A Test for Product Design Features Perceived as Sustainable to Drive Online Purchasing Decisions

Nasreddine El Dehaibi
Mechanical Engineering,
Stanford University,
Stanford, CA 94305
e-mail: ndehaibi@stanford.edu

Aiyanna Herrera[1]
Mechanical Engineering,
Stanford University,
Stanford, CA 94305
e-mail: aiyannah@stanford.edu

Disney Rattanakongkham[1]
Mechanical Engineering,
Stanford University,
Stanford, CA 94305
e-mail: drattana@stanford.edu

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

Designers are challenged to create sustainable products that succeed in the marketplace, often relying on life cycle analyses to identify engineered sustainable features while neglecting perceived-as-sustainable (PAS) features. PAS features may not contribute to engineered sustainability but are identified by customers as sustainable. In previous papers, we proposed methods for extracting PAS features from online reviews using machine learning techniques and validating them using collage placement techniques. We demonstrated our methods using French presses (and other products). In this paper, we combined design and marketing approaches to test previously extracted PAS features in terms of purchasing products that include PAS features, as compared to others that do not. We built a simulated Amazon shopping experience using incentive alignment and constructed a within-subject, fractional factorial design with a variety of product features and physical appearances. We collected data on purchase intent, willingness to pay, and sustainability rating. We found that participants opted to purchase products with PAS features more often than products with features that are not PAS, termed "dummy" features. Participants also indicated they were willing to pay more for products with PAS features and rated those products as more sustainable, despite the features not contributing to engineered sustainability. Our findings demonstrate the potential value of identifying and including PAS features in sustainable products and a new application for shopping simulation experiments in design research. We recommend that sustainable designers include both engineered and PAS features in sustainable products to align with customer needs, drive purchasing decisions, and potentially increase profitability. [DOI: 10.1115/1.4054873]

Keywords: artificial intelligence, cognitive-based design, data-driven design, design automation, design for humans, design for the environment, design methodology, design process, design theory, design theory and methodology, design validation, life cycle analysis and design, machine learning, product design, product development, sustainable design, user-centered design

1 Introduction

Customers often reveal gaps between their stated purchase intentions and their real purchase decisions. For example, while 66% of global consumers state they are willing to pay more for sustainable products, this has not translated into real market success [1]. This gap can especially be found with sustainable products where factors like social desirability bias can influence what customers share about their purchase intentions. In a paper towel survey, for example, MacDonald et al. found that 87% of participants stated they would not buy non-recycled paper, but also indicated they bought from brands with non-recycled paper the last time they went shopping [2]. This mismatch challenges designers to create successful sustainable products in a market with an apparent growing demand from customers, but a lack of sustainability knowledge among customers.

Designers typically focus on the hard facts of sustainability when designing sustainable products. They use tools such as life cycle analysis (LCA) to prioritize design goals, such as energy usage and recyclability.[2] Unfortunately, much of this hard work is hidden within the final product, and unless customers know the right questions to ask, think to ask these questions, and know where to find the relevant information, many sustainable design efforts are never known to the customer. Much of what the customer perceives as related to sustainability is what they can see on the surface of the product, for example, see Ref. [3].

To address this challenge, designers have created methods for identifying design cues that can help customers form product perceptions and boost their intentions to purchase sustainable products. She and MacDonald demonstrated that perceived sustainable features termed "sustainability triggers" led participants to prioritize hidden sustainability features in a realistic decision scenario for toaster prototypes [4]. For example, a sustainability trigger like an embossed leaf pattern correlated with a prioritization of engineered sustainability features like energy usage in survey questions on purchase intention. In another paper, MacDonald et al. demonstrated that customer preferences are constructed, in-part, based on the context of the purchase decision and are not necessarily innate in people. The authors provided participants with slightly modified versions of discrete choice surveys for paper towels and found inconsistent preferences between them [5]. Seemingly superfluous features can therefore help customers value engineered (real) sustainability information about a product that aligns with their perceptions.

In a previous paper, we proposed a method for designers to extract perceived-as-sustainable (PAS) features from online reviews using a natural language processing machine learning algorithm combined with human annotators [6]. While these features

[1]These authors contributed equally.
[2]http://www.sustainableminds.com/

Contributed by the Design Theory and Methodology Committee of ASME for publication in the Journal of Mechanical Design. Manuscript received August 5, 2021; final manuscript received June 12, 2022; published online July 21, 2022. Assoc. Editor: Scott Ferguson.

 Copyright © 2022 by ASME

may not contribute to engineered sustainability, meaning they do not decrease the life cycle impact of the product, the features aid in communicating the purpose of the product to the customer. The inclusion of PAS features supports existing sustainable design methods like LCAs in that designers can create sustainable products that meet both engineered and perceived sustainability requirements.

As a case-study for our previously proposed method, we extracted salient PAS features from French press online reviews that drove positive and negative sentiment. We demonstrated that there is a gap between engineered and PAS features, highlighting the importance of accounting for both in design. In a subsequent paper, we used a novel collage approach to validate that users identified the PAS features as sustainable despite these features not necessarily contributing to engineered sustainability [7]. The collage consisted of a two-by-two axis where participants dragged and dropped products and selected features from a dropdown menu to describe the products.

In this paper, we conduct a strong test of PAS features by investigating how they can drive purchasing decisions for sustainable products in a simulated Amazon shopping experience (Fig. 1). Our goal is to provide a validated method for designers to create sustainable products that resonate with customers and drive purchasing decisions that are based, in-part, on valuing sustainability. We test the French press features extracted from our previous paper using a within-subject fractional factorial experiment design to demonstrate how PAS features influence purchase decisions. The rest of the paper is organized as follows: the background is presented in Sec. 2, followed by an overview of the propositions and hypotheses in Sec. 3. In Sec. 4 we describe the methods, in Sec. 5 we present the results, and in Sec. 6 we discuss the findings. Finally, we conclude the paper in Sec. 7.

2 Background

There is an extensive body of research from design and marketing on investigating customer choices and purchase decisions. Our work utilizes approaches from both research areas to identify how PAS features may influence online purchasing decisions. Common design approaches include conjoint and discrete choice analyses to tease-out preference of product features presented in different combinations of options. Designers can model and predict which product configurations are the most valued by customers based on these preferences [8]. Marketing approaches typically rely on historical data to model factors that influence purchase decisions. This section provides an overview of customer preference modeling from design and marketing and provides an overview of our previous papers that we build off for this paper.

2.1 Customer Preference Modeling in Design. This section presents an overview of recent customer preference modeling research in design. Suryadi and Kim proposed an automated method to construct choice sets using online product information and customer reviews [9]. The authors mined Amazon product data for 84 laptops and 46,194 verified customer reviews. From the data, they clustered products using X-means on the attributes, clustered customers using vector representation similarity of the reviews, and then constructed choice sets using a multinomial logit model. Using Kullback–Lieber (KL) divergence, the authors showed that the generated choice sets have higher preference predictive ability compared to a baseline random constructed choice set.

Fig. 1 Current paper builds off work from previous papers

Goucher-Lambert et al. used functional magnetic resonance imaging (fMRI) to investigate how customers make multi-attribute product decisions when considering sustainability [10]. The authors recruited participants to complete a within-subject conjoint analysis inside an fMRI. Participants were presented with two product options at a time with information on their form, function, and price. In the control condition, participants were also shown Poisson's ratio while in the test condition participants were shown information on the environmental impact. Using empirical fMRI results, the authors found that participants prioritized function while deprioritizing visual appearance when given environmental information about products. This work validated findings of a previous conjoint analysis study [11].

Tovares et al. proposed a method for incorporating experiences into consumer preference modeling [12]. The authors used virtual reality to provide participants with the ability to interact with products before indicating their preferences. Two within-subject experiments were conducted: one explored layout preference using a truck cab dashboard and the other explored form preference using mugs. For each experiment, participants completed a conjoint analysis with an experiential setup and a non-experiential (standard) setup. In addition, for the mug experiment, participants completed a "real" conjoint analysis where they interacted with real mugs before indicating preferences. The authors found that the experiential conjoint analyses did not provide better preference predictive capabilities than the visual conjoint analysis, although the results from the experiential and real mug conjoint analyses were statistically similar.

