Exhibit 13

Project Title: Consumer Accuracy at Identifying Plant-based and Dairy-based Milk Items

Authors: Silke Feltz, University of Oklahoma
Adam Feltz, University of Oklahoma

Executive Summary:

On June 14, 2017, the European Parliament ruled that 'milk' terms could no longer be used to describe plant-based products. There are similar policy debates in the United States (e.g., the Dairy Pride Act). The motivation for the ruling and policies centers on two key claims: (1) using 'milk' terms to describe plant-based and animal-based products would cause confusion about which products are plant-based and animal-based, and (2) using 'milk' terms would cause confusion about the nutritional content of plant-based and animal-based products. We conducted 8 studies testing the extent to which people display the confusions indicated in 1 or 2. Overall, participants behaved as if confusions indicative of 1 and 2 are not pervasive.

We focused on confusion concerning milk and cheese products because of their ubiquity. We conducted a series of studies to determine if participants could accurately identify plant and animal-based cheese and milk products as plant or animal-based. A different series of studies was conducted to determine if people could accurately identify general nutritional differences between plant-based and animal-based milk and cheese products. Also, we aimed to develop an objective, knowledge-based measure of differences between plant-based and animal-based milk products. The measure was designed to help predict accuracy on the product and nutrition identification tasks. Finally, we replicated in a separate study all of the findings in a national sample.

Here, we report meta-analytically combined results. On average, participants were not measurably worse at identifying plant-based products than they were at identifying animal-based-products. Participants accurately identified the source of animal-based milk products 84% of the time, plant-based milk products 88% of the time, animal-based cheese products 81% of the time, and plant-based cheese products 74% of the time.

Participants accurately identified nutritional differences 62% of the time for milk products and 50% of the time for cheese products. The relatively low correct answer rate for nutritional differences should be interpreted as a lower bound estimate since participants who responded "I don't know" were not coded as answering the question correctly. “I don’t know” responses may accurately reflect an individual’s assessment of their knowledge and not an error. Treating “I don’t know” responses as correct increases milk nutrition accuracy to 73% and cheese nutrition accuracy to 75%.

The objective Milk Literacy Scale that we developed successfully predicted performance on the identification tasks (mean $r = .2$, 95% CI .08 - .32), suggesting those who knew more about differences between plant-based and animal-based products were better at the identification tasks. The national sample had a large enough sample size to construct path models estimating relations among variables. The path models suggested that those who are more numerate tend to be more milk literate and know more about general nutrition and about animals used as food. Milk literacy, general nutrition knowledge, and knowledge of animals used as food predicted performance on the nutrition identification tasks. Consistently, participants performed better on the subscale measuring differences between plant and animal-based milk products ($M = 3.47$ out of 6) than the subscale measuring differences among animal-based milk products ($M = 1.89$ out of 6).

These results suggest that generally, people are fairly accurate at identifying plant-based from animal-based products. To the extent that they are not, the path models suggest that some educational interventions would likely be successful in increasing consumer accuracy (via knowledge).

Consumer Accuracy at Identifying Plant-based and Dairy-based Milk Items

Silke Feltz[1] and Adam Feltz[1]

[1]University of Oklahoma

Address for Correspondence:

Word Count (excluding abstract and references):

Adam Feltz
Center for Applied Social Research &
Department of Psychology
University of Oklahoma
455. W. Lindsey St.
Dale Hall Tower, Room 705
Norman, OK 73019

Abstract

Recent debates have centered on whether people are product literate enough to make informed decisions about plant-based and animal-based milk products. In 8 studies, we provide evidence that consumers do not make mistakes indicative of pervasive lack of product literacy. Overall, people were accurate at identifying plant-based and animal-based milk and cheese products as being plant or animal-based (74% - 84% of the time). Participants were also generally accurate at identifying nutritional differences between plant-based and animal-based milk and cheese products (50%-62% accuracy). We developed the Milk Literacy Scale, which is a 12-item, validated, knowledge-based instrument that measures knowledge of differences among plant-based and animal-based milk products. The Milk Literacy Scale predicted accuracy in the identification tasks (meta-analytically estimated $r = .2$). All results were replicated in a large, national sample ($N = 1054$). These results suggest that people are generally product literate about milk products and offer some insights into what kinds of interventions would help make people more product literate about milk.

Consumer Accuracy at Identifying Plant-based and Dairy-based Milk Items

On June 14, 2017, the European Parliament ruled that producers of plant-based milk items could no longer use the terms ‘cheese’ or ‘milk’ to describe their products. Those terms (along with the related terms like ‘whey’, ‘cream’, ‘butter’, ‘buttermilk’, and ‘yogurt’) are to be exclusively used for items that contain animal milk. Among the major reasons for this decision was the risk of confusion for consumers if terms traditionally used to designate animal-based milk (e.g., cheese, cream, whey) items were also used for plant-based milk items (e.g., soy milk, soy cheese). But do people make mistakes when identifying plant-based and dairy-based food items? Are people product literate enough to reliably distinguish plant-based milk products from dairy-based milk products?

We present 8 studies to help address these questions. Studies 1-3 present a short, efficient, 12-item measure of plant-based milk and animal-based milk knowledge. Studies 4-5 provide evidence that people are generally accurate at identifying plant-based from animal-based milk products, and this accuracy is predicted by increased knowledge of milk products and general nutrition knowledge. Studies 6-7 suggest that people are also able to accurately identify nutritional differences between plant-based and animal-based milk products. Again, accuracy was predicted by increased knowledge of milk products and general nutrition knowledge. Study 8 replicated the results from studies 1-7 on a larger, more nationally representative sample. Study 8 also allowed testing path models indicating that while knowledge of milk products and general nutrition knowledge were prominent predictors of accuracy, numeracy was also related to increased identification accuracy. These results suggest that generally people are product literate enough to identify key differences between plant-based and animal-based milk products and offer avenues for helping those who are not.

Consumer Product Literacy

Nearly all plant-based milk items are labeled as such (e.g., soy milk, vegan cheese). On the face of it, then, a conscientious consumer should have no trouble distinguishing plant-based and dairy-based milk items simply by reading the label. However, whether consumers use information on product labels is context sensitive and depends on individual motivations and backgrounds (for reviews, see Hall and Osses (2013), and Hess, Visschers, and Siegrist (2012)). By some estimates, nearly everyone uses product labels especially if the labels have the following features: graphs or symbols, adjective labels with minimal numerical information, and information on the front of the package (Campos, Doxey, & Hammond, 2011).[1] Consequently, nutrition labels can contain "highly credible sources information, and many consumers report using nutrition labels to guide their selection of food products" (Campos et al., 2011, pp. 6-7).

One of the justifications for restricting the use of milk terms is that using terms like 'cheese' and 'milk' for plant-based products will lead to consumer confusion. Hence, even if people use labels in making decisions, those people will not generally be able to make accurate decisions about the milk products they buy. There are many examples of people not understanding information on product labels. For instance, while people generally understand the term 'calorie' they have difficulty linking calories to other concepts like energy and have trouble converting numerical information about calories meaningfully (Cowburn & Stockley, 2005). In this light, one could argue that using 'milk' terms in an environment that includes plant-based products would lead to misunderstandings like those seen with the term 'calorie.' People might in principle understand the difference between plant-based and animal-based milk products, but they would not be able to meaningfully translate that understanding into an accurate buying decision.

These examples call into question consumer *product literacy* about milk items. Product literacy can be defined as "the degree to which consumers have the capacity to locate, evaluate, apply, and communicate basic information needed to make appropriate product related decisions" (Kopp, 2012, p. 196). For some products, people do not have adequate product literacy (e.g., with calories). However, in other instances, people are very product literate. For example, people tend to be able to make simple comparisons and understand some vocabulary that is presented on labels (Cowburn & Stockley, 2005).

Given this background, our central question is: Are people product literate enough to accurately identify plant-based and dairy-based milk products? There are at least two different current arguments that have been offered for thinking that people generally are not product literate enough about milk products. The first argument is that people might be confused about which products are animal-based and which products are plant-based. For example, according to the European Court's ruling, "In the absence of such limits, those designations would not enable products with particular characteristics related to the natural composition of animal milk to be

[1] This same systematic review suggested that nearly 75% of Americans report using nutrition labels at least sometimes when they make a buying decision.

identified with certainty" (Case C-422/16, §44). Mistaking the nature of a product is the most basic kind of mistake that consumers could make. It stands to reason that if a person cannot accurately identify what a product is, then that person will also not be able to reliably identify other relevant properties of that item (e.g., nutritional content and differences with other products, environmental impacts).

A different argument holds that even if people are product literate enough to *identify* plant-based from animal-based products, they are not product literate enough to understand nutritional differences between plant-based and animal-based products. For example, according to the Dairy Pride Act (*Dairy Pride Act*, 2017), there is the risk that consumers would mistakenly assume nutritional equivalency between plant-based and animal-based 'milk' products. As the argument goes, the proper protection given such mistakes is to create or enforce legislation that bans using 'milk' terms for anything other than animal-based products.

To our knowledge, all of these claims about dairy confusion are not empirically tested. We set out to test them in 8 studies. The studies proceeded in three stages. The first stage was aimed to develop a research instrument to measure *Milk Literacy*. The Milk Literacy Scale (MLS) was designed to be an objective, general knowledge based instrument measuring what people know about differences between (a) animal-based and plant-based milk products and (b) different animal-based milk products (Studies 1-3). The second stage was designed to test whether people could accurately identify milk and cheese products as being animal or plant-based (Studies 4-5). The third stage was designed to test the degree to which people can accurately understand nutritional differences between dairy-based and plant-based milk items (Studies 6-7). Study 8 was designed to replicate findings of Studies 1-7 in a large, more nationally representative sample.

