile glass	Does not last	Too expensive
Too small	Rusts easily	Not worth the money

Positive features perceived as socially sustainable were mainly intangible, such as "nice gift" or "good for family," while negative features were mostly tangible relating to safety issues, such as "glass breaks easily" or "sharp corners." Positive features perceived as environmentally sustainable were more tangible, such as "stainless steel" or "no plastic" while negative features related to the durability of the product, such as "glass breaks easily" or "too fragile." Positive features perceived as economically sustainable related to the product being of good value while negative features related to the product not being worth the price. It is important to consider context when assessing perceived features. For example, "easy to use" for social aspects may relate to safety while for environmental aspects it may relate to reliability.

Tables 1 and 2 show that material is a salient perceived environmental concern for French press carafes while energy and water consumption features are not. In reality, energy and water consumption have a much higher environmental impact for French press carafes [2]. This demonstrates the gap between perceived and engineered sustainability and highlights the importance for designers to consider both when creating sustainable products.

The literature has yet to explore how extracted features from online reviews that are perceived as sustainable resonate with customers when thinking of various sustainability concerns. We aim to fill the research gaps by proposing a method to validate if extracted perceived features in Tables 1 and 2 will resonate with customers. The goal is to provide designers a method for validating perceived features to communicate sustainability more accurately to customers and better differentiate sustainable products in the market.

2.3 Evaluating Products Using a Collage. A collage in design research is a set of two axes that range on specific criteria. For example, one axis might range from relaxing to not relaxing while the other axis ranges from like to dislike. Participants then place items on the collage to evaluate based on the criteria. Using the responses designers can identify customer perceptions for the selected criteria. In this section, we cover two applications in design research for using the collage approach to identify customer perceptions.

Guyton was one of the first to use a collage as a systematic method for designers to gain insights into creating sustainable products that resonate with customers [18]. Rather than evaluating products on a single axis, the collage consisted of two axes ranging from "unsustainable to "sustainable" and "dislike" to "like." To validate the method, participants placed images of products on the collage and selected words from a vocabulary list to describe the products. Several products were tested including spatulas, mugs, and boots. Based on the product placements and vocabulary used, Guyton demonstrated the collage activity as an effective method for capturing sustainability perceptions of products.

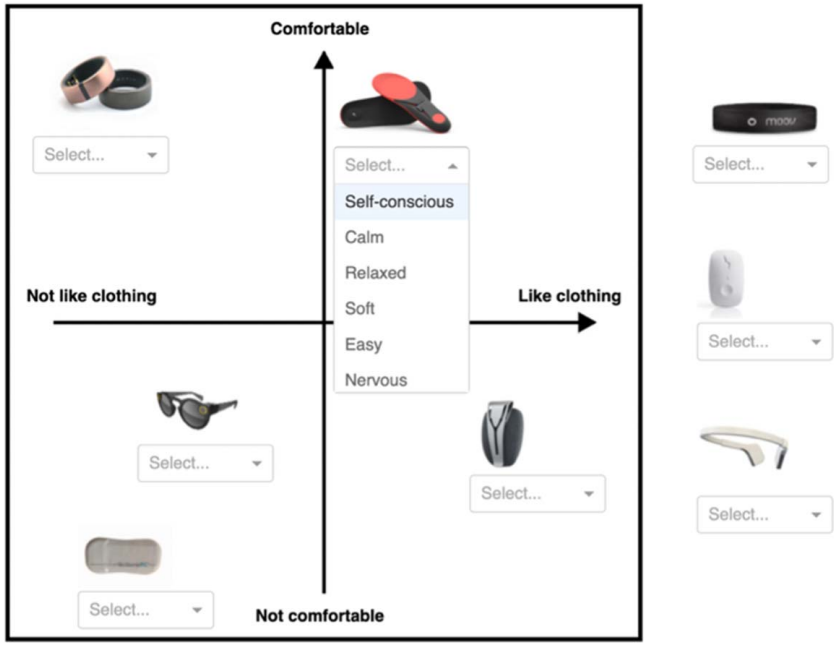

Fig. 2 Example of a collage tool from Liao et al. [19]

Liao and MacDonald build on this approach by using the collage as an evaluation tool for participants to express their product preferences [19]. The authors explored how users perceive products based on a product's form and visible characteristics using a two-axis interactive collage tool to evaluate eight wearable products. The authors evaluated comfort, delight, and usefulness on three separate collages, respectively. They generated a list of emotional descriptive words to identify perceptions. The authors developed the collage activity into a webapp where participants drag and drop product images and select descriptive words from a dropdown list (Fig. 2). Based on responses from 400 participants, the tool revealed relationships between product characteristics and user perceptions. For example, wearables that resembled clothes were perceived as more delightful and comfortable. Moreover, the authors found a high correlation between likeability and the other axes of the collage: comfort, user delight, and usefulness.

Based on these previous studies, we used the collage in this study as an engaging way for participants to evaluate product sustainability. The collage tool has proven itself as an intuitive way to get input from users and to measure the relationship between products and user emotions. By allowing participants to actively choose one or more features without drawing attention to them, we can determine if participants resonate with those features. Moreover, evaluating the placement of products on two axes enables us to separately study how participants like a product and their perceived sustainability of the product.

3 Research Propositions and Hypotheses

This work proposes a method to validate whether perceived sustainable features extracted from online reviews resonate with customers using a collage tool. The two axes used in the collages are likeability and sustainability. While perceived sustainability features may not contribute to engineered sustainability, they can register as sustainable to customers and help them create cognitive alignment with sustainable products. The selection and position of the features on the collage are recorded and tested to validate that perceived sustainable features are evaluated as sustainable. We also validate whether perceived sustainability is associated with likeability by recording and testing the position of the products on the collage. The following propositions and hypotheses are tested.

Proposition 1: Designing-in perceptions can help customers create an alignment between perceived sustainability and sustainable products. Based on this, we propose that customers will resonate with perceived sustainable features as being sustainable.

Hypothesis 1: Participants evaluating product sustainability on a collage will select features perceived as sustainable for products that they place higher on the "sustainability" axis of the collage.

Proposition 2: Customers tend to like products that create cognitive alignment for them, and perceptions can help them achieve that. We therefore propose that perceptions of product sustainability contribute to how much customers like a sustainable product. This is motivated by prior research that 73% of millennials are willing to pay more for sustainable products [20].

Hypothesis 2: A statistically significant relationship exists between the placement of a product on the "sustainability" axis of the collage, and the "like" axis of the collage.

4 Method

To test the hypotheses, we designed an activity for 1200 respondents from Amazon Mechanical Turk (MTurk) to evaluate French press products based on sustainability criteria. MTurk is a crowdsourcing platform to recruit workers for completing tasks. We refer to the respondents as participants in this paper (see Sec. 4.4 for more information on participants). The activity consisted of three parts: (1) a pre-survey to learn about the sustainability criteria and get familiarized with the products, (2) a collage tool that was adapted from previous work by Guyton [18] and Liao and MacDonald [19], and (3) a post-survey where participants answered follow-up questions about the products and demographics (Fig. 3).

We designed three versions of this activity to evaluate products on the three sustainability aspects separately: social, environmental, and economic (Fig. 4). We randomly assigned participants to one of the three versions. Our choice to have participants focus on one sustainability aspect was motivated by our previous study where a pilot test showed that this led to better clarity for participants and more usable responses [2]. While focusing on one aspect is not realistic for a purchasing scenario, it provides participants clarity for evaluating products on the collage which is crucial for our study.

The pre-survey, collage task, and post-survey are described in detail below.