Maccioni et al. investigated preferences for sustainable products using a combination of stated preferences and biological measurements [13]. The authors recruited 43 participants to evaluate 20 baseline products and 20 eco-friendly products. Participants wore eye-tracking equipment and a device that measures galvanic skin response while evaluating products. The authors found that participants perceived eco-friendly products as more innovative while they perceived the baseline products as more functional and reliable. No significant results were found from biological measurements.

2.2 Customer Preference Modeling in Marketing Using Online Reviews. This section presents some relevant papers from marketing research that focus on customer preference modeling using online reviews.

Chevalier and Mayzlin studied the impact of online reviews on sales using data from Amazon and Barnes and Nobles [14]. For a sample of books, the authors compared differences in the number of reviews and their ratings over three time points from both websites and determined their relationship with relative sales rank. Using a linear model, the authors found that positive reviews on one site correspond to higher sales relative to the other site. Moreover, they found that the decrease in sales from a negative review is greater than the increase in sales from a positive review.

Chen et al. disaggregated the impacts of online reviews and recommendations on online sales rank [15]. The authors used digital cameras as a case-study and collected information on number of reviews, ratings, recommended cameras in terms of purchase percentage, and sales rank. Using a linear model, the authors found that a negative review had a much greater impact on sales than a positive one. Moreover, they found that positive recommendations (high purchase percentages, for example, "86% of users ultimately purchase this product") have a positive effect on sales while negative recommendations (low purchase percentages) have no effect on sales.

Liu studied the impact of Yahoo movie reviews on Box Office revenues [16]. The author found that reviews are most active during the prerelease of a movie and more critical after the release. Moreover, using a linear model, Liu found that the volume of reviews around the time of release correlates with Box office revenues and not the valence of the reviews. Dhar and Chang built on this by studying the impact of blog posts and social networking sites on music sales [17]. The authors collected the volume of blog posts for an album, the number of friends an artist has on Myspace (a social media platform), and the number and ratings of online album reviews. The authors used a linear model to study the impact of the data on music sales four weeks before and after the album release. Data for 108 albums were collected. Album sales were computed based on sales ranks from Amazon.com. The authors found a positive correlation between the volume of blog posts with future album sales.

A limitation of the above approaches is that they do not study how specific product features may be driving online sales. Our work utilizes approaches from design and marketing to test PAS design features and provide actionable insights for designers on driving purchasing decisions for sustainable products. An overview of our previous work is included below to provide context on PAS design features.

2.3 Extracting and Testing Features Perceived-as-Sustainable From Online Reviews. We briefly summarize two previous papers here, as this paper builds off them. The first paper developed a semi-automated approach to extract features PAS from online reviews using crowdsourced annotations of online reviews and a natural language processing algorithm [6]. As a case-study, we used French presses and collected 1474 reviews to extract PAS features from. We recruited 900 annotators from Amazon Mechanical Turk (MTurk), a website for hiring crowdsourced workers, to highlight phrases in reviews they perceived-as-sustainable. Annotators were trained and assigned to one of the three sustainability pillars: social, environmental, and economic. Using a logistic classifier model and for each sustainability pillar, we then extracted salient PAS features with positive and negative sentiment based on the beta parameters of the model. Table 1 shows the positive salient features extracted. A subset of these features is selected for this paper (see Sec. 4.2.1).

The features in Table 1 include a combination of visual and descriptive, and tangible and intangible features. Notably, energy and water consumption were not identified as salient environmental PAS features although they are important engineered sustainability requirements for French presses. To investigate this further, we conducted a life cycle analysis and found the use phase (where energy and water consumption contribute) had one of the largest environmental impacts (Fig. 2). This gap between engineered and PAS features highlights the importance for designers to consider both in sustainable design.

The second paper tested the extracted PAS features to determine if users would identify them as sustainable and how the features might relate to users liking a product [7]. We designed a novel collage approach where participants dragged and dropped products on a set of two axes, sustainability and likeability, and selected features from a dropdown menu. The collage consisted of multiple versions so that we may study each sustainability pillar (social, environmental, and economic) separately. By studying each pillar separately, we gained a deeper understanding on the nuances of customer perceptions of sustainability. Figure 3 shows an example of a social sustainability version of the collage where a product is being placed on the collage with features to select from. The placement of products and features on the collage validated that participants identified PAS features as sustainable and that the collage is an effective tool for testing customer perceptions.

Table 1 Positive features of French presses perceived-as-sustainable [6]

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

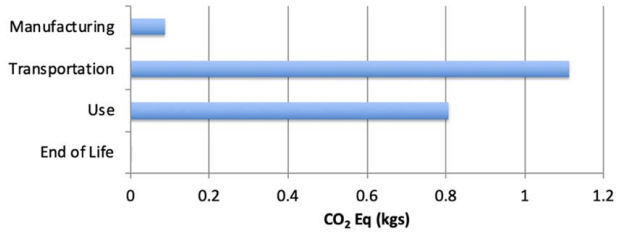

Fig. 2 Life cycle analysis of French press

3 Research Propositions and Hypotheses

In this research, we leverage PAS French press features identified in the previous work. These features do not contribute to real engineered sustainability [6]. We test these features in a simulated Amazon shopping experience where we modified images, descriptions, and reviews according to PAS features as well as alternative features that are not PAS, termed "dummy" features. We tested the influence of the PAS features on purchase decisions in a within-subject study. The following propositions and hypotheses are tested.

Proposition 1. *Online customers rely on product descriptions to guide their purchasing decisions. Based on this, we propose that designers can modify descriptions to drive purchasing decisions for sustainable products.*

Hypothesis 1. *Participants are more likely to opt to purchase a product when the description is combined with features extracted from online reviews that are perceived-as-sustainable versus dummy features.*

Proposition 2. *Online customers rely on product descriptions to learn about products. Based on this, we propose that designers can modify product descriptions so that customers resonate more with sustainable products.*

Hypothesis 2a. *Participants will rate a product as more desirable to purchase when the description is combined with perceived-as-sustainable features extracted versus dummy features.*

Hypothesis 2b. *Participants will rate a product as more sustainable when the description is combined with perceived-as-sustainable features versus dummy features.*

4 Method

To test the hypotheses, we designed and conducted a simulated Amazon shopping experience for 200 Amazon Mechanical Turk workers (referred to as participants, see Sec. 4.4 for more information). Participants browsed between product options based on fractional factorial design and selected to "purchase" a product as if they were making real purchase decisions. Participants also completed a post-experiment-survey, in which they rated the products based on purchase desirability and sustainability. In the following sections, we provide an overview of the experiment design and the products and features used in the shopping simulation. The experiment contained base and dummy features that we created