Milk Literacy Scale

We used Item Response Theory (IRT) to develop the MLS. IRT analyses measure latent traits. Latent traits are unobserved yet assumed to be causally responsible for a pattern of responses. In this case, the latent trait is knowledge of milk products. Unlike classical test theory, IRT can provide item-level analyses. In particular, IRT methods can estimate the probability that people of different levels of knowledge will answer a question correctly. If one plots the probabilities of correct answers among people with different levels of knowledge, the resulting plot forms an *S* curve (from low probability of correct answer for low-knowledge people to high probability of correct answer for high-knowledge people). This *S* curve is called the *item characteristic curve*. Difficulty and discrimination are two important properties of item characteristics curves. An item's characteristics curve can be located on a scale of how difficult the item is. Items that have better discrimination will have sharper up-slopes on the *S* curve (i.e., the item does a better job discriminating among low and high ability at that ability location). Ideally for our purposes, the knowledge test should have items with strong discrimination and a variety of difficulties so that different ability levels can be estimated by the test. Study 1 was designed to identify those properties of desirable items. Study 1 was the first in a planned series of studies to find the set of items with desired properties.

Study 1

Study 1 was designed to test an initial battery of items to measure objective knowledge about milk products. The goal of Study 1 was to identify item-level properties of those knowledge based items. Using these analyses, we planned to identify empirically desirable items to retain for subsequent studies.

Participants

Two hundred and twenty-eight participants were recruited from Amazon's Mechanical Turk. Amazon's Mechanical Turk data are generally taken to have acceptable quality, especially in comparison to typical subject pools (e.g., university undergraduate subjects pools) (M. Buhrmester, Kwang, & Gosling, 2011; M. Buhrmester, Talifar, & Gosling, 2018; Crump, McDonnell, & Gureckis, 2013; Mason & Suri, 2012; Paollacci, Chandler, & Ipeirotis, 2010; Rouse, 2015). Demographics for the participants (for all studies) are reported in Table 1.

Materials

We developed 23 items that had face validity concerning aspects of soy milk and whole milk (see Appendix A). Twelve items dealt specifically with nutritional differences between whole cow milk and soy milk (e.g., whole cow milk has more cholesterol than fortified soy milk). These nutritional differences were based on an analysis by Vanga and Raghavan (2018). Call this the *Soy* subscale of the MLS. Ten items dealt specifically with differences between whole cow milk and skim cow milk (e.g., whole cow milk has more protein than skim cow milk). Call this the *Milk* subscale of the MLS. We also include one general question concerning whether soy milk is made with any cow milk (Item 23). Participants were asked to rate the statements as either being true or false. Finally, we collected basic demographic information.

Results and Discussion

Analyses proceeded on the assumption that each of the Soy and Milk subscales of the MLS measured only one latent variable (see subsequent studies for evidence for this assumption). So, two sets of IRT analyses were conducted on each set of items (item 23 was included in the Milk subscale). All IRT analyses were conducted using R (R Core Team, 2018) with the LTM package (Rizopoulos, 2006). A 2-parameter model was used for each set of items.[2] As expected, some items did not have desirable properties. Two Milk subscale items had negative discrimination (i.e., as one knows more, one is less likely to answer the item correctly). Four items of the Soy subscale items had reverse discrimination, and one item was exceedingly easy and had little discrimination (see Appendix A for details of each item).

Study 2

The results of Study 1 suggested several advantageous modifications to the MLS. First, the IRT analyses from Study 1 showed that some items had reverse discrimination (items 2, 6, 8, 13, 15, and 16). These items were eliminated for Study 2. Also, Item 1 had very low discrimination (0.17) and was very easy (-9.1), so Item 1 was also eliminated from subsequent studies. We also randomly selected items to change their true-values (i.e., taking the opposite truth-value) to ensure that the items behaved roughly the same with different truth-values. Finally, the discrimination was relatively low for many of the items. While we did not have direct evidence for this, we suspected that many people guessed at answers they did not know. This would likely result in getting some answers correct by chance and thereby reducing discrimination. To help alleviate this problem, we included an “I don’t know” option in this and subsequent studies. The primary goal of Study 2 was to verify the item-level properties found in Study 1 using the revised MLS.

Participants

Two hundred and twenty-six participants were recruited from Amazon’s Mechanical Turk.[3] Because we planned on a series of studies that drew on Amazon’s Mechanical Turk’s

---

[2] A two-parameter model is different from a 1-parameter model. One-parameter models only estimate item difficulty and assume that the discrimination for each item is the same. Three-parameter models include a pseudo-guessing parameter in addition to estimating difficulty and discrimination that helps to control for people getting items correct simply by guessing (Baker, 2004).

[3] A coding mistake prevented demographic data from being collected in Study 2.

participants pool, we kept track of those who participated in previous studies. This was an effort to help ensure naiveté and non-repeated responses. No participants were allowed to take part in more than 1 study.

Materials

Participants received the modified MLS (see Appendix B).

Results and Discussion

Two separate IRT analyses using 2-parameter models were conducted, one for the Milk subscale and one for the Soy subscale of the MLS. The modified scale largely had acceptable discrimination along with a range of difficulties (see Table 2).

However, the results of Study 2 suggested that further refinements of the MLS were possible. First, two items in the Soy subscale had very low discrimination (Items 1 and 4, .23 and .29 respectively). Four items in the Milk subscale had similar difficulty (Items 10, 11, 13, and 15; 0.5, 0.61, 0.53, and 0.54 respectively). So, some of those four items could be eliminated without loss of information from the scale.

Study 3

Study 3 was designed to replicate the IRT results of Study 2 with the modifications suggested by Study 2. Study 3 was also designed to demonstrate that the MLS was multidimensional consisting of two unidimensional subscales Soy and Milk. Finally, since there was good reason to think that the 12-item Milk Literacy scale was going to have acceptable formally IRT properties, Study 3 afforded the opportunity to begin to display convergent and divergent validity. To help establish convergent validity, we included the Nutrition Knowledge Scale (Dickson-Spillmann, Siegrist, & Keller, 2011). If the MLS measures food knowledge, then the MLS score should be related to general nutrition knowledge. To help establish divergent validity, participants responded to a general personality inventory. If the MLS measures knowledge, then it should be largely unrelated to general personality traits.

Participants

Two-hundred and thirty participants were recruited from Amazon's Mechanical Turk.

Materials

The Milk Literacy Scale (MLS). A modified, 12-item version of the MLS was used. In particular, Items 1 and 4 were eliminated from the Soy subscale because they had very low discrimination. Because 4 items in the Milk subscale had similar properties, some of those items could be eliminated without much loss of information. To that end, we eliminated Items 11 and 13 because they had the lowest discrimination of the 4 items. The Soy and Milk subscales of the MLS had 6-items each. Each statement was rated as being either true, false, or the participants could respond that they did not know. Correct answers were coded as '1' and incorrect or "I don't know" responses were coded as '0'. A total correct answer score for each of the two subscales was calculated.

The Nutrition Knowledge Scale (Dickson-Spillmann et al., 2011). The Nutrition Knowledge scale is a 20-item scale with general statements about nutrition (e.g., “Brown sugar is much healthier than white sugar). Response options were true, false, or "I don’t know." Correct answers were coded as '1' and incorrect or “I don’t know” responses were coded as '0'. A total correct answer score was calculated for the Nutrition Knowledge Scale.

The Ten-Item Personality Inventory (TIPI) (Gosling, Rentfrow, & Swann, 2003). The TIPI is a 10-item measure of the Big Five Personality traits. Each of the Big Five traits is measured by rating how much pairs of adjectives describe one’s self (e.g., “extraverted, enthusiastic”) on a 7- point Likert scale (Disagree strongly to Agree strongly). Scores for each of the Big Five are calculated by averaging ratings from two pairs of adjectives.

Results and Discussion

A test of unidimensionality was conducted on the entire 12-item MLS. The test for unidimensionality tested whether the eigenvalue for the second factor is greater than would be theoretically expected. If the second eigenvalue is greater than would be expected, then one can reject unidimensionality. The theoretical eigenvalue based on 200 Monte Carlo samples was 1. The eigenvalue of the second factor in the data was 2.4, significantly greater than the theoretically derived eigenvalue (p = .005). Unidimensionality could be rejected for the full MLS. Unidimensionality tests were done for each of the Milk and Soy subscales using the same method. In each case, unidimensionality could not be rejected: Soy observed second eigenvalue = .38, average eigenvalue of 200 Monte Carlo samples = .52, p = .89; Milk observed second eigenvalue = 0.82, average of second eigenvalue in 200 Monte Carlo samples = .78, p = .31.

A series of IRT analyses were conducted on each subscale of the MLS. The first set of analyses concerned the Soy subscale of the MLS. Planned analyses compared a constrained 1-parameter model to an unconstrained 1-parameter model. The unconstrained 1-parameter (AIC = 1543.26, BIC = 1567.39) model was a better fit to the data than the constrained model (AIC =

1559.08, BIC = 1579.76), $p < .001$. The unconstrained 1-parameter model had good fit to the data, passing a goodness of fit test, $p = .17$ and having acceptable residuals on the margins (all chi squared values < 1.31). A 2-parameter model (AIC = 1544.33, BIC 1585.69) was not significantly better than the 1-parameter unconstrained model, $p = .11$. However, the 2-parameter model also had acceptable fit to the data with all chi square values of residuals on the margins less than .76.

We performed the same series of analysis on the Milk subscale of the MLS. A 1-parameter unconstrained model (AIC = 1463.17, BIC = 1487.3) was a better fit to the data than a 1-parameter constrained model (AIC = 1466.87, BIC = 1487.55), $p = .02$. A 2-parameter (AIC = 1434.11, BIC 1475.47) model was a better fit to the data than a 1-parameter unconstrained model, $p < .001$. The 2-parameter model also had acceptable residuals on the margins for the items, all chi squared values < 1.24.

The IRT analysis suggested that the 12-item version of the MLS had acceptable internal properties. Convergent, divergent, and criterion validity remained to be demonstrated. While we planned to establish criterion validity in subsequent studies, some evidence for convergent and divergent validity could be provided in the current study. Correlations were calculated between the variables gathered (see Table 3). As expected, both of the MLS subscales were moderately to strongly related to the Nutrition Knowledge Scale, suggesting convergent validity. The Soy subscale of the MLS was also moderately related to the global personality trait conscientiousness. This somewhat unexpected finding makes sense in the context that conscientious people are likely to be more engaged and vigilant about what they eat (Lunn, Nowson, Worsley, & Torres, 2014). The Soy and Milk subscales were weakly related to other personality traits suggesting divergent validity.