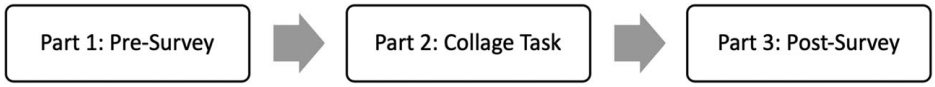

Fig. 3 Breakdown of the three parts of the activity

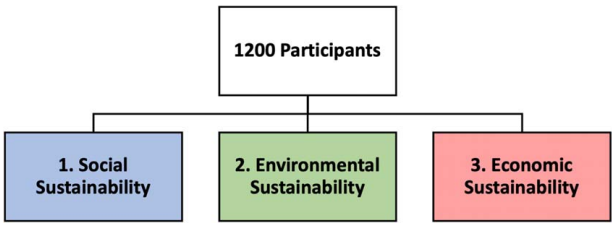

Fig. 4 Participants distributed across three activity versions

4.1 Pre-Survey. In the pre-survey, participants were screened for eligibility (check Sec. 4.4 for information on participants), trained and tested on sustainability criteria, and familiarized with the products and the collage activity. Sections 4.1.1 and 4.1.2 provide more information on the criteria and products, respectively.

4.1.1 Sustainability Criteria. Participants were trained on sustainability criteria during the pre-survey to prepare them for evaluating product sustainability in the collage task. Each version of the activity had a customized training portion based on one of the three sustainability aspects. To train participants we displayed to them sustainability evaluation criteria as defined in El Dehaibi et al. [2], shown in Fig. 5. We chose these definitions because they provide participants with guidelines on how to evaluate products without overriding their personal opinions. For example, while the social sustainability definition includes health and safety as a criterion, participants can determine on their own what makes a product healthy or safe. We trained participants to focus on a specific sustainability aspect and ignore others. As an example, if participants were working on the social sustainability version of the activity, they were trained to focus on social sustainability and not to consider environmental and economic aspects. We chose this approach after a pilot study revealed that participants were still evaluating products using mixed sustainability criteria when we specified what criteria to focus on only.

We did not provide training on how to evaluate "likeability" of the products. While sustainability is a multi-faceted concept, likeability is more of a feeling that humans can detect. We therefore let participants decide how to evaluate likeability and what it means for them.

Following the training portion, participants completed a test that they had to pass to make sure they had understood the training. The test consisted of two multiple choice questions: the first question asked participants to select factors they will evaluate according to their sustainability aspect, and the second question asked participants to select factors they will not evaluate according to their sustainability aspect. For example, in the social sustainability version, participants could choose "family and culture" for the first question and "energy and water consumption" for the second question.

Six French presses were used in this activity (shown in Table 3). We chose French press products to build off our previous work [2]. French presses are ubiquitous and likely to contain features related to sustainability. We selected the French presses in Table 3 for this study based on their varying aesthetic features and materials that cover the design space (e.g., stainless steel, plastic, glass, wood, etc.) as well as their varying price points, number of ratings, and reviews. Our goal was to provide participants with enough variety so they can evaluate products differently on the collage. Participants were presented with direct links to the Amazon product pages in the pre-survey so that they could get familiarized with the products. The order of the products was randomized between each participant. We asked the participants to consider the product features, price, and reviews for each product. The Amazon links opened as popups instead of new windows or tabs to reduce the number of participants we may lose from completing the activity. The Amazon popup was equivalent to opening a browser in "Incognito mode" so that participants' prior browsing history did not influence dynamic content shown on the pages, such as recommended Amazon products or reviews. This facilitated having a common baseline of live Amazon pages between participants. Other factors may still bias the dynamic content shown such as geographic location of participants, but the random choice of the MTurk participants likely limited the influence of these factors. If participants skipped any of the Amazon links or proceeded too quickly, they were shown the product links again and on the third time, they exited from the activity. We included this step to make sure that participants were familiar with the products before evaluating them.

4.2 Collage Activity. After completing the pre-survey successfully, participants started the collage task to evaluate products. We asked them to drag and drop products on the collage and select features from a dropdown list to describe each product based on the sustainability criteria. Participants could select the features before or after dragging the product, and we did not provide additional guidance on what features to select. Moreover, participants were able to modify their product placements and feature selections up until they completed the collage activity. While participants may alter their feature selections to justify their product placement and vice versa, their justifications would still validate that they evaluated perceived sustainable features as more sustainable. We tested

Fig. 5 Sustainability aspect definitions and training

Table 3 List of products

						
Product name	Chef	Frielin	Madrid	Melbo	Brookl	Terra
Price	$14.99	$56.44	$35.00	$39.99	$29.99	$19.99

hypothesis 1 based on the placement of the selected features on the collage, and we tested hypotheses 2 based on the placement of the products on the collage.

4.2.1 Interface. We used the same collage tool interface from the webapp used in Liao and MacDonald [19] but modified it to include a two-by-two grid, a set of products on the right side, and an evaluation criteria button on the left side. The *x*-axis on the grid ranges from "Dislike" to "Like" while the *y*-axis depends on the version of the activity, ranging from "Not XX Sustainable" to "XX Sustainable," where "XX" represents one of the three sustainability aspects: social, environmental, and economic. We chose a two-by-two grid in this study for three reasons: first, it allows us to differentiate between what participants like about a product and what they determine is sustainable; second, it is a tested tool that has been used in literature for evaluating product perceptions [18,19], and third, it is an engaging way for participants to evaluate product sustainability. Figure 6 shows an example of the interface for the social sustainability version.

Clicking on the sustainability criteria button on the left opens a popup with information from the training section of the pre-survey. An example from the social sustainability version is shown in Fig. 7.

Each product image has an Amazon icon on the top right corner which links to a popup of the product's Amazon page when clicked on. An example is shown in Fig. 8. Like the pre-survey, Amazon popups in the collage opened like a browser in "incognito mode" so that past browsing history did not influence contents shown to participants. For both the sustainability criteria button and Amazon icon, participants could close the popup by clicking outside of the popup to return to the collage.

To evaluate the products, participants dragged and dropped the products onto the collage and then selected features from a dropdown list to describe each product based on the criteria (Fig. 9). See Sec. 4.2.2 for information on the features we provided. We asked participants to place all products on the grid and select at least one feature from the dropdown list for each product to proceed. The dropdown list was randomized between participants, and they could select as many features as they liked.

Participants could also relocate the images on the collage until they proceeded to the next page. We recorded the location of the center of the product image as a float number. While reducing the location of the image to one point adds uncertainty, the impact is negligible since our focus is on the relative placement of products and features. On the next page, we asked participants to rate how relevant each of the features they selected are to sustainability on a 5-point Likert scale. The scale labels included "Not at all related" for a 1 out of 5 rating, "Somewhat related" for a 3 out of 5 rating, and "Very related" for a 5 out of 5 rating. This gave us insight on which features were selected based solely on likeability versus perceived sustainability. After rating the features, participants were able to submit their evaluations and received a password to proceed with the post-survey.

4.2.2 Interface. Below we discuss two sets of features that we presented to participants: (1) positive and negative features perceived as sustainable and (2) positive features perceived as sustainable and features not perceived as sustainable.