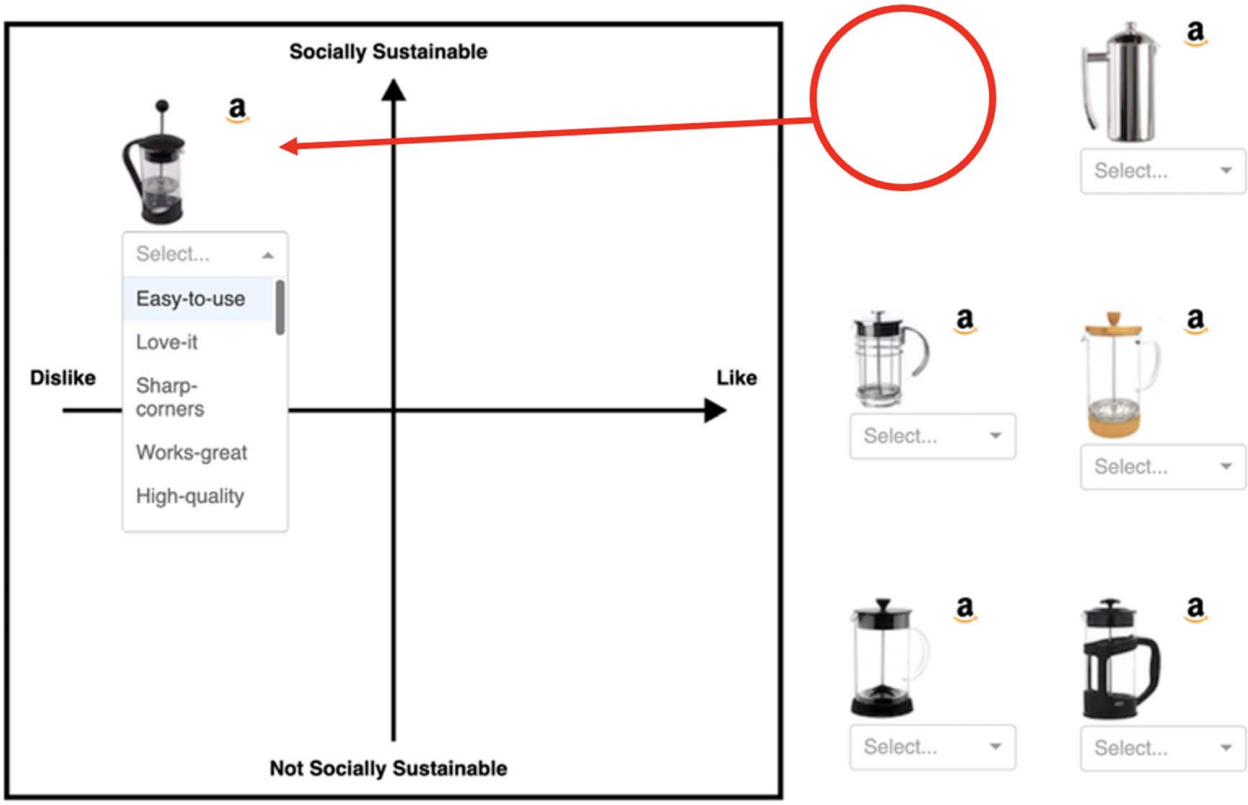

Fig. 3 Dragging and dropping products on collage and selecting at least one feature to describe each product

for this experiment, and PAS features extracted from a previous paper [6]. Then we discuss how we tested the products and features in a simulated Amazon shopping experience. Finally, we provide an overview of the post-survey and the participants in the experiment.

4.1 Experiment Design Overview. The experiment compared how participants made purchasing decisions when given products with dummy features in a control condition versus products with PAS features in a test condition (see Fig. 4). The stimuli included base products to create a reference point between both conditions. We used a within-subject experiment design to assess how PAS features can influence purchasing decisions.

The experiment was conducted via a Qualtrics survey with instructions about the activity, a test to make sure participants understood the task, and links to the shopping simulations. The test questions asked participants to recall from the instructions how many product options were included in each shopping simulation, the type of product they were shopping for, and the number of shopping simulations they were completing.

After passing the test, participants received links that led them to the control and test condition shopping simulations. Participants always completed the control condition first to limit the chance of social desirability bias influencing participants' choices in the following condition. The goal was to provide the least advantage for PAS features to rigorously test their ability to drive purchase decisions in the test conditions. Each shopping simulation condition displayed four products to browse from. In the control condition, two base products and two products with dummy features were displayed. In the test condition, the same two base products and two products with PAS features were displayed. Participants had to spend a minimum of 5 min on each simulation before they could

Fig. 4 Within-subject experiment design

Table 2 Breakdown of product features

Feature category	Feature name	Level 1	Level 2
Base	Handle shape	Circular	Rectangular
Base	Spout	Filter	Easy-pour
Base	Lid	Button	Lift
Dummy	Hourglass timer	Present	Not present
Dummy	Ventilated lid	Present	Not present
PAS	Material	Stainless steel	Plastic
PAS	Glass	Strong glass	Not present
PAS	Clean	Easy to clean	Not present
PAS	Quality	High quality	Not present
PAS	Gift	Perfect gift	Not present

proceed to the next one. With four products per simulations, we selected a 5-min minimum to ensure participants adequately evaluated all options. To incentivize participants to evaluate products carefully, we used incentive alignment as part of their reward (see Sec. 4.4).

Following the completion of the shopping simulation, participants received a password to a Qualtrics post-survey. Participants entered which product they selected for purchase and rated all products on a 5-point Likert scale based on their purchase desirability and based on their sustainability. As a proxy for purchase desirability, the survey asked participants to rate their willingness to purchase (WTP) a product. Participants also selected from a list the main driving factor in their shopping decisions online. Lastly, participants entered demographic information before the completion of the survey.

4.2 Products. The experiment focused on French press coffee brewers, building off our findings from previous papers [6,7], because they are popularly reviewed products, sold with various esthetic and practical design features. Additionally, the sustainability concerns associated with French presses present an opportunity to study consumers' perceptions of sustainability, such as how quickly a glass exterior might break or how much energy can be saved with a heat-insulated press. The experiment showed four presses that had a variety of features in each condition to simulate a realistic shopping experience. In this section, we describe the features, images, descriptions, and reviews that we used in the paper.

4.2.1 Product Features. The presses had three feature categories: base features, dummy features, and PAS features (see Table 2). Each category had two levels that could be displayed. Note that French presses had dummy features in the control condition, and PAS features in the test condition, never both.

All products shown included all three base features at one level, consisting of core functional features of a French press, including a handle, spout, and lid. For the handle, the varieties included either a circular or rectangular shape; for the spout, the varieties included filtered or easy-pour; and for the lid, these included a button or lift mechanism. The purpose of these features was to provide a baseline to compare purchasing decisions between the two experimental conditions. The base features pilot tested as neutral and did not impact customer sentiment significantly, see Sec. 4.3.2 for details.

Dummy features are intended to appeal to customers for their functionality but are not strongly related to perceived sustainability. These included a built-in handle hourglass timer for proper brewing time, and a ventilated lid to help steam escape. The dummy features were included in the control conditions only. The goal of the dummy features was to challenge and assess the popularity of products with PAS features. Pilot-testing aided in selecting dummy features that were on par with the PAS features in terms of purchase desirability (see Sec. 4.3.2).

A previous paper developed a method to extract PAS features from online reviews, demonstrating the method using French presses and identifying a gap between engineered and PAS features (Table 1) [6]. We selected a subset of PAS features for this experiment, shown in Table 2. We selected this subset because it includes both visual and descriptive features and is representative of perceptions from the three sustainability pillars: social, environmental, and economic.

For the material feature, products were made of plastic or stainless steel (with steel being a PAS feature). For the remaining PAS features, the products either had or did not have them, e.g., strong glass, easy to clean, high quality, and perfect gift—all PAS. The PAS features were included in only the test condition.

4.2.2 Fractional Factorial Experiment Design. With the available features in Table 2, we used a fractional factorial experiment design to account for different combinations and created 12 different products—each participant saw eight of these products. The features per product are shown in Table 3. For each product we created images, descriptions, and reviews to include in the Amazon shopping simulation as described below.

4.2.2.1 Product Images. We rendered images of the 12 products for this study using the computer-aided design software Fusion 360 (Fig. 5). Products 1 and 2 include base features only and are shown in every experiment condition. Products 3 and 4 have dummy features and are only shown in the control condition. Products 5–12 have PAS features; we randomly created five random pairs and assigned participants to a test condition with one of five pairs. The other two products in the test condition were the base products.

Table 3 Features per product

		Base Features						Perceieved Sustainable Features					Extra Features	
		Handle		Spout		Lid		Stainless	Strong	Easy to	High quality	Perfect	Hourglass	Ventilated
French Press #	Condition	Circular	Rectangular	Filter	Easy-pour	Button	Lift	steel	glass	clean	construction	gift	timer	lid
1	C&T	■		■		■								
2	C&T		■		■	■								
3	C		■	■			■						■	
4	C	■			■		■							■
5	T	■		■		■		■						
6	T		■		■	■			■					
7	T		■	■			■			■				
8	T	■			■		■				■			
9	T	■		■		■						■		
10	T		■		■	■		■	■					
11	T		■	■			■	■		■				
12	T	■			■		■	■			■			

Fig. 5 Product image renderings

Each French press was designed to closely resemble the others, as well as those on the market. All products were shown on a white background from the same angle. Additionally, each product had a close-up image of the handle and top.