Of note, the Milk subscale of the MLS ($M = 1.95$, $SD = 1.5$) was more difficult than the Soy subscale ($M = 4$, $SD = 1.7$), $t(231) = 15.7$, p , .001, $d = 1.03$. This result suggests that people are less knowledgeable about the differences among animal-based 'milk' products then they are about differences between animal-based and plant-based milk products, at least as measured by the MLS.

Production Identification

One of the main arguments for forbidding the use of 'milk' terms for plant-based products is that the usage would cause confusion among consumers. Studies 4 and 5 were designed to determine how good people are at correctly identifying animal based and plant-based milk products. Two types of milk products were selected because of their general ubiquity and availability. The first set of items (Study 4) was milk items (e.g., soy milk and whole milk). The second set of items (Study 5) was cheese items.

Study 4

Participants

One hundred and twenty-five participants were recruited from Amazon's Mechanical Turk.

Materials

The materials were a set of images from commercially available milk products. We selected 4 images of animal-based milk products and 4 images of plant-based milk products. There was one between-subjects condition. In one condition, participants received 4 animal-based and 2 plant-based images. In the other condition, participants received 2 animal-based and 4 plant-based images. The plant-based images included almond milk, coconut milk, rice milk, and soy milk. The animal-based images included 1% milk, 2% milk, skim milk, and whole milk. An example image for each animal-based and plant-based products is included in Appendix C (all images are available from the authors upon request). All six images were presented at once on the screen. Participants were instructed to select the items that were made with real cow's milk by clicking on the image. After completing the product identification task, participants completed the MLS and basic demographic information was gathered.

Results and Discussion

The different number of images did not reliably influence accuracy of plant-based images F (1, 124) = 0.8, $p = .78$, $\eta^2 < .001$ or animal-based images F (1, 124) = 1.82, $p = .18$, $\eta^2 = .02$. Because there was no reliable difference with respect to the number of images used, we did not include the number of images as a factor in subsequent analyses. Correct scores for the product identification task were combined for each of the plant-based and animal-based products for analyses.

We analyzed whether participants were reliably different from chance at identifying products (chance = 0.5). Participants were substantially better than chance at identifying animal-based products (77%, $t(1,124) = 8.83$, $p < .001$, $d = 0.79$) and plant-based products (94%, $t(1, 124) = 30.74$, $p < .001$, $d = 2.75$.

Participants were reliably better at identifying plant-based based items compared to animal-based items, $t(1, 124) = 6.76$, $p < .001$, $d = .61$. Participants were also reliably better on the Soy subscale of the MLS ($M = 3.54$, $SD = 1.73$) than they were on the Milk subscale of the MLS ($M = 2.05$, $SD = 1.33$), $t(125) = 8.32$, $p < .001$, $d = 0.74$.

The correlations among the variables are reported in Table 4. The Soy subscale of the MLS was a significant predictor of product identification accuracy, suggesting criterion validity for the Soy subscale.

Study 5

Participants

One hundred and twenty-five participants were recruited from Amazon's Mechanical Turk.

Materials

The procedure used in Study 4 was used in Study 5. Participants were presented with either 4 or 2 animal-based cheese items along with 2 or 4 plant-based cheese items at one time on a screen. The plant-based images included vegan cheddar cheese, vegan cream cheese, vegan nacho sauce, and vegan cheese slices. The animal-based images included cheddar cheese, cheese dip, cream cheese, and swiss cheese. An example item is included in Appendix C (all images available upon request). Participants were asked to identify which of the 6 images were made from "real cow's milk" by clicking on the image of the product. After completing the product identification task, participants answered the 12-item MLS and basic demographic information was collected.

Results and Discussion

The number of images did not reliably influence accuracy for animal-based items (t (1, 123) = .01, $p = .99$, $d = .002$) or plant-based (t (1, 123) = 0.53, $p = .6$, $d = 0.1$) products. Number of images was therefore excluded as a factor in subsequent analyses. A total correct answer score was calculated for each of the plant-based and animal-based products.

Participants were reliably better than chance (= 0.5) at identifying plant-based cheese items (90% accurate, t (1, 124) = 22.87, $p < .001$, $d = 2.05$) and animal-based cheese items (64% accurate, $t(1, 124) = 5.43$, $p < .001$, $d = .49$). Participants were reliably better at identifying plant-based compared to animal-based cheese items, t (1, 124) = 8.08, $p < .001$, $d = 0.72$. Participants were also better on the Soy subscale of the MLS ($M = 2.95$, $SD = 1.33$) than they were on the Milk subscales of the MLS ($M = 1.82$, $SD = 1.36$), $t(125) = 6.89$, $p < .001$, $d = 0.62$.

Correlations were calculated (see Table 4). In this case, there were no reliable predictors of performance on the cheese product identification task.

Nutrition Identification

A separate concern about consumer product literacy is whether using 'milk' terms for both animal and plant-based products causes nutritional confusion. Studies 6 and 7 were designed to test the extent to which people are confused about the nutritional content of plant-based and animal-based milk items. If using 'milk' terms for both kinds of items causes confusion, then there should be substantial errors when people compare the nutritional content of plant-based and animal-based milk products.

Study 6

Study 6 was designed to see how well participants could identify simple nutritional information comparing animal-based to plant-based milk items.

Participants

One hundred and twenty-five participants were recruited from Amazon's Mechanical Turk.

Materials

We selected two paradigmatic images representing plant-based and animal-based milk: almond milk and whole cow Milk (see Appendix C for images). These images were selected because they clearly display what kind of product they are to help minimize the chances product confusion. We then selected several nutrition questions that were easily identified on the label of the product. We did not present nutritional labels to participants because we were interested in native nutritional knowledge of the products. The nutritional questions and instructions were (correct answers in parentheses):

> Please answer the following questions about these two products. PLEASE DO NOT LOOK UP ANSWERS ONLINE. If you do not know the answer, please respond that you do not know.

1. Which product has more calories? (Milk)
2. Which product has more fat? (Milk)
3. Which product has more cholesterol? (Milk)

4. *Which product has more sodium? (Almond)*
5. *Which product has more protein? (Milk)*
6. Which product has more fiber? (Almond)
7. Which product has more sugars? (Milk)

We also used three environmental impact questions. These questions were used to estimate the extent to which people know about the relative contribution to environmental problems of each product. While these are not explicitly about nutrition, they are related to general health concerns that people might have (see for more information, see Ho, Maradiaga, Martin, Nguyen, and Trinh (2016)).

8. *Which product uses more water? (Almond)*
9. Which product generates more waste? (Milk)
10. Which product has a larger carbon footprint? (Milk).

For each question, participants were allowed to select one of the two images and were also allowed to select that they did not know. Correct answers were coded as 1. Incorrect answers and "I don't know" responses were coded as 0.

Participants then completed the MLS, the General Nutrition Scale used in Study 3, and basic demographic information was gathered. The General Nutrition scale was used in this experiment because participants were asked specifically about the nutritional content of plant-based and animal-based milk products. To further help establish validity of the MLS, we intended to estimate whether the MLS predicted accuracy on the Nutrition Identification task beyond knowledge estimated by the General Nutrition Scale.

Results and Discussion

IRT analysis indicated that 3 items had reverse discrimination (Items 4, 5, and 8). Those items were eliminated from analysis. Another IRT analysis was conducted on the remaining 7 items. All items had acceptable discrimination (> 0.43) and a range of difficulties (-1.74 to 0.1). A 2-parameter model had an acceptable fit to the data (all residuals on the margin had chi-squared < 3.5). So, a composite score of the 7-items were calculated. The resulting scale was roughly normal (M = 4.47, SD = 1.76, *skewness* = -.51, *kurtosis* = -0.24). On average, participants could answer 64% the questions correctly. Item-level correct answers were: Calories 71%, Fat 85%, Cholesterol 70%, Fiber 67%, Sugars 48%, Waste 64%, and Carbon Footprint 61%. Additionally, participants were reliably better at Soy subscale of the MLS (M = 3.3, SD =

1.84) than they were at the Milk subscale of the MLS (M = 1.94, SD = 1.47), $t(125) = 7.52$, $p <$ .001, d = .67.

We were also interested in predicting performance on the Milk Nutrition Identification task. To do so, we calculated correlations among the variables (see Tale 5). Again, the Soy subscale of the MLS was a reliable predictor of correct responses to the Nutrition Identification Task suggesting criterion validity. To determine the unique predictive ability of the Soy subscale of the MLS, we used a stepwise multiple regression with the total score on the Nutrition Identification Task as the outcome variable and MLS Soy, MLS Dairy, Nutrition Knowledge, Sex, Age, and Politics as predictor variables (see table 6). The Soy subscale of the MLS was the strongest predictor of correct responses to the Nutrition Identification Task.

Study 7

Study 7 was designed to estimate how well people could identify nutritional information about plant-based and dairy-based cheese items.

Participants

One hundred and thirty-four participants were recruited from Amazon's Mechanical Turk.

Materials

The same general approach that was used in Study 6 was used in Study 7 except that cheese images were used instead of milk images. We selected two paradigmatic images that represent animal-based and plant-based cheese items. One image depicted a Daiya plant-based cheese product and the other image depicted an animal-based cheese product (See appendix C). Participants answered the following questions about each pair of images:

> Please answer the following questions about these two products. PLEASE DO NOT LOOK UP ANSWERS ONLINE. If you do not know the answer, please respond that you do not know.

1. *Which product has more calories per slice?* (Daiya)
2. *Which product has more fat per slice?* (Daiya)
3. Which product has more cholesterol per slice? (Milk)
4. Which product has more sodium per slice? (Milk)
5. Which product has more protein per slice? (Milk)
6. Which product has more calcium per slice? (Milk)

7. Which product has more sugars per slice? (Milk)
8. Which product uses more water per slice? (Milk)
9. Which product generates more waste per slice? (Milk)
10. Which product has a larger carbon footprint per slice? (Milk).

Participants could select one of the two images or indicate that they did not know. Participants also complete the MLS, the Nutrition Knowledge scale, and basic demographic information was gathered.