Positive and Negative Features Perceived as Sustainable: In the collage task, we provided participants with a set of 20 features in a dropdown list to select from for each product. These features were extracted from online reviews of French presses in our previous work and are shown in Tables 1 and 2 for each of the sustainability aspects [2]. While our original study extracted 40 features for each sustainability aspect, we selected a subset of 20 features for this study to account for overlaps as well as to have a similar quantity of features as in previous collage experiments [19]. Our goal in this study is to validate that these features resonate with customers as sustainable using the collage approach. The selected features consist of the ten most positive and ten most negative features perceived as sustainable, according to review annotations and a machine learning algorithm. While the extracted features are perceived as sustainable, they may not actually contribute to engineered sustainability. Each sustainability aspect has a

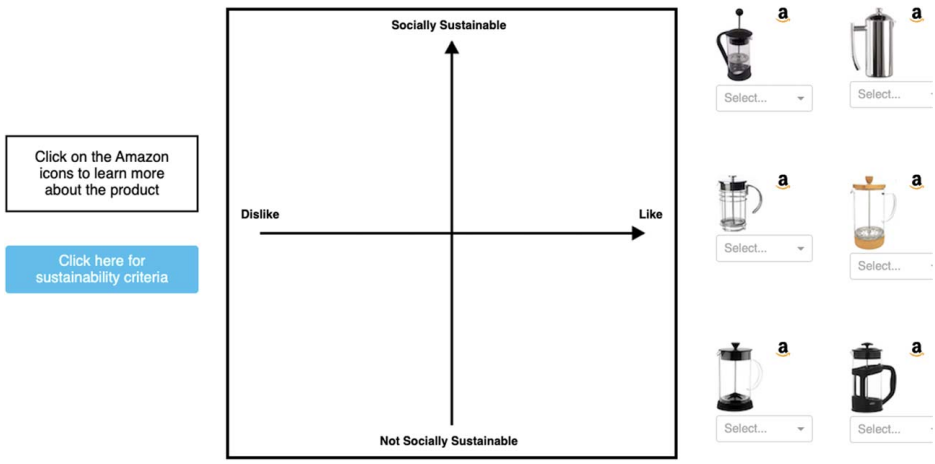

Fig. 6 Collage tool interface for social sustainability

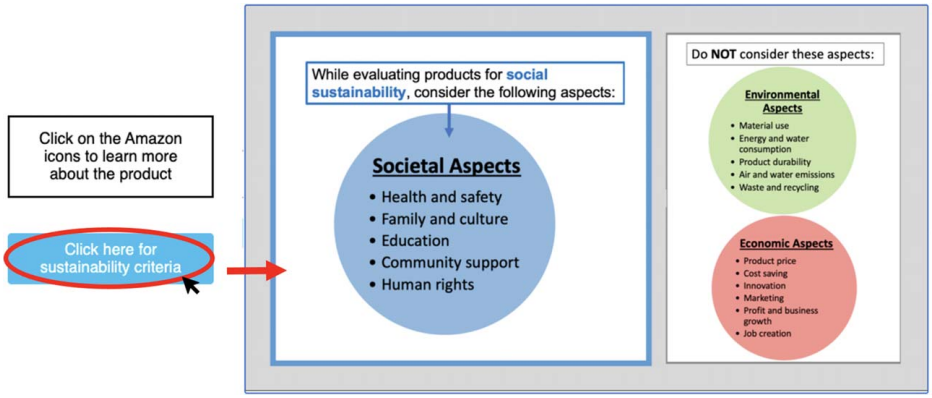

Fig. 7 Evaluation criteria button for social sustainability

corresponding set of 20 features. Participants actively elected to select at least one of these features for each product they placed on the collage and could select the features before or after placing the products. We did not provide additional guidance on which features to select. Out of the features participants selected, we analyzed the features that they rated as three out of five or higher on relevance to sustainability. Our goal was to filter out features that were selected solely based on liking the product since we were interested in the features that participants perceived as sustainable. We investigated how participants resonated with the features perceived as sustainable based on where the features were placed on the collage. Moreover, we investigated the relationship between the

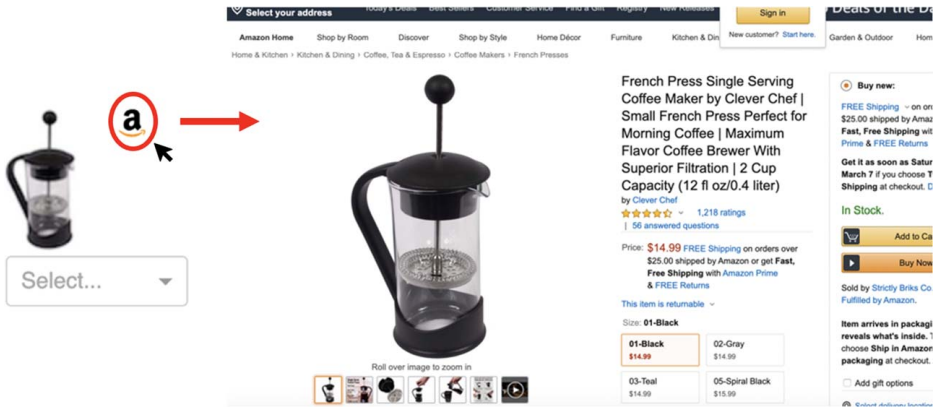

Fig. 8 Amazon product page popup example

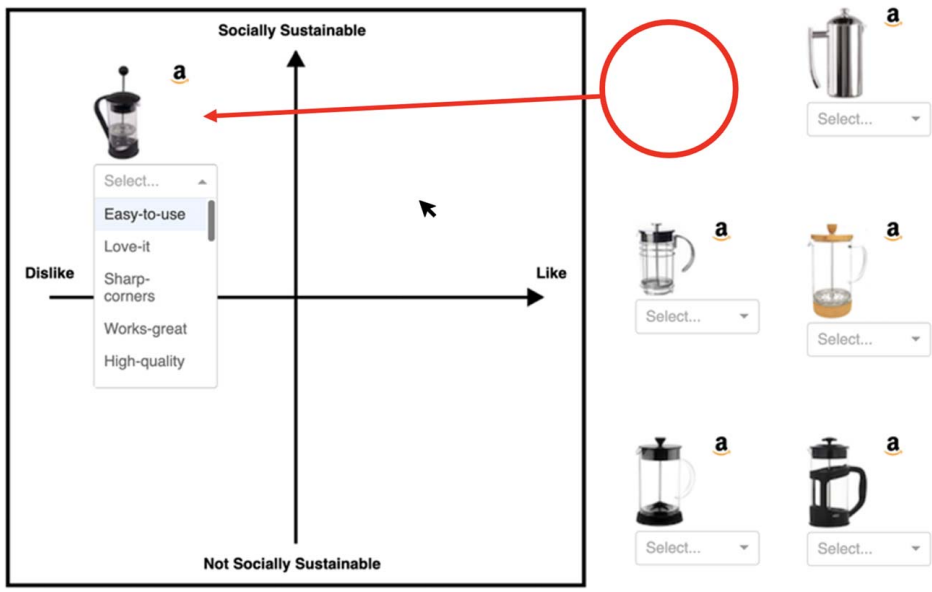

Fig. 9 Dragging and dropping products on collage and selecting at least one feature to describe each product

Table 4 Features not related to sustainability

Used for environmental aspects only
Course ground coffee
Typical French press
Wonderful beverage
Single use
Daily coffee
Black product
Tight part
Light weight
Bitter coffee
French way
Moisture problem

two axes of the collage, perceived sustainability and likeability of a product, based on the placement of the products on the collage.

Positive Features Perceived as Sustainable and Features not Perceived as Sustainable: We chose to test hypothesis 1 with a more challenging set of features to validate our findings. These features are more challenging because they are closer in sentiment. We tested how participants evaluate products on the collage when provided with ten positive features perceived as sustainable and ten features not perceived as sustainable. We opted to test these features on the environmental aspect of sustainability. For the ten positive features, we used the same list as before for positive features of environmental sustainability in Table 1. For the ten features not perceived as sustainable, we derived a list of features using data from our previous study where we asked participants to annotate parts of reviews that were relevant to sustainability [2]. We assumed that the unannotated parts of the reviews are not perceived as sustainable, combined them into one text, and identified adjectives and noun phrases from them using a part-of-speech tagger. We then randomly matched ten adjectives to ten noun phrases to generate the features in Table 4.

Since this approach aims to identify features perceived as not sustainable, some of the features may contribute to engineered sustainability. For example, "single use" might contribute to engineered sustainability but was identified to not be perceived as sustainable using our approach. Moreover, our automated approach of identifying the features may result in noise. Our primary goal here was to automate how we select features perceived as not sustainable to limit potential biases from ourselves selecting the features.