Pilot-testing ensured that the rendered product images were equally esthetically pleasing (see Sec. 4.3.2), aiming to minimize the effect of other potential variables on purchasing decisions. It is important to note that some features in Table 2 cannot be shown visually, for example "easy to clean." These features are included in descriptions on the Amazon simulated product page.

4.2.2.2 Product Descriptions. Each product had a corresponding description that outlined the product's features in a bulleted list. Descriptions were written with the goals of brevity, maintaining Amazon's organizational structure, and emphasizing the feature. Each description ranged between 25 and 30 words, and the feature was described at the beginning of each list item (see Fig. 8 in Sec. 4.3.1 for an example).

Products were also titled using their features, a technique that is commonly used on Amazon. For example, Product 1 was named "French press Coffee Maker with button for lid removal, filtered spout, and circular grip handle." These descriptions made the product features noticeable. Pilot-testing ensured that the descriptions were equally readable and understandable (see Sec. 4.3.2).

4.2.2.3 Product Reviews. Each product was shown with five unique reviews: three 5-star reviews and two 3-star reviews. All products were rated as 4.2 stars overall. Having products with equally positive reviews mitigated the influence of reviews on purchasing decisions. Additionally, the reviews did not mention any of the features in the experiment and did not mention sustainability. We created the review text as follows. First, we selected initial candidate reviews from Amazon listings of French presses and then removed specific details so that the reviews were applicable to any French press. Each review had only two to three sentences total. Pilot-testing ensured that the five reviews for each product were equally positive (see Sec. 4.3.2).

4.3 Amazon Shopping Experience. In this section, we present the shopping experience flow as well as measures taken to normalize web content between products.

4.3.1 Simulation Flow. Participants were able to freely click and browse between three types of pages: the product search page, the product pages, and the checkout page (Fig. 6). The experience was a controlled simulation without distracting content, such as advertisements and hyperlinks to other web pages on Amazon.com.

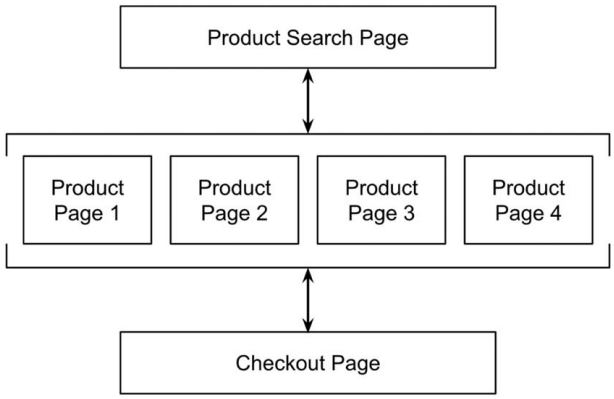

Fig. 6 Simulated Amazon flow

The product search page showcased four French presses on the participant's screen (Fig. 7). This page was intended to replicate the results of a consumer searching "French presses" on Amazon.com. The products shown on this page depended on which experiment condition the participants were taking (see Sec. 4.2).

From the product search page, participants can click on the product's image, price, title, or reviews to access the product information page (Fig. 8). This page provides details on the product's features, five reviews from past consumers, and three images of the French press product. The participants can zoom in on the image for a closer look at the French press product. All links to external pages were deactivated to prevent the participant from navigating away from our survey. From the Product Page, participants can click on the "Buy Now" button on the right side to access the checkout page or go back to the product search page to read about another product.

The checkout page was intended to model the experience of officially purchasing a product on Amazon.com (Fig. 9). All data entry queries were removed so that the user did not enter any personal information to proceed with buying the French press of their choice. Clicking the "Place your order" button ended the shopping experience.

4.3.2 Normalizing Web Content. We took careful steps to normalize web content between products and control the influence of external variables on purchasing behavior. External variables include brand, price, number of reviews, review ratings, review content, description content, number of images, and image quality.

Prior to launching the full shopping simulation, we conducted a pilot study to measure the equality of images, descriptions, and reviews between products used in the study. The goal was to control variables so that only product features had a significant difference between products. The pilot study asked participants to indicate their level of agreement on a range of statements using a 5-point Likert scale. For the images, the pilot study asked participants to rate how esthetically pleasing the images are and their level of quality. For the descriptions, the pilot study asked participants to rate how easy they are to read and understand. For the reviews, the pilot study asked participants to rate a set of reviews in terms of the clarity and sentiment of the text. Participants rated images, descriptions, and reviews separately. We modified product information and ran several rounds of the pilot study to achieve no statistically significant differences in ratings between products across the board.

Additional measures we took include controlling for branding and prices. To prevent branding or previous knowledge of a brand from impacting purchasing decisions, we removed brand names from the product titles, descriptions, and images. Product prices had a $2 range between $27.98 and $29.99. We decided on

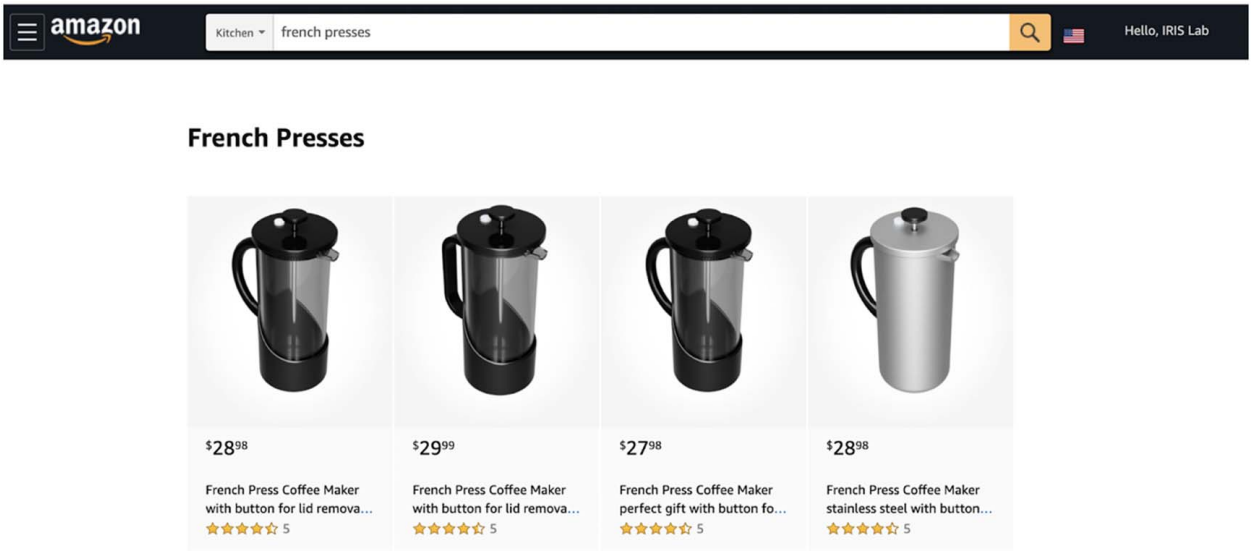

Fig. 7 Product search page

this range so that the different prices can portray a realistic shopping experience, but at the same time have a negligible effect on purchase behavior.

4.4 Participants. We recruited a total of 200 participants from MTurk to complete the shopping experience which took 25 min on average; participants were compensated $6 each for their time. To create an incentive alignment, we also entered participants into a lottery for a product of similar or less value to the product they chose for purchase in the experiment. We opted to recruit from MTurk instead of in-person participants to accommodate for COVID-19 restrictions and to quickly collect many responses. Moreover, MTurk demographics are likely to match online users and therefore online shopper demographics more closely [18].

To increase the quality of data collected, we screened for participants on MTurk using features on the MTurk platform as well as screening questions in Qualtrics. We required that participants should have a 97% prior approval rating and are based in the United States; literature shows that respondents in the United States tend to deliver better quality responses [19].