Results and Discussion

The IRT analysis showed that two of the Cheese Nutrition Identification items had reverse discrimination (Items 1 and 2). Those items were eliminated from analyses. After excluding those items, a 2-parameter IRT model was an acceptable fit to the data (all residuals on the margins had chi-squared < 3.5). A total score for the remaining 8 items was calculated and used in analysis. On average, participants knew the correct answer for 55% of the statements (M = 4.42, SD = 2.09). Item level descriptive statistics were: Cholesterol 62%, Sodium, 49%, Protein 49%, Calcium 50%, Sugars 50%, Water 43%, Waste 64%, Carbon Footprint 60%. In addition, participants were reliably better at the Soy subscale of the MLS (M = 3.37, SD = 1.78) than they were at the Milk subscale of the MLS (M = 1.74, SD = 2.05), $t(134) = 6.92$, $p < .001$, d = 0.6.

Correlations among the variables are reported in Table 4. The Soy subscale of the MLS was a reliable predictor of performance on the Cheese Nutrition Identification task, suggesting criterion validity. We also performed a stepwise regressions using performance on the Cheese Nutrition Identification task as the outcome variable and using the MLS Soy, MLS Dairy, Nutrition Knowledge, Sex, Age, and Politics as predictor variables. The only significant predictor of performance on the Cheese Nutrition Identification task was the Soy subscale of the MLS, so we do not include the regression analysis here.

National Sample

The final in the planned series of studies was to replicate the findings of Studies 1-7 in a more nationally representative sample drawn from a different sampling service. MTurk data is generally reliable for many tasks, but there are some known issues with data collected from MTurk including non-naiveté and inattentiveness (M. Buhrmester et al., 2011; M. D. Buhrmester, Talaifar, & Gosling, 2018; Chandler, Mueller, & Paolacci, 2014; Thomas &

Clifford, 2017). To help alleviate worries associated with biases in MTurk samples, we collected a sample using Qualtrics (see Qualtrics.com for more information).

Study 8

Participants

One thousand one hundred and eighty participants were recruited from Qualtrics testing service. For analyses, 126 participants were excluded for straight-lining responses (see below) leaving 1054 participants.

Materials

We used all of the finalized instruments from Studies 1-7 with some slight modifications. For the product identification tasks, we used 8 milk images (4 plant-based and 4 animal-based) and 8 cheese images (4 plant-based and 4 animal-based). Participants were given each set of images on 2 separate screens and the images were presented in random order. The participants were given the following instructions: "Please drag the items made with cow's milk into the 'milk' box and the items not made with cow's milk into the 'non-milk' box." There were two boxes on screen labeled "cow's milk" or "non-cows' milk." Participants were required to drag the images to one of the two boxes. Number of correct responses was calculated for each of cow's milk and non-cow's milk. Participants completed the modified Nutrition Identification Task from studies 6 and 7 (i.e., eliminating the items that had reverse discrimination). The 4 identification tasks were counterbalanced for order.

Participants were given the final version of the MLS, the Nutrition Knowledge, BNT, and the TIPI. Participants were also given the Knowledge of Animals as Food scale (KAFS) (Feltz & Feltz, submitted). The KAFS is a 9-item measure of how much people know about animals used as food. The KAFS has been shown to be related to general food decisions and related to a reduction in consuming animal products. Consequently, it was hypothesized that the KAFS would predict accuracy in the product identification. BNT was included as a general measure of numeracy (Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012). Numeracy has been related to normatively correct decisions in a host of domains (Ghazal, Cokely, & Garcia-Retamero, 2014; Petrova et al., 2017). We hypothesized that the BNT would be positively related to correct responses in the identification tasks. Finally, basic demographic information was gathered.

Results and Discussion

A visual inspection of the descriptive statistics revealed some problematic aspects of the data. In particular, there were a large number of zeros for the total score for the Nutrition Knowledge Scale. Otherwise, the distribution of results for the Nutrition Knowledge Scale was normal. This pattern of results was unlike the results obtained in the instrument's original validation and unlike the pattern of results observed in previous studies we conducted. Further investigation of this deviant pattern revealed a number of "straight-lined" responses—many participants answered "Don't know" to all of the nutrition questions, even questions that were very easy based on previous research. Those who straight-lined responses to the Nutrition Knowledge Scale also appeared to straight-line responses on other instruments. This suggests that some participants were not attentive or rushed through the survey. Consequently, following established practice after identifying straight-lined response (Leiner, 2016), we eliminated those who answered every question of the Nutrition Knowledge Scale "I don't know" (N = 126).[4]

The ***MLS Scale***

A test of unidimensionality on the full MLS suggested that unidimensionality could be rejected: observed second eigenvalue = 1.97, average second eigenvalue in 200 Monte Carlo samples = 0.58, p = .005. Tests for unidimensionality were conducted on each of the MLS subscales: Soy observed second eigenvalue = .55, average eigenvalue of 200 Monte Carlo samples = .30, p = 0.005; Milk observed second eigenvalue = 0.68, average of second eigenvalue in 200 Monte Carlo samples = .57, p = .03. While the test for unidimensionality was significant for the two subscales, the second eigenvalues were substantially less the second eigenvalue observed for the full scale. Plus, with the increased power of the study, conventionally significant results are likely to be detected even if the second eigenvalues were small. So, we proceeded by assuming that the MLS consisted of two unidimensional subscales.

We first performed IRT analyses on the Soy subscale of the MLS. A one-parameter unconstrained model (AIC = 7656.49, BIC 7691.21) was a better fit to the data that a one-parameter constrained model (AIC = 7660.99, BIC 7690.75), p = .01. A 2-parameter model (AIC 7647.52, BIC = 7707.06) was a better fir to the data than a 1-parameter unconstrained model, p = .002 (item difficulty and discrimination are provided in Table 2). The 2-parameter model had

[4] We conducted analyses without excluding participants who straight-lined. As expected, including those participants did not change the results drastically, but they did mute effects making some of the effects more difficult to detect. This pattern is exactly what would be expected given straight-lined responses.

largely good fit to the data—the chi squared values for the residuals of the margins were largely in the acceptable range (< 3.5). Two items were involved with chi-squared values larger than 3.5—items 5 and 3 (= 4.36) suggesting that in this study, the model did not fit those items particularly well.

The same analyses were conducted on the Milk subscale of the MLS. A one-parameter unconstrained model (AIC = 6755.58, BIC 6790.3) was a better fit to the data that a one-parameter constrained model (AIC = 6777.58, BIC 6807.34), p < .001. A 2-parameter model (AIC 6639.88, BIC = 6699.4) than a 1-parameter unconstrained model, $p < .001$ (item difficulty and discrimination are provided in Table 2). The 2-parameter model had largely good fit to the data—the chi-squared values for the residuals of the margins were largely in the acceptable range (< 3.5). One item was involved in residuals greater than 2.5 (item 10). ICCs and the TIF are presented in Figures 3 and 4.

As observed in previous studies, participants performed better on the Soy subscale (M = 3.62, SD = 1.61) than the Milk subscale (M = 1.69, SD = 1.52), t (1, 1053) = 32.59, $p < .001$, d = 1.

Product Identification Tasks

The identification tasks were counterbalanced for order (i.e., each task occurred in only 1 of the 1, 2, 3, of 4th spot). The first step in the analysis was to test for orders effects. Each of the product identification tasks was entered as the dependent variable and the order of presentation was used as the independent variable. There were no order effects for the two milk product identification tasks: Plant-based milk $F(1, 1050) = 2.02$, $p = .1$, $\eta^2 = .006$, Animal-based milk $F(1, 1050) = 1.86$, $p = .17$, $\eta^2 = .005$. There were statistically significant order effects for the two cheese product identification tasks: plant-based cheese $F(1, 1050) = 2.86$, $p = .04$, $\eta^2 = .008$, animal-based cheese $F(1, 1050)) = 4.47$, $p = .004$, $\eta^2 = .01$. Even though responses to the two cheese based product identification tasks displayed a statistically significant order effect, the magnitude of the effect was small. So, because of the small effect sizes and for simplicity of analyses, we did not include order as a factor in subsequent analyses.

Participants were better than chance (= 2) at identifying all products: Animal-based milk ($M = 3.65$, $SD = 0.64$ $t(1, 1053) = 83.38$, $p < .001$, $d = 2.57$; Plant-based milk ($M = 3.24$, SD = 1.25) $t(1, 1053) = 32.36$, $p < .001$, $d = 0.99$; plant-based cheese ($M = 2.82$, $SD = 1.08$) $t(1, 1053)$ = 24.55, $p < .001$, $d = 0.76$; and animal-based cheese ($M = 3.31$, $SD = 0.85$) $t(1, 1053) = 49.78$, p

$< .001$, $d = 1.53$. In this study, participants were better at identifying animal-based milk products $t(1, 1053) = 9.85$, $p < .001$, $d = 0.3$ and animal based cheese products $t(1, 1053) = 12$, $p < .001$, $d = 0.37$.

Finally, we calculated correlations among the variables gathered. Correlations are reported in Table 7.

Nutrition Identification

No order effect was found for either the milk nutrition identification task ($F(3, 1050) = 0.83$, $p = .48$, $\eta^2 = .002$) or cheese Nutrition Identification Task ($F(3, 1050) = 0.05$, $p = .99$, $\eta^2 = 0$). People were generally better at identifying milk nutrition items ($M = 4.29$, $SD = 1.9$, 61% correct) compared to cheese nutrition items ($M = 3.61$, $SD = 2.21$, 45% correct), $t(1053) = 19.29$. $p < .001$, $d = .59$.

Item level statistics were calculated for each question. For the milk nutrition items, the following were the percent of correct responses: Calories 69%, Fat 22%, Cholesterol 40%, Fiber 64%, Sugars 44%, Waste 47%, and Carbon Footprint 47%. For the cheese questions, the following percent responded correctly: Cholesterol 62%, Sodium, 51%, Protein 37%, Calcium 46%, Sugars 44%, Water 30%, Waste 46%, Carbon Footprint 43%.

Correlations among the dependent variables are reported in Table 7.

Path Analyses

We used path analyses to estimate relations among key variables measured. The primary outcome variable of interest was the performance on the identification tasks. We randomly divided the data into 2 groups: a test set and a validation set. We formulated path models based on the correlations observed in the studies and then tested and refined the models on the test set. The modified path models were then tested again on the validation set. Path models for the test and validation sets for the product identification Tasks are in Figures 5 and 6. Path models for the nutrition identification tasks are in Figure 7. All but one of the models passed conventional fit criteria. The models had the following test statistics.

