Big Business Returns on B Corp?
Growing with Green & Lean as any Label is a Good Label

Valdimar Sigurdsson,[a] Nils Magne Larsen,[b] Michał Folwarczny,[c] Freyja Thoroddsen Sigurdardottir,[a] R. G. Vishnu Menon,[d, a] and Asle Fagerstrøm[e]

[a]Department of Business Administration, Reykjavik University, Menntavegur 1, 102 Reykjavik, Iceland

[b]UiT The Arctic University of Norway, Campus Harstad, N-9480 Harstad, Norway

[c]Discipline of Marketing, J. E. Cairnes School of Business and Economics, University of Galway, University Road,
Galway H91 TK33, Ireland

[d] School of Communication, Journalism and Marketing, Massey University, PO Box 756, Wellington 6140, New Zealand

[e] School of Economy, Innovation and Technology, Department of Technology, Kristiania University College, Prinsensgate 7–9, 0107 Oslo, Norway

Author Note

The authors thank The Icelandic Centre for Research (RANNIS) for partially funding this study (grant number 218235-051).

Please send correspondence to Valdimar Sigurdsson, Department of Business Administration, Reykjavik University, Menntavegur 1, 102 Reykjavik, Iceland (valdimars@ru.is)

Big Business Returns on B Corp?

Growing with Green & Lean as any Label is a Good Label

Abstract

This current research contributes to the concept of consumer-based food label equity (CBFLE) by testing the predictive validity of a scale developed by Coderre et al. (2022) in the sustainability and health domains of seafood products. In Study 1 (N = 301; between-within subjects), we found that scores on all subscales, except the (Dis)honesty subscale, were significantly related to willingness to buy fish fillets without a label in comparison with the B Corp sustainability label and a fictitious label. There were no differences between labels. In Study 2 (N = 200; within-subjects), we found similar results for fillets with a health-related label: the American Heart Association Heart-Check. However, scores on the awareness subscale were not significantly associated with willingness to buy fish fillets. Overall, our results suggest that the CBFLE and the scale predict WTB in the context of sustainability and health signaling.

Keywords: Consumer-based food label equity, Health certification, Sustainability certification, CBFLE

1 Introduction

Presently, the retail landscape of health and sustainability food labels is increasingly crowded with hundreds of different certifications (Ecolabel Index, n.d.), including the ones made by retailers. For example, Sam's Club (n.d.) and Whole Foods Market (2021) have created labels to reinforce product sustainability to consumers familiar with the original certifications. Consumers' awareness and understanding of certified food labels are limited due to multifarious factors such as visual and feature complexity (Donato & Adıgüzel, 2022) and the existence of several competing signals and stimuli at the point of purchase (Sigurdsson et al., 2020), all of which leads to low consumer-based label equity (Coderre et al., 2022). From an industry standpoint, the use of third-party labels can be expensive in terms of submission fees and yearly fees/licenses. For example, the annual recertification fee for the sustainability certification "B Corp" (signaling that the firm is meeting high standards of social and environmental performance) ranges from $2,000 to over $50,000 per year based on annual sales from their certified clients (B Lab United States & Canada, n.d.). To justify the cost, certified labels should be more effective than unverified claims to ensure the certification itself is acquired, retained, and further developed. They must provide clear commercial value for firms using them to cover the certification cost. This analysis points to a need for research to identify the factors determining the effectiveness of a food label on consumers' buying behavior and the scrutinization of specific food labels to evaluate the commercial value derived from consumers' perception of the food label (the consumer-based label equity).

A common approach to test the effectiveness of food labels is to examine them concerning brand equity (e.g., Larceneux et al., 2012). However, researchers are increasingly using consumer-based label equity (Carpenter & Larceneux, 2008) instead of traditional

consumer-based brand equity (e.g., Buil et al., 2008; Keller, 1993, 2001; Netemeyer et al., 2004; Pappu et al., 2005; Till et al., 2011; Yoo & Donthu, 2001) for academic advancements and to determine the critical success factors and guidelines for professional practice. So far, extant research on the determining factors of effective food labeling has been fragmented and relatively unexplored (Carpenter & Larceneux, 2008; Coderre et al., 2022). Recently, Coderre et al. (2022) developed and validated a consumer-based food label equity (CBFLE) scale to understand the structure of food label equity.

The academic literature on CBFLE is sparse, with limited insights into different facets of label equity that help food labels achieve expected outcomes. Coderre et al. (2022) addressed this gap by introducing subscales (Visibility, Awareness, Clarity, Design, Relevance, and Credibility) that measure the antecedents of CBFLE. In this study, we aim to test the predictive ability of these subscales under the assumption that adherence to them increases the equity (such as willingness to buy WTB) of a food label. Therefore, one of the main objectives of this paper is to test the subscales put forth by Coderre et al. (2022). To the best of our knowledge, we are the first to test and confirm the predictive ability of the Coderre et al. (2022) FLE scale for WTB. We also test whether the scale discriminates between fish fillets with different food labels. To select the labels for our study, we conducted pilot studies to understand the usage of labels in retail scenarios through consumer surveys and by observing salmon filet products, shelves, and sites. We identified “B Corp,” “MSC” (Marine Stewardship Council), and “ASC” (Aquaculture Stewardship Council) as prevalent sustainability certification labels, and the “American Heart Association Heart Check” (AHA H-C) as the most prevalent health certification label. Based on the pilot studies, we chose B Corp and AHA H-C as target labels. Companies leverage the B Corp certification to demonstrate their commitment to socially and environmentally sustainable

practices as measured by B Lab (B Corp Certification, n.d.-a). As of late 2022, over 6,000 companies in 158 industries from 86 countries have become certified B Corps (B Corp Certification, n.d.-b), earning the right to use the B Corp label—the letter B in a circle, an icon-only logo. On the other hand, the American Heart Association Heart-Check (AHA H-C) has a narrower focus, certifying only food and beverage products that fit within specified nutritional requirements (American Heart Association, n.d.). The AHA H-C label is in the shape of a shield with the words "Meets criteria for heart-healthy food" and "Certified." In addition, we also created a new fictitious "clean" label (CSAP) to assess greenwashing possibilities. Extant research shows that consumers are willing to pay more for a seafood product with an unverified sustainability claim (e.g., a simple tag stating sustainability) compared to a similar product with the widely used Marine Stewardship Council (MSC) certified eco-label (Sigurdsson et al., 2022). This labeling can lead to greenwashing (e.g., see de Freitas Netto et al., 2020). Greenwashing occurs when companies hide facts from consumers or communicate misinformation as facts to make consumers believe that the company's environmental practices are better than actuality. Similar to greenwashing, *leanwashing* (Karnani et al., 2017) is a term used to describe the activities of a firm that misleads consumers through marketing, making them believe a product is beneficial for their health due to specific attributes that are present (or not present) when, in fact, the product does not provide any health benefits. This point emphasizes the importance of an accessible way to evaluate the effectiveness of labels and certification schemes.

Our data shows that the subscales can predict WTB for ten different farmed fish fillets when using the established B Corp label and the fictitious CSAP label as sustainability signals. Using the AHA H-C label as the health signal, we found significant relationships between the subscales and WTB. We also found that consumers do not perceive much difference in WTB

between established and new (fake) sustainability certifications, indicating the danger of greenwashing. We also tested the predictive validity of the CBFLE scale by investigating the relationship between its subscales and the relationship between a single global equity construct, also proposed by Coderre et al. (2022), and willingness to buy fish fillets. We also tested whether the scale's predictive validity depended on the domain. Specifically, we investigated the scale's predictive validity of WTB in the case of fish fillets with health tags, sustainability tags, and unlabeled ones. We found that all subscales (Visibility, Awareness, Clarity, Design, Relevance, Credibility, and (Dis)honesty) correlate with the single global equity (SGE) construct. Furthermore, the SGE construct predicted WTB for different seafood fillets in the two experimental studies even after controlling for pro-environmental consumption (Haws et al., 2014) and social desirability values (Stöber, 2001). Our paper contributes to the retail landscape by providing a pathway for certifiers to leverage the Coderre et al. (2022) FLE scale using different signals to create better certification labels that speak to consumers' health and sustainability values, thereby predicting their WTB for sustainable products. The key findings of this research on health and sustainability labeling are listed below.

- We contribute to the existing literature by further validating the FLE scale by showing that the indirect measure of SGE used by Coderre et al. (2022) to assess the concurrent validity of their FLE scale also predicts consumers' evaluations of products bearing the food labels in a simulated choice-based situation.
- There is little to no difference in WTB based on different labels. We show that the B Corp label, the AHA H-C label, and the label we constructed generally do not differ significantly in their label equity. This result indicates that consumers do not see much difference between labels and/or that labels do not influence them in a choice-based

situation. However, when they deemed the labels low in equity, they showed lower WTB for products with our constructed label than the B Corp label.

- Using food labels as guidance, the subscales of the Coderre et al. (2022) FLE scale predict WTB in the context of sustainability and health signaling.