Out of the 200 participants that completed their task, we approved 162 based on two requirements: (1) completing the activity in time (t) that is within 1 standard deviation(s) of the average time to complete the activity (μ) or longer (i.e., $t \geqslant \mu - s$) and (2) correctly answering the check question, "What is the capital city of the United States?" which we asked in the post-survey. We did not analyze results if they did not meet one or both criteria. We used similar approval criteria in previous papers [6,7].

5 Analysis and Results

This section is split into two parts: first, we present participant data and demographics, and second, we present the shopping experience results that test the hypotheses from Sec. 3.

5.1 Demographics. The demographics of the 162 approved participants are summarized in Fig. 10. Our participants were mostly young, white, educated, working full-time, about 60% male, and making about an average US income. The demographics of our participants are like those of the Amazon Mechanical Turk respondents in our previous paper [7] and are in-line with demographic analyses of Amazon Mechanical Turk respondents from the literature [19]. While this demographic is not representative of the general US population, it is closer to typical online users and is ideal for studying online purchasing decisions [7].

Figure 11 shows the most important factor participants reported for making purchases on Amazon. Note that participants reported what factors are generally most important to drive their purchase decisions, and not specifically what drove their choice for this experiment. Reviews, brand, and price were the highest three factors, followed by product description. In our shopping simulation, we normalized all factors besides product description to isolate the influence of different product features on purchasing decisions.

5.2 Shopping Simulation. This section is split into three parts: first, we present the results on purchasing decisions that test hypothesis 1 (products with PAS features are more likely to be purchased than those with dummy features), second, we present the results on purchase desirability ratings that test hypothesis 2a (products with PAS features are rated as more desirable to purchase than those with dummy features), and third, we present the results on sustainability ratings that test hypothesis 2b (products with PAS features are rated as more sustainable than those with dummy features).

5.2.1 Products Selected for Purchase. Figure 12 shows the raw counts of products selected for purchase in the control and test conditions. More participants selected to purchase products with PAS features in the test condition than products with dummy features in the control condition, suggesting that products with PAS features can drive purchasing decisions.

To determine the influence of the experiment conditions relative to the base products, Fig. 13 shows the fraction of products selected for purchase with PAS features in the test condition versus with dummy features in the control condition. Approximately 80% of products selected in the test condition were products with PAS features while 71% of products selected in the control condition where products with dummy features (as opposed to only having base features), supporting hypothesis 1.

We tested if the difference in the fraction of products selected for purchase between the control and test conditions was statistically significant using a t-test, shown in Table 4. The difference was significant at the 0.05 level, supporting hypothesis 1. The findings therefore indicate that participants are more likely to select to purchase a product when the description is combined with PAS features rather than dummy features.

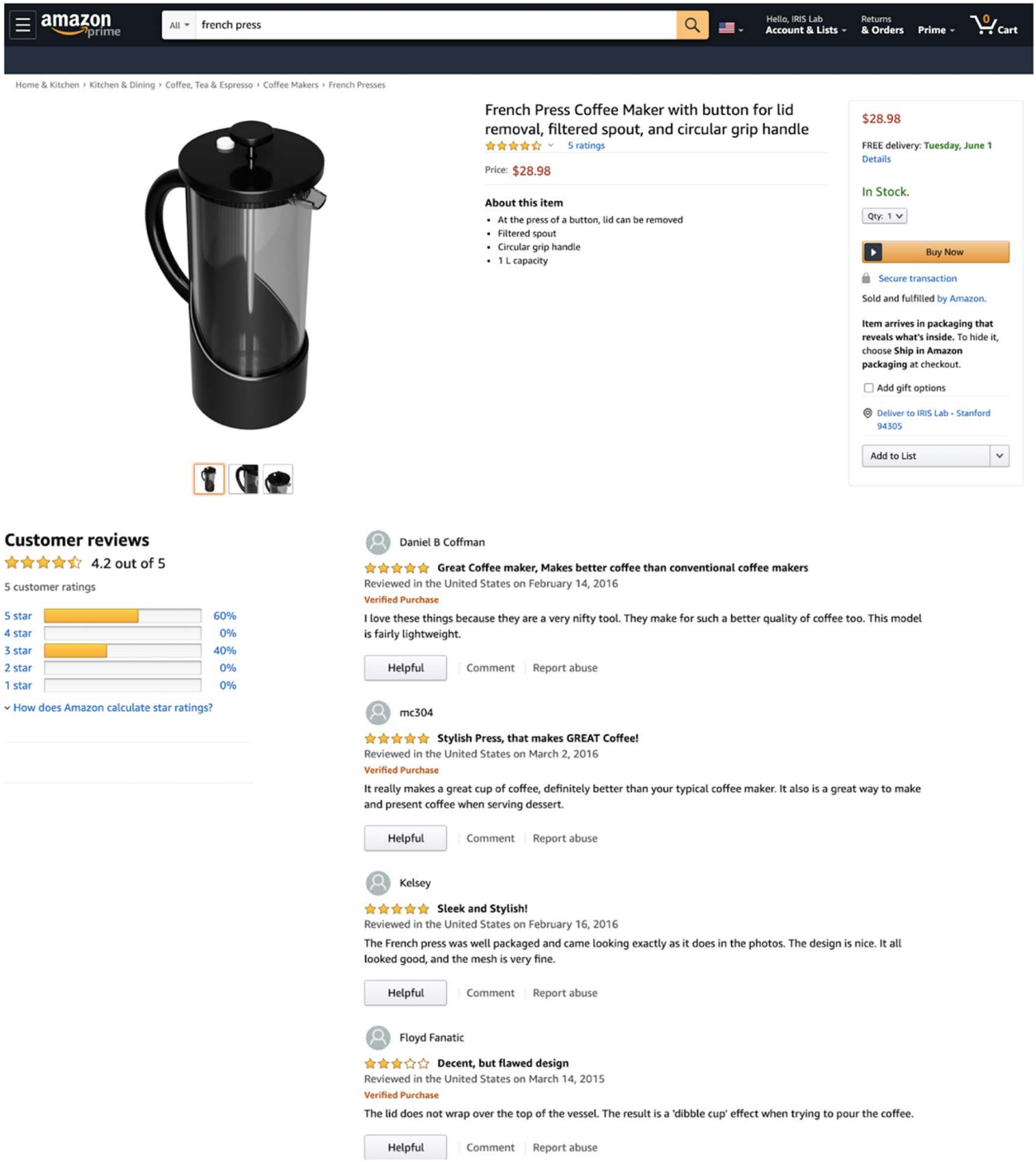

Fig. 8 Product information page

5.2.2 Purchase Desirability Ratings. As a proxy for purchase desirability, the survey asked participants to rate their WTP products on a 5-point Likert scale. Figure 14 shows the mean willingness to pay ratings for products in the control and test conditions. The WTP for products with PAS features in the test condition is slightly higher than the WTP for products with dummy features in the control condition.

To determine the influence of the experiment conditions on WTP relative to the base products, Fig. 15 shows the mean difference in WTP in the test conditions versus the control condition. In the control condition, participants rated WTP products with dummy features 0.34 higher than the base products on a 5-point Likert scale. In the test condition, participants rated WTP products with PAS features 0.52 higher than the base products on a 5-point Likert scale. The greater difference in WTP in the test condition supports hypothesis 2a.

The *t*-test results for mean difference in WTP are included in Table 5, showing that the difference between conditions is statistically significant at the 0.05 level. The findings therefore indicate that participants rate products as more desirable to purchase when the description is combined with PAS features versus dummy features.

5.2.3 Sustainability Ratings. The survey asked participants to rate products on their sustainability using a 5-point Likert scale.