4.3 Post-Survey. After completing the collage activity, participants were directed to a post-survey where we asked them to rate the quality of product images, product descriptions, and the overall quality of the products based on the respective Amazon product pages. We then asked participants about their purchasing behavior on Amazon to check if they are target customers who buy home and kitchen items on Amazon. Finally, we asked participants basic demographic questions to check for any other significant variables.

4.4 Participants. We recruited a total of 1200 participants from MTurk to complete the activity which took 20 min on average; participants were compensated $5 each for their time. We opted to recruit from MTurk over in-person participants so that we could quickly collect many responses. Moreover, the demographics of respondents on MTurk align closely with the online population [21], and therefore better fits a target Amazon customer. This is ideal for our study since participants were likely familiar with Amazon and comfortable to interact with the Amazon popup pages.

To ensure quality in the responses, we screened for participants on MTurk that have at least a 97% prior approval rating and are based in the United States. We set these as requirements on the MTurk platform and then validated participant location using screening questions in the survey. In accordance with literature, respondents in the United States consistently deliver better quality responses [22]. We also conducted the activities during weekday mornings Pacific Time as this was reported to help improve data quality [21]. Furthermore, the collage interface was compatible only with desktop devices with screens larger than 10 in. We therefore screened participants for eligible devices. Participants self-reported their screen size.

Out of the 1200 participants that completed their task, we approved 935 based on two requirements: (1) completing the activity in time (t) that is within 1 standard deviation (σ) of the average time to complete the activity (μ) or longer (i.e., $t \geqslant \mu - \sigma$) and (2) correctly answering the check question, "Which sustainability criteria were you evaluating for?" which we asked in the post-survey. The first criteria aimed to filter for responses from participants that were going with their gut feeling rather than justifying their responses, and the second criteria served as an attention check. We excluded responses from the results if they did not meet one or both criteria. These requirements are like the ones used to approve responses in our previous study [2].

5 Analysis and Results

This section is split into three parts: in the first we analyze the participant demographic pool, in the second we analyze the placement of the features associated with testing hypothesis 1, and in the third we analyze the placement of the products associated with testing hypothesis 2.

5.1 Participant Demographics. Our demographic pool of 935 participants includes a broad representation of age groups, education levels, and incomes, with the gender distribution slightly skewed toward more male than female (Fig. 10). Most participants were young, white, educated, and used, with many having above-average incomes. This is in line with a prior demographics study of MTurk [22]. While the distributions are not representative of the general population, it is representative of the online population.

The participants' purchasing habits show that over 90% of them have shopped on Amazon within the past year and that the majority are subscribed to Amazon Prime (Fig. 11). This indicates that participants are familiar with Amazon's user interface and were comfortable interacting with the Amazon popups in the study.

In addition, most participants purchase home and kitchen items from Amazon monthly or more (Fig. 12). This is ideal for our study because it indicates that participants can provide evaluations that are like a potential French press customer.

5.2 Feature Analysis. Below we present our analysis for testing hypothesis 1 based on the placement of features in the collage task. Each of the 935 participants placed multiple features on the collage. After filtering for features that were rated less than three out of five as relevant to sustainability, we had a total of 7263 location points of features on the collage. We excluded an additional 373 data points that were not moved from their starting location when the collage activity launched (starting locations are shown in Fig. 6).

5.2.1 Positive and Negative Features Perceived as Sustainable. In this section, we present the results and analyses for testing hypothesis 1 using positive and negative features perceived as sustainable. This hypothesis considers how the perceived features resonate with participants when evaluating product sustainability. Table 5 summarizes the information on the features selected.

Participants selected positive features for products more often than negative in all activity versions. This was most apparent with social sustainability, followed by environmental sustainability, and then economic sustainability and shows how participants resonate differently with features based on the sustainability criteria.

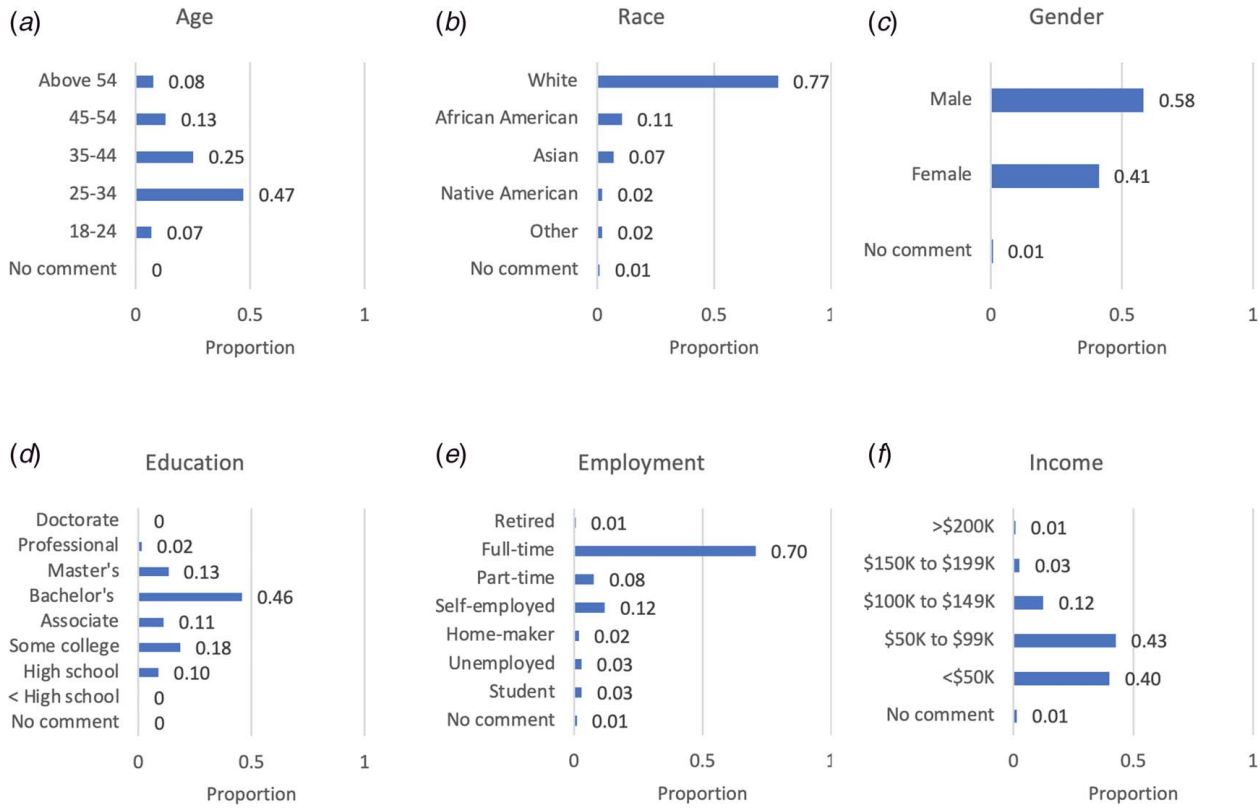

Fig. 10 Participant demographics

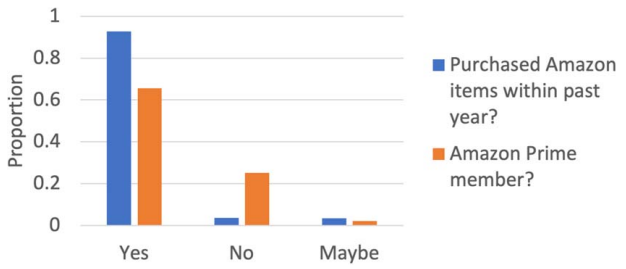

Fig. 11 Distribution of participants that are Amazon customers

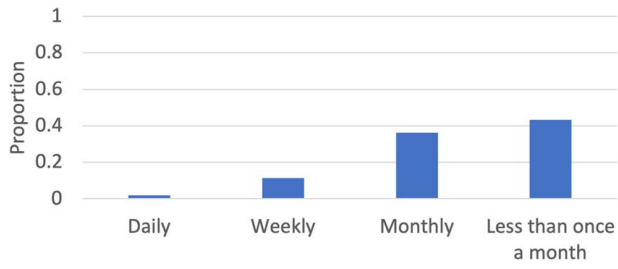

Fig. 12 Distribution of participant purchase frequency from Amazon's home and kitchen department

The most used positive feature across all three sustainability criteria was "Reasonable price" while the most commonly used negative feature was "too much plastic." This illustrates the features that resonated most with the participants and aligns with our previous findings on extracted French Press features perceived as sustainable from online reviews [2]. For social sustainability, "looks flimsy" is not intuitively relevant but is likely related to being perceived as unsafe for use.