2 Conceptual framework and hypotheses

2.1 Signaling theory

Research on sustainable buying behavior tends to rely on the theory of planned behavior (Frommeyer et al., 2022). However, the current paper draws on signaling theory as its overarching theoretical perspective. Food labels have emerged due to asymmetric information in exchange situations (Boulding & Kirmani, 1993). Their primary role is to signal product attributes that cannot be assessed properly by consumers at the point of sale or when consumed but can have a positive impact on product choice and the price consumers are willing to pay if signaled. In the theory of information economics, such attributes are categorized as credence attributes (Schrobback et al., 2023; Holland, 2016). Sustainability credence attributes include, among others, the environmental sustainability of production processes, such as eco-friendliness, carbon neutrality, and organic (Schrobback et al., 2023). Health credence attributes include fat, fiber, salt, and sugar content (Ballen et al., 2021). Food labels aim to reduce information asymmetry stemming from the presence of such credence attributes (Larceneux et al., 2012). Food labels draw attention to otherwise unobservable information on environmental or health impact, enabling consumers to distinguish more sustainable and healthy products from less sustainable and healthy alternatives (Johnston & Roheim, 2006).

2.2 Consumer-based food label equity (CBFLE)

The current paper advances consumer-based food label equity. It has been clearly distinguished from brand equity, one of marketing's most important and studied concepts (see Lang et al., 2022, who also recently studied low-involvement products). Carpenter & Larceneux (2008) argue that CBFLE generates positive associations about a product's quality and that labels with sufficient credibility increase overall perceived quality. However, research shows that CBFLE encompasses more than consumers' trust regarding the food label (Coderre et al., 2022). Therefore, it is essential to define CBFLE broadly such that the commercial value is derived from consumers' perception of the food label rather than the product or brand name itself. A common approach to measure CBFLE is to examine the impact of specific food labels on, for instance, perceived quality, choice, WTB, and willingness to pay (WTP), which emphasizes the food labels' "global equity" (Coderre et al., 2022). Since WTB and WTP normatively should yield the same valuation order (Lu & Hsee, 2019), we focus in the current research on WTB as a measure of commercial value to elicit consumers' underlying valuations of a food label.

Certain facets of CBFLE have also been examined in the literature, most commonly consumers' awareness, understanding, and trust (Chen et al., 2015; Feucht & Zander, 2014; Grunert et al., 2014). Furthermore, Sigurdsson et al. (2022) examined how familiarity with a food label mediates the effect of the label on WTP, and Donato and Adigüzel (2022) studied how label design affects consumers' evaluations of sustainable products. Zepeda et al. (2013) point to other dimensions of CBFLE by showing that labels are compelling not just because of their design and source but also because of how consumers perceive the label message and how well the label fits with their own values. Sigurdsson et al. (2023) take this point further by examining multiple facets of CBFLE (familiarity, understanding, and trust) using single-item constructs and

demonstrate that consumers have a higher WTP for products with labels having higher CBFLE (in terms of familiarity, understanding, and trust) in a choice-based situation.

According to Coderre et al. (2022), the traditional brand equity models are unsuitable for food labels because food labels focus on specific quality dimensions and are often collectively owned. Furthermore, the differential effects are unclear compared to traditional branding, as food labels are co-branded with other food labels and brands. Food labels are also based on credence attributes and production standards defined in specifications and audited by monitoring agencies. Moreover, food label awareness differs from brand awareness since the former refers to understanding the meaning of the label. In contrast, the latter, as shown by Keller (1993), refers to the ability to recognize or recall a brand (Coderre et al., 2022). Additionally, existing studies on food labels have not provided a comprehensive perspective on the facets of CBFLE. Recognizing this gap, Coderre et al. (2022) developed a CBFLE scale with three second-order dimensions (Label Familiarity, Informational Value, and Trust) measured by seven first-order sub-dimensions: Visibility, Awareness, Clarity, Design, Relevance, Credibility, and Honesty (or dishonesty as it is a reversed scale).

Among these sub-dimensions, Visibility, Awareness, and Clarity measure *Label Familiarity*, with Visibility referring to the consumer's familiarity with the label and the visibility of the label in the marketplace, Awareness referring to how well aware the consumer is of what the label stands for and understands the meaning of the label, and Clarity referring to how easy it is to spot, remember, and understand the logo. Design and Relevance measure *Informational Value*, with Design referring to how the logo is perceived (pretty, attractive, nice design) and Relevance to the relevancy of the criteria the label signals for the consumer in a buying situation. Finally, Credibility and (Dis)honesty measure *Trust*, with Credibility referring

to how the consumer perceives the organization's trustworthiness behind the certification, and (Dis)honesty to consumers' beliefs that the label is a scam or misleading.

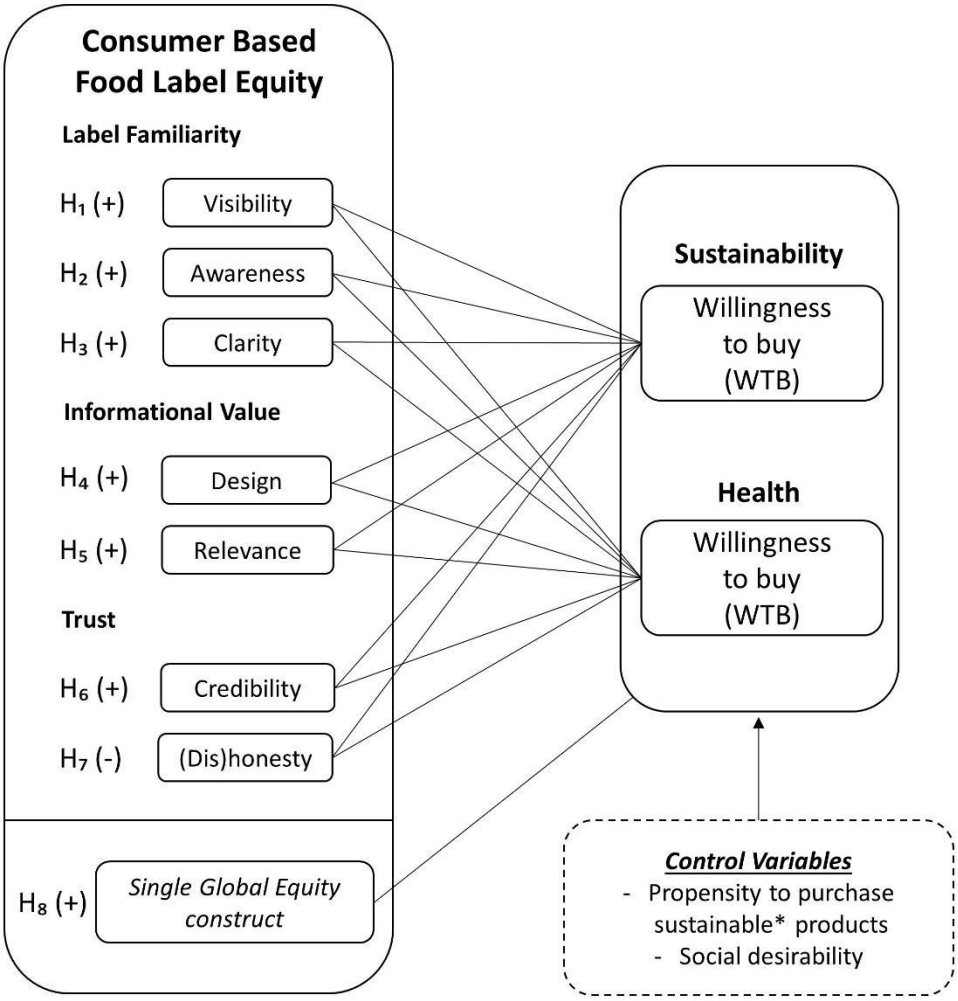
Coderre et al. (2022) further propose a single global equity construct (SGE) consisting of five items reflecting consumers' opinions and preferences for products displaying the label, the extent to which the label prompts consumers to buy and recommend the product to their friends, and consumers' WTP ten percent more for products displaying the label. This construct is an indirect global measure of the consequences of CBFLE and is used by Coderre et al. (2022) to assess the concurrent validity of their FLE scale. Our research takes on their future research advice to assess the FLE scale using more direct measurements of real consequences, such as when consumers indicate their WTB for products with the label in a choice-based situation.

2.3 Research foundations and hypotheses

In contrast to Coderre et al. (2022), who developed a psychometric scale to examine the structure of latent variables, we focused on developing a framework for researchers, food manufacturers, and retailers to evaluate food labels regarding their influence on WTB. Consequently, we examine the predictability of the scores on the seven subscales of the CBFLE on WTB. Given retailers' limited resources for research, we also explored whether the simpler five-item SGE construct (Coderre et al., 2022) was appropriate for evaluating the impact of a certification label on WTB for a product bearing that label. The following figure (Figure 1) presents our research conceptually.

Figure 1

The conceptual framework for evaluating food labels' impact on WTB.



Note. *Sigurdsson et al. (2023) found green consumption values (Haws et al., 2014) predictive of WTP ($) for fish fillets with sustainability tags and health certifications.

One of the main research propositions we put forth is that consumers will be more willing to buy products with sustainability or health labels with higher (than with lower) scores on the subscales of the Coderre et al. (2022) FLE scale. The theoretical argument is that a food label should inform consumers about sustainability and health credence attributes related to a product that can potentially increase consumer value (Carpenter and Larceneux, 2008). This value increases with

label familiarity, informational value, and consumer trust (Coderre et al., 2022). Below, we introduce the subscales, their relations, and their direct linear effects on key marketing variables such as willingness to buy grocery products.

2.3.1 Label Familiarity

Coderre et al. (2022) have shown that Visibility, Awareness, and Clarity are subsets of Label Familiarity.