Animal-Based Milk Product Identification: The test set model passed all conventional fit criteria: $\chi^2(2) = 1.31$, $p = .52$, *RMSEA* = 0, *90% CI* = 0-.08, *pclose* = .81, *CFI* = 1, *TLI* = 1. The validation set model also passed all conventional fit criteria: $\chi^2(2) = 0.74$, $p = .69$, *RMSEA* = 0, *90% CI* = 0-0.6, *pclose* = .9, *CFI* = 1, *TLI* = 1. All indirect paths were significant ($p < .05$).

Plant-based Milk Product Identification: The test set model passed all conventional fit criteria: χ^2 (2) = 2.35, p = .31, *RMSEA* = .02, 90% CI 0-0.9, *pclose* = .66, *CFI* = 1, *TLI* = .99. The validation set model also passed all conventional fit criteria: χ^2 (2) = 0.76, p = .69, *RMSEA* = 0, 90% CI 0-0.6, *pclose* = .9, *CFI* = 1, *TLI* = 1. All indirect paths were significant ($p < .05$).

Animal-based Cheese Product Identification: The test set model passed all conventional fit criteria: χ^2 (2) = 0.71, *p* = .7, *RMSEA* = 0, *90% CI* = 0-.07, *pclose* = .9, *CFI* = 1, *TLI* = 1. The validation set model also passed all conventional fit criteria: χ^2 (2) = 1.11, *p* = .57, *RMSEA* = 0, *90% CI* = 0-.07, *pclose* = .85, *CFI* = 1, *TLI* = 1. All indirect paths were significant ($p < .05$).

Plant-based Cheese Product Identification: The test set model passed all conventional fit criteria: χ^2 (2) = 0.65, *p* = .72, *RMSEA* = 0, *90% CI* = 0-0.6, *pclose* = .91, *CFI* = 1, *TLI* = 1. The validation set model also passed all conventional fit criteria: χ^2 (2) = 0.75, *p* = .69, *RMSEA* = 0, *90% CI* = 0-.06, *pclose* = .9, *CFI* = 1, *TLI* = 1. All indirect paths were significant ($p < .05$).

Milk Nutrition Identification: The test set model passed all conventional fit criteria: χ^2 (2) = 0.8, *p* = .67, *RMSEA* = 0, *90% CI* = 0-.07, *pclose* = .88, *CFI* = 1, *TLI* = 1. The validation set model also passed all conventional fit criteria: χ^2 (2) = 1.49, *p* = .48, *RMSEA* = 0, *90% CI* = 0-.08, *pclose* = .8, *CFI* = 1, *TLI* = 1. All indirect paths were significant ($p < .05$).

Cheese Nutrition Identification: The test set model passed all conventional fit criteria: χ^2 (2) = 4.93, *p* = .08, *RMSEA* = .05, *90% CI* = 0-.12, *pclose* = .36, *CFI* = .99, *TLI* = .95. The validation set model, however, did not pass all conventional fit criteria: χ^2 (2) = 10.44, *p* = .005, *RMSEA* = .09, *90% CI* = .04-.14, *pclose* = .09, *CFI* = .97, *TLI* = .84. All indirect paths were significant ($p < .05$) except for the BNT→ Cheese Nutrition indirect path, *p* = .06.

General Discussion

The results of the eight studies suggested that people are often product literate enough to reliably distinguish plant-based products from animal-based products. People also generally understand nutritional differences among plant-based and animal-based products.

To help further illustrate people's general product literacy, there were some differences among the studies. Studies 4 and 5 suggested that people are generally better at identifying plant-based milk products and study 8 suggested that they were better at identifying animal-based milk products. To help address the conflicting results, we meta-analytically combined the percentage of correct responses and tested for differences using plant-based v. animal-based products as a moderator. Concerning milk products, there was no significant moderator effect between animal-based (84%) and plant-based (88%) accuracy, $z = 0.37$, $p = .71$. A similar pattern was found for cheese product identification accuracy. People we no better at identifying animal-based (81%) cheese products compared to plant-based (74%) cheese products, $z = 0.52$, $p = .6$.

The evidence suggests that participants in general have the ability to identify plant-based and animal-based 'milk' products. As the meta-analytic estimates suggest, people identify products correctly between 74% and 88% of the time. While this is not 100% accuracy, it is unreasonable to expect 100% accuracy. There are many reasons why one would make a mistake including simple performance errors (clicking on the wrong item) and inattentiveness. These sources of error do not reflect a deep, systematic ignorance. If an element of product literacy is that consumers are able to understand and articulate differences among products, then it appears that people are generally product literate enough to at least distinguish plant-based from animal-based milk products.

Concerning nutritional differences, participants' accuracy was worse than their performance on product identification. To illustrate, we again meta-analytically combined the nutrition accuracy for Studies 5, 6, and 8. We used the mean correct scores for each of the tasks and used them as a moderator whether the nutrition was being identified for cheese or milk products. There as no overall moderator effect between the cheese and milk tasks, $z = 0.96$, $p = .34$. Overall mean correct score for milk was 4.37 (62%), *95% CI* = 3.81 - 4.94 and for cheese was 3.98 (50%), *95% CI* = 3.42 - 4.55.Hence, people tended to be roughly as good at identifying nutritional differences among milk and cheese items.

One might think that the overall scores for the nutrition identification tasks supports the argument that using 'milk' terms for plant and animal-based products causes nutritional confusion. But such support should be tempered for at least 3 reasons. First, the way that we coded responses to the nutritional task was that only correct answers were scored as correct and all other answers were scored as being incorrect. We adopted this scoring strategy partially to follow previous research (e.g., Cowburn and Stockley (2005)) and partially to provide the strongest test of consumer product literacy (i.e., only counting as correct those who knew the correct response). However, one could argue that if one knows that one does not know, then that does not indicate confusion. Rather, that reflects that one honestly does not know and not that one believes something that is false. So, people may not be making a mistake when they respond that they do not know. If we include those who responded that they do not know as being correct, then the percentage of correct responses increases dramatically. In Study 6, 86 of 1000 (~8.6%) response were "I don't know." Similarly for Study 7, the total number of "I don't know" response was 291 of 1072 (~20%). If we included those "incorrect" responses in the "correct" response category, then people correctly responded to 72.6% of the milk nutrition questions and 75% of the cheese nutrition questions. Consequently, the results for the nutrition identification tasks should be taken as a lower bound of accuracy.

Second, participants were more knowledgeable about differences between plant-based and animal-based products than they were about differences between animal-based products (i.e., whole milk v. skim milk) as measured by the MLS. We meta-analytically combined the results the MLS from studies 3-8 and found that on average, people were more knowledgeable on the Soy subscale ($M = 3.47$, $SE = 0.15$, *95% CI* = 3.18 - 3.76) than they were for the Milk subscale ($M = 1.89$, $SE = 0.07$, *95% CI* = 1.76 - 2.02), $d = 0.79$ ($SE = 0.08$) $z = 9.67$, $p < .001$, *95% CI* = 0.62 - 0.95. This result is consistent with the general ignorance about the nutritional differences among animal-based milk products (Finnell & John, 2017). The results from the meta-analysis of the MLS and previous research weaken the claim that there is widespread confusion about the nutritional difference between plant-based and animal-based milk products. Or, by parity of reasoning, one should be concerned about the nutritional ignorance surrounding animal-based milk products. According to our studies, people knew about ¾ of a standard deviation more about plant-based compared to milk-based products. If this is right, then, if anything, having plant-based products labeled as milk will make people *more* knowledgeable about nutritional differences among milk products.

Third, our data indicate that people are not perfectly knowledgably about milk products. But that leaves open what the best interventions are for those who need help. The structural models provide some important clues about how to help people make better consumer decisions. In the broadest terms, those who knew more about milk and nutrition were better at the identification tasks. That means that there are some fairly clear interventions that would likely help people become more product literate. For example, the links between animal welfare knowledge, general nutrition knowledge, and milk specific knowledge suggest that simply educating people about the facts of milk would help people make better, more informed decisions about milk products. Indeed, given the evidence presented here, simply forbidding the use of some language will not rectify issues of knowledge concerning milk products since many people are ignorant about some facts concerning animal-based milk.

Additionally, the links with numeracy suggest that providing some simple visual aids (e.g., on packages or in supermarkets) could help people make better, more informed decisions. Those who are more numerate tend to make more normative correct choices in general (Cokely, Garcia-Retamero, Ghazal, Allan, & Feltz, in press). In this instance, those who were more numerate did a better job on the product and nutrition identification tasks but also tended to be more knowledgeable in general (e.g., the MLS, KAFS, Nutrition Knowledge). In related research, providing simple visual aides have been shown to help those who are less numerate make choices that are more like those who are highly numerate (Garcia-Retamero, Petrove, Feltz, & Cokely, in press). So, there are likely to be some simple visual aides that would help people to make more correct choices.

There are a number of limitations with the current series of studies. First, the choices concerning product identification were somewhat artificial. Participants were shown images of products and asked to make decisions about them. In real environments like grocery stores, a different pattern of results may have been observed. Moreover, there could be some other, even more subtle confusions that using 'milk' terms for both animal-based and plant-based products could cause (e.g., overestimating the nutritional of quality of plant-based 'milk' products). This kind of confusion is ultimately best addressed empirically. However, given the results of our studies, the more subtle confusions about plant-based milk will also likely be present in animal-based milk products. So, we are skeptical that making the argument any more nuanced will help support the central empirical claim of those who favor restricting the use of 'milk' terms only for animal products. Finally, it is important to estimate the effectiveness of educational interventions versus policy level prohibitions on consumer product literacy about milk products.

Depending on one’s perspective, these results do little to support the general claim that people are confused about animal-based and plant-based food products. Recall the main concern from the European Court’s decision along with the Dairy Pride Act is that using ‘milk’ terms for plant-based products would cause confusion. We see little evidence that either kind of confusion exists—or that kind of confusion does not exist in any greater degree than it would exist if there were only animal-based milk products labeled using the term ‘milk’.

Appendix A

Items used in Study 1. Correct answer in parentheses. Difficulty and discrimination, respectively, in brackets.