To visualize the data, we plotted the average placement of features by the participants in the collage tasks. Figures 13–15 show

Table 5 Summary of features selected in collage

	Social sustainability		Environmental sustainability		Economic sustainability		Combined	
	Positive features	Negative features	Positive features	Negative features	Positive features	Negative features	Positive features	Negative features
Number of participants	253		268		241		762	
Observations	1073	439	1328	613	1165	753	3566	1805
Average features per participant	4.24	1.74	4.96	2.29	4.83	3.12	4.68	2.37
Average features per product	178.83	73.17	221.33	102.17	194.17	125.5	594.33	300.83
Average features per product per participant	0.71	0.29	0.83	0.38	0.81	0.52	0.78	0.39
Most common feature selected	Good for my family	Looks flimsy	Well made	Too much plastic	Reasonable price	Too expensive	Reasonable price	Too much plastic

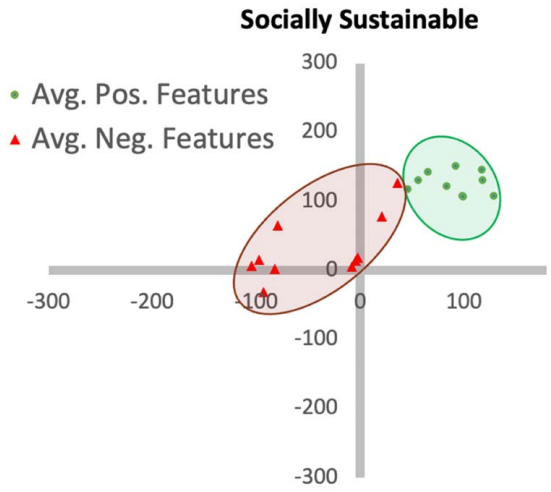

Fig. 13 Average placement of positive and negative features perceived as socially sustainable on collage

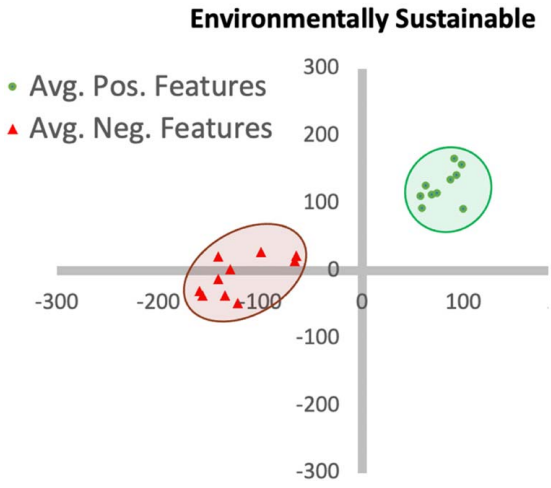

Fig. 14 Average placement of positive and negative features perceived as environmentally sustainable on collage

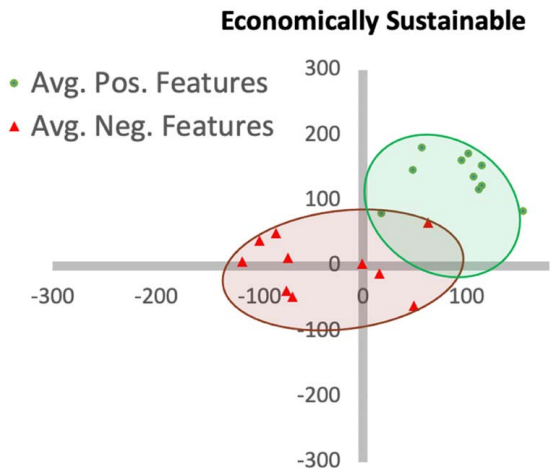

Fig. 15 Average placement of positive and negative features perceived as economically sustainable on collage

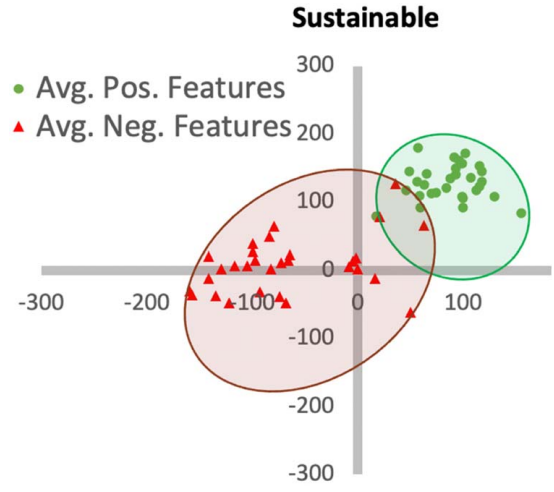

Fig. 16 Average placement of positive and negative features perceived as sustainable for all criteria on collage

the results for social, environmental, and economic criteria, respectively. Each plot is color-coded to differentiate between positive and negative features. In each of the figures we see distinct clusters between the average placement of positive features (green-shaded) and negative features (red-shaded) along the axes of the collage. The distinct clusters across the *y*-axis (measure of perceived sustainability) support hypothesis 1. Figure 15 for economic sustainability shows some overlap between positive and negative clusters which is likely attributed to machine learning model noise output from our previous study [2]. The machine learning model for economic sustainability had the weakest performance due to an imbalance between positive and negative annotations. Figure 16 shows the average placement of all features combined from the three sustainability criteria. We again see distinct clusters between positive and negative perceived sustainability features with a slight overlap attributed to the economic sustainability results. It is important to note that there is variance across both the *x* and *y* axes. While the location of features is represented as dots, the placement of products on the collage occupies a larger space and can introduce additional variance.

To test if each of the two groups of features are statistically different, we conducted a two-sample *t*-test assuming unequal variances using the *y*-coordinate values of features for each sustainability criteria. We assumed unequal variance based on Levene's test showing that the variances of the positive and negative feature locations are statistically different in each of the collage activities. Table 6 shows the results for social, environmental, economic, and combined criteria, respectively. In all cases, we see that there is a significant difference along the *y*-axis (measure of perceived sustainability) where positive features are placed higher than negative features, supporting hypothesis 1.

We further investigated if the statistical significance holds when considering repeated measures from participants. We conducted a multivariate analysis of variance (MANOVA) using the *x* and *y* coordinates of the features as the dependent variables and the participant demographics as the independent variables. Based on scatter plots showing nonlinear patterns in our data, we chose to work with the Pillai criterion because it is the most powerful and robust statistic when assumptions of linearity and homogeneity of variances are not met [23]. The results are shown in Table 7 for social, environmental, economic, and combined criteria, respectively. In each case, we see that the phrases are highly significant. We can therefore state with statistical significance that participants more often selected features perceived as sustainable when placing products higher on the sustainability axis (i.e., fail to reject hypothesis 1).

Table 6 Two-sample *t*-test between positive and negative features perceived as sustainable

	Social sustainability		Environmental sustainability		Economic sustainability		Combined	
	Positive features	Negative features	Positive features	Negative features	Positive features	Negative features	Positive features	Negative features
Mean *y*-coordinate	129.02	24.65	132.08	−13.16	144.84	4.91	135.33	3.57
Observations	1073	439	1328	613	1165	753	3566	1805
$P(T \leq t)$ one-tail	<0.001[a]		<0.001[a]		<0.001[a]		<0.001[a]	
t Critical one-tail	1.65		1.65		1.65		1.65	

[a]Significant at $p = 0.001$.