Visibility refers to the consumer's familiarity with the label and the visibility of the label in the marketplace (Coderre et al., 2022). Although studies indicate that familiarity may impact skepticism/trust toward the label (Sirieix et al., 2013; Teisl et al., 2002), Coderre et al. (2022) have established that familiarity is not an antecedent of trust. The branding literature shows that familiar brands are better liked than unfamiliar brands (e.g., Colombo and Morrison, 1989). This point also applies to food labels, as repeated exposure can enhance consumers' attitudes toward a food label (Zajonc, 1968). Research on food labels specifically shows that the market penetration of a label (Aprile & Punzo, 2022) and label familiarity positively impact consumer preferences (Sigurdsson et al., 2022, 2023). Therefore, we hypothesize

H1: Visibility is positively related to WTB.

Awareness refers to consumer awareness and understanding of a label and its inherent meaning (Coderre et al., 2022). When consumers lack knowledge and understanding of the issues conveyed by a food label, the label is of less help to them (Feucht & Zander, 2014). Similarly, preferences for a labeled product increase when consumers have proper knowledge about the meaning and content of the label (Aprile & Punzo, 2022; Carpenter and Larceneux, 2008; Grunert, 2014). Therefore, we propose

H2: Awareness is positively related to WTB.

Clarity refers to how easy it is to spot, remember, and understand the logo (Coderre et al., 2022). Food labels come in the form of logos with textual and/or visual design elements with different descriptiveness levels. The more self-explanatory, the greater the ability of the label to communicate its meaning (Grunert et al., 2014). Thus, a label should be able to capture a clear connotation of its benefits and be easy to process (Donato & Adıgüzel, 2022; Luffarelli et al., 2019; Nikolova & Inman, 2015). Since the clarity of the logo increases consumers' ability to spot, remember, and understand the food label, we propose:

H3: Clarity is positively related to WTB.

2.3.2 Informational Value

Coderre et al. (2022) have shown that Design and Relevance are subsets of Informational value.

Label design refers to the appealing nature of the label in terms of being pretty, nice, and attractive (Coderre et al., 2022). Zepeda et al. (2013) demonstrate that whether a consumer finds a label compelling depends on label design, such as the label's look, how pretty the label is perceived, and the use of bright colors. Previous research has found that logo characteristics impact recognition and consumers' affective responses (Henderson & Cote, 1998; Machado et al., 2015). Thus, we hypothesize

H4: Design is positively related to WTB.

Relevance refers to the pertinence of food label criteria in consumers' consideration of food labels (Coderre et al., 2022). For instance, the effects of labeling vary depending on consumers' concern for the attribute signaled by the food label (Majer et al., 2022; Grunert et al., 2014). Since relevance affects consumers' motivation to engage in processing information, they will pay more attention to a label if their values and concerns are congruent with the attributes the label communicates (Ghvanidze et al., 2017; Grunert et al., 2014; Taufique et al., 2017; Zepeda et al., 2013). The more relevant the label attribute, the higher the consumer value and willingness to buy. Therefore,

H5: Relevance is positively related to WTB.

2.3.3 Trust

Consumers form trust-related judgments through interacting with food labels (Tonkin et al., 2016). Coderre et al. (2022) show that Credibility and (Dis)honesty are a subset of Trust.

Credibility refers to the consumer perception of the organization's trustworthiness behind the label (Coderre et al., 2022). As consumers cannot immediately verify whether a product meets the label's criteria (Anisimova & Sultan, 2014), trust in the source is needed for a label to be effective. The more credible the source, the more influential the label is in consumer purchasing situations (Janssen & Hamm, 2012; Sigurdsson et al., 2023; Taufique et al., 2017; Vecchio et al., 2016; Zepeda et al., 2013). Thus, we propose

H6: Credibility is positively related to WTB.

(Dis)honesty refers to a consumer's perception of a label as fraudulent, misleading, and developed purely for commercial reasons. (Coderre et al., 2022). Consumers feel deceived and distrustful when they unearth inappropriate or manipulative marketing communication (Campbell, 1995). Research shows that perceived deception harms product attitude and purchase intention (An et al., 2019; Campbell, 1995; Darke & Ritchie, 2007; Newell et al., 1998). Therefore, we propose

H7: (Dis)honesty is negatively related to WTB.

2.3.4. Single Global Equity (SGE)

Coderre et al. (2022) have shown that the overall FLE measure (a summated scale comprising their seven subscales) directly impacts the SGE construct. In this study, we test our assumption that Visibility, Awareness, Clarity, Design, Relevance, and Credibility positively correlate with the SGE construct. At the same time, a negative correlation exists between the (Dis)honesty and the SGE. Since the SGE construct represents a short and global approach to FLE (Coderre et al., 2022) reflecting consumers' relative preference, purchase intention, willingness to recommend, and WTP, we test the ability of this *indirect* measurement to predict purchase behavior in a simulated choice-based situation. Thus, we hypothesize

H8: The single global equity construct is positively related to WTB.

3 Empirical studies

We wanted to see whether the seven Coderre et al. (2022) FLE scale subscales predicted WTB farmed fish fillets. We also used the GREEN scale (Haws et al., 2014) and the social desirability scale (Stöber, 2001) to find whether these consumption values affected WTB. Study 1 focused on sustainability, where we compared two labels (B Corp and CSAP, B Corp being an ever-

growing sustainability certification and CSAP being a label made up specifically for this study) and a control group exposed to no label. In Study 2, we examined an established health certification (the AHA H-C) using the same scales as in Study 1 and compared it with the results from a control group where no label was displayed. All analyses were performed in R version 4.1.0 (R Core Team, 2021). The experiments were coded in PsyToolkit version 3.4.0 (Stoet, 2010, 2017), which also stores participants' completion time. This allowed us to ensure that the available data did not include careless responses, i.e., responses completed in an unrealistically short time.

Recent research has shown that green consumption values (Haws et al., 2014) moderate WTP for fish fillets with various labels such as MSC or fictitious sustainability tags (Sigurdsson et al., 2022, 2023). The influence of environmental sustainability labels depends on personal relevance and motivation to act on such information, as consumers engage in green consumption only if they value the environment and/or want to contribute to environmental protection (Hoek et al., 2017). Consumers' identity-related aspects, such as environmental concern and green consumption values, have played an important role in understanding environmentally sustainable consumption (e.g., Haws et al., 2014) and consumers' responses to green marketing communication (Bailey et al., 2018). Sigurdsson et al. (2022) found that green consumption values were positively related to WTP for seafood with an eco-label. They also found green consumption to predict label familiarity. Thus, we conducted the analysis controlling for participants' green consumption values to rule out the possible influence of this variable on the outcome measure. Social desirability bias, that is, participants' tendency to respond in a way that increases their chances of being seen in a positive light (Stöber, 2001), may be another potentially influential confounder in studies measuring propensity to consume goods in an

environmentally friendly manner (for a related discussion, see Otterbring & Folwarczny, 2022). Therefore, we conducted another analysis in which we added the social desirability measure (Stöber, 2001) as a covariate.

3.1 Study 1

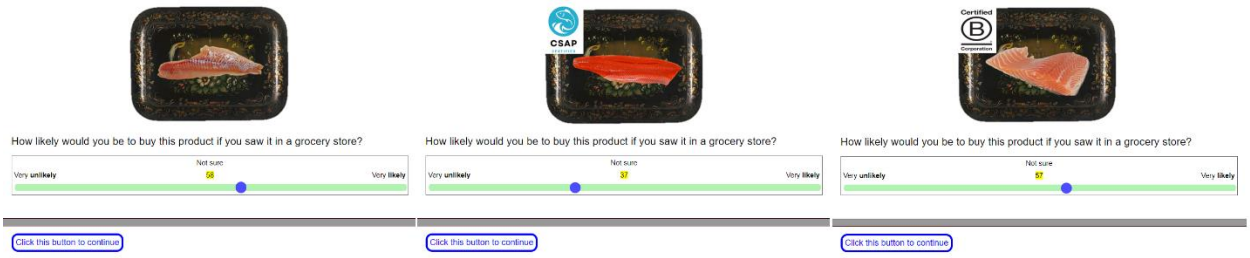
3.1.1 Participants and Procedure

We recruited 301 US participants on Prolific Academic (Mean age = 41.9, *SD* = 13.6, 44.2% women) who were primary grocery shoppers in their households with no diet restrictions. We chose Prolific Academic because of recent concerns about data quality on some crowdsourcing platforms, which can lead to a large proportion of responses being rejected (see, e.g., Webb & Tangney, 2022). Prolific Academic provides high data quality on key criteria such as participant attention, comprehension, honesty, and reliability of responses, with results often significantly outperforming other crowdsourcing platforms (Peer et al., 2022). We used the same prescreening criteria across all studies described below, and no data has been excluded. Participants provided informed consent and were assigned to one of the three experimental conditions (the study used a between-within-subjects design) as presented in Figure 2, in which they indicated their WTB ("How likely would you be to buy this product if you saw it in a grocery store?") the ten fish fillets on a 100-point sliding scale (0 = *Very unlikely*; 100 = *Very likely*). In the control condition (*n* = 103), the participants saw the fish fillets without labeling. In the B Corp condition (*n* = 94), they saw fillets with the B Corp label in the upper left corner of each product image, whereas, in the CSAP condition (*n* = 104), they saw the CSAP label in the same location as in the case described above. Our CSAP label shows a teal-colored fish and the outlines of waves. The color was chosen to connect with previously established labels such as MSC (dark blue) and ASC (aqua green). After this task, they indicated their agreement on a seven-point scale (1 = *strongly*

disagree; 7 = *strongly agree*) to the six items (e.g., "I am concerned about wasting the resources of our planet.") from the GREEN scale, which captures consumers' propensity to purchase sustainable products (Haws et al., 2014). Next, the participants who saw B Corp or CSAP labeled fillets evaluated them on the Coderre et al. (2022) FLE and SGE scales (Coderre et al., 2022). Specifically, on a five-point scale, they indicated their agreement (1 = *Disagree Strongly*; 5 = *Agree Strongly*) with the 26 items (e.g., "This logo is attractive.") that capture these two constructs (21 items captured the seven FLE subscales, and five items captured SGE). Three items belonging to Trust have been negatively worded: *"This label is misleading," "This label is a scam," and "This label was developed for purely commercial reasons."* Finally, participants provided their demographic data (age and sex) and were asked if they guessed the hypothesis ("What do you think we were trying to find out in this study?"). At the conclusion of the study, participants indicated whether each of the 16 items (e.g., "I would never live off other people") from the social desirability scale, measuring their tendency to respond as they thought the researcher would want them to (Stöber, 2001), applied to them (1 = *true*) or not (0 = *false*).