Soy Subscale

1. Whole cow milk has more cholesterol than fortified soy milk. (T) [-9.1, 0.17]
2. Whole cow milk has more protein than fortified soy milk. (F) [-0.23, -1.5]
3. Whole cow milk has more Vitamin C than fortified soy milk. (T) [0.31, 1.62]
4. Whole cow milk has more calories than fortified soy milk (T) [-3.4, 0.5]
5. Whole cow milk has more fat than fortified soy milk. (T) [-4.54, 0.47]
6. Whole cow milk has more fiber than fortified soy milk. (F) [0.44, -1.58]
7. Whole cow milk has more sodium than fortified soy milk. (T) [-0.37, 0.82]
8. Whole cow milk has more iron than fortified soy milk. (F) [-0.03, -2.4]
9. Whole cow milk has more saturated fat than fortified soy milk. (T) [-6.61, 0.3]
10. Whole cow milk has more calcium than fortified soy milk. (T) [-0.33, 1.66]
11. Whole cow milk has more carbohydrates than fortified soy milk. (T) [-0.92, 0.85]
12. Whole cow milk has more lactose than fortified soy milk. (T) [-4.86, 0.29]
13. Cow milk and fortified soy milk have all the same nutrients. (F) [1.63, -0.61]

Milk Subscale

14. Whole cow milk has more protein than skim cow milk. (F) [0.71, 1.48]
15. Whole cow milk has more fat than skim cow milk. (T) [12.85, -0.18]
16. Whole cow milk has more calories than skim cow milk. (T) [14.22, -0.18]
17. Whole cow milk has more calcium than skim cow milk. (F) [0.36, 2.62]
18. Whole cow milk has more Vitamin C than skim cow milk. (F) [-0.19, 2.36]
19. Whole cow milk has more sodium than skim cow milk. (F) [0.26, 1.59]
20. Whole cow milk has more fiber than skim cow milk. (F) [0.4, 2.11]
21. Whole cow milk has more cholesterol than skim cow milk. (F) [2.71, 0.81]
22. Whole cow milk has more iron than skim cow milk. (F) [0.39, 2.11]
23. Fortified soy milk is made with some cow milk. (F) -1.51, 0.62]

Appendix B

Items used in Studies 2. Items removed from Studies 3-8 in italics.

<u>Soy Subscale</u>

1. *Whole cow milk has more Vitamin C than fortified soy milk. (T)*
2. Whole cow milk has more calories than fortified soy milk (T)
3. Whole cow milk has more fat than fortified soy milk. (T)
4. *Whole cow milk has more sodium than fortified soy milk. (T)*
5. Whole cow milk has less saturated fat than fortified soy milk. (F)
6. Whole cow milk less more calcium than fortified soy milk. (F)
7. Whole cow milk has fewer carbohydrates than fortified soy milk. (F)
8. Whole cow milk has less lactose than fortified soy milk. (F)

<u>Milk Subscale</u>

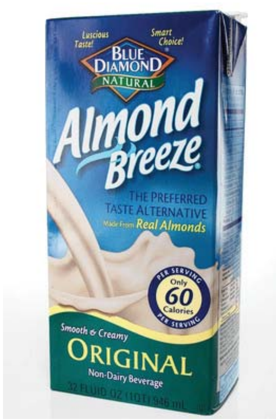
9. Whole cow milk has more protein than skim cow milk. (F)
10. Whole cow milk has more calcium than skim cow milk. (F)
11. *Whole cow milk has more Vitamin C than skim cow milk. (F)*
12. Whole cow milk has more sodium than skim cow milk. (F)
13. *Whole cow milk has more fiber than skim cow milk. (F)*
14. Whole cow milk has more cholesterol than skim cow milk. (F)
15. Whole cow milk has more iron than skim cow milk. (F)
16. Fortified soy milk is made with some cow milk. (F)

Appendix C

Sample Product Identification Milk Items

Sample Product Identification Cheese Items

Milk Nutrition Identification Items



Cheese Nutrition Identification Items


NEW!
daiya
deliciously dairy-free
AMERICAN STYLE
SLICES
NET WT 7.8 OZ (220g)
DAIRY FREE
SOY FREE
GOOD SOURCE OF CALCIUM
GLUTEN FREE


NO ARTIFICIAL PRESERVATIVES or FLAVORS
Kraft
Singles
AMERICAN
PASTEURIZED PREPARED CHEESE PRODUCT
24

Table 1: Demographics for Studies 1, 3-8

Study #		**1**	**3**	**4**	**5**	**6**	**7**	**8**
Age	*M*	36.18	38.41	35.7	37.04	35.54	35.76	45.64
	SD	11.02	12.61	12.4	12.53	11.22	11.49	17.85
Male		51.3%	51%	48.8%	57.6%	58.4%	46.3%	34.7%
Religion								
	Catholic	17.5%	20.3%	21.6%	17.6%	30.4%	27.6%	--
	Protestant	30.7%	31.9%	25.6%	28%	27.2%	26.9%	--
	Mormon	0.9%	0.9%	0.8%	2.4%	1.6%	1.5%	--
	Muslim	2.6%	0.4%	2.4%	2.4%	1.6%	1.5%	--
	Jewish	1.3%	1.3%	5.6%	1.6%	1.6%	1.5%	--
	Atheist	19.3%	18.1%	17.6%	20.8%	12%	19.4%	--
	Agnostic	19.3%	21.1%	16.8%	18.4%	15.2%	17.2%	--

	Preferred not to indicate	8.3%	6%	9.6%	8.8%	10.4%	4.5%	--
Education								
	Grammar school	0%	0%	0%	0.8%	0.8%	0%	3.3%
	High School	10.5%	9.1%	13.6%	3.2%	10.4%	7.5%	25.8%
	Vocational	5.3%	3.9%	2.4%	3.2%	4%	3.7%	11.1%
	Some College	25%	31.5%	19.2%	27.2%	19.2%	32.1%	23.3%
	Bachelor’s	38.6%	40.9%	50.4%	47.2%	45.6%	32.1%	23.6%
	Master’s	17.5%	11.2%	8.8%	12%	14.4%	17.2%	--
	PhD	0.9%	0.9%	3.2%	3.2%	0.8%	3.7%	2.1%
	Professional	2.2%	2.6%	2.4%	0%	4.8%	3.7%	10.7%
Ethnicity								
	Arab	0%	0%	0%	0%	0%	0.7%	--

	Asian/Pacific Islander	5.7%	4.7%	10.4%	7.2%	7.2%	9.7%	9.1%
	Black	8.3%	6%	4.8%	9.6%	5.6%	*8.2%*	17.3%
	Caucasian/White	76.3%	78%	72.8%	73.6%	77.6%	74.6%	65%
	Hispanic	5.3%	6%	3.2%	6.4%	5.6%	3.7%	--
	Indigenous	0%	0%	0.8%	0%	0.8%	0.7%	1.5%
	Latino	0.9%	0.4%	3.2%	0%	0%	0.7%	--
	Multiracial	2.6%	3.9%	4.8%	0.8%	2.4%	0.7%	--
	Would rather not say	0.9%	0.9%	0%	2.4%	0.8%	0.7%	7.1%
Marital Status								
	Divorced	5.7%	9.5%	8%	5.6%	5.6%	7.5%	--
	Cohabitation	14.5%	10.8%	12.8%	9.6%	15.2%	4.5%	--

	Married	43.9%	39.2%	38.4%	49.6%	38.4%	44%	--
	Separated	1.8%	1.3%	0.8%	32.8%	0.8%	0.7%	--
	Single	34.2%	36.2%	37.6%	0.8%	36.8%	40.3%	--
	Widowed	0%	1.3%	1.6%	0.8%	3.2%	3%	--
	preferred not to respond	0%	1.7%	0.8%	1.6%	0%	0%	--
Income								
	<$10,000	3.1%	6.5%	8%	5.6%	4.8%	3%	--
	$10,000-19,999	7%	9.5%	9.6%	3.2%	6.4%	8.2%	--
	$20,000-29,999	16.2%	11.2%	9.6%	13.6%	14.4%	12.7%	--
	$30,000-39,999	14%	11.6%	11.2%	7.2%	14.4%	11.2%	--
	$40,000-49,999	12.7%	12.1%	11.2%	15.2%	7.2%	12.7%	--

	$50,000-74,999	24.1%	22.8%	22.4%	29.6%	24%	16.4%	--
	$60,000-99,999	12.7%	13.4%	14.4%	12%	19.2%	15.7%	--
	$100,000-150,000	5.7%	7.3%	5.6%	9.6%	3.2%	11.2%	--
	> $150,000	3.5%	1.7%	6.4%	3.2%	5.6%	7.5%	--
	Preferred not to respond	0.9%	3.9%	1.6%	0.8%	0.8%	1.5%	--
Living area								
	Urban	33.8%	30.6%	36%	0.8%	36%	37.3%	--
	Suburban	46.9%	52.2%	47.2%	33.6%	48%	47%	--
	Rural	19.3%	17.2%	16%	45.6%	16%	15.7%	--
	Preferred not to respond	0%	0%	0.8%	20%	0%	0%	--

Table 2: Descriptive IRT Statistics for the KAFS in Studies 2, 3, and 8

Item	**% correct**			**Difficulty**			**Discrimination**		
Study #	**2**	**3**	**8**	**2**	**3**	**8**	**2**	**3**	**8**
Knowledge 1	25			4.89			0.23		
Knowledge 2	77	73	73	-1.01	-.81	-1.12	1.7	1.95	1.06
Knowledge 3	80	80	79	-0.96	-1.02	-1.12	3.0	2.43	1.75
Knowledge 4	36			1.95			0.29		
Knowledge 5	65	71	52	-0.74	-0.93	-0.08	1	1.22	1.36
Knowledge 6	52	52	52	-0.09	-0.08	-0.13	1.19	1.06	0.82
Knowledge 7	40	48	41	0.57	0.07	0.42	0.85	1.91	1.01
Knowledge 8	70	77	65	-1.02	-1.22	-0.71	0.95	1.26	1.09
Knowledge 9	27	26	26	1	0.81	0.83	1.28	2.14	2.17
Knowledge 10	35	31	26	0.5	0.58	0.8	2.28	2.79	2.49
Knowledge 11	34			0.61			1.65		
Knowledge 12	26	27	25	1.2	1.42	1.31	1.06	0.81	1
Knowledge 13	37			0.53			1.48		
Knowledge 14	12	11	16	2.64	2.75	2.03	0.86	0.85	0.94