5.2.2 Demographic Interactions. In this section, we present results related to the demographics of participants. We investigated the demographics as independent variables in the MANOVA analysis to gain more insight into the data. From the MANOVA results in Table 7, we see that certain demographics variables are significant for different cases. To understand how the variables influence each other, we performed an analysis of variance (ANOVA) on each dependent variable separately. The results are shown in Table 8 for social, environmental, and economic criteria, respectively.

The ANOVA results reveal how the demographics variables interact with the dependent variables. For social sustainability, we see that race has a statistical significance for participants identifying product sustainability. In terms of how much participants like a product, demographics have no significance. For environmental sustainability, education and employment are significant for participants identifying product sustainability. In terms of how much participants like a product, race and education are significant. Moving on to economic sustainability, income is significant for participants identifying product sustainability. In terms of how much participants like a product, education, and gender are significant.

While the results in Table 8 provide preliminary insights on demographic interactions with sustainability perceptions, the participant demographics are not representative. A deeper study on demographics is needed to identify stronger insights.

Finally, the results for when we combine the data from all three sustainability criteria are shown in Table 9. None of the demographics variables are significant for participants identifying product

Table 7 MANOVA output with positive and negative features perceived as sustainable

	Social sustainability			Environmental sustainability			Economic sustainability			Combined		
	Pillai	∼*F*	Pr(>*F*)	Pillai	∼*F*	Pr(>*F*)	Pillai	∼*F*	Pr(>*F*)	Pillai	∼*F*	Pr(>*F*)
Product	0.095	12.02	<0.001[c]	0.357	82.1	<0.001[c]	0.173	28.18	<0.001[c]	0.155	78.0	<0.001[c]
Criteria	–	–	–	–	–	–	–	–	–	0.004	5.18	<0.001[c]
Feature	0.348	13.30	<0.001[c]	0.293	17.0	<0.001[c]	0.488	25.29	<0.001[c]	0.398	21.40	<0.001[c]
Age	0.007	0.85	0.581	0.003	0.58	0.832	0.004	0.65	0.768	0.001	0.54	0.869
Race	0.012	1.51	0.128	0.011	2.05	0.025[a]	0.008	1.13	0.338	0.003	1.44	0.156
Gender	0.005	1.57	0.179	0.001	0.44	0.779	0.006	2.31	0.056	0.001	1.32	0.260
Educ.	0.014	1.38	0.170	0.031	5.02	<0.001[c]	0.018	1.95	0.018[a]	0.008	2.21	0.004[b]
Emplo.	0.009	0.79	0.678	0.013	1.77	0.037[a]	0.009	0.98	0.472	0.004	1.34	0.177
Income	0.012	1.47	0.145	0.006	1.21	0.279	0.008	1.20	0.288	0.002	0.83	0.595

[a]Significant at $p = 0.05$.
[b]Significant at $p = 0.01$.
[c]Significant at $p = 0.001$.

Table 8 ANOVA output for social, environmental, and economic sustainability

	Social sustainability				Environmental sustainability				Economic sustainability			
	Sustainability (*y*-axis)		Like (*x*-axis)		Sustainability (*y*-axis)		Like (*x*-axis)		Sustainability (*y*-axis)		Like (*x*-axis)	
	F	Pr(>*F*)	*F*	Pr(>*F*)	*F*	Pr(>*F*)	*F*	Pr(>*F*)	*F*	Pr(>*F*)	*F*	Pr(>*F*)
Product	5.19	<0.001[c]	21.9	<0.001[c]	101	<0.001[c]	112	<0.001[c]	27.4	<0.001[c]	28.8	<0.001[c]
Feature	8.70	<0.001[c]	27.3	<0.001[c]	14.3	<0.001[c]	31.0	<0.001[c]	24.7	<0.001[c]	42.0	<0.001[c]
Age	0.79	0.556	0.71	0.618	0.56	0.730	0.63	0.673	0.68	0.639	0.62	0.685
Race	2.22	0.049[a]	1.25	0.282	1.49	0.189	2.33	0.041[a]	0.80	0.553	1.48	0.195
Gender	1.02	0.361	2.36	0.095	0.85	0.428	0.02	0.982	0.35	0.704	4.23	0.015[a]
Education	0.85	0.528	1.71	0.114	6.47	<0.001[c]	3.18	0.004[b]	1.54	0.160	2.30	0.025[a]
Employment	1.02	0.416	0.53	0.810	2.08	0.043[a]	1.78	0.087	1.07	0.379	0.86	0.541
Income	1.51	0.183	1.04	0.393	0.73	0.604	1.62	0.152	2.25	0.047[a]	0.14	0.982

[a]Significant at $p = 0.05$.
[b]Significant at $p = 0.01$.
[c]Significant at $p = 0.001$.

Table 9 ANOVA output for combined sustainability criteria

	Sustainability (*y*-axis)		Like (*x*-axis)	
	F value	Pr(>*F*)	*F* value	Pr(>*F*)
Product	39.61	<0.001[c]	132.63	<0.001[c]
Criteria	6.91	<0.001[c]	4.81	0.008[b]
Feature	20.18	<0.001[c]	37.38	<0.001[c]
Age	0.80	0.550	0.24	0.947
Race	0.42	0.836	2.59	0.024[a]
Gender	0.81	0.446	2.08	0125
Education	1.39	0.194	3.04	0.002[b]
Employment	0.82	0.573	2.07	0.043[a]
Income	1.44	0.207	0.27	0.931

[a]Significant at $p = 0.05$.
[b]Significant at $p = 0.01$.
[c]Significant at $p = 0.001$.

sustainability, although the criteria variable has a strong significance. Race, education, and employment are significant variables for participants liking a product.

5.2.3 Positive Features Perceived as Sustainable and Features Not Perceived as Sustainable. In this section, we present results for the collage activity with a more challenging set of features, including features perceived as sustainable and features not perceived as sustainable. These sets of features are closer in sentiment. Figure 17 shows the scatterplot for the average placement of features, color-coded by positive features and features not perceived as sustainable.

The distinct clusters are less prominent along the *y*-axis (measure of perceived sustainability), and we also see less of a horizontal spread in the average positive features likely due to the list of features having a closer range of sentiment in this activity. Table 10 shows the results for the two-sample *t*-test assuming unequal variances for the *y*-coordinate values features. We see that there is a significant difference along the *y*-axis (measure of perceived sustainability) between the two groups of features, which supports our initial findings with hypothesis 1.

We conducted a MANOVA shown in Table 11 and the results confirmed that features are highly significant. Therefore, even with the more challenging set of features, we found that participants evaluating product sustainability on a collage selected features perceived as sustainable for products that they placed higher on the sustainability axis of collage (i.e., fail to reject hypothesis 1).

5.3 Product Analysis. In this section, we present the results and analyses for testing hypothesis 2. The hypothesis considers how perceived sustainability of a product and liking that product

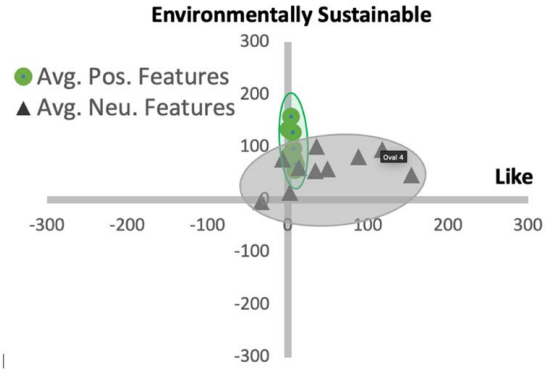

Fig. 17 Average placement of positive features perceived as sustainable and features not perceived as sustainable

Table 10 Two-sample *t*-test between positive features perceived as environmentally sustainable and features not perceived as sustainable

	Positive features	Features not perceived as sustainable
Mean	100.00	43.46
Observations	1652	901
Number of participants		262
Average features per participant	6.31	3.44
Average features per product	275.33	150.17
Average features per product per participant	1.05	0.57
$P(T \le t)$ one-tail	<0.001[a]	
t Critical one-tail	1.65	
$P(T \le t)$ two-tail	<0.001[a]	
t Critical two-tail	1.96	

[a]Significant at $p = 0.001$.