Figure 2

The sample tasks across the three experimental conditions in Study 1



3.1.2 Results and Discussion

Descriptive results and analytic approach: The Coderre et al. (2022) FLE scale. We first created indexes by averaging responses to the seven subscales and the SGE. Internal consistency

reliability for the Visibility (α = .83, *M* = 2.20, *SD* = 0.99), Clarity (α = .69, *M* = 3.54, *SD* = 0.84), Awareness (α = .94, *M* = 2.13, *SD* = 1.22), Design (α = .91, *M* = 3.42, *SD* = 1.05), Dishonesty (α = .73, *M* = 2.64, *SD* = 0.80), Credibility (α = .85, *M* = 2.96, *SD* = 0.88), Relevance (α = .87, *M* = 3.09, *SD* = 0.89), and the SGE construct (α = .90, *M* = 2.81, *SD* = 0.91) subscales ranged from acceptable to excellent.

We also averaged responses to the GREEN scale (Haws et al., 2014), showing excellent internal consistency and reliability (α = .95, *M* = 4.97, *SD* = 1.41). The social desirability (Stöber, 2001) index created similarly (α = .85, *M* = 0.54, *SD* = 0.26) had good reliability. For reference, we followed the same procedure to create the WTB index, that is, by averaging responses to the ten corresponding questions; the index showed good reliability (α = .85, *M* = 49.53, *SD* = 16.74; B-Corp: *M* = 50.94, *SD* = 18.09; CSAP: *M* = 47.69, *SD* = 17.21; Control: *M* = 50.12, *SD* = 14.86).

Table 1 shows the correlations between the variables included in the main study. To account for the nested data structure, with ten measures of WTB for fish fillets per participant, we fitted linear mixed models to our data using the *lme4* package for R (Bates et al., 2015) and obtained *p* values using the *lmerTest* package (Kuznetsova et al., 2017). We added random intercepts for participants and fillets in all models described below. As indicated by the bar graphs in Table 1, the distribution of our dependent variable was nearly normal; therefore, we performed parametric tests when analyzing the data.

Table 1

Correlations between the variables included in Study 1

	WTB	Visibility	Clarity	Awareness	Design	Dishonesty	Credibility	Relevance	SGE	SDS
Visibility **α = .83**	.19**									
Clarity **α = .69**	.18*	.51***								
Awareness **α = .94**	.21**	.82***	.53***							
Design **α = .91**	.10	.27***	.43***	.21**						
Dishonesty **α = .73**	-.01	-.31***	-.33***	-.42***	-.28***					
Credibility **α = .85**	.25***	.73***	.55***	.77***	.34***	-.40***				
Relevance **α = .87**	.26***	.67***	.51***	.73***	.35***	-.39***	.82***			
SGE **α = .90**	.28***	.72***	.53***	.77***	.34***	-.43***	.81***	.88***		
GREEN **α = .95**	.15**	.25***	.17*	.34***	.05	-.16*	.34***	.44***	.44***	
SDS **α = .85**	.13*	.11	.08	.11	.03	-.03	.06	.09	.10	.18**

Note. This table presents Pearson correlation coefficients. * $p < .05$, ** $p < .01$, *** $p < .001$.

Subscales' and SGE relations to WTB: We began our analysis by testing whether the subscales of the consumer-based food label equity scale and the SGE (Coderre et al., 2022) predicted WTB scores for the ten fish fillets, controlling for the effects of experimental condition assignment (i.e., control, B Corp, and CSAP). Thus, we added the visibility index as the first predictor and the effect of experimental conditions as a covariate to the model. The Visibility

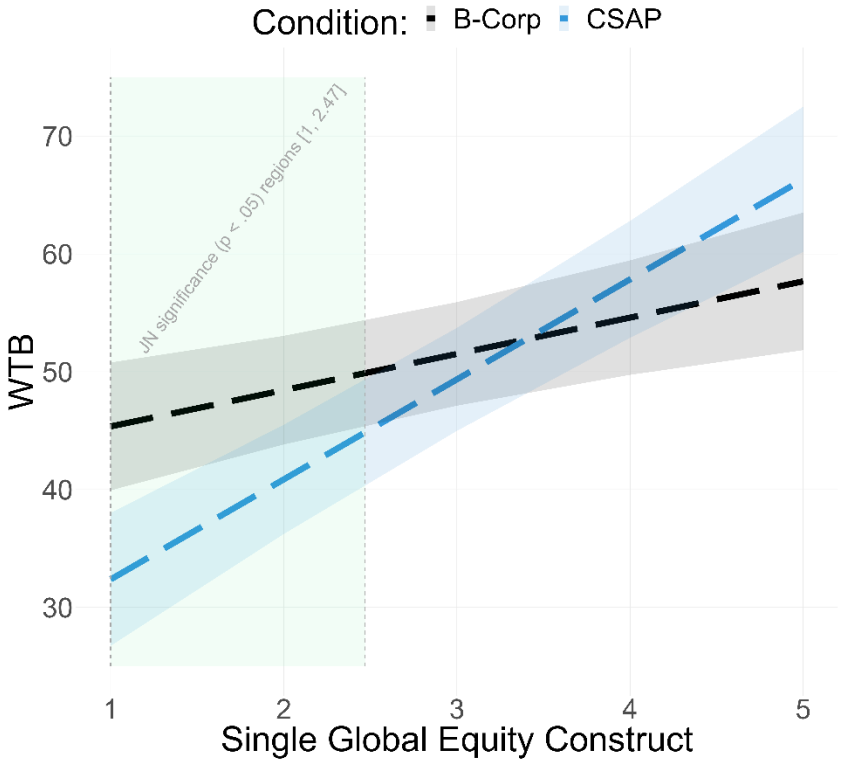
subscale was positively related to WTB, b = 3.21, 95% CI [0.77, 5.66], t = 2.57, p = .011. A similar analysis revealed the Clarity subscale's significant and positive effect on WTB, b = 3.62, 95% CI [0.72, 6.51], t = 2.45, p = .015. Using awareness as a predictor in the model revealed a significant and positive relationship between the Awareness subscale and WTB, b = 2.99, 95% CI [0.99, 4.99], t = 2.93, p = .004. The Design subscale was also positively related to WTB, b = 3.05, 95% CI [0.40, 5.69], t = 2.26, p = .025. Further, we found a significant and positive relationship between the Credibility subscale and WTB, b = 4.84, 95% CI [2.12, 7.57], t = 3.49, p < .001. The model results indicated a positive association between the Relevance subscale and WTB for fillets, b = 5.12, 95% CI [2.44, 7.80], t = 3.74, p < .001. Finally, the SGE construct was positively related to WTB, b = 5.46, 95% CI [2.85, 8.06], t = 4.10, p < .001. However, the (Dis)honesty subscale was unrelated to the dependent measure, b = -0.37, 95% CI [-3.46, 2.71], t = -0.24, p = .813.

Experimental conditions; different labels: We tested whether consumers showed a higher WTB for fillets with either CSAP or B Corp labels than for fillets without any labeling—a reference condition. Specifically, we included condition as a predictor of WTB in the model without any covariates. We found no main effect for B Corp, b = 0.82, 95% CI [-3.86, 5.50], t = -0.34, p = .732, or CSAP, b = -2.43, 95% CI [-6.99, 2.13], t = -1.05, p = .297. Consumers also showed no greater WTB for B Corp fillets than CSAP, b = -3.25, 95% CI [-7.92, 1.42], t = -1.36, p = .173. We conducted additional analysis on a standardized version of the data by testing for the potential interaction between the SGE construct that adequately captures food label equity (Coderre et al., 2022) and experimental conditions to test whether the effects of experimental conditions were significant at different levels of the moderator—the SGE construct. This analysis revealed no main effects for experimental conditions, β = -.11, 95% CI [-.28, .05], t = -1.34, p =

.183, or the SGE construct on WTB, β = -.25, 95% CI [-.66, .17], t = -1.17, p = .244. However, we found a significant interaction between experimental conditions and the SGE construct, β = .17, 95% CI [.01, .34], t = 2.04, p = .043. With an alpha level of 0.05, we further explored this interaction to estimate the exact levels of the SGE construct at which the effect of experimental conditions became significant. For this, we calculated Johnson-Neyman (JN) intervals using the raw (nonstandardized) version of the data (for readability). As can be seen in Figure 3, the effect of experimental conditions was significant when consumers deemed the labels to be relatively low in equity, that is, within the range of 1.00 and 2.47 on a scale of 1 to 5, or 0.37 times *SD* below the mean (M = 2.81, SD = 0.91). Consumers showed lower WTB for fillets labeled with CSAP than with B Corp within this interval.