Knowledge 15	32	29	25	0.54	0.69	0.88	4.39	2.39	2.84
Knowledge 16	65	71	51	-1.08	-2.51	-0.1	0.62	0.37	0.44

*Table 3: Correlations among dependent variables in Study 3. * p < .05, ** p < .01*

	1	2	3	4	5	6	7	8	9	10
1. Soy	1									
2. Milk	.23**	1								
3. Nutrition	.52**	.3**	1							
4. Extraversion	-.09	.04	-.03	1						
5. Agreeableness	.13*	.05	.16*	.19**	1					
6. Conscientiousness	.22**	.07	.3**	.12	.3*	1				
7. Emotional Stability	.16*	.03	.19**	.29**	.35**	.49**	1			
8. Openness to Experience	.1	.07	.19**	.34**	.35**	.26**	.23**	1		
9. Age	.06	.17**	.21**	.14*	.25**	.27**	.24**	.06	1	
10. Sex	.06	-.04	-.01	-.07	-.18**	-.06	.08	-.16**	-.05	1
11. Politics	.12	.07	-.07	-.03	-.1	-.13	.07	-.24**	.01	.19**

*Table 4: Correlations from Study 4 and 5. * p < .05, ** p < .01*

	Study #	1	2	3	4	5	6
1. Plant-based ID	4	1					
	5	1					
2. Animal-based ID	4	.52**	1				
	5	-.04	1				
3. Soy MLS	4	.24**	.22*	1			
	5	-.12	-.14	1			
4. Milk MLS	4	.08	.05	.16	1		
	5	.01	-.09	.08	1		
5. Age	4	.08	.27**	.1	.11	1	
	5	.04	.05	.13	.12	1	
6. Sex	4	-.12	-.14	-.17	.01	-.13	1
	5	.02	.11	-.2*	-.21*	-.06	1
7. Politics	4	.37**	-.28**	-.18*	-.05	-.22*	.14
	5	.03	.21*	-.34**	-.1	-.05	.25**

Table 4: Correlations from Study 6 and 7

	Study #	1	2	3	4	5	6	7
1. Product Nutrition ID	6	1						
	7	1						
2. MLS Soy	6	.5**	1					
	7	.24**	1					
3. MLS Milk	6	.26**	.28**	1				
	7	.16	.21*	1				
4. Nutrition	6	.45**	.58**	.33**	1			
	7	.18*	.47**	.33**	1			
5. Numeracy	6	-.02	-.09	.02	.19*	1		
	7	0	0	.13	.13	1		
6. Age	6	0	.08	.07	.27**	.02	1	
	7	0	.1	-.01	.16	-.06	1	
7. Sex	6	-.36**	-.24**	-.02	-.19*	.13	-.1	1
	7	.07	-.13	-.04	-.21*	.07	-.09	1
8. Politics	6	-.2*	-.17	0	-.17	-.04	.06	.11
	7	-.1	-.09	.07	-.25**	-.07	-.02	.14

*Table 5: Stepwise Regression from Study 6. ** p < .01, * p < .05.*

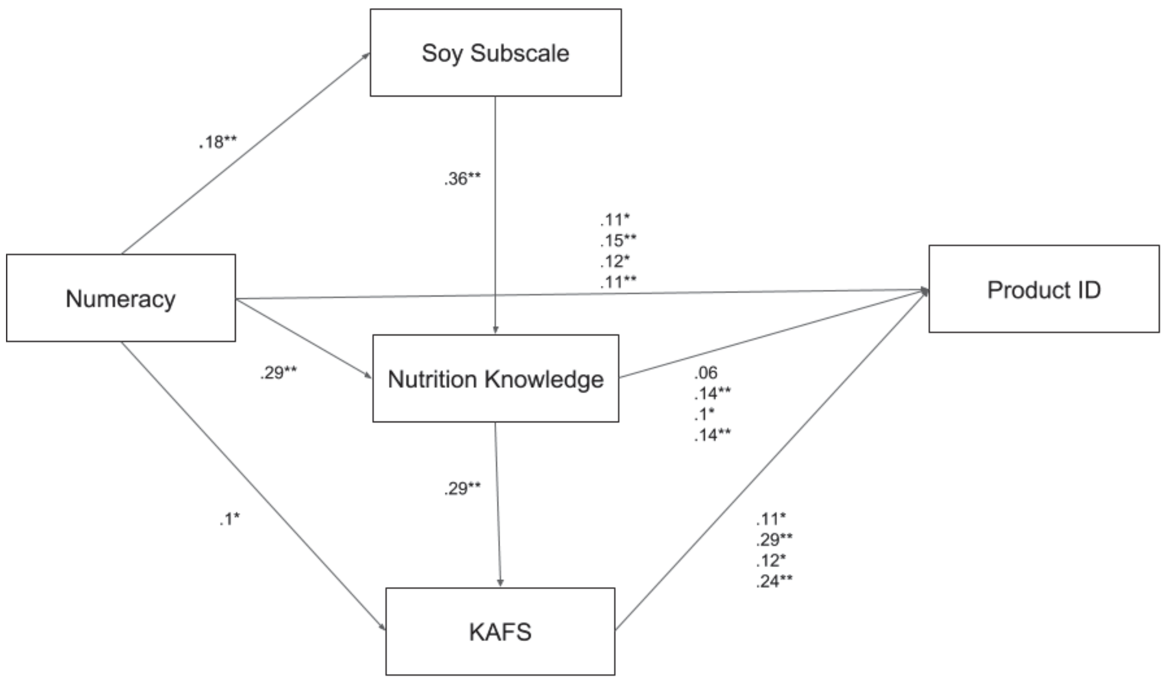
Model #	Predictor	β	Adjusted R^2	F	P	R^2_{Change}	F_{change}	P F_{change}
1	MLS Soy	.48**	.22	36.42	< .001	..22	36.42	< .001
2	MLS Soy	.31**	.27	24.17	< .001	.06	9.42	.003
	Nutrition	.29**						
3	MLS Soy	.28**	.29	18	< .001	.03	4.34	.04
	Nutrition	.27**						
	Sex	-.16*						

Table 6: Correlations from Study 8

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. Animal Milk ID	1																	
2. Plant Milk ID	.11**	1																
3. Animal Cheese ID	.28**	.18**	1															
4. Plant Cheese ID	.16**	.42**	.08**	1														
5. Milk Nutrition	.08*	.14**	.07*	.14**	1													
6. Cheese Nutrition	-.05	-.01	-.05	..51**	.5**	1												
7. MLS Soy	.09**	.16**	0	.1**	.39**	.33**	1											
8. MLS Dairy	0	-.05	.16**	0	.13**	.13**	.24**	1										
9. Nutrition	.16**	.3**	.16**	.22**	.27**	.21**	.4**	.29**	1									
10. KAFS	.17**	.34**	.19**	.27**	.18**	.04	.17**	-.02	.35**	1								
11. Extraversion	0	-.02	-.06	-.08**	.03	.06	.07*	.05	.05	-.04	1							
12. Agreeableness	.08*	.15**	.03	.13**	.07	.01	.11**	-.04	.15**	.18**	-.08*	1						
13. Conscientious	.06*	.18**	.04	.12**	.1**	.04	.11**	-.05	.18**	.19**	.11**	.32**	1					
14. Emotional	.03	.05	0	.04	.06*	.04	.09**	.02	.11**	.05	.1**	.36**	.39**	1				
15. Openness	.06*	.13**	.05	.08**	.13**	.09**	.09**	-.01	.12**	.16**	.16**	.32**	.28**	.2**	1			
16. Age	.06*	.21**	.07*	.15**	-.03	-.08**	.12**	-.03	.25**	.18**	.04	.26**	.28**	.28**	.01	1		
17. Gender	-.1**	-.1**	-.06*	.06	-.04	.05	-.05	.02	-.13**	-.14**	-.02	-.09*	-.03	.05	-.07*	.02	1	
18. Politics	.05	.05	.01	.05	.04	-.01	.01	-.08*	-.03	-.06*	-.01	0	.07*	.01	-.1**	.12**	-.01	1

19. Numeracy	.17**	.23**	.19**	.19**	.07*	-.04	.13**	.05	.31**	.22**	-.05	.1**	.15**	.07*	.06*	.1**	.13**	.01

*Figures 1 and 2. Path Model of Product Identification in the Test Sample (first figure) and validation sample (second figure). The paths with one path coefficient were identical in all models because the same sample was used. The paths predicting the product identification were in the following order from first to last path coefficient: Animal-based Milk Product ID, Plant-based Milk Product ID, Animal-based Cheese Product ID, Plant-based Cheese Product ID. All significant paths were are marked * < . 05, ** < .01.*