Table 11 MANOVA output with positive features perceived as sustainable and features not perceived as sustainable

	Pillai	~*F*	num Df	den Df	Pr(>*F*)
Product	0.245	60.24	10	4318	<0.001[b]
Feature	0.083	4.92	38	4318	<0.001[b]
Age	0.010	2.10	10	4318	0.021[a]
Race	0.008	1.78	10	4318	0.059
Gender	0.004	2.39	4	4318	0.049[a]
Education	0.017	2.59	14	4318	<0.001[b]
Employment	0.013	2.01	14	4318	0.014[a]
Income	0.024	5.24	10	4318	<0.001[b]

[a]Significant at $p = 0.05$.
[b]Significant at $p = 0.001$.

are related. Based on the MANOVA results in Table 7, we see that there is a significant difference in products across the *y*-axis (measure of perceived sustainability) and *x*-axis (measure of how much the product is liked) in each of the sustainability versions, which supports hypothesis 2. To investigate this further, we ran a multiple linear regression model with the Like values of the product placement as the dependent variable and the sustainability values of product placement and demographics as the independent variables. Table 12 shows the *p*-values from this model for each sustainability aspect; the sustainability values were significant while the demographic variables were not.

We wanted to study the relationship between perceived sustainability and liking a product further, so we measured the correlation between the sustainability and like values of the product placements for the different sustainability aspects. We used a repeated measures

Table 12 Multiple linear regression for liking the product versus perceived sustainability and demographics

	Social	Environmental	Economic	Combined
Sustainability	<0.001[a]	<0.001[a]	<0.001[a]	<0.001[a]
Age	0.27	0.67	0.41	0.41
Race	0.69	0.28	0.39	0.36
Gender	0.58	0.81	0.12	0.46
Education	0.26	0.21	0.43	0.21
Employment	0.47	0.36	0.60	0.50
Income	0.38	0.47	0.47	0.43

[a]Significant at $p = 0.001$.

Table 13 Repeated measures correlation between perceived sustainability of a product and liking the product

	Social	Environmental	Economic	Combined
Repeated measure correlation	0.28	0.38	0.24	0.31
p-value	<0.001[c]	<0.001[c]	<0.001[c]	<0.001[c]

[a]Significant at $p = 0.05$.
[b]Significant at $p = 0.01$.
[c]Significant at $p = 0.001$.

correlation to determine the relationship between perceived sustainability and likeability while controlling for between-participant variance [24]. We chose this measure because it considers that multiple data points on the collage can be attributed to the same participant. The results are shown in Table 13. The correlations range from 0.24 to 0.38. These correlations are low despite being significant, suggesting that perceived sustainability plays a small role in likeability. Based on the correlations, perceived sustainability of a product accounts for 24–38% to liking a product. The *p*-values are highly significant; therefore, a statistically significant relationship exists between the placement of a product on the sustainability and like axes of the collage (i.e., we fail to reject hypothesis 2).

6 Discussion and Limitations

The patterns of evaluating product sustainability reveal essential insights for sustainable design. While it is important for designers to meet engineered sustainability criteria in products, the products need to also resonate with their intended customers. Therefore, designers must address both the perceptions and the engineered challenges of sustainability. Here, we detail the value of using features perceived as sustainable to resonate with customers and the effectiveness of the collage as an evaluation tool to validate sustainability perceptions.

First, we found that participants chose features perceived as sustainable over features that are not perceived as sustainable to describe products they identify as sustainable (Tables 7 and 11). To reiterate, perceived features used in this study may or may not contribute to engineered sustainability, yet they resonated with users as sustainable. Participants used positive features perceived as sustainable to describe products they identified as more sustainable and negative features to describe products they identified as less sustainable. Therefore, while perceived features might be different from engineered sustainability, they may also lead customers to learn more accurate information about a product.

Second, we measured a significant (yet low) correlation between participants liking a product and how sustainable they think the product is (Table 13). This relationship suggests that perceived sustainability plays a small role in how customers like a product, among other factors. When looking at sustainability as a whole, about 31% of why participants liked a product can be attributed to how sustainable they identified the product to be. The correlation was highest with environmental sustainability at 38%, while it was lowest for economic sustainability at 24%. The low correlations indicate that likeability and sustainability of a product can be measured separately and demonstrate the effectiveness of the collage as a tool in this context. This contrasts with Liao and MacDonald's study on smart products where other attributes like user delight and comfort had correlations of above 70% with likeability [19].

Third, we found that demographics can be a significant factor in how participants identify aspects of sustainability in a product (Table 8). For environmental sustainability, we found that education and employment had significant effects on how participants view sustainability. This suggests that education is an important factor in how participants perceive sustainability. For economic sustainability, income had a significant effect, while for social sustainability, race had a somewhat significant effect. These are both intuitive findings since race is an important social factor while income is an important economic factor. The effect of feature perceptions can therefore be enhanced by personalizing different sustainability aspects based on the target customer demographic. Interestingly, when looking at sustainability holistically, we found that demographics did not have any significant effect (Table 9). This suggests that the effects from specific sustainability aspects average out when combined. These results reveal potentially impactful insights, but it is important to note that our participant demographic was skewed and that a deeper analysis is needed for meaningful conclusions on demographic interactions with sustainability perceptions. For example, gender or age may also have a significant effect on how participants perceive social sustainability although our results did not reflect this.

Our findings have direct real-world implications for designers. First, designers can use the collage method to validate features perceived as sustainable from online reviews. This enables sustainable designers to confidently consider perceived sustainable features in their products. Second, we showed that features perceived as sustainable that are extracted from online reviews resonate with participants as sustainable despite the features not contributing to engineered sustainability. Designers should therefore combine perceived and engineered sustainability features to differentiate sustainable products with customers. We recommend that designers use the perceived sustainable features to improve how they communicate sustainability to customers in both their existing and next iteration products. For existing products, designers can adapt their product designs to include both features perceived as sustainable and engineered sustainability. For next iteration products, designers can use features perceived as sustainable to guide their design decisions in combination with engineered sustainability tools like life cycle analyses. For example, they can determine the color or texture of the material in their product to align more closely with customer sustainability perceptions while also meeting engineered sustainability criteria. For niche products, we recommend that designers consider the demographics of their customers to refine the perceptions based on the different aspects of sustainability.

These insights have crucial implications for designers but come with a few limitations. First, we did not test for generalizability, therefore the findings might not apply to other products. The features used in this study were extracted from online reviews of French press products [2], and we tested how participants used the features to describe French presses when evaluating sustainability. Second, our participant demographics were skewed which may mean that our demographics findings are not repeatable for evaluating product sustainability. For example, certain demographic variables that we found significant in our study might not turn out significant, or certain variables that were not significant (like gender for perceiving social sustainability) may be significant in a repeated study. Therefore, while the results show that demographics can have a significant role in sustainability perceptions, we recommend a deeper study to confirm how these variables interact with perceived sustainability. Third, while we demonstrated that there is a significant relationship between perceived sustainability and likeability of a product, this does not necessarily indicate purchase behavior. For example, the results showed perceptions of economic sustainability contribute just 24% to why participants like a product, but intuitively we know that price (an economic sustainability factor) plays a large role in a customer purchasing a product. In other words, a sustainable product may resonate with customers, but if the price point is too high it is unlikely that customers will purchase the product. This is supported by literature showing that intent to purchase a product does not equal making a purchase decision [25]. Other factors might also influence real online purchasing decisions that are not assessed here, for example the aesthetics of the product images or review ratings. We therefore recommend conducting an in-depth study on how features perceived as sustainable can influence purchase decisions of sustainable products.