Figure 3

Johnson-Neyman significance regions



Note. The green shaded area represents the Johnson-Neyman significance regions. The confidence bands around the regression slopes show the standard errors.

Control scales' relations to WTB. We tested whether the GREEN scale capturing pro-environmental consumption values (Haws et al., 2014) and the social desirability scale (Stöber, 2001) predicted WTB for fish fillets, also accounting for the effects of experimental conditions (we fit the two separate models with either the social desirability or pro-environmental consumption values as predictors). We found a significant positive effect of the green consumption index on WTB, $b = 1.75$, 95% CI [0.41, 3.08], $t = 2.56$, $p = .011$. Interestingly, within the second model, we also found a positive effect of the social desirability index on WTB, $b = 8.72$, 95% CI [1.49, 15.95], $t = 2.36$, $p = .019$. We completed our analysis by testing whether the effects of pro-environmental consumption values and social desirability on WTB were robust

to the inclusion of the SGE construct index in the model. The association between pro-environmental consumption values and WTB in this model became nonsignificant ($p = .703$), and the association between social desirability values and WTB became marginally significant ($p = .076$), while the SGE construct remained positively related to WTB ($p < .001$).

Overall, our results suggest that all subscales of the Coderre et al. (2022) FLE scale, except the (Dis)honesty subscale, were positively associated with WTB for fish fillets, even when controlling for the possible effect of experimental condition assignment. We found that the effect of experimental conditions was significant at relatively low levels of the SGE construct, that is, less than 0.37 times the standard deviation below the mean (see Figure 3). Consumers showed lower WTB for this range of fillets labeled with CSAP than with B Corp. In addition, green or pro-environmental consumption (Haws et al., 2014) and social desirability values (Stöber, 2001) were positively related to WTB measures; however, the former association disappeared when the SGE index was included in the model. However, the SGE index was positively associated with WTB for fish fillets, even when the above variables were included in the same model.

3.2 Study 2

Except for the (Dis)honest subscale, Study 1 found positive associations between the subscales of the FLE scale and WTB fish fillets and between the SGE construct and WTB. The labels used in Study 1 were about environmental sustainability. However, it remains unclear whether these effects exhibit cross-domain generalizability. Therefore, to test the external validity of our results from Study 1, we conducted Study 2 to investigate the efficacy of the Coderre et al. (2022) scale to predict WTB fish fillets with health label, namely the American Heart Association Heart-Check (AHA H-C) label.

3.2.1 Participants and Procedure

Using the same prescreening criteria as in Study 1, we recruited 200 US participants on Prolific Academic (Mean age = 37.0, *SD* = 13.9, 58.0% women). No data were excluded from the analysis. Study 2 used a within-subjects design. The procedure used was similar to Study 1 except for the logo; for Study 2, all the participants were exposed to fillets with the AHA H-C logo, as shown in Figure 4.

[INSERT FIGURE 4 HERE]

3.2.2 Results and Discussion

Descriptive results and analytic approach. We mirrored the analytical approach used in Study 1 and created index measures for all the subscales of the Coderre et al. (2022) FLE scale and the SGE scale (Coderre et al., 2022). We found that the internal consistency reliability of Visibility (α = .84, *M* = 3.40, *SD* = 0.98), Clarity (α = .74, *M* = 4.09, *SD* = 0.70), Awareness (α = .78, *M* = 3.42, *SD* = 0.90), Design (α = .87, *M* = 3.40, *SD* = 0.85), (Dis)honesty (α = .81, *M* = 2.48, *SD* = 0.85), Credibility (α = .67, *M* = 3.60, *SD* = 0.75), Relevance (α = .80, *M* = 3.29, *SD* = 0.88), and the SGE construct (α = .84, *M* = 2.86, *SD* = 0.79) ranged from acceptable to excellent. We also averaged responses to the GREEN scale (Haws et al., 2014), with its index showing excellent reliability (α = .94, *M* = 4.74, *SD* = 1.39). The social desirability (Stöber, 2001) index created in the same manner as the scale above had acceptable reliability (α = .75, *M* = 0.53, *SD* = 0.21). Further, for illustrative purposes, we created the WTB index by averaging responses to the ten questions capturing this variable and found its index to have excellent reliability (α = .90, *M* = 43.49, *SD* = 19.49). Table 2 shows the correlations between the variables included in Study 2.

Table 2

Correlations between the variables included in Study 2

	WTB	Visibility	Clarity	Awareness	Design	Dishonesty	Credibility	Relevance	SGE	SDS
Visibility **α = .84**	.12									
Clarity **α = .74**	.19**	.35***								
Awareness **α = .78**	.10	.45***	.50***							
Design **α = .87**	.20**	.23**	.47***	.33***						
Dishonesty **α = .81**	-.13	-.12	-.35***	-.50***	-.31***					
Credibility **α = .67**	.19**	.22**	.55***	.50***	.42***	-.53***				
Relevance **α = .80**	.27***	.27***	.51***	.52***	.51***	-.45***	.67***			
SGE **α = .84**	.29***	.26***	.42***	.50***	.51***	-.45***	.61***	.84***		
GREEN **α = .94**	.18*	-.02	.04	.05	.09	-.07	.21**	.23***	.25***	
SDS **α = .75**	.10	.03	-.01	.05	.14*	.05	.02	.11	.16*	.02

Note. This table presents Pearson correlation coefficients. * $p < .05$, ** $p < .01$, *** $p < .001$.

We fit models with the same random effects structure as in Study 1. We used the averaged scores of the corresponding subscales of the Coderre et al. (2022) FLE scale as predictors. We used the WTB scores for the ten fish fillets as a dependent variable.

Subscales' and SGE relations to WTB. The Visibility subscale was marginally and positively related to WTB (b = 2.49, 95% CI [-0.27, 5.24], t = 1.77, p = .078). Likewise, we found a significant and positive effect of the Clarity subscale on WTB (b = 5.35, 95% CI [1.54,

9.17], t = 2.76, p = .006). Further, the Design (b = 4.52, 95% CI [1.38, 7.67], t = 2.82, p = .005), Credibility (b = 5.07, 95% CI [1.51, 8.63], t = 2.80, p = .006), and Relevance (b = 6.03, 95% CI [3.04, 9.02], t = 3.96, p < .001) subscales and the SGE construct (b = 7.07, 95% CI [3.79, 10.36], t = 4.23, p < .001) were positively related to WTB. On the other hand, the (Dis)honesty subscale (b = -2.87, 95% CI [-6.04, 0.29], t = -1.78, p = .076) was marginally and negatively related to WTB. Unlike Study 1, we found no association between the Awareness subscale (b = 2.09, 95% CI [-0.91, 5.09], t = 1.37, p = .174) and WTB for the ten fish fillets. Following the analytical approach from Study 1, we tested whether the GREEN scale measuring pro-environmental consumption values (Haws et al., 2014) and the social desirability scale (Stöber, 2001) predicted WTB for fish fillets. Once again, we found a significant positive effect of the pro-environmental consumption values (b = 2.51, 95% CI [0.58, 4.44], t = 2.55, p = .011) on WTB. To test the robustness of the association between pro-environmental consumption values and WTB, we added the SGE construct as a covariate to the model and, similar to Study 1, found that the positive association between pro-environmental consumption values and WTB was no longer significant (p = .107) when the SGE index was added to the model. Like Study 1, the SGE index was positively associated with WTB (p < .001). Unlike in Study 1, the results from Study 2 revealed no association between the social desirability index (b = 9.59, 95% CI [-3.12, 22.29], t = 1.48, p = .140) and WTB.

The results of Study 2 largely mirrored the patterns observed in the previous study. All subscales of the Coderre et al. (2022) FLE scale were positively associated with WTB fish fillets, except the Awareness subscale, which showed no association with WTB. The (Dis)honesty subscale, on the other hand, was marginally and negatively associated with WTB. Importantly, we found a positive association between the SGE index covering a broad range of underlying

constructs (Coderre et al., 2022) and WTB fish fillets with AHA H-C labeling, replicating the results from Study 1. In addition, we found that pro-environmental consumption values (Haws et al., 2014) were positively related to WTB. However, similar to Study 1, this association was nonsignificant when the SGE construct was added to the model. The SGE construct remained positively related to WTB. In contrast with Study 1, Study 2 found no association between social desirability and WTB. Table 3 summarizes the findings from Studies 1 and 2.