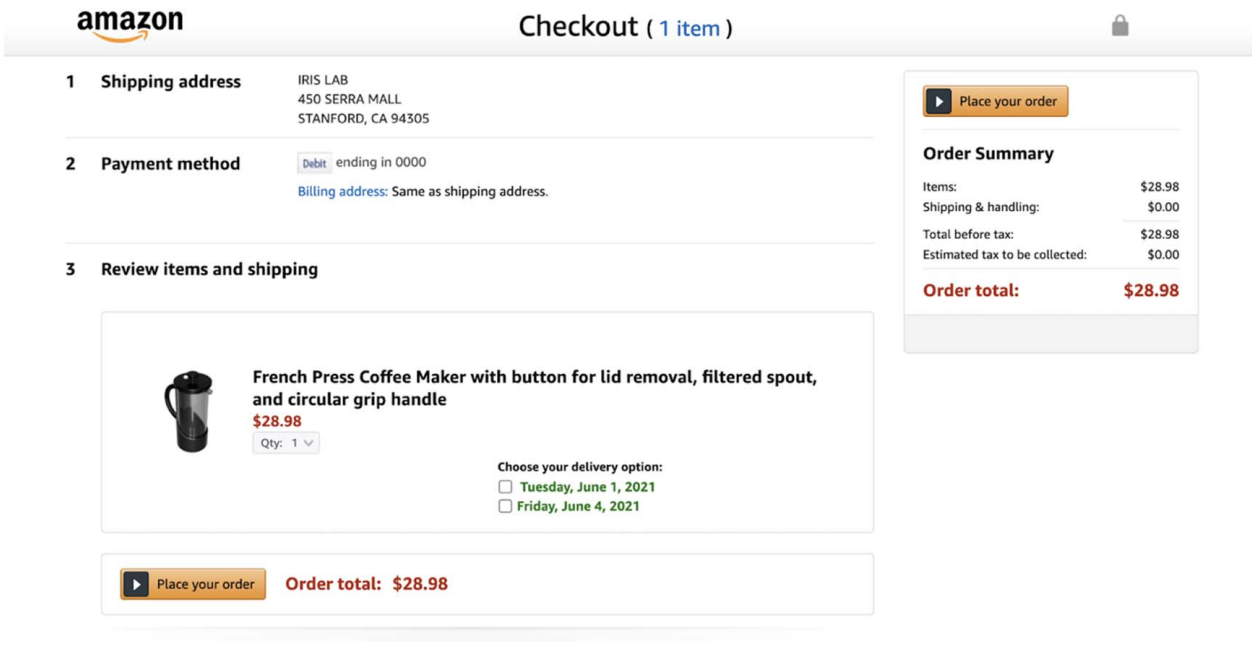

Fig. 9 Checkout page

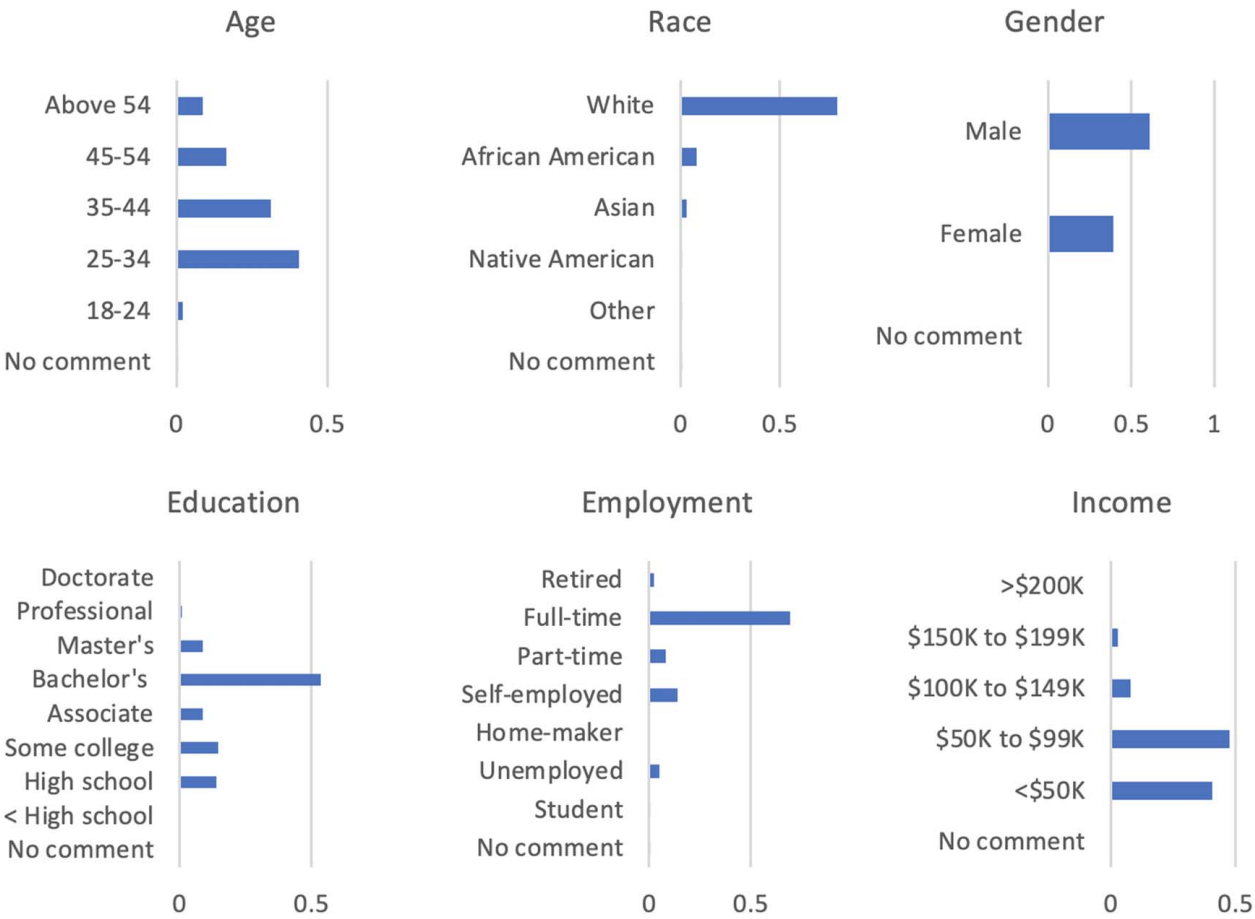

Fig. 10 Participant demographics

Figure 16 shows the mean sustainability ratings for products in the control and test conditions. On average, the sustainability rating for products with PAS features in the test condition is higher than the sustainability rating for products with dummy features in the control condition.

To determine the influence of the experiment conditions on sustainability ratings relative to the base products, Fig. 17 shows the mean difference in sustainability rating of the base products under the control conditions versus the test conditions. In the control conditions, participants rated products with dummy features