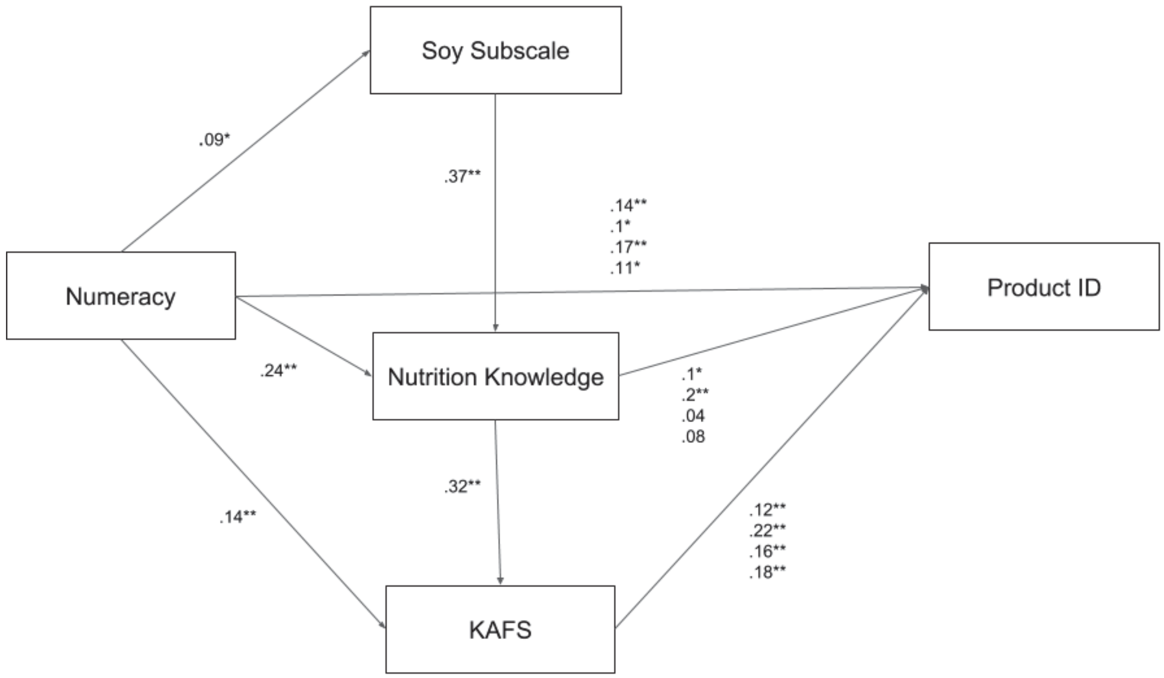
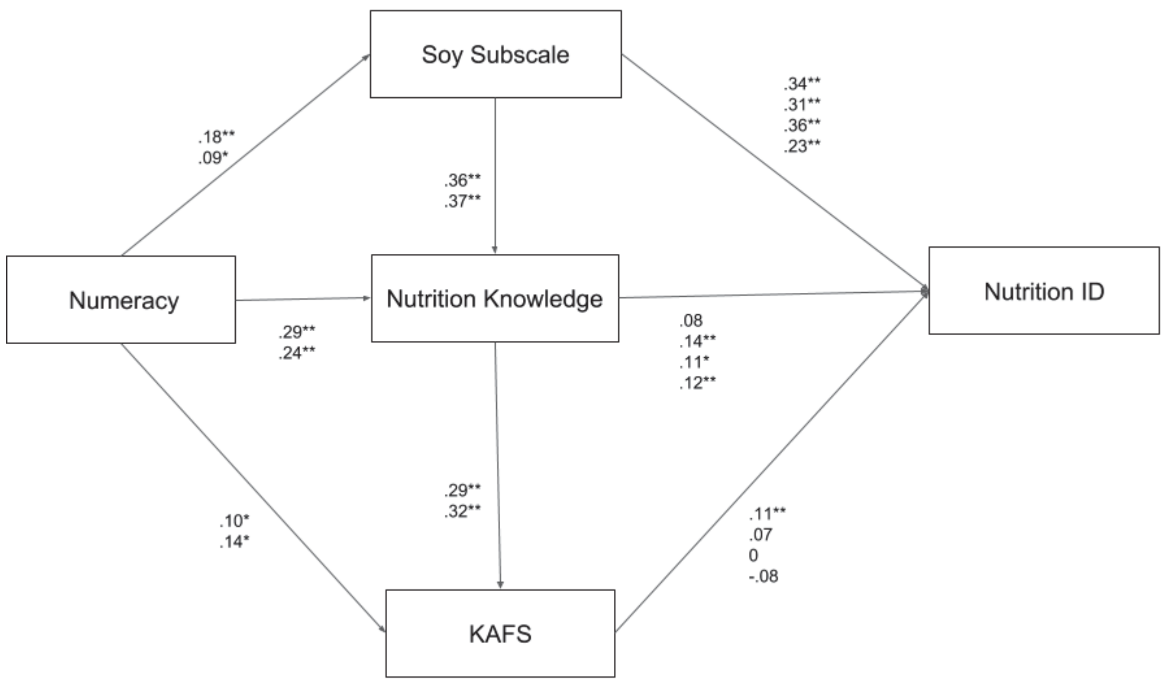

Soy Subscale
.09*
.37**
.14**
.1*
.17**
.11*
Numeracy
Product ID
.24**
Nutrition Knowledge
.1*
.2**
.04
.08
.32**
.14**
.12**
.22**
.16**
.18**
KAFS

Figure 3. *Path Model of the Nutrition Identification in the Test Sample and validation samples from Study 8. The paths with two path coefficient reflect the values from the test set and validation set (respectively). The paths with 4 path coefficients reflect the results in order: Milk nutrition identification test set, validation set, cheese nutrition identification test set, and validation set. All paths from the hypothesized models were included, significant paths were are marked * < . 05, ** < .01.*



References

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*(1), 3-5. doi:Doi 10.1177/1745691610393980

Buhrmester, M., Talifar, S., & Gosling, S. D. (2018). An evaluation of Amazon’s Mechanical Turk, its rapid rise, and its effective use. *Perspectives on Psychological Science, 13*(2), 149-154.

Buhrmester, M. D., Talaifar, S., & Gosling, S. D. (2018). An Evaluation of Amazon's Mechanical Turk, Its Rapid Rise, and Its Effective Use. *Perspectives on Psychological Science, 13*(2), 149-154. doi:10.1177/1745691617706516

Campos, S., Doxey, J., & Hammond, D. (2011). Nutrition labels on pre-packaged foods: a systematic review. *Public Health and Nutrition, 14*(8), 1496-1506. doi:10.1017/S1368980010003290

Chandler, J., Mueller, P., & Paolacci, G. (2014). Nonnaivete among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. *Behavior Research Methods, 46*(1), 112-130. doi:10.3758/s13428-013-0365-7

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring Risk Literacy: The Berlin Numeracy Test. *Judgment and Decision Making, 7*(1), 25-47.

Cokely, E. T., Garcia-Retamero, R., Ghazal, S., Allan, J., & Feltz, A. (in press). General decision making skill. In *Cambridge Handbook of Expertise*.

Cowburn, G., & Stockley, L. (2005). Consumer understanding and use of nutrition labelling: A systematic review. *Public Health and Nutrition, 8*, 21-28.

Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *Plos One, 8*(3). doi:ARTN e57410
DOI 10.1371/journal.pone.0057410

Dairy Pride Act. (2017).

Dickson-Spillmann, M., Siegrist, M., & Keller, C. (2011). Development and validation of a short, consumer-oriented nutrition knowledge questionnaire. *Appetite, 56*(3), 617-620. doi:10.1016/j.appet.2011.01.034

Finnell, K. J., & John, R. (2017). Research to Understand Milk Consumption Behaviors in a Food-Insecure Low-Income SNAP Population in the US. *Beverages, 46*, 1-12.

Garcia-Retamero, R., Petrove, Feltz, A., & Cokely, E. (in press). Measuring graph literacy: A systematic review and meta-analysis. In *Encyclopedia of Health and Risk Message Design and Processing*. Cambridge: Cambridge University Press.

Ghazal, S., Cokely, E. T., & Garcia-Retamero, R. (2014). Predicting biases in very highly educated samples: Numeracy and metacognition. *Judgment and Decision Making, 9*(1), 15-34.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37*(6), 504-528. doi:Doi 10.1016/S0092-6566(03)00046-1

Hall, C., & Osses, F. (2013). A review to inform understanding of the use of food safety messages on food labels. *International Journal of Consumer Studies, 37*, 422-432.

Hess, R., Visschers, V. H., & Siegrist, M. (2012). The role of health-related, motivational and sociodemographic aspects in predicting food label use: a comprehensive study. *Public Health and Nutrition, 15*(3), 407-414. doi:10.1017/S136898001100156X

Ho, J., Maradiaga, I., Martin, J., Nguyen, H., & Trinh, L. (2016). Almond Milk vs. Cow Milk Life Cycle Assessment. Retrieved from https://www.ioes.ucla.edu/wp-content/uploads/cow-vs-almond-milk-1.pdf

Kopp, S. (2012). Defining and conceptualizing product literacy. *Journal of Consumer Affairs, 46*, 190-203.

Leiner, D. (2016). Too fast, too straight, too weird: Post-hoc identification of meaningless data in internet surveys. Retrieved from https://www.researchgate.net/profile/Dominik_Leiner/publication/258997762_Too_Fast_Too_Straight_Too_Weird_Post_Hoc_Identification_of_Meaningless_Data_in_Internet_Surveys/links/59e4596baca2724cbfe85921/Too-Fast-Too-Straight-Too-Weird-Post-Hoc-Identification-of-Meaningless-Data-in-Internet-Surveys.pdf

Lunn, T. E., Nowson, C. A., Worsley, A., & Torres, S. J. (2014). Does personality affect dietary intake? *Nutrition, 30*(4), 403-409. doi:10.1016/j.nut.2013.08.012

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods, 44*(1), 1-23. doi:Doi 10.3758/S13428-011-0124-6

Paollacci, G., Chandler, J., & Ipeirotis, P. (2010). Running experiements using Amazon Mechanial Turk. *Judgment and Decision Making, 5*, 411-419.

Petrova, D., Garcia-Retamero, R., Catena, A., Cokely, E., Carrasco, A. H., Moreno, A. A., & Hernandez, J. A. R. (2017). Numeracy Predicts Risk of Pre-Hospital Decision Delay: a Retrospective Study of Acute Coronary Syndrome Survival. *Annals of Behavioral Medicine, 51*(2), 292-306. doi:10.1007/s12160-016-9853-1

Rizopoulos, D. (2006). ltm: An R package for latent variable modeling and item response theory analyses. *Journal of Statistical Software, 17*(5).

Rouse, S. (2015). A reliability analysis of Mechanical Turk data. *Computers in Human Behavior, 43*(1), 304-307.

Team, R. C. (2018). R: A language and environment for statistical computing. In *R Foundation for Statistical Computing*. Vienna, Australia.

Thomas, K. A., & Clifford, S. (2017). Validity and Mechanical Turk: An assessment of exclusion methods and interactive experiments. *Computers in Human Behavior, 77*, 184-197. doi:10.1016/j.chb.2017.08.038

Vanga, S. K., & Raghavan, V. (2018). How well do plant based alternatives fare nutritionally compared to cow's milk? *J Food Sci Technol, 55*(1), 10-20. doi:10.1007/s13197-017-2915-y