Table 3

Summary of Results

Hypotheses		Sustainability labels	B Corp		CSAP		Health labels	AHA H-C	
			M	SD	M	SD		M	SD
H_1	Visibility +	Supported	2.27	1.06	2.14	0.94	Supported	3.40	0.98
H_2	Awareness +	Supported	2.29	1.34	1.99	1.08	Not supported	3.42	0.90
H_3	Clarity +	Supported	3.60	0.88	3.49	0.80	Supported	4.09	0.70
H_4	Design +	Supported	2.89	1.05	3.89	0.80	Supported	3.40	0.85
H_5	Relevance +	Supported	3.08	0.96	3.11	0.82	Supported	3.29	0.88
H_6	Credibility +	Supported	3.01	0.96	2.92	0.80	Supported	3.60	0.75
H_7	(Dis)honesty -	Not supported	2.67	0.83	2.62	0.78	Supported	2.48	0.85
H_8	Single Global Equity Construct +	Supported	2.81	0.99	2.80	0.83	Supported	2.86	0.79

4 General discussion and conclusions

The current research aimed to test the subscales of the FLE scale developed by Coderre et al. (2022) and determine the critical success factors for food label equity in the context of sustainability (Study 1) and health (Study 2). We propose a conceptual framework operationalizing CBFLE based on the subscales developed by Coderre et al. (2022) and assessed their connections with consumers' WTB. As part of the framework, we also assessed the association between an outcome-based measure (Coderre et al., 2022) of CBFLE and WTB. Table 3 summarizes the findings. Studies 1 and 2 found acceptable to high internal consistency

reliability across the subscales of the Coderre et al. (2022) FLE scale, and the subscales were correlated with one another, including the SGE construct. Study 1 confirmed hypotheses 1 to 6 and hypothesis 8 but not hypothesis 7. These results align with the research proposition that consumers will be more willing to buy products with sustainability labels with higher subscale scores on the Coderre et al. (2022) FLE scale, even when controlling for the possible effect of the experimental conditions. Consumers generally did not show a higher WTB for products labeled with either CSAP or B Corp than for unlabeled products. They only showed lower WTB for fillets with CSAP compared to the B Corp label when it was deemed relatively low in equity (Study 1). Study 2 tested this thesis further in the health category by exposing consumers to fillets with the AHA H-C label. The findings confirmed hypotheses 1 and 3 to 7, but not hypotheses 2. Again, this result showed that consumers were, overall, willing to base their WTB on how they perceive the label as measured through the subscales of the CBFLE scale.

In summary, the results of all subscales, except for the dis(honesty) subscale of the CBFLE scale (Coderre et al., 2022), were positively associated with WTB for sustainability labels (B Corp and CSAP). These results were similar in the case of the relationship of the subscales to WTB for health labels (AHA H-C). However, the awareness subscale was not associated with WTB for products with health labels. The single global equity construct was positively associated with WTB for sustainability and health label products. These results suggest that both the CBFLE subscales and the single global equity construct can be used to measure the influence of labels on WTB for products. Because the latter tool contains relatively few items, the single global equity construct may be particularly useful when lengthy data collection is undesirable.

4.1 Discussion of Results

According to Coderre et al. (2022), the FLE scale is for generic food labels incorporating attributes such as product origin, expertise, religious conformity precepts, and adoption of safety standards. In line with this notion, we found that the FLE scale did not distinguish seafood products with sustainability labels (Study 1) and those without labels. Specifically, in all the experimental conditions in Study 1, FLE subscales were associated with WTB fish fillets. The current research empirically verifies the scale's predictive validity in two related categories: sustainability (Study 1) and health (Study 2). Our results show that for Study 1, except the (Dis)honesty subscale, all the remaining subscales (Visibility, Awareness, Clarity, Design, Relevance, and Credibility) were positively associated with consumers' WTB.

Similarly, for Study 2, except for the Awareness scale, all other scales were positively associated with consumers' WTB. Our research provides a framework (see Figure 1) to help academics, food producers, and retailers navigate the crowded labeling schemes marketplace. Academics and practitioners evaluating the effectiveness of food labels need a systematic tool to measure improvements to marketing performance (such as WTB) attributable to the certifications. They need to know that the benefits outweigh the costs and that the benefits are more than doing nothing. Our findings show that the Coderre et al. (2022) FLE scale has the potential to be used both when creating a new scale (such as our CSAP) and as an evaluation tool for different established scales as we show by scrutinizing two established labels (B Corp and AHA H-C).

4.2 Theoretical Implications

4.2.1 The Consumer-Based Food Label Equity Scale

Our findings confirm that the Coderre et al. (2022) FLE scale can be used in sustainability and health. The scale's internal consistency is at least acceptable. Cronbach's alpha values in the two

studies were at least .70, or slightly below the higher criterion by DeVellis (2012). We found all the subscales (Visibility, Awareness, Clarity, Design, Relevance, Credibility, and (Dis)honesty) to be associated with the global equity construct (consumer opinion, preference, buying, paying, and recommendation). Furthermore, the SGE construct (Coderre et al., 2022) predicted WTB for different seafood fillets in both studies, even after controlling for pro-environmental consumption (Haws et al., 2014) and social desirability values (Stöber, 2001).

4.2.2. Subscales' & global equity relations to WTB

Coderre et al. (2022) argue that the scale is developed to utilize a more focused micro-investigation of food label equity. Our results confirm that the factors (subscales) tend to be correlated with WTB, strengthening its predictive validity. The connection, though, tends to be different between categories and labels, indicating that different subscales/factors need to be emphasized for different categories, labels, and strategies, such as for a broad strategy (such as B Corp), specific category (such as only for fish), the development of a new scale, or for different categories (such as sustainability and health). While Coderre et al. (2022) point to label familiarity—comprising Visibility, Awareness, and Clarity—as the most important component of FLE, we found relevance, particularly, being the strongest predictor of WTB in a simulated choice-based situation. This result suggests that the relevancy of the label signals' criteria can be more important for the consumer in a buying situation than Coderre et al. (2022) suggested.

In our two studies, the measure of the SGE construct correlated with each of the seven subscales, demonstrating the appropriateness of the model proposed by Coderre et al. (2022). Considering that these correlations were moderate to high in most cases, it is plausible that the SGE construct can be used as a stand-alone measure of FLE. Crucially, the SGE construct (Coderre et al., 2022) predicted WTB for seafood fillets, even after controlling for pro-

environmental consumption (Haws et al., 2014) and social desirability values (Stöber, 2001). That suggests that this brief measure can be used to assess various pro-environmental consumer outcomes. Thus, researchers may consider the instrument when capturing environment-related variables, as environmental friendliness is a strategy that consumers sometimes use to present themselves positively in the eyes of others (see, e.g., White et al., 2019). The current research confirms the FLE scale's concurrent validity, as the indirect measure of the SGE used by Coderre et al. (2022) also predicts consumers' evaluations of products on which the food labels appear in a simulated choice-based situation.

4.2.3 Identifying CBFLE for different labels

B Corp was unrelated to higher WTB as a main effect, compared to CSAP or no intervention. Study 1 showed that only consumers who rate a fictitious label (CSAP) and a real label (B Corp) as moderately high in the SGE construct (Coderre et al., 2022), that is, 0.37 times standard deviation below the mean report similar WTB for fillets with these two labels. Our findings reveal how it is possible to use the CBFLE scale to do a critical fine-grained analysis of different labels. The B Corp sustainability label is weak in Visibility, Awareness, (Dis)honesty, and Design. For the latter, the average score for B Corp was 2.89, compared to 3.89 for CSAP, a new scale with no investment or intervention. Concerning validity, it should be lower as it is impossible that consumers have seen it (item 1), it is not widely used (item 2), and it cannot be familiar (item 3). Using the scale can, therefore, give clear ideas regarding what factors should be worked on.

Coderre et al. (2022) conclude that familiarity is not an antecedent of the other facets of FLE, contrary to the Brand Resonance Model (Jung Jung et al., 2014; Keller, 2001), as a part of their argument for the differentiation of CBFLE vs CBBE. Our results support this notion that

previous exposure is unnecessary for high CBFLE. Despite having relatively low Visibility and Awareness, CSAP showed the same main effect as B Corp and the “no label" condition but a high average value for the Design factor. This result shows that previous consumer experience, learning, or knowledge is not necessarily a prerequisite for other factors.

4.3 Practical implications

The results show that the labels have low equity. The current research points to a need for organizations promoting certified food labels to discern and articulate better signals that are more visual and, therefore, easier to spot, clear (easier to understand), familiar, and trustworthy. The Coderre et al. (2022) FLE scale can be used as a screening tool to examine the strengths and weaknesses of different food labels and schemes and to enhance label quality. Our findings suggest that either counterfeit (CSAP) or genuine but not yet widely known (B Corp) or genuine and known (AHA H-C) labels do not differ significantly in their label equity. These factors may indicate that consumers do not see much difference; any label is a good label to them. That means a serious danger of green/leanwashing can damage consumers’ trust in certified labeling. Instead of focusing so much on creating new labels or improving the certification process (more transparency), perhaps producers and retailers should drop some of them (and use the Coderre et al. (2022) FLE scale as one of the tools to trim) and invest more time and money in creating labels with higher equity and educate consumers on how to read and interpret these labels.

The Coderre et al. (2022) FLE subscales are a fine-grained tool, but in the fast pace of the retail world, quicker methods are sometimes needed. Therefore, we propose also using the SGE construct measure (Coderre et al., 2022) as a potentially cost-effective tool to assess consumer WTB-labeled seafood. Because of its relatively short format, this instrument is particularly relevant to marketing practice. In summary, rather than focusing on finding the most prestigious

certifications, retailers should consider labels with at least a moderately high value in an SGE construct.

4.4 Limitations and Future Research

Our main conclusion is that there is limited difference between the labels. As we tested only three labels, we recommend further studies to examine more different types of certifications and testing tags (unverified claims—see Sigurdsson et al., 2022, 2023). Although we asked participants to rate ten fillets to make our research design relatively realistic, these experimental conditions simplified a real-world shopping process. We recommend doing online and in-store retailing experiments in the future (similar to Sigurdsson et al., 2020—validating laboratory findings with experimentation in retailing).