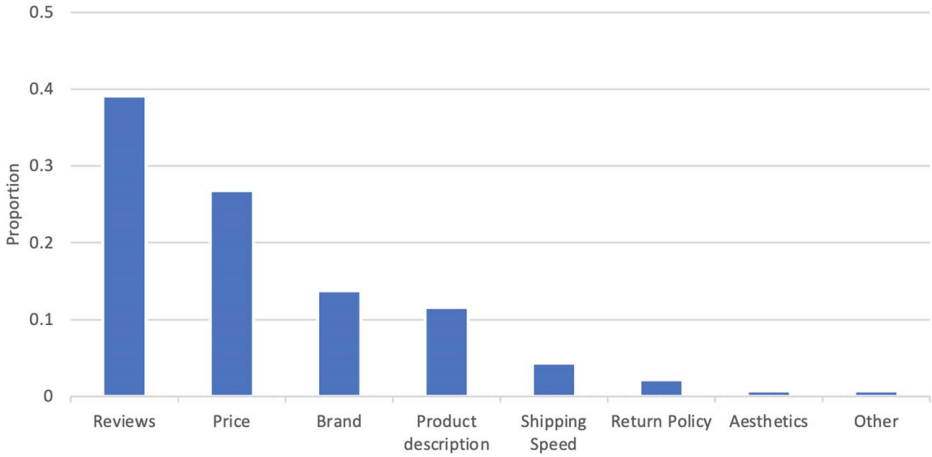

Fig. 11 Self-reported important factors for purchasing on Amazon by participants

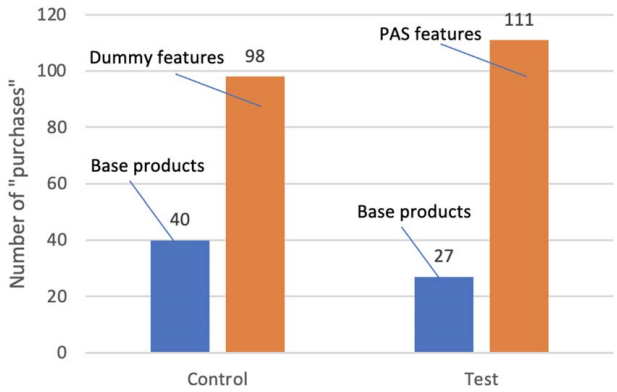

Fig. 12 Number of purchases for base, dummy, and PAS products

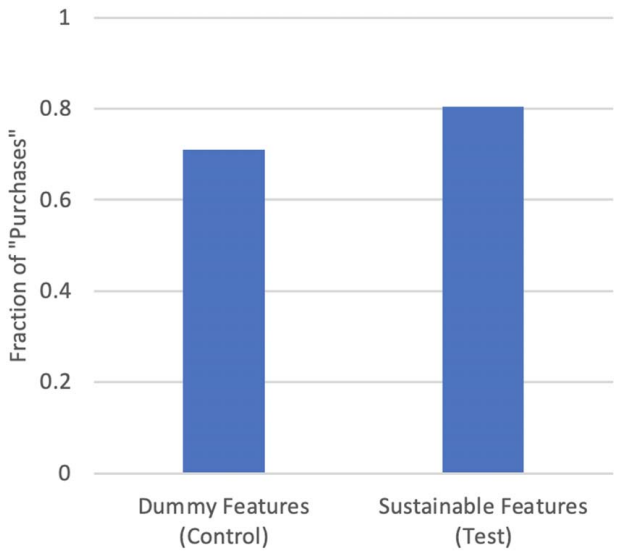

Fig. 13 Fraction of products selected for purchase in the control condition versus the test condition

0.21 higher than base products on a 5-point Likert scale. In the test conditions, participants rated products with PAS features 0.48 higher than the base products on a 5-point Likert scale. The greater difference in mean sustainability rating in the test condition supports hypothesis 2b.

Table 4 Two sample *t*-test between control and test conditions for fraction of products selected for purchase

	Dummy features (control)	Sustainable features (test)
Mean fraction of purchases	0.71	0.80
Variance	0.21	0.16
Observations	162	
$P(T \leq t)$ one-tail	0.026*	
t-Critical one-tail	1.66	

Note: *significant at $p = 0.05$, **significant at $p = 0.01$, ***significant at $p = 0.001$.

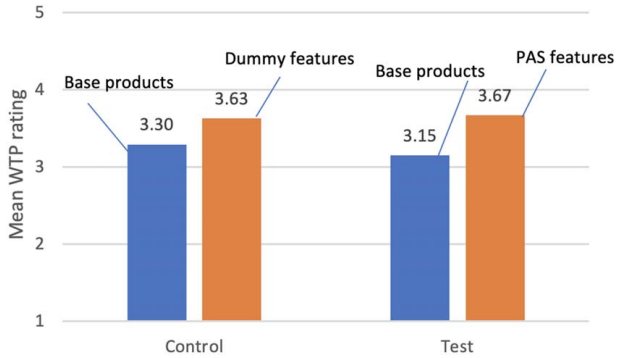

Fig. 14 Willingness to pay rating for base, dummy, and PAS products

The *t*-test results are included in Table 6 and show that the difference between the two conditions is statistically significant at the 0.001 level. The findings therefore strongly indicate that participants will rate products as more sustainable when the description is combined with PAS features versus dummy features. This also validates our previous work with assessing PAS features using a collage approach [7].

6 Discussion

The results provide actionable insights for designers on how to make sustainable products more successful online. The experiment approximated purchase decisions using a simulated shopping experience with incentive alignment and measured both customer preferences and purchase decisions. In this section, we discuss the

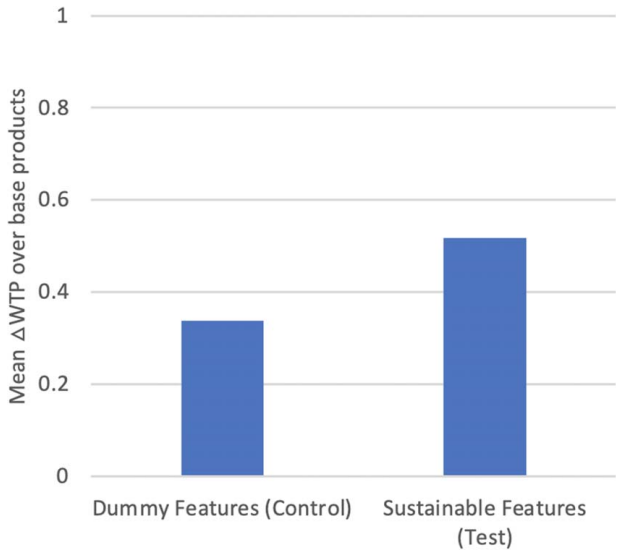

Fig. 15 Mean △WTP in the control condition versus the test condition

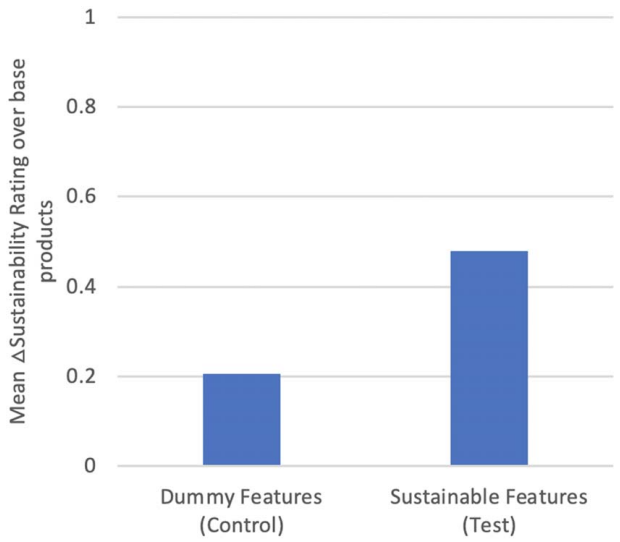

Fig. 17 Mean △sustainability rating in the control condition versus the test condition

Table 5 Two sample *t*-test between control and test conditions for mean △WTP

	Dummy features (control)	Sustainable features (test)
Mean △WTP	0.34	0.52
Variance	1.95	2.23
Observations	324	
$P(T \leq t)$ one-tail	0.039*	
t-Critical one-tail	1.65	

Note: *significant at $p = 0.05$, **significant at $p = 0.01$, ***significant at $p = 0.001$.

Table 6 Two sample *t*-test between control and test conditions for mean △sustainability rating

	Dummy features (control)	Sustainable features (test)
Mean △sustainability rating	0.21	0.48
Variance	0.87	1.58
Observations	324	
$P(T \leq t)$ one-tail	0.001***	
t-Critical one-tail	1.65	

Note: *significant at $p = 0.05$, **significant at $p = 0.01$, ***significant at $p = 0.001$.

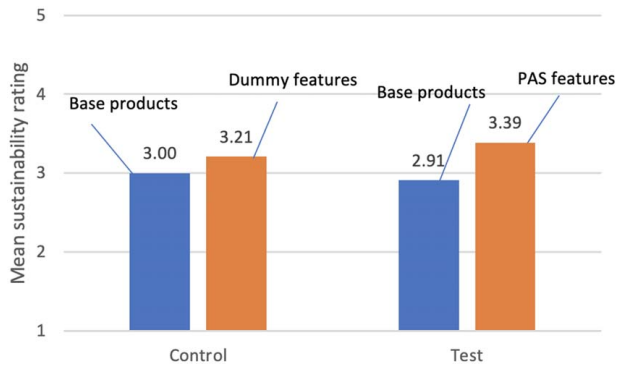

Fig. 16 Sustainability rating for base, dummy, and PAS products

value of using PAS features in design as well as using simulated shopping experiences for customer preference modeling.