7 Conclusion and Future Work

While perceived sustainability features may or may not contribute to engineered sustainability, this study validates that perceived sustainability features extracted from online reviews can help customers resonate more with product sustainability than features not related to sustainability. We designed a set of collage activities for participants to evaluate French press products on the three aspects of sustainability: social, environmental, and economic, and on how they like the products. We provided a list of features to describe the products including positive and negative features perceived as sustainable as well as features not perceived as sustainable. The features used in this study were extracted from online reviews of French press products in a previous study using annotations and a natural language processing algorithm [2].

Our findings point to important directions for sustainable design. First, designers can more effectively communicate product sustainability to customers using features perceived as sustainable, even if the features may not contribute directly to engineered sustainability. Participants placed features perceived as sustainable higher on the perceived sustainability axis of the collage than features not related to sustainability. Second, we measured a significant (yet low) correlation between the collage axes, perceived sustainability and likeability. This demonstrates that perceived sustainability plays a small role in liking a product, and the low correlation demonstrates that perceived sustainability and likeability can be measured separately. Moreover, it highlights the value of the collage approach for validating customer perceptions. Third, designers can use demographics to identify relevant feature perceptions in niche markets; however our findings suggest that feature perceptions can be generalized across demographics.

For future, we recommend closely investigating how perceived and engineered sustainable features interact with each other, the role of demographics on perceived sustainability, and how these factors can influence purchasing decisions to drive purchases for sustainable products.

Acknowledgment

We would like to thank Qatar National Research Fund (QNRF) for supporting this work. This work was funded by QNRF under the Qatar Research Leadership Program (QRLP).

We would also like to thank Dr. Conrad Tucker and Dr. Noah Goodman for their guidance while developing the method for this study. We would also like to thank Kesler Tanner for developing the original collage webapp.

This paper was also submitted to the 2021 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (IDETC/CIE 2021).

Conflict of Interest

There are no conflicts of interest.

Data Availability Statement

The datasets generated and supporting the findings of this article are obtainable from the corresponding author upon reasonable request. The authors attest that all data for this study are included in the paper.

References

[1] de Medeiros, J. F., Ribeiro, J. L. D., and Cortimiglia, M. N., 2016, "Influence of Perceived Value of Purchasing Decisions of Green Products in Brazil," J. Cleaner Prod., **110**(1), pp. 158–169.

[2] El Dehaibi, N., Goodman, N. D., and MacDonald, E. F., 2019, "Extracting Customer Perceptions of Product Sustainability From Online Reviews," ASME J. Mech. Des., **141**(12), p. 121103.

[3] Sheth, J. N., Sethia, N. K., and Srinivas, S., 2011, "Mindful Consumption: A Customer-Centric Approach to Sustainability," J. Acad. Mark. Sci., **39**(1), pp. 21–39.

[4] Peterson, M., and Brockhaus, S., 2017, "Dancing in the Dark: Challenges for Product Developers to Improve and Communicate Product Sustainability," J. Cleaner Prod., **161**(1), pp. 345–354.

[5] MacDonald, E. F., and She, J., 2015, "Seven Cognitive Concepts for Successful Eco-Design," J. Cleaner Prod., **92**, pp. 23–36.

[6] O'Rourke, D., and Ringer, A., 2015, "The Impact of Sustainability Information on Consumer Decision Making," J. Ind. Technol., **20**(4), pp. 882–892.

[7] She, J., and MacDonald, E. F., 2018, "Exploring the Effects of a Product's Sustainability Triggers on Pro-Environmental Decision-Making," ASME J. Mech. Des., **140**(1), p. 011102.

[8] Joung, J., and Kim, H. M., 2020, "Automated Keyword Filtering in LDA for Identifying Product Attributes From Online Reviews," ASME J. Mech. Des., **143**(8), p. 084501.

[9] Hou, T., Yannou, B., Leroy, Y., and Poirson, E., 2019, "Mining Changes in User Expectation Over Time From Online Reviews," ASME J. Mech. Des., **141**(9), p. 091102.

[10] Goldstein, E. B., and Brockmole, J. R., 2016, *Sensation and Perceptions*, 10th ed., Cengage Learning, Boston, MA, pp. 3–22.

[11] Papista, E., and Krystallis, A., 2013, "Investigating the Types of Value and Cost of Green Brands: Proposition of a Conceptual Framework," J. Bus. Ethics, **115**(1), pp. 75–92.

[12] MacDonald, E., Gonzalez, R., and Paplambros, P., 2007, "The Construction of Preferences for Crux and Sentinel Product Attributes," International Conference on Engineering Design, Paris, France, Aug. 28–31, pp. 609–626.

[13] Borin, N., Cerf, D. C., and Krishnan, R., 2011, "Consumer Effects of Environmental Impact in Product Labeling," J. Consum. Mark., **28**(1), pp. 76–86.

[14] Maccioni, L., Borgianni, Y., and Basso, D., 2019, "Value Perception of Green Products: An Exploratory Study Combining Conscious Answers and Unconscious Behavioral Aspects," Sustainability, **11**(5), p. 1226.

[15] Steenis, N. D., van Herpen, E., van der Lans, I. A., Ligthart, T. N., and van Trijp, H. C. M., 2017, "Consumer Response to Packaging Design: The Role of Packaging Materials and Graphics in Sustainability Perceptions and Product Evaluations," J. Cleaner Prod., **162**(1), pp. 286–298.

[16] Catlin, J. R., Luchs, M. G., and Phipps, M., 2017, "Consumer Perceptions of the Social vs. Environmental Dimensions of Sustainability," J. Consum. Policy, **40**(1), pp. 245–277.

[17] Rai, R., 2012, "Identifying Key Product Attributes and Their Importance Levels From Online Customer Reviews," ASME 2012 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference, Chicago, IL, Aug. 12–15, pp. 533–540.

[18] Guyton, A. A., 2006, "Developing Sustainable Product Semantics for Consumer Products: A Sustainable Designer's Guide," Ph.D. thesis, Georgia Institute of Technology, Atlanta, GA.

[19] Liao, T., and MacDonald, E., 2019, "Revealing Insights of Users' Perceptions: An Approach to Evaluate Wearable Products Based on Emotions," International Conference on Engineering Design, Delft, Netherland, Aug. 5–8, p. E14.

[20] McCaskill, A., 2015, "Consumer-Goods' Brands That Demonstrate Commitment to Sustainability Outperform Those That Don't," Nielsen (online), https://www.nielsen.com/us/en/press-room/2015/consumer-goods-brands-that-demonstrate-commitment-to-sustainability-outperform.html

[21] Goodman, J. K., and Paolacci, G., 2017, "Crowdsourcing Consumer Research," J. Consum. Res., **44**(1), pp. 196–210.

[22] Paolacci, G., and Chandler, J., 2014, "Inside the Turk: Understanding Mechanical Turk as a Participant Pool," Curr. Dir. Psychol. Sci., **23**(3), pp. 184–188.

[23] Maxwell, S. E., and Delaney, H. D., 2003, *Designing Experiments and Analyzing Data: A Model Comparison Perspective*, 3rd ed., Taylor and Francis, New York, pp. 724–745.

[24] Bakdash, J. Z., and Marusich, L. R., 2017, "Repeated Measures Correlation," Front. Psychol., **8**(456), pp. 1–13.

[25] Luchs, M. G., Naylor, R. W., Irwin, J. R., and Raghunathan, R., 2010, "The Sustainability Liability: Potential Negative Effects of Ethicality on Product Preference," J. Mark., **74**(5), pp. 18–31.