We focused on WTB, as increased buying usually means additions in market share, benefiting the retailer and leading to market access for suppliers using effective labels. This market share tends to give the retailer better terms with the supplier and can, therefore, also benefit consumers through better pricing and healthier, sustainable food products. We recommend studying other consumer variables related to CBFLE, such as willingness to pay and recommend. There is an ongoing debate in the marketing literature about how WTB and WTP should be measured, with different approaches being appropriate in different situations (Hofstetter et al., 2021; Miller et al., 2011). We asked participants to indicate their WTB fish fillets on a sliding scale, so we cannot determine how much variance can be attributed to our way of capturing the focal dependent variable. We recommend multi-methods. Further research could, for example, be based on consumer choice with choice-based conjoint (CBC) analysis and using an open-ended question format to see things better from consumers' perspectives (see Miller et al., 2011). Further, future studies could include in-store experiments in physical or

online stores. That could prove useful to get a more realistic insight into the consumers' choices and preferences based on these different stimuli, similar to what was done by Sigurdsson et al. (2020).

The labels we used across experimental conditions covered a relatively narrow range in the SGE construct subscale (Coderre et al., 2022), raising questions about the distribution of the latent variable. Because the above scale may not cover a broad spectrum of the construct, further studies may be needed to validate this instrument and possibly revise it with other psychometric approaches, such as a Mokken Scale Analysis (MSA), which does not assume that the latent construct is continuous (for related procedures, see, e.g., Dima, 2018; Folwarczny et al., 2021). Although the five-item measure capturing the SGE construct was generally highly correlated with most subscales of the Coderre et al. (2022) FLE scale, we found weak to moderate correlations between this measure and some subscales, and these correlations differed substantially between Studies 1 and 2. Therefore, further studies should investigate whether the SGE construct measure can be used as a stand-alone instrument.

Future research could include testing and comparing several made-up label designs and measuring consumers' responses to the labels based on the Coderre et al. (2022) FLE scale. This research could, for example, be done through lab experiments, then taking the results further and designing a choice experiment or doing a choice-based conjoint analysis to get clear information on the attributes that matter most to respondents. Correct vs false recognition, as defined by Henderson & Cote (1998), is something that would be interesting to further connect to the food label literature, seeing as how important it is to prevent green/leanwashing of food products, where false recognition could help greenwashers accomplish their goals by mimicking labels with higher FLE.

References

American Heart Association. (n.d.) *Heart-Check Certification.* Retrieved November 17 2022 from https://www.heart.org/en/healthy-living/company-collaboration/heart-check-certification

An, S., Kerr, G., & Jin, H. S. (2019). Recognizing native ads as advertising: Attitudinal and behavioral consequences. *Journal of Consumer Affairs*, 53(4), 1421–1442.

Anisimova, T. & Sultan, P. (2014). The role of brand communications in consumer purchases of organic foods: A research framework. *Journal of Food Products Marketing, 20*(5), 511–532. http://doi.org/10.1080/10454446.2012.728984#.VF9dcGdadZs

Aprile, M. C., & Punzo, G. (2022). How environmental sustainability labels affect food choices: Assessing consumer preferences in southern Italy. *Journal of Cleaner Production*, 332, 130046. https://doi.org/10.1016/j.jclepro.2021.130046

Ballen, F. H., Evans, E., & Parra-Acosta, Y. K. (2021). Consumer preferences for green skin avocados in the US market: the role of experienced quality attributes, credence attributes, and demographic factors. *Journal of Agricultural & Food Industrial Organization*, 20(1), 15–23. https://doi.org/10.1515/jafio-2020-0006

B Corp Certification. (n.d.-a). *About B Corp certification. Measuring a company's entire social and environmental impact.* Retrieved November 17 2022 from https://www.bcorporation.net/en-us/certification

B Corp Certification. (n.d.-b). *Make business a force for good.* Retrieved November 17 2022 from https://www.bcorporation.net/en-us

B Lab United States & Canada. (n.d.). *Process, requirements & fees.* Retrieved November 15 2022 from https://usca.bcorporation.net/process-requirements-fees/

Bailey, A. A., Mishra, A. S., & Tiamiyu, M. F. (2018). Application of GREEN scale to understanding US consumer response to green marketing communications. *Psychology & Marketing, 35*(11), 863–875. https://doi.org/10.1002/mar.21140

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Boulding, W., & Kirmani, A. (1993). A consumer-side experimental examination of signaling theory: do consumers perceive warranties as signals of quality? *Journal of consumer research, 20*(1), 111–123.

Buil, I., de Chernatony, L., & Martínez, E. (2008). A cross-national validation of the consumer-based brand equity scale. *Journal of Product & Brand Management, 17*(6), 384–392. https://doi.org/10.1108/10610420810904121

Campbell, M. C. (1995). When attention-getting advertising tactics elicit consumer inferences of manipulative intent: The importance of balancing benefits and investments. *Journal of Consumer Psychology*, 4(3), 225–254. https://doi.org/10.1207/s15327663jcp0403_02

Carpenter, M., & Larceneux, F. (2008). Label equity and the effectiveness of values-based labels: An experiment with two French Protected Geographic Indication labels. *International Journal of Consumer Studies, 32*(5), 499–507. https://doi.org/10.1111/j.1470-6431.2008.00704.x

Chen, X., Alfnes, F., & Kyrre, R. (2015). Consumer preferences, eco-labels, and effects of negative environmental information. *AgBioForum, 18*(3), 327–336.

Coderre, F., Sirieix, L., & Valette-Florence, P. (2022). The facets of consumer-based food label equity: Measurement, structure and managerial relevance. *Journal of Retailing and Consumer Services, 65*, 102838. https://doi.org/10.1016/j. Jretconser.2021.102838

Colombo, R.A. and Morrison, D.G. (1989). Consumer preference formation and pioneering advantage, *Journal of Marketing Research*, 26(3), 285–98.

Darke, P. R., & Ritchie, R. J. (2007). The defensive consumer: Advertising deception, defensive processing, and distrust. *Journal of Marketing Research*, 44(1), 114–127.

DeVellis, R. (2012). *Scale development: Theory and applications.* SAGE

Dima, A. L. (2018). Scale validation in applied health research: Tutorial for a 6-step R-based psychometrics protocol. *Health Psychology and Behavioral Medicine, 6*(1), 136–161. https://doi.org/10.1080/21642850.2018.1472602

Donato, C., & Adıgüzel, F. (2022). Visual complexity of eco-labels and product evaluations in online setting: Is simple always better?. *Journal of Retailing and Consumer Services*, 67, 102961. https://doi.org/10.1016/j.jretconser.2022.102961

Ecolabel Index. (n.d.). *All ecolabels.* https://www.ecolabelindex.com/ecolabels/

Feucht, Y., & Zander, K. (2014). Consumers' knowledge and information needs on organic aquaculture. In Rahmann, G. & Aksoy, U. (Eds.), *Building organic bridges, Proceedings of the 4th ISOFAR Scientific Conference, 13–15 Oct., Istanbul, Turkey (pp. 375–378).* ISOFAR.

Folwarczny, M., Li, N. P., Sigurdsson, V., Tan, L. K., & Otterbring, T. (2021). Development and psychometric evaluation of the Anticipated Food Scarcity Scale (AFSS). *Appetite, 166*, 105474.

de Freitas Netto, S. V., Sobral, M. F. F., Ribeiro, A. R. B., & de Luz Soares, G. R. (2020). Concepts and forms of greenwashing: A systematic review. *Environmental Sciences Europe, 32*(19). https://doi.org/10.1186/s12302-020-0300-3

Frommeyer, B., Wagner, E., Hossiep, R., & Schewe, G. (2022). The utility of intention as a proxy for sustainable buying behavior - A necessary condition analysis. *Journal of Business Research, 143,* 201–213. https://doi.org/10.1016/j.jbusres.2022.01.041

Ghvanidze, S., Velikova, N., Dodd, T., & Oldewage-Theron, W. (2017). A discrete choice experiment of the impact of consumers' environmental values, ethical concerns, and health consciousness on food choices: A cross-cultural analysis. *British Food Journal*, 119(4), 863–881. https://doi.org/10.1108/BFJ-07-2016-0342

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

Haws, K. L., Winterich, K. P., & Naylor, R. W. (2014). Seeing the world through GREEN-tinted glasses: Green consumption values and responses to environmentally friendly products. *Journal of Consumer Psychology*, *24*(3), 336–354. https://doi.org/10.1016/j.jcps.2013.11.002

Henderson, P. W., & Cote, J. A. (1998). Guidelines for selecting or modifying logos. *Journal of Marketing, 62,* 14–30.