First, the results showed that participants selected to purchase products with PAS features more than they did with dummy features (Fig. 15 and Table 4). This supports our first proposition that designers can modify descriptions to drive purchasing decisions for sustainable products. It is important to note that PAS features may or may not contribute to engineered sustainability [6]. We demonstrated that despite the importance of LCAs to inform engineered sustainability features, our proposed method to extract PAS features can drive purchasing decisions for sustainable products.

Second, the results showed that participants are willing to pay more for products with PAS product features compared with dummy features based on 5-point Likert scale ratings (Fig. 15 and Table 5). This supports our second proposition that designers can modify product descriptions so that customers resonate with sustainable products. The preferences that participants stated in the Likert ratings matched with their purchase decisions in the simulated shopping experience, indicating the value of using simulated shopping experiences in design to model customer preferences. Moreover, the finding supports previous literature that participants are willing to pay more for products they perceive as sustainable [1].

Third, the results showed that participants rated products with PAS features as more sustainable compared with dummy features based on 5-point Likert scale ratings (Fig. 17 and Table 6). It is important to note that none of the PAS, dummy, or base features contribute to engineered sustainability. The finding supports our previous work that PAS features resonate with participants as more sustainable [7], emphasizing the value of using PAS features to communicate product sustainability to customers.

The results in this paper demonstrate that designers should use PAS design features in addition to engineering sustainable features to align sustainable products with customer perceptions. In doing so, designers can create products that are both engineered to be sustainable as well as successful in the marketplace. For example, an LCA might indicate that choosing plastic is a more sustainable manufacturing option [6] but adding some stainless steel elements to a product might be worth the trade-off to drive online sales. An LCA could determine if dropping plastic entirely in favor of metal can

actually be beneficial to the environment, due to the promotion of other "invisible" design features, such as energy-savings or shipping.

With the knowledge that not all PAS features align with engineered sustainability features, it is important to consider ethical implications of this work. Although this research has focused on the significance of both engineered sustainability and perceived sustainability, the results demonstrate that the two might not always be aligned in practice. If used with malintent, the findings of this work could be used to create products that customers perceive are sustainable but are not in reality. The intent of this research, however, is to shed light on the difference between perceived sustainability and engineered sustainability. It is up to the designers and sellers to encourage ethical practices when designing their products. Similarly, consumers should be aware that there can be a disconnect between perceived sustainability and engineered sustainability. This research benefits consumers by helping them make more informed decisions about their purchases, since it is not always the case that a product they perceive as sustainable is sustainable.

While the responsibility of making informed purchase decisions ultimately lies on the customer, Amazon could use the findings in this work to facilitate and guide informed purchase decisions. For example, Amazon could monitor PAS features mentioned in online reviews using natural language processing techniques proposed previously [6]. Moreover, Amazon could allow users to flag reviews that might be spreading misinformation. Ideally, the findings of this work can enable both designers and e-commerce platforms to build an informed customer base that can bridge the gap between engineered and PAS features, and drive purchases for sustainable products.

There are important limitations to keep in mind with the findings in this paper. First, while we carefully designed the simulated shopping experience to be as realistic as possible, the activity did not involve real purchasing decisions. We included incentive alignments to approximate real purchase decisions, but the results may differ in a real shopping environment with real products. Second, the shopping simulation was a controlled environment with variables kept constant except for the product features. In reality, customers are exposed to varying types of images, descriptions, prices, and reviews when shopping online. The interactions between these variables and how they might influence purchase decisions were not studied in this work. Third, our experiment used French press products as a case-study, building off our previous papers, but does not study purchase decisions for different types of products. We recommend conducting an additional study to investigate the generalizability of our findings, ideally using real products and purchase decisions.

7 Conclusions and Future Work

This paper shows that PAS features can help designers drive purchase decisions for sustainable products. We created a simulated Amazon shopping experience to control what is shown to participants and investigate purchase decisions. We studied how PAS features can influence online purchase decisions compared to dummy features in a within-subject fractional factorial experiment. We built on findings from our previous work where we extracted salient PAS features from online product reviews of French presses [6] and demonstrated that these features resonate with participants as sustainable despite not contributing to engineered sustainability [7].

During each of the control and test conditions, participants selected a product to purchase from four options: in the control condition we included two base products and two products with dummy features, and in the test condition we included two base products and two products with PAS features. We also asked participants to rate products in terms of willingness to pay and sustainability. The results showed that more participants selected to purchase products with PAS features in the test condition than with dummy features in the control condition. Moreover, participants indicated that they are willing to pay more for products with PAS features and rated them as more sustainable too.

The findings indicate that designers should include both engineered sustainable features (from tools like an LCA) and PAS features (from our proposed method) to drive purchasing decisions for sustainable products. Moreover, the findings demonstrate the value of conducting online shopping simulations in design research. Next steps for this work include testing the findings in a real purchasing environment as well as testing how the findings generalize with different 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).

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.

References

[1] McCaskill, A., 2015, "Consumer-Goods' Brands That Demonstrate Commitment to Sustainability Outperform Those That Don't," Nielsen [Online], https://www.nielsen.com/wp-content/uploads/sites/3/2019/04/Global20Sustainability20Report_October202015.pdf

[2] MacDonald, E. F., Gonzalez, R., and Papalambros, P. Y., 2009, "Preference Inconsistency in Multidisciplinary Design Decision Making," ASME J. Mech. Des., **131**(3), p. 031009.

[3] Johnstone, M.-L., and Tan, L. P., 2015, "Exploring the Gap Between Consumers' Green Rhetoric and Purchasing Behaviour," J. Bus. Ethics, **132**(2), pp. 311–328.

[4] She, J., and MacDonald, E. F., 2017, "Exploring the Effects of a Product's Sustainability Triggers on Pro-Environmental Decision-Making," ASME J. Mech. Des., **140**(1), p. 011102.

[5] MacDonald, E. F., Gonzalez, R., and Papalambros, P., 2009, "The Construction of Preferences for Crux and Sentinel Product Attributes," J. Eng. Des., **20**(6), pp. 609–626.

[6] 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.

[7] El-Dehaibi, N., Liao, T., and MacDonald, E. F., 2021, "Validating Perceived Sustainable Design Features Using a Novel Collage Approach," ASME 2021 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference/Design Automation Conference, Virtual, Aug. 17–20, p. 031404.

[8] Green, P. E., and Rao, V. R., 1971, "Conjoint Measurement for Quantifying Judgmental Data," J. Mark. Res., **8**(3), p. 355.

[9] Suryadi, D., and Kim, H. M., 2019, "A Data-Driven Methodology to Construct Customer Choice Sets Using Online Data and Customer Reviews," ASME J. Mech. Des., **141**(11), p. 111103.

[10] Goucher-Lambert, K., Moss, J., and Cagan, J., 2017, "Inside the Mind: Using Neuroimaging to Understand Moral Product Preference Judgments Involving Sustainability," ASME J. Mech. Des., **139**(4), p. 041103.

[11] Goucher-Lambert, K., and Cagan, J., 2015, "The Impact of Sustainability on Consumer Preference Judgments of Product Attributes," ASME J. Mech. Des., **137**(8), p. 081401.

[12] Tovares, N., Boatwright, P., and Cagan, J., 2014, "Experiential Conjoint Analysis: An Experience-Based Method for Eliciting, Capturing, and Modeling Consumer Preference," ASME J. Mech. Des., **136**(10), p. 101404.

[13] 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.

[14] Chevalier, J. A., and Mayzlin, D., 2006, "The Effect of Word of Mouth on Sales: Online Book Reviews," J. Mark. Res., **43**(3), p. 10.

[15] Chen, Y., Wang, Q., and Xie, J., 2011, "Online Social Interactions: A Natural Experiment on Word of Mouth Versus Observational Learning," J. Mark. Res., **48**(2), pp. 238–254.

[16] Liu, Y., 2006, "Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue," J. Market., **70**(3), p. 17.

[17] Dhar, V., and Chang, E. A., 2009, "Does Chatter Matter? The Impact of User-Generated Content on Music Sales," J. Interact. Mark., **23**(4), pp. 300–307.

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

[19] 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.