Hoek, A. C., Pearson, D., James, S. W., Lawrence, M. A., & Friel, S. (2017). Healthy and environmentally sustainable food choices: Consumer responses to point-of-purchase actions. *Food Quality and Preference, 58,* 94–106. https://doi.org/10.1016/j.foodqual.2016.12.008

Hofstetter, R., Miller, K. M., Krohmer, H., & Zhang, Z. J. (2021). A de-biased direct question approach to measuring consumers' willingness to pay. *International Journal of Research in Marketing, 38*(1), 70–84. https://doi.org/10.1016/j.ijresmar.2020.04.006

Holland, S. (2016). Lending credence: Motivation, trust, and organic certification. *Agricultural and Food Economics, 4*, 14. https://doi.org/10.1186/s40100-016-0058-5

Janssen, M., & Hamm, U. (2012). Product labeling in the market for organic food: Consumer preferences and willingness-to-pay for different organic certification logos. *Food Quality and Preference, 25*(1), 9─22. https://doi.org/10.1016/j.foodqual.2011.12.004

Johnston, R. J., & Roheim, C. A. (2006). A battle of taste and environmental convictions for ecolabeled seafood: A contingent ranking experiment. *Journal of Agricultural and Resource Economics, 31*(2), 283–300. https://www.jstor.org/stable/40987319

Jung Jung, H., Lee, Y., Kim, H., & Yang, H. (2014). Impacts of country images on luxury fashion brand: Facilitating with the brand resonance model. *Journal of Fashion Marketing and Management, 18*(2), 187─205. https://doi.org/10.1108/JFMM-10-2013-0113

Karnani, A., McFerran, B., & Mukhopadhyay, A. (2017). Corporate leanwashing and consumer beliefs about obesity. *Current Nutrition Reports, 6*, 206–211. https://doi.org/10.1007/s13668-017-0210-1

Keller, K. L. (1993). Conceptualizing, measuring, and managing customer-based brand equity. *Journal of Marketing, 57*(1), 1-22. https://doi.org/10.1177/002224299305700101

Keller, K. L. (2001). *Building customer-based brand equity: A blueprint for creating strong brand* (Working paper, report no. 01-107). Marketing Science Institute, pp. 1─31

Kohli, C., Suri, R., & Thakor, M. (2002). Creating effective logos: Insights from theory and practice. *Business Horizons, 45*(3), 58–64. http://dx.doi.org/10.1016/S0007-6813(02)00203-3

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lang, L. D., Lim, W. M., & Guzmán, F. (2022). How does promotion mix affect brand equity? Insights from a mixed-methods study of low involvement products. *Journal of Business Research, 141,* 175-190. https://doi.org/10.1016/j.jbusres.2021.12.028

Larceneux, F., Benoit-Moreau, F., & Renaudin, V. (2012). Why might organic labels fail to influence consumer choices? Marginal labelling and brand equity effects. *Journal of Consumer Policy*, 35(1), 85–104.

Lu, Z. Y., & Hsee, C. K. (2019). Less willing to pay but more willing to buy: How the elicitation method impacts the valuation of a promotion. *Journal of Behavioral Decision Making*, 32(3), 334-345. https://doi.org/10.1002/bdm.2115

Luffarelli, J., Mukesh, M., & Mahmood, A. (2019). Let the logo do the talking: The influence of logo descriptiveness on brand equity. *Journal of Marketing Research, 56*(5), 862–878. https://doi.org/10.1177/0022243719845000

Machado, J. C., de Carvalho, L. V., Torres, A., & Costa, P. (2015). Brand logo design: Examining consumer response to naturalness. *Journal of Product & Brand Management, 24*(1), 78–87. https://doi.org/10.1108/JPBM-05-2014-0609

Majer, J. M., Henscher, H. A., Reuber, P., Fischer-Kreer, D., & Fischer, D. (2022). The effects of visual sustainability labels on consumer perception and behavior: A systematic review of the empirical literature. *Sustainable Production and Consumption*, *33*, 1–14. https://doi.org/10.1016/j.spc.2022.06.012

Miller, K. M., Hofstetter, R., Krohmer, H., & Zhang, Z. J. (2011). How should consumers' willingness to pay be measured? An empirical comparison of state-of-the-art approaches. *Journal of Marketing Research, 48*(1), 172–184.

Netemeyer, R. G., Krishnan, B. C., Pullig, C., Wang, G., Yagci, M., Dean, D. H., Ricks, J. M., & Wirth, F. (2004). Developing and validating measures of facets of customer-based brand equity. *Journal of Business Research, 57*(2), 209–224. https://doi.org/10.1016/S0148-2963(01)00303-4

Newell, S. J., Goldsmith, R. E., & Banzhaf, E. J. (1998). The effect of misleading environmental claims on consumer perceptions of advertisements. *Journal of Marketing Theory and Practice*, 6(2), 48–60. https://doi.org/10.1080/10696679.1998.11501795

Nikolova, H. D., & Inman, J. (2015). Healthy choice: The effect of simplified point-of-sale nutritional information on consumer food choice behaviour. *Journal of Marketing Research*, 52(6), 817–835. https://doi.org/10.1509/jmr.13.0270

Otterbring, T., & Folwarczny, M. (2022). Firstborns buy better for the greater good: Birth order differences in green consumption values. *Personality and Individual Differences, 186*, 111353. https://doi.org/10.1016/j.paid.2021.111353

Pappu, R., Quester, P. G., & Cooksey, R. W. (2005). Consumer-based brand equity: Improving the measurement - empirical evidence. *Journal of Product & Brand Management, 14*(3), 143–154. https://doi.org/10.1108/10610420510601012

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, *54*(4), 1643–1662. https://doi.org/10.3758/s13428-021-01694-3

R Core Team. (2021). R: A l*anguage and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.Rproject.org/

Sam's Club. (n.d.). *Helping you make smart choices about your seafood.* https://www.samsclub.com/content/seafood-sustainability

Schrobback, P., Zhang, A., Loechel, B., Ricketts, K., & Ingham, A. (2023). Food Credence Attributes: A Conceptual Framework of Supply Chain Stakeholders, Their Motives, and Mechanisms to Address Information Asymmetry. *Foods*, *12*(3), 538. https://doi.org/10.3390/foods12030538

Sigurdsson, V., Larsen, N. M., Alemu, M. H., Gallogly, J. K., Menon, R. G. V., & Fagerstrøm, A. (2020). Assisting sustainable food consumption: The effects of quality signals stemming from consumers and stores in online and physical grocery retailing. *Journal of Business Research, 112,* 458–471. https://doi.org/10.1016/j.jbusres.2019.11.029

Sigurdsson, V., Larsen, N. M., Folwarczny, M., Fagerstrøm, A., Menon, R. G. V., & Sigurdardottir, F. T. (2023). The importance of relative customer-based label equity when signaling sustainability and health with certifications and tags. *Journal of Business Research, 154,* 113338. https://doi.org/10.1016/j.jbusres.2022.113338

Sigurdsson, V., Larsen, N. M. Pálsdóttir, R. G., Folwarczny, M. Menon, R. G. V., & Fagerstrøm, A. (2022). Increasing the effectiveness of ecological food signaling: Comparing sustainability tags with eco-labels. *Journal of Business Research, 139,* 1099–1110. https://doi.org/10.1016/j.jbusres.2021.10.052

Sirieix, L., Delanchy, M., Remaud, H., Zepeda, L., & Gurviez, P. (2013). Consumers' perceptions of individual and combined sustainable food labels: a UK pilot investigation. *International Journal of Consumer Studies*, 37(2), 143–151. https://doi.org/10.1111/j.1470-6431.2012.01109.x

Stöber, J. (2001). The social desirability scale-17 (sds-17): Convergent validity, discriminant validity, and relationship with age. *European Journal of Psychological Assessment, 17*(3), 222–232. https://doi.org/10.1027//1015-5759.17.3.222

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods, 42*(4), 1096–1104. https://doi.org/10.3758/BRM.42.4.1096

Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, *44*(1), 24–31. https://doi.org/10.1177/0098628316677643

Taufique, K. M. R., Vocino, A., & Polonsky, M. J. (2017). The influence of eco-label knowledge and trust on pro-environmental consumer behaviour in an emerging market. *Journal of Strategic Marketing*, 25(7), 511–529. https://doi.org/10.1080/0965254X.2016.1240219

Teisl, M. F., Peavey, S., Newman, F., Buono, J., & Hermann, M. (2002). Consumer reactions to environmental labels for forest products: A preliminary look. *Forest Products Journal*, 52(1), 44–50.

Till, B. D., Baack, D. W., & Waterman, B. (2011). Strategic brand association maps: Developing brand insight. *Journal of Product & Brand Management, 20*, 92–100. https://doi.org/10.1108/10610421111121080

Tonkin, E., Meyer, S. B., Coveney, J., Webb, T., & Wilson, A. M. (2016). The process of making trust related judgements through interaction with food labelling. *Food policy*, *63*, 1–11. https://doi.org/10.1016/j.foodpol.2016.06.007

Vecchio, R., Van Loo, E. J., & Annunziata, A. (2016). Consumers' willingness to pay for conventional, organic and functional yogurt: evidence from experimental auctions. *International Journal of Consumer Studies*, 40(3), 368–378.

Webb, M. A., & Tangney, J. P. (2022). Too Good to Be True: Bots and Bad Data From Mechanical Turk. *Perspectives on Psychological Science*, 17456916221120027. https://doi.org/10.1177/1745691622112002

White, K., Habib, R., & Hardisty, D. J. (2019). How to SHIFT consumer behaviors to be more sustainable: A literature review and guiding framework. *Journal of Marketing*, *83*(3), 22–49. https://doi.org/10.1177/0022242919825649

Whole Foods Market. (2021, April 20). *Whole Foods Market Seafood Department Quality Standards overview for farm-raised seafood*. https://assets.wholefoodsmarket.com/www/missions-values/seafood-sustainability/Quality%20Standards%20for%20Farm%20Raised%20Species_5.1.21%20wfm-1.pdf

Yoo, B., & Donthu, N. (2001). Developing and validating a multidimensional consumer-based brand equity scale. *Journal of Business Research*, *52*(1), 1–14. https://doi.org/10.1016/S0148-2963(99)00098-3

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *92*(2, Pt.2), 1–27. https://doi.org/10.1037/h0025848

Zepeda, L., Sirieix, L., Pizarro, A., Corderre, F., & Rodier, F. (2013). A conceptual framework for analyzing consumers' food label preferences: An exploratory study of sustainability labels in France, Quebec, Spain and the US. *International Journal of Consumer Studies*, 37(6), 605